



# Towards a quantitative bond valence description of coordination spheres – the concepts of valence entropy and valence diversity coordination numbers

Piotr A. Guńka\* and Janusz Zachara\*

Received 13 August 2018

Accepted 18 December 2018

Edited by J. Lipkowski, Polish Academy of Sciences, Poland

This paper is dedicated to the memory of Prof. Andrzej W. Górski (1920–2017).

**Keywords:** coordination number; entropy; diversity; bond valence.

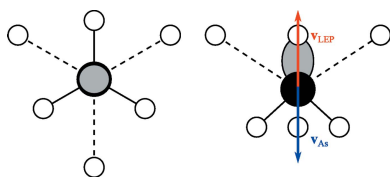
**Supporting information:** this article has supporting information at journals.iucr.org/b

Faculty of Chemistry, Warsaw University of Technology, Noakowskiego 3, Warszawa, 00-664, Poland. \*Correspondence e-mail: piogun@ch.pw.edu.pl, janzac@ch.pw.edu.pl

Two novel definitions of chemical coordination numbers – valence entropy coordination number  ${}^{\circ}\text{VECN}$  and valence diversity coordination number  ${}^{\circ}\text{VDCN}$  – are proposed. Their originality stems from the fact that they are the first definitions based solely on bond valences. The expressions for them are derived from their definitions and their properties are studied. The unexpected close relationship of  ${}^{\circ}\text{VECN}$  to Shannon entropy and  ${}^{\circ}\text{VDCN}$  to diversity are revealed and the names of the new coordination numbers are taken therefrom. Finally, as an example, a study of arsenic(III) lone electron pair stereoactivity with respect to  $\text{As}^{\text{III}}$  coordination number is carried out to demonstrate the usefulness and advantages of the new definitions as well as to compare them with the existing ones.

## 1. Introduction

Coordination number (CN) is one of the basic concepts used in chemistry and like other fundamental chemical ideas such as molecule, chemical bond or atomic core it lacks a precise and straightforward definition. This was pointed out quite some time ago by Hoppe (1970) in his engrossing paper where he dubbed CN an ‘inorganic chameleon’. Defining the basic ideas remains not only problematic within the scope of chemistry but pertains also to crystallography as exemplified by a rather fuzzy definition of a crystal being any solid that has essentially discrete diffraction pattern (IUCr, 1992). According to *IUPAC Compendium of Chemical Terminology* (Gold Book) the CN of a specified atom in a chemical species is the number of atoms directly linked to that specified atom or, in an inorganic coordination entity, CN is the number of  $\sigma$ -bonds between ligands and the central atom (IUPAC, 1997). It is explicitly said in the IUPAC definition that ‘the term is used in a different sense in the crystallographic description of crystals’. And it is in the extended solids that problems in defining CN start to crop up. Intuitively, it is often accepted that CN equals the number of nearest neighbours in a crystal structure (Giacovazzo *et al.*, 2011). For instance, every Xe atom in solid Xe has 12 closest neighbours and consequently its CN is 12; every  $\text{Na}^+$  cation in sodium chloride is surrounded by six closest  $\text{Cl}^-$  anions which gives  $\text{Na}^+$  a CN of 6 in this compound. Finding the closest neighbours seems to be fairly simple, but it is only such in high-symmetry structures. As the symmetry is lowered the distances between coordination centres and ligands become varied and the distinction between the closest and not closest but very close ligands gets unclear. The CN is often expressed in such cases as a sum of two



natural numbers *e.g.* in body-centred cubic (b.c.c.) structures CN is written as 8+6. The International Union of Crystallography Commission on Crystallographic Nomenclature Subcommittee on the Nomenclature of Inorganic Structure Types proposed another definition: CN is equal to the number of coordinating atoms and their choice depends on the bonding model used and the type of calculations (Lima-de-Faria *et al.*, 1990), which is not a very precise definition either.

Even this short introduction shows that defining CN is not as straightforward as it may seem at a first glance. Therefore, we propose new approaches to the issue based on bond valence model and define two original CN concepts aiming at easing some of the above mentioned difficulties. The paper is organized as follows. Section 2 gives historical background with a brief description of some CN concepts presented in the literature so far which in our opinion are the most significant ones. Our CN definitions are given in Section 3 together with an inspection of their properties, their relationship to the notions of diversity and entropy and with a comparison to Hoppe's (1979) effective coordination number. We proceed in Section 4 to apply the proposed definitions to the analysis of arsenic(III) coordination sphere in oxocompounds.

## 2. Historical background

The presented list does not pretend to be exhaustive and only the most prominent ideas in the authors' view are described. The ideas are introduced in the chronological sequence but where necessary there are some departures from it in favour of the logical order. Generally, CN definitions can be grouped into two classes – geometrical and chemical (Allig & Trömel, 1992). Geometrical CNs are based solely on geometrical parameters such as distances between atoms, coordination polyhedra and solid angles subtended by the faces of these polyhedra. Chemical definitions, on the other hand, also take into account some chemical knowledge as to which interactions are significant. The latter concepts are much less popular in the literature and they usually lead to lower CNs than the geometrical ones.

The first definition that we mention presented by Frank & Kasper (1958) is a geometrical one and relies on the ideas of atomic domains and neighbours. Atomic domain of a central atom is the space in which all points are nearer to the centre of that atom than to any other atom. The concept of atomic domain is equivalent to the idea of Voronoi polyhedron. Neighbour is an atom whose atomic domain shares a face with the atomic domain of the atom in question and, finally, CN is the number of neighbours of the atom. Another approach to define the neighbours of the central atom was proposed by Brunner & Schwarzenbach (1971) and it involves the analysis of histograms in which the number of neighbours of the atom is plotted as a function of their distances. The CN is defined by a maximum 'gap' in the histogram which is the largest difference in distances between consecutive shells of neighbours. This approach has its drawbacks and gives counterintuitive results even for very simple nets, *e.g.* for a planar square net it gives a CN of 20. This was noticed by Brunner (1977) and he

later proposed to examine the histogram with the number of neighbours of the central atom plotted as a function of reciprocal distances. Similarly, as in the previous definition the largest 'gap' in the histogram is used to evaluate CN.

The Frank–Kasper CN suffers from some problems and one of them was pointed out by O'Keeffe (1979). Namely, the atomic domain for an atom in b.c.c. structure is the Archimedean truncated octahedron and, consequently, the Frank–Kasper CN is 14, even though there are only eight closest neighbours of any atom. Furthermore, the Frank–Kasper approach yields a CN of 16(!) for diamond. CNs higher than that of closest packing (12) do not seem to be reasonable. O'Keeffe proposed to weight the contribution of atomic domain faces to CN by the solid angle that they subtend at the centre of the polyhedron. Such an approach gives CNs of 10.16 and 4.54 for the b.c.c. and diamond structures, respectively. O'Keeffe's definition has not been widely applied in the literature, probably because calculating of solid angles is quite complicated, especially for low-symmetry structures. Interestingly, O'Keeffe (1979) was convinced that in order to be useful, the concept of CN must be based on solely geometrical principles rather than on weighting schemes relying on the ideas about bond strengths.

Another disadvantage of the Frank–Kasper definition is that CN is inherently greater than or equal to 4. In order to overcome this problem (but not only this one) Allig & Trömel (1992) proposed a chemical CN which essentially consists in reducing the geometrical CN by excluding from consideration the atoms that have the same sign of oxidation state (number) as the central atom. Carter (1978) proposed yet another definition of chemical CN:

$$\frac{1}{\text{CN}_w} = \sum_i^N \left( \frac{w_i A_i}{\sum_j^N w_j A_j} \right)^2, \quad (1)$$

where  $A_i$  is a measure of central atom interaction with its  $i$ th neighbour and  $w_i$  are finite weighting factors (indices  $i$  and  $j$  enumerate neighbours of the central atom). He proposed bond strengths, bond energies, force constants for the bond-stretching motion, overlap integrals or bond orders as the reasonable measures  $A_i$  but noted that they could also be based on some geometric parameters. Additionally, he put a few important constraints on CN which his definition fulfilled: (a) CN is dimensionless and  $\geq 1$  if any neighbours with non-zero  $A_i$  exist; (b) CN is a continuous function of the  $A_i$  but its slope not necessarily so; (c) if  $N$  interactions exist such that  $A_1 = A_2 = \dots = A_N$  for all neighbours with non-zero  $A_i$ , then  $\text{CN} = N$ ; (d) if some of the  $A_i$  are unequal then  $\text{CN} < N$ ; (e) if  $m$  of the  $A_i$  are equal and large and  $(N - m)$  are equal and small, then  $N > \text{CN} > m$  (Carter, 1978).

Last but not least, we mention Hoppe's definition of effective coordination number (ECoN) which we consider particularly important for structural chemists and crystallographers. It is based on the idea of fictive ionic radii (FIR) defined for every pair of atoms in a crystal structure:

$$\text{FIR}(h \rightarrow i)_j = d(h \rightarrow i)_j \frac{R(h)}{R(h) + R(i)}, \quad (2)$$

where  $d(h \rightarrow i)_j$  is the bond length between coordination centre  $h$  and ligand  $i$  in a  $j$ th given shell, and  $R(h)$  and  $R(i)$  are ionic radii of  $h$  and  $i$ . Interatomic contacts are sorted in an ascending order and grouped into shells indexed by subscript  $j$ . All contacts/bonds in a shell exhibit the same value of FIR.

FIR are averaged to obtain mean fictive ionic radius (MEFIR). The weighted arithmetic average is applied with the following weights:

$$n(h \rightarrow i)_j \exp \left[ 1 - \left( \frac{\text{FIR}(h \rightarrow i)_j}{\text{FIR}(h \rightarrow i)_1} \right)^6 \right], \quad (3)$$

where  $n(h \rightarrow i)_j$  is the number of distances between coordination centre  $h$  and ligands  $i$  in a  $j$ th shell.

The ECoN is calculated using this formula:

$$\text{ECoN}(h \rightarrow i) = \sum_{j=1}^{\infty} n(h \rightarrow i)_j \exp \left[ 1 - \left( \frac{\text{FIR}(h \rightarrow i)_j}{\text{MEFIR}(h \rightarrow i)} \right)^6 \right]. \quad (4)$$

The exponent was arbitrarily set to 6 so that ECoN would give intuitive CNs of 4, 6 and 12 for diamond,  $\alpha$ -Po and c.c.p. of spheres, respectively. There are also iterative variants of  $^n\text{MEFIR}$  and  $^n\text{ECoN}$  described very well in the original paper (Hoppe, 1979). They differ significantly from the non-iterative variant only if there are large variations in consecutive FIR. It is worth noting that in principle all of the described geometrical CN concepts may be used to study intermolecular interactions and ‘molecular’ CNs in solids. That is, they are not restricted to atoms as the central and ligating entities. This was also noticed and explicitly stated by Hoppe (1979) with reference to ECoN in his paper.  $^n\text{ECoN}$  has been used not only for the geometrical analysis of crystal structures but also to calculate charge distribution (CHARDI) in extended solids. This was introduced by Hoppe *et al.* (1989) and later developed by Nespolo and co-workers (Nespolo *et al.*, 1999, 2001; Eon & Nespolo, 2015; Nespolo, 2016). CHARDI may be used for the prediction of bond strengths along with other approaches, such as resonance bond number (Rutherford, 1998).

### 3. Valence diversity and valence entropy coordination numbers

#### 3.1. Definitions

The proposed definitions belong to the class of chemical ones and are based on the valences  $s_i$  of bonds/contacts between the coordination centre and ligands (Brown, 2009, 2016; Brown & Poeppelmeier, 2014). Let us define valence diversity coordination number of order  $n$  ( $^n\text{VDCN}$ ,  $k_{\text{VDCN},n}$ ) such that:

$$\frac{\sum_{i=1}^N s_i}{k_{\text{VDCN},n}} = \left( \frac{\sum_{i=1}^N s_i^n}{k_{\text{VDCN},n}^n} \right)^{1/n}, \quad (5)$$

where  $s_i$  is the bond valence of  $i$ th bond, the summation goes from 1 to  $N$  bonds taken into account and  $n$ , called the VDCN order, is a positive number (although in principle it could be any real number except for 0).

Such a choice of name and the restriction to positive orders will be justified in §3.2. The left-hand side of equation (5) brings to mind weighted arithmetic mean with all ‘weights’ equal to the inverse of  $^n\text{VDCN}$ . However, the weights do not sum to unity, clearly indicating the expression is not a weighted arithmetic mean. Similarly, the right-hand side of this equation looks like weighted power mean.  $^n\text{VDCN}$  is the inverse of such weight for which the two expressions resembling means are equal. It should also be mentioned that the left-hand side of equation (5) is the average ligand bond order mentioned by Carter (1978).

After a few algebraic transformations one gets the following formula for  $^n\text{VDCN}$ :

$$k_{\text{VDCN},n} = \left[ \sum_{i=1}^N \left( \frac{s_i}{S} \right)^n \right]^{1/(1-n)}, \quad (6)$$

where  $S = \sum_{i=1}^N s_i$  is the bond valence sum (BVS) for the coordination centre.

Equation (6) permits for the calculation of  $^0\text{VDCN}$ , but becomes indeterminate for  $n = 1$ . It can be shown though that its limit at  $n \rightarrow 1$  exists and equals:

$$k_{\text{VDCN},1} = \prod_{i=1}^N \left( \frac{s_i}{S} \right)^{-s_i/S} \quad (7)$$

or

$$\ln k_{\text{VDCN},1} = - \sum_{i=1}^N \left( \frac{s_i}{S} \right) \ln \left( \frac{s_i}{S} \right). \quad (8)$$

Now let us define the valence entropy coordination number of order  $n$  ( $^n\text{VECN}$ ,  $k_{\text{VECN},n}$ ) such that:

$$\left( \frac{\sum_{i=1}^N s_i^n}{k_{\text{VECN},n}^n} \right)^{\frac{1}{n}} = \lim_{\delta \rightarrow 0} \left( \frac{\sum_{i=1}^N s_i^{n+\delta}}{k_{\text{VECN},n}^{n+\delta}} \right)^{1/(n+\delta)}. \quad (9)$$

This yields the following formula for calculating  $^n\text{VECN}$  (see Appendix A for derivation):

$$\ln k_{\text{VECN},n} = - \sum_{i=1}^N \left( \frac{s_i^n}{S_n} \right) \ln \left( \frac{s_i^n}{S_n} \right), \quad (10)$$

where  $S_n = \sum_{i=1}^N s_i^n$ . For  $n \neq 1$   $S_n$  does not correspond to the BVS of coordination centre. Again,  $n$  can, in principle, be any real number. We suppose, however, that only  $^n\text{VECN}$  of positive orders will be of practical importance. Note that  $^1\text{VDCN}$  and  $^1\text{VECN}$  are identical.

It can be shown that the same expression for  $^n\text{VECN}$  results from a different definition (see Appendix B for derivation).  $^n\text{VECN}$  is such a number that:

$$\frac{\partial \left[ \left( \frac{\sum_{i=1}^N s_i^n}{k_{\text{VECN},n}} \right)^{1/n} \right]}{\partial n} = 0. \quad (11)$$

It means that  ${}^n\text{VECN}$  is such a ‘weight’ that minimizes the expression resembling power mean at an *a priori* given positive  $n$  or that maximizes the expression for a negative  $n$  (see Fig. S1 for an illustration of this).

### 3.2. Relationship with diversity and Shannon entropy

It is noteworthy that the expression for  ${}^n\text{VDCN}$  is identical to the expression for species diversity of index  $q$  ( ${}^qD$ ) used in ecology (Hill, 1973; Jost, 2006):

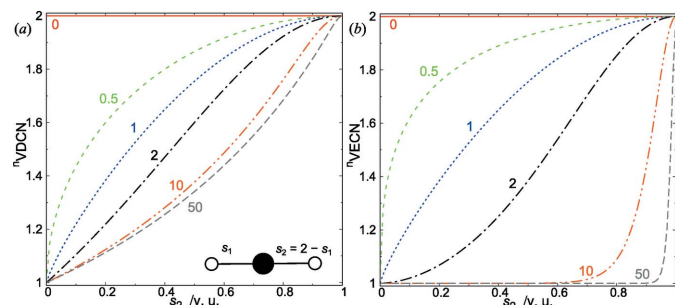
$${}^qD = \left( \sum_{i=1}^N p_i^q \right)^{\frac{1}{1-q}} = \exp(H_q), \quad (12)$$

where  $p_i$  is the abundance of species  $i$  expressed as the probability of finding an individual of species  $i$  in the whole ecosystem. This fact is the reason for calling the CN defined by equation (1)  ${}^n\text{VDCN}$ . In the case of  ${}^n\text{VDCN}$ ,  $p_i$  can be treated as the relative contribution of a certain bond to the total valence  $S$  of central atom. The diversity of order  $q$  is closely related to Rényi entropy of order  $q$  ( $H_q$ ) as indicated in equation (12).

It is noteworthy that diversity of index 1 equals to  $e$  raised to the power of Shannon entropy  $H_1$  (Jost, 2006):

$${}^1D = \exp\left(-\sum_{i=1}^N p_i \ln p_i\right) = \exp(H_1). \quad (13)$$

It follows that  ${}^n\text{VECN}$  of any order and  ${}^1\text{VDCN}$  are closely related to Shannon entropy expressed in natural units of information (nat). It is worth mentioning here that Lalik (2005) proposed to use entropy to define a measure of distortion for coordination polyhedra. However, he did not notice that it can be used as a CN or at least he did not mention this in his publication (Lalik, 2005). Information theory has been recently harnessed to the study of disordered solid materials and probably will become increasingly more important in the field (Varn & Crutchfield, 2016).



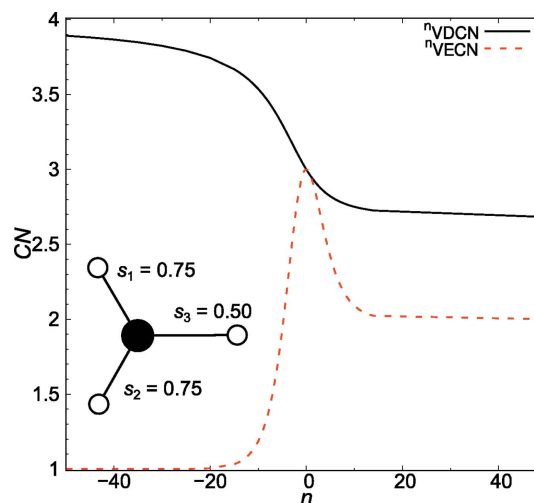
**Figure 1**  
(a)  ${}^n\text{VDCN}$  and (b)  ${}^n\text{VECN}$  as functions of various orders on bond valences in an  $AX_2$  system. Coordination centre of total valence 2 denoted with a black circle is surrounded by two ligands illustrated with open circles. The bond valences of the bonds between them are  $s_1$  and  $s_2$ . The CN orders are given next to the corresponding curves.

We want to stress, nonetheless, that the relationship of the herein defined  ${}^n\text{VDCN}$  and  ${}^n\text{VECN}$  with information theory in general as well as diversity and entropy in particular came as a surprise and is a consequence of the adopted definitions. This link has, however, profound consequences as one may envisage a whole new suite of CNs based on the diversity with the  $p_i$  probabilities expressed in terms of different quantities than bond valences which measure bond strengths. For instance, given that bond energies are roughly proportional to squares of bond valences (Brown, 2009)  ${}^2\text{VECN}$  can be treated as a first-order entropic CN based on bond energies.

### 3.3. Properties and discussion

The relationship of  ${}^n\text{VDCN}$  with diversity of index  $q$  permits for the straightforward translation of the properties of diversity to the properties of  ${}^n\text{VDCN}$ . Namely,  ${}^0\text{VDCN}$  is equal to  $N$  no matter what the values of  $p_i$  and, consequently,  $s_i$  are. The contribution of weak interactions *i.e.* bonds with low valences decreases with growing  $n$ . For  $n = 1$  the contributions of both weak and strong bonds to  ${}^1\text{VDCN}$  (and to  ${}^1\text{VECN}$  as well) are perfectly balanced (Jost, 2006). In the special case that all bonds have exactly the same bond valence, both  ${}^n\text{VDCN}$  and  ${}^n\text{VECN}$  are equal to  $N$  for all the values of  $n$ . All of these trends are illustrated in Fig. 1. Besides,  ${}^n\text{VECN}$  falls more quickly than  ${}^n\text{VDCN}$  with the growing disproportion between bond valences for CN orders higher than 1 as illustrated in Fig. 1. The contrary holds for orders between 0 and 1. Importantly,  ${}^n\text{VDCN}$  fulfils all the restrictions put forth by Carter (1978).

Moreover, one may say, paraphrasing Hill (1973), that  ${}^n\text{VDCN}$  is figuratively a measure of how many bonds there are in the coordination sphere of a particular coordination centre, if we examine the bonds sorted in the descending order according to their valences down to a certain depth. If we check superficially (*e.g.* by using  ${}^2\text{VDCN}$ ) we will see the



**Figure 2**  
 ${}^n\text{VDCN}$  and  ${}^n\text{VECN}$  as functions of the order  $n$  for an  $AX_3$  system. The  $A$  and  $X$  species are drawn as black and open circles, respectively.

stronger bonds. If we look deeply (*e.g.* by using  ${}^0\text{VDCN}$ ) we will count all the bonds. Therefore, the  ${}^n\text{VDCN}$ s of various orders have a natural intuitive interpretation, albeit a rather vague one (Hill, 1973). Negative  ${}^n\text{VDCN}$  orders were excluded because they may lead to greater values of CN than the number of ligands around coordination centre which is physically impossible (see Fig. 2).

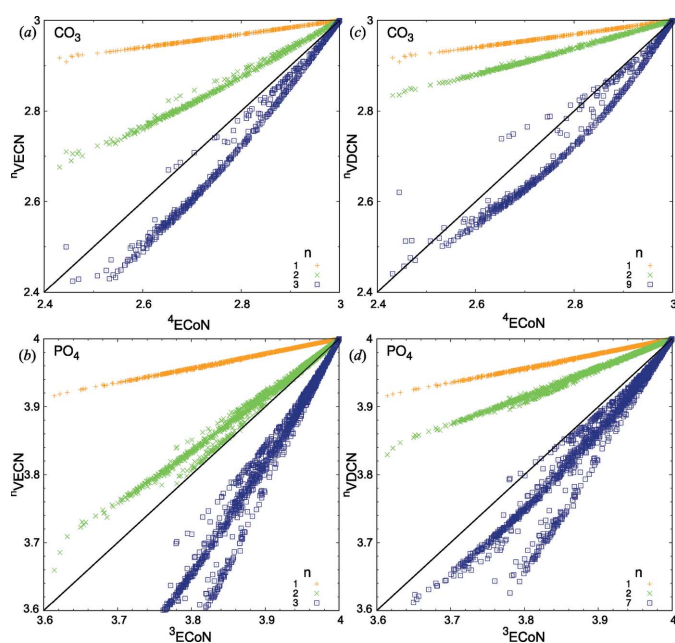
One may show that  ${}^n\text{VECN}$  converges for  $n \rightarrow +\infty$  and  $n \rightarrow -\infty$  to the number of bonds with highest and lowest bond valence in a given coordination sphere, respectively (see Fig. 2 for an illustration and Appendix C for a proof). This indicates that for  $n = 1$  there is a perfect balance between weak and strong bonds *i.e.* neither group is favoured, for  $n > 1$  strong bonds are favoured and for  $n < 1$  weak ones. It also follows that  $k_{\text{VECN},n} \leq N$  and the equality occurs only if the VECN order is 0 or if all bonds in the coordination sphere have the same bond valence (see Fig. 1). It must be stressed that  ${}^n\text{VECN}$ , similarly to  ${}^n\text{VDCN}$ , fulfils all the restrictions imposed on CNs by Carter (1978).

Last but not least, it can be proven for  $A_xB_y$  compounds crystallizing with one symmetry independent atom  $A$  and one symmetry independent atom  $B$  that (see Appendix D for proof):

$$x k_{\text{VECN},nA} = y k_{\text{VECN},nB}. \quad (14)$$

${}^n\text{VECN}$  and  ${}^n\text{ECoN}$  are to the best of our knowledge the only CNs which may adopt non-integer values and satisfy this condition.

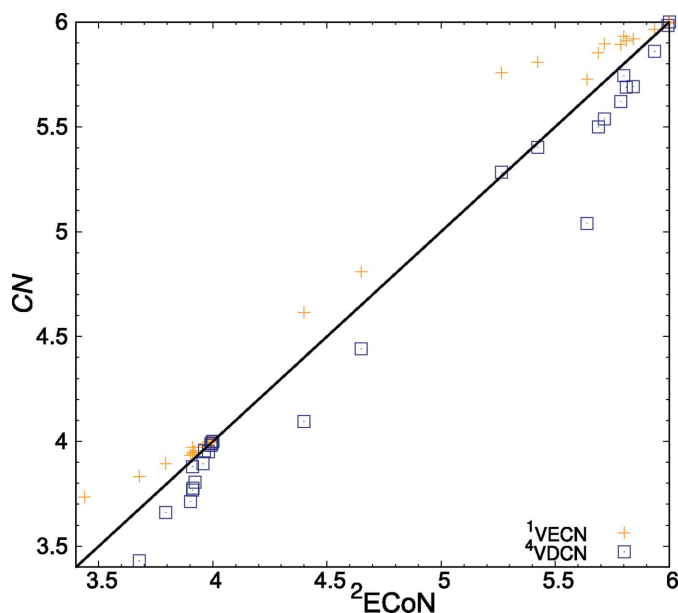
One may argue that the dependence of  ${}^n\text{VDCN}$  and  ${}^n\text{VECN}$  on  $n$  is their weakness. However, we are confident that the opposite is true. Different properties that depend on CN, such



**Figure 3**  
The correlation between iterative  ${}^n\text{ECoN}$  and  ${}^n\text{VECN}$  as well as  ${}^n\text{VDCN}$  of various orders for (a, c)  $\text{CO}_3$  and (b, d)  $\text{PO}_4$  moieties. The black lines denote the  $y = x$  function.

as the stereoactivity of lone electron pair or the chemical shift in NMR, may be affected more significantly either by weak or strong bonds. Consequently, the choice of appropriate CN order  $n$  may be critical in correlating and/or predicting the relationship of a particular property with CN.

Let us compare  ${}^n\text{VDCN}$  and  ${}^n\text{VECN}$  with  $\text{ECoN}$  before proceeding to an application of the newly defined CNs to an exemplary problem in structural chemistry. Organic carbonates containing  $\text{CO}_3$  moiety and organic phosphates(V) with a  $\text{PO}_4$  moiety were chosen as model systems with conventional CNs of 3 and 4, respectively. The CSD (version 5.39 with updates, August 2018) searches were restricted to the structures containing  $\text{CO}_3$  or  $\text{PO}_4$  skeleton, respectively. Only structures containing neither errors nor disorder and that with an  $R$  factor of less than 0.05 and with the mean  $\sigma(\text{C}-\text{C})$  smaller than  $0.005 \text{ \AA}$  were taken into account. The bond valence parameters of Gagné and Hawthorne (2015) were used. The BVS rule was checked and structures for which BVS deviated by more than 0.5 from the ideal values of 4 and 5 for carbon and phosphorus, respectively, were rejected. There were four such structures both for  $\text{CO}_3$  and for  $\text{PO}_4$  moieties. The resulting data sets comprise 569 (1047) entries with 688  $\text{CO}_3$  (1688  $\text{PO}_4$ ) fragments. Iterative  ${}^n\text{ECoN}$  was calculated for both data sets and it was found that it does not change significantly after four and three iterations for carbonates and phosphates(V), respectively. The difference indicates that C–O bonds are more varied in the  $\text{CO}_3$  fragments than P–O bonds in phosphates(V). Then, the iterative  ${}^n\text{ECoN}$  was correlated with  ${}^n\text{VECN}$  and  ${}^n\text{VDCN}$  of various orders (see Fig. 3). It was found that  $\text{ECoN}$  correlates linearly with  ${}^1\text{VECN}$  (and, of course, with  ${}^1\text{VDCN}$ ). However, the slopes of the linear relationship between them is much smaller than one (0.15 and 0.22 for  $\text{CO}_3$  and  $\text{PO}_4$  coordination spheres,



**Figure 4**  
The correlation between iterative  ${}^2\text{ECoN}$  and  ${}^1\text{VECN}$  as well as  ${}^4\text{VDCN}$  for  $\text{ML}_n$  moieties where  $M = \text{Al, Ga, In}$  and  $L = \text{O, S, Se}$ . The black line denotes the  $y = x$  function

respectively). The slope becomes close to one for  ${}^n\text{VECN}$  and  ${}^n\text{VDCN}$  orders significantly higher than 1 which indicates that ECoN disproportionately favours strong bonds over weaker ones. It is noteworthy that for  ${}^n\text{VECN}$  and  ${}^n\text{VDCN}$  orders higher than one the data lie on distinct lines which permits their separation into classes. The values of  ${}^n\text{VECN}$  and  ${}^n\text{VDCN}$  change slightly when different bond valence parameters are used (see Fig. S2 for an illustration).

Similar comparison of the herein proposed CNs with ECoN was conducted for inorganic compounds. Oxides, sulfides and selenides of aluminium, gallium and indium were chosen as model systems and structures of their polymorphs were retrieved from ICSD (Belsky *et al.*, 2002; Hellenbrandt, 2004). The structures were checked and the limit on coordination was decided manually for every coordination centre. Then  ${}^1\text{VECN}$ ,  ${}^4\text{VDCN}$  and  ${}^2\text{ECoN}$  were calculated. The results are given in Table S1. In this case similar trends are observed as for organic carbonates and phosphates, but the sample is much smaller (see Fig. 4). There are two distinct curves for  ${}^1\text{VECN}$  – one starting from point with coordinates (4, 4) and the other from point (6, 6) – and there are two additional points which would belong to the curve starting from (5, 5), were there more structures with five ligands around coordination centre in the database.

## 4. Application to the analysis of arsenic(III) coordination sphere in oxocompounds

### 4.1. Introduction and methodological details

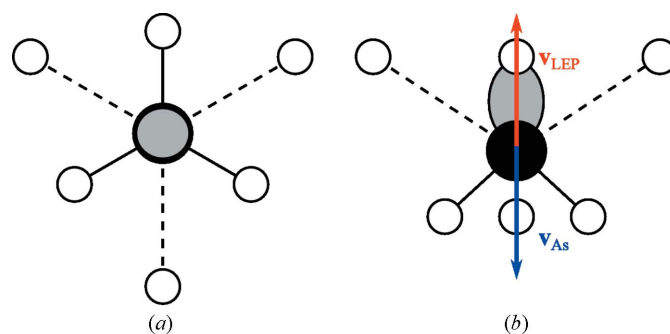
In order to prove the usefulness of the proposed CN definitions and to show one of their possible applications  ${}^1\text{VECN}$  has been applied to study the correlation between the stereoactivity of  $\text{As}^{\text{III}}$  lone electron pair (LEP) and arsenic CN. There are numerous studies of LEPs stereoactivity of  $p$ -block elements (Wang & Liebau, 1996; Sidey, 2008; Christy & Mills, 2013; Gagné & Hawthorne, 2018, 2018a) and Rahm has recently proposed a novel descriptor to quantify the nature of LEPs – HELP – high-ELF localization domain population (Rahm & Christe, 2013; Rahm, 2015). Arsenic(III) was chosen as the stereoactivity of its LEP has not been studied as extensively as for other elements, because of our long-term interest in this element and for the sake of simplicity. Namely,  $\text{As}^{\text{III}}$  is heavy enough for its coordination sphere to be extended to contain both primary and secondary bonds, but it is relatively easy to place a boundary as to which  $\text{As}\cdots\text{O}$  contacts should be taken into account and which can be neglected. In the following paragraphs we will use the Alcock's (1972) concept of secondary bonds which are weak interactions of coordination centre possessing LEP with ligands located *trans* with respect to the strong primary bonds of the coordination centre (see Fig. 5).

The study was restricted to inorganic compounds of arsenic(III) with only oxygen atoms as ligands. Structural data were collected from the ICSD by searching for all compounds containing  $\text{As}^{\text{III}}$  and oxygen ligands up to 3.6 Å from the arsenic (Belsky *et al.*, 2002; Hellenbrandt, 2004). *PLATON*

and bash scripts were used for the extraction of As–O bond lengths and bond angles from the cif files retrieved from the database (Spek, 2009). The data were then carefully analysed using *LibreOffice Calc* and structures with disorder, refined to  $R_1$  values higher than 10%, with incorrectly assigned arsenic oxidation state as +3 and with unusually short or long As–O bonds were removed from the data set. This led to a dataset containing 546 coordination spheres of arsenic determined under ambient conditions and 141 coordination spheres determined in high-pressure (up to 30 GPa) experiments of ours which were treated separately (Guńka, Dziubek *et al.*, 2015; Guńka, Dranka *et al.*, 2015). The exponential relationship between bond valence and bond length was utilized:

$$s_i = \exp\left(\frac{R_0 - r_i}{B}\right), \quad (15)$$

where  $r_i$  is bond length,  $R_0$  is the bond length of unit valence and  $B$  is the parameter describing the bond softness. The  $R_0$  and  $B$  parameters of 1.775 and 0.423 Å, respectively, calculated by Gagné & Hawthorne (2015) were used. For the high-pressure data the dependence of  $R_0$  on pressure proposed by Brown was utilized (Brown *et al.*, 2003). The analysis was restricted up to six oxygen ligands *i.e.* the coordination sphere of As was restricted to three primary As–O bonds and up to three secondary As $\cdots$ O bonds. The stereoactivity of As LEP was assessed by the magnitude of the resultant bond valence vector (BVV) of arsenic atom using the approach of Zachara (2007). For the computations arsenic coordination centre was treated as  $\text{As}^{5+}$  and LEP was treated as an additional pseudoligand with a bond valence of  $s_{\text{LEP}} = -2$ . The BVV associated with LEP was taken as the negative of the arsenic resultant BVV calculated as the sum of BVVs associated with real ligands [see Fig. 5(b)]. If the LEP were fully stereoactive, the magnitude of the BVV associated with it should be  $\|\mathbf{v}_{\text{LEP}}\| = s_{\text{LEP}}(1 - s_{\text{LEP}}/Q_{\text{As}}) = 2(1 - 2/5) = 1.2$  v.u. The decrease in the LEP stereoactivity causes lowering of this value. Again the  $\text{As}^{5+}$  bond valence parameters from Gagné & Hawthorne (2015) of 1.765 and 0.352 Å were used. The source code for the *LibreOffice* functions to calculate  ${}^n\text{VECN}$ ,

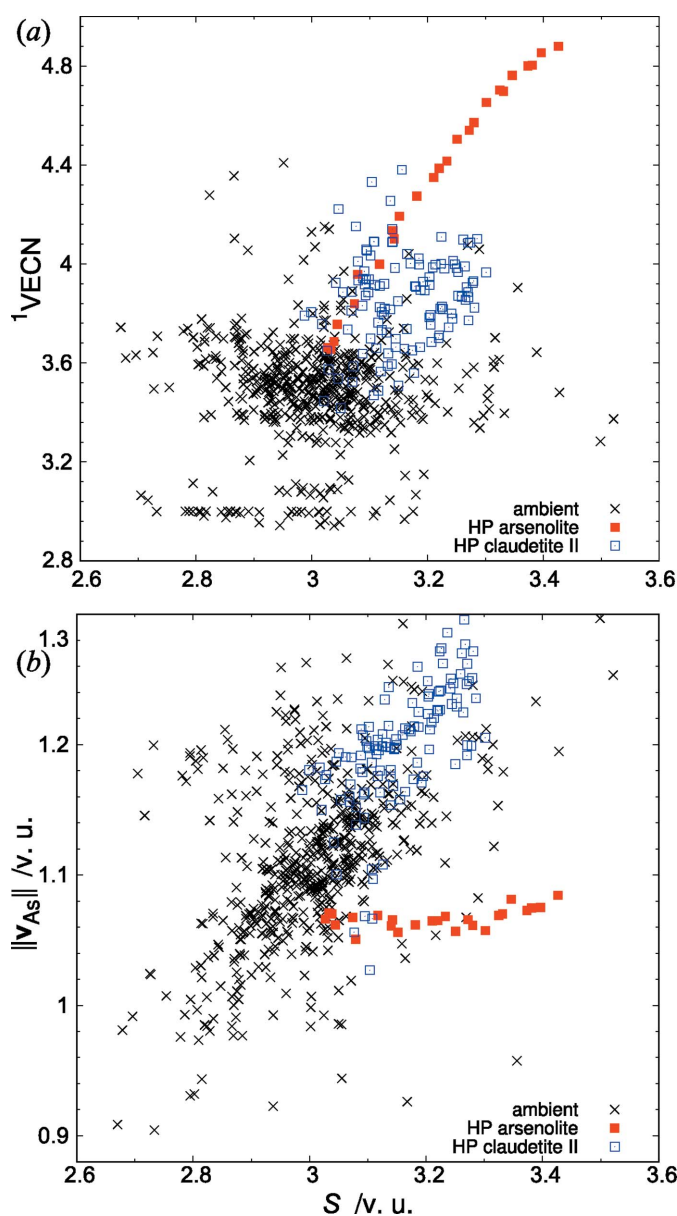


**Figure 5**  
(a) Top and (b) side view of the idealized  $\text{As}^{\text{III}}$  coordination sphere with secondary  $\text{As}\cdots\text{O}$  interactions located *trans* with respect to primary As–O bonds. As and O atoms are marked as black and open circles, respectively. LEP is represented as a grey circle and ellipse in (a) and (b), respectively. Additionally, the resultant BVV vector of arsenic and the BVV associated with LEP are drawn as blue and red arrows, respectively.

<sup>1</sup>VDCN and <sup>1</sup>ECoN are available in the supporting information and an instruction how to import them is detailed in §S1.

#### 4.2. Results and discussion for ambient-pressure structures

The histogram of bond valence sums (BVS, denoted *S*) for As<sup>III</sup> coordination spheres shows a normal distribution centred on 3 (see Fig. S3 for the histogram and for the normal probability plot). This indicates that the adopted cut-off for the As···O distances of 3.6 Å is correct. As expected there is no correlation between the BVS and <sup>1</sup>VECN and only a weak one between BVS and LEP stereoactivity described by the resulting bond valence vector denoted  $\|\mathbf{v}_{\text{As}}\|$  (see Fig. 6). The Pearson correlation coefficients *r* are 0.02 and 0.47, respec-



**Figure 6**  
(a) <sup>1</sup>VECN and (b) the resultant BVV magnitude  $\|\mathbf{v}_{\text{As}}\|$  plotted as a function of bond valence sum *S* for As<sup>III</sup> coordination spheres determined at ambient pressure and at high pressure (HP) up to 30 GPa.

tively. The correlation between <sup>1</sup>VECN and LEP stereoactivity is stronger (*r* = −0.57; see Fig. 7). Two clusters of data points can be discerned in the plot. The first one concentrated around <sup>1</sup>VECN of 3 and  $\|\mathbf{v}_{\text{As}}\|$  of 1.2 v.u. corresponding to AsO<sub>3</sub>  $\psi$ -tetrahedra in which arsenic atom does not form any significant secondary bonds. The second one, corresponding to coordination spheres comprising both primary and secondary bonds, contains many more data points, is larger and concentrated around <sup>1</sup>VECN of 3.5 and  $\|\mathbf{v}_{\text{As}}\|$  of 1.1 v.u.

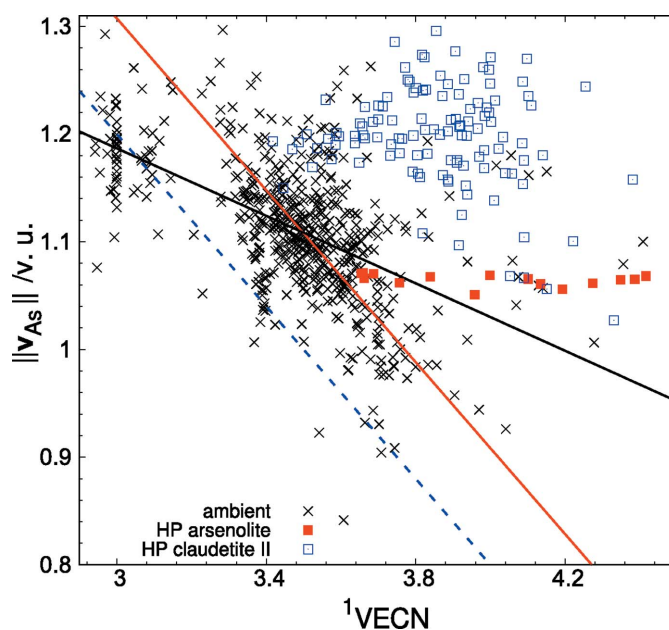
Interestingly, the correlation between <sup>1</sup>VECN and  $\|\mathbf{v}_{\text{As}}\|$  follows roughly a crude model derived as follows. Suppose there is an isolated AsO<sub>3</sub>  $\psi$ -tetrahedron. There are only three primary As–O bonds and LEP stereoactivity is maximal with the corresponding  $\|\mathbf{v}_{\text{As}}\|$  value of 1.2 v.u. The LEP stereoactivity would be minimal *i.e.*  $\|\mathbf{v}_{\text{As}}\|$  = 0 when As would form six As–O bonds of equal bond valences (see Fig. 5). <sup>1</sup>VECN is for the latter case six and for the former three. Assuming a linear relationship one gets the following dependence of  $\|\mathbf{v}_{\text{As}}\|$  on <sup>1</sup>VECN:

$$\|\mathbf{v}_{\text{As}}\| = -0.40k_{\text{VECN},1} + 2.40. \quad (16)$$

This function is represented by a dashed blue line in Fig. 7. It is slightly off the data points and the following equation describes the experimental points much better (solid red line in Fig. 7):

$$\|\mathbf{v}_{\text{As}}\| = -0.40k_{\text{VECN},1} + 2.50. \quad (17)$$

This indicates that LEP stereoactivity is enhanced for higher <sup>1</sup>VECNs. The origin of this effect remains unclear. For the sake of completeness, the line fitted to data points by least-squares method is drawn as solid black line in Fig. 7. It devi-



**Figure 7**  
The resultant BVV magnitude  $\|\mathbf{v}_{\text{As}}\|$  plotted as a function of <sup>1</sup>VECN. The dashed blue line and solid red line correspond to the crude models relating  $\|\mathbf{v}_{\text{As}}\|$  and <sup>1</sup>VECN described in the main text whereas the solid black line is the least-squares fit to the data points. Only the ambient pressure data were taken into account.





(see Fig. S4). VECN is not suitable for this as there are no As–As bonds in arsenolite.

### 5. Summary and conclusions

Basing on the bond valence model we presented two novel definitions of chemical CNs which unexpectedly turned out to be related to the information theory concepts of diversity and entropy. The valence entropy and valence diversity coordination numbers are flexible as their definitions include an additional parameter  $n$  called order. This makes it possible to choose such a CN order that balances properly the contributions of higher and lower bond valences to CN and, as a result, makes the CN most suitable for correlating a particular chemical property with it. Both  ${}^n\text{VECN}$  and  ${}^n\text{VDCN}$  fulfil the restrictions imposed on CNs by Carter (1978). On one hand, our results imply that geometrical CNs, the most sophisticated of which is Hoppe's (1979) ECoN, are more useful than chemical CNs for studying strictly geometrical changes in crystal structures. On the other hand, we have shown that  ${}^n\text{VECN}$  and  ${}^n\text{VDCN}$  operating with bond valences are intrinsically better suited than geometrical CNs for studying structural relationships and transitions which involve interplay between primary and secondary bonds. The analysis of systems with other CNs are the subject of further studies in our group.

#### APPENDIX A

##### Derivation of the expression for ${}^n\text{VECN}$ from definition – equation (9)

Before calculating the limit in equation (9) let us perform some transformations:

$$\left(\frac{\sum_{i=1}^N s_i^n}{k_{\text{VECN},n}}\right)^{\frac{1}{n}} = \left(\frac{\sum_{i=1}^N s_i^{n+\delta}}{k_{\text{VECN},n}}\right)^{1/(n+\delta)} \quad (18)$$

Let us set  $S_n \equiv \sum_{i=1}^N s_i^n$  and take natural logarithm of the equation (18):

$$\frac{1}{n}(\ln S_n - \ln k_{\text{VECN},n}) = \frac{1}{n+\delta}(\ln \sum_{i=1}^N s_i^{n+\delta} - \ln k_{\text{VECN},n}) \quad (19)$$

Now let us multiply both sides of equation (19) by  $n(n+\delta)$  and sort the resulting terms. It follows that:

$$\ln k_{\text{VECN},n} = \ln S_n - \frac{n}{\delta} \ln \frac{\sum_{i=1}^N s_i^{n+\delta}}{S_n} \quad (20)$$

Let us now calculate the limit for  $\delta \rightarrow 0$  of the second term on the right-hand side of equation (20) using L'Hospital's rule (the 0/0 indeterminate form):

$$\begin{aligned} \lim_{\delta \rightarrow 0} \frac{n}{\delta} \ln \frac{\sum_{i=1}^N s_i^{n+\delta}}{S_n} &= \lim_{\delta \rightarrow 0} \frac{n \frac{S_n}{\sum_{i=1}^N s_i^{n+\delta}} \frac{1}{S_n} \sum_{i=1}^N s_i^{n+\delta} \ln s_i}{1} \\ &= \frac{n}{S_n} \sum_{i=1}^N s_i^n \ln s_i \end{aligned} \quad (21)$$

Inserting this into equation (20) gives

$$\ln k_{\text{VECN},n} = \ln S_n - \frac{n}{S_n} \sum_{i=1}^N s_i^n \ln s_i, \quad (22)$$

which after additional simple but tedious transformations:

$$\begin{aligned} \ln k_{\text{VECN},n} &= \ln S_n - \sum_{i=1}^N \frac{s_i^n}{S_n} \ln \frac{s_i^n}{1} \\ &= \ln S_n - \sum_{i=1}^N \frac{s_i^n}{S_n} \ln \left(\frac{s_i^n}{S_n} \cdot S_n\right) \\ &= \ln S_n - \sum_{i=1}^N \frac{s_i^n}{S_n} \left(\ln \frac{s_i^n}{S_n} + \ln S_n\right) \\ &= \ln S_n - \sum_{i=1}^N \frac{s_i^n}{S_n} \ln \frac{s_i^n}{S_n} - \sum_{i=1}^N \frac{s_i^n}{S_n} \ln S_n \\ &= \ln S_n - \sum_{i=1}^N \frac{s_i^n}{S_n} \ln \frac{s_i^n}{S_n} - \ln S_n \sum_{i=1}^N \frac{s_i^n}{S_n} \\ &= \ln S_n - \sum_{i=1}^N \frac{s_i^n}{S_n} \ln \frac{s_i^n}{S_n} - \ln S_n \end{aligned} \quad (23)$$

leads to

$$\ln k_{\text{VECN},n} = - \sum_{i=1}^N \frac{s_i^n}{S_n} \ln \frac{s_i^n}{S_n} \quad (24)$$

#### APPENDIX B

##### Derivation of the expression for ${}^n\text{VECN}$ from the differential expression

Let us notice first that:

$$[\ln f(x)]' = \frac{1}{f(x)} f'(x) \Rightarrow f'(x) = f(x) [\ln f(x)]'. \quad (25)$$

Let us make use of equation (25) in order to differentiate the function from the left-hand side of equation (11):

$$f(n) := \left(\frac{\sum_{i=1}^N s_i^n}{k}\right)^{1/n} \quad (26)$$

$$\begin{aligned} \frac{\partial f(n)}{\partial n} &= \left(\frac{\sum_{i=1}^N s_i^n}{k}\right)^{1/n} \left[\frac{1}{n} \ln \left(\frac{\sum_{i=1}^N s_i^n}{k}\right)\right]' \\ &= \left(\frac{\sum_{i=1}^N s_i^n}{k}\right)^{1/n} \left(-\frac{1}{n^2} \ln \frac{\sum_{i=1}^N s_i^n}{k} + \frac{1}{n} \frac{k}{\sum_{i=1}^N s_i^n} \frac{\sum_{i=1}^N s_i^n \ln s_i}{k}\right) \end{aligned} \quad (27)$$

$$f'(n) = \left(\frac{\sum_{i=1}^N s_i^n}{k}\right)^{1/n} \left(\frac{1}{n} \frac{\sum_{i=1}^N s_i^n \ln s_i}{\sum_{i=1}^N s_i^n} - \frac{1}{n^2} \ln \frac{\sum_{i=1}^N s_i^n}{k}\right) \quad (28)$$

$$f'(n) = \left(\frac{S_n}{k_{\text{VECN},n}}\right)^{1/n} \frac{1}{n} \left(\frac{\sum_{i=1}^N s_i^n \ln s_i}{S_n} - \frac{1}{n} \ln \frac{S_n}{k_{\text{VECN},n}}\right) = 0 \quad (29)$$

$$\frac{\sum_{i=1}^N s_i^n \ln s_i}{S_n} - \frac{1}{n} \ln \frac{S_n}{k_{\text{VECN},n}} = 0 \quad (30)$$

$$-\frac{1}{n} \ln \frac{S_n}{k_{\text{VECN},n}} = \frac{\sum_{i=1}^N s_i^n \ln s_i}{S_n} \quad (31)$$

$$\ln S_n - \ln k_{\text{VECN},n} = n \frac{\sum_{i=1}^N s_i^n \ln s_i}{S_n} \quad (32)$$

Equation (32) leads to the same expression for  ${}^n\text{VECN}$  as equation (22). Note that  $k$  is used instead of  $k_{\text{VECN},n}$  in equations (26)–(28) as the latter is a special value of the former for which  $f'(n) = 0$ .

The analysis of equation (29) indicates that for positive values of  $n$   $f'(n)$  changes sign from negative to positive indicating that  $f(n)$  has a minimum at particular  $n$  determined by the order of  $k_{\text{VECN},n}$ . Conversely, for negative  $k_{\text{VECN},n}$  orders  $f'(n)$  changes sign from positive to negative indicating that  $f(n)$  has a maximum for this particular value of  $n$ .

### APPENDIX C

#### Derivation of VECN of positive and negative infinite orders

Let us assume without loss of generality that:

$$\begin{aligned} s &= s_1 = s_2 = \dots = s_m < s_{m+1} < \dots < s_{N-M+1} = s_{N-M+2} = \dots \\ &= s_N = S, \end{aligned} \quad (33)$$

where  $m$  and  $M$  are positive natural numbers. This assumption means that there are  $m$  bonds/contacts with lowest valence  $s$  and  $M$  bonds with highest valence  $S$  in some coordination sphere. Let us notice in the beginning that:

$$\lim_{n \rightarrow +\infty} \left(\frac{s_i}{S}\right)^n = \begin{cases} 1, & i = N - M + 1, N - M + 2, \dots, N \\ 0, & i = 1, 2, \dots, N - M \end{cases} \quad (34)$$

Besides, it can be shown using L'Hospital's rule (the  $\infty/\infty$  indeterminate form) that for  $i = 1, 2, \dots, N - M$ :

$$\lim_{n \rightarrow +\infty} n \left(\frac{s_i}{S}\right)^n = \lim_{n \rightarrow +\infty} \frac{n}{\left(\frac{s_i}{S}\right)^{-n}} = \lim_{n \rightarrow +\infty} \frac{1}{\left(\frac{s_i}{S}\right)^{-n} \ln \left(\frac{s_i}{S}\right)} = 0. \quad (35)$$

We may now proceed to calculate the  ${}^{+\infty}\text{VECN}$ :

$$\begin{aligned} \ln k_{\text{VECN},+\infty} &= \lim_{n \rightarrow +\infty} - \sum_{i=1}^N \frac{s_i^n}{S_n} \ln \frac{s_i^n}{S_n} \\ &= - \sum_{i=1}^N \lim_{n \rightarrow +\infty} \frac{\left(\frac{s_i}{S}\right)^n}{\sum_{j=1}^N \left(\frac{s_j}{S}\right)^n} \ln \frac{\left(\frac{s_i}{S}\right)^n}{\sum_{j=1}^N \left(\frac{s_j}{S}\right)^n}. \end{aligned} \quad (36)$$

Using the values of limits given in (34) and (35) it follows that:

$$\begin{aligned} \ln k_{\text{VECN},+\infty} &= - \sum_{i=1}^{N-M} \lim_{n \rightarrow +\infty} \frac{\left(\frac{s_i}{S}\right)^n}{M} \ln \frac{\left(\frac{s_i}{S}\right)^n}{M} - \\ &\quad \sum_{i=N-M+1}^N \frac{1}{M} \ln \frac{1}{M} \\ &= 0 - M \frac{1}{M} \ln \frac{1}{M} = \ln M. \end{aligned} \quad (37)$$

Thus,  $k_{\text{VECN},+\infty} = M$ . It can be analogously shown that  $k_{\text{VECN},-\infty} = m$  by dividing all terms by  $s$  instead of  $S$  and taking a limit to  $-\infty$  in (36). One has to notice that:

$$\lim_{n \rightarrow -\infty} \left(\frac{s_i}{s}\right)^n = \begin{cases} 1, & i = 1, 2, \dots, m \\ 0, & i = m + 1, m + 2, \dots, N \end{cases} \quad (38)$$

### APPENDIX D

#### Proof that the ratio of VECN for two coordination centres for highly symmetric structures is stoichiometric

We will prove that for binary compound  $A_xB_y$  crystallizing with one symmetry independent atom  $A$  and one symmetry independent atom  $B$  the ratio of  ${}^n\text{VECN}$  for  $A$  and  $B$  is stoichiometric. Firstly, let us notice that bonds/contacts between atoms  $A$  and  $B$  can be divided into shells according to the  $AB$  separation or the respective bond valence. For atom  $A$  there are  $N_{1A}, N_{2A}, \dots, N_{mA}$  bonds/contacts of valences  $s_1, s_2, \dots, s_m$ , respectively, in  $m$  shells. Similarly, for atom  $B$  there are  $N_{1B}, N_{2B}, \dots, N_{mB}$  of the same valences  $s_1, s_2, \dots, s_m$ , respectively. Of course,  $\sum_{j=1}^m N_{jA} = N_A$  and  $\sum_{j=1}^m N_{jB} = N_B$  where index  $j$  is used for enumerating shells. Index  $i$  will be used, just as everywhere else in the manuscript, for enumerating bonds/contacts. The stoichiometry of the compound enforces that

$$\forall j \frac{N_{jA}}{N_{jB}} = \frac{y}{x} \quad (39)$$

and consequently

$$\frac{N_A}{N_B} = \frac{y}{x}. \quad (40)$$

Besides

$$S_{nB} = \sum_{i=1}^m s_i = \sum_{j=1}^m N_{jB} s_j = \sum_{j=1}^m \frac{x}{y} N_{jA} s_j = \frac{x}{y} S_{nA}. \quad (41)$$

Equations (39) and (41) indicate that

$$\frac{N_{jA}}{S_{nA}} = \frac{N_{jB}}{S_{nB}}. \quad (42)$$

Let us now transform the expression for VECN of atom  $B$ :

$$\begin{aligned} \ln k_{\text{VECN},nB} &= - \sum_{i=1}^{N_B} \frac{s_i^n}{S_{nB}} \ln \frac{s_i^n}{S_{nB}} \\ &= - \sum_{j=1}^m N_{jB} \frac{s_j^n}{S_{nB}} \ln \frac{s_j^n}{S_{nB}} \\ &= - \sum_{j=1}^m N_{jA} \frac{s_j^n}{S_{nA}} \ln \frac{s_j^n}{\frac{x}{y} S_{nA}} \\ &= - \sum_{i=1}^{N_A} \frac{s_i^n}{S_{nA}} \left( \ln \frac{s_i^n}{S_{nA}} - \ln \frac{x}{y} \right) \\ &= - \sum_{i=1}^{N_A} \frac{s_i^n}{S_{nA}} \ln \frac{s_i^n}{S_{nA}} + \sum_{i=1}^{N_A} \frac{s_i^n}{S_{nA}} \ln \frac{x}{y} \\ &= \ln k_{\text{VECN},nA} + \ln \frac{x}{y}, \end{aligned} \quad (43)$$

$$\ln \frac{k_{\text{VECN},nB}}{k_{\text{VECN},nA}} = \ln \frac{x}{y} \quad (44)$$

$$\frac{k_{\text{VECN},nA}}{k_{\text{VECN},nB}} = \frac{y}{x}, \quad (45)$$

QED.

### Funding information

The following funding is acknowledged: Politechnika Warszawska.

### References

- Alcock, N. W. (1972). *Advances in Inorganic Chemistry and Radiochemistry*, Vol. 15, edited by H. J. E. and A. G. Sharpe, pp. 1–58. Academic Press.
- Allig, H. & Trömel, M. (1992). *Z. Kristallogr.* **201**, 213–222.
- Belsky, A., Hellenbrandt, M., Karen, V. L. & Luksch, P. (2002). *Acta Cryst.* **B58**, 364–369.
- Brown, I. D. (2009). *Chem. Rev.* **109**, 6858–6919.
- Brown, I. D. (2016). *The Chemical Bond in Inorganic Chemistry: The Bond Valence Model*. Oxford University Press.
- Brown, I. D., Klages, P. & Skowron, A. (2003). *Acta Cryst.* **B59**, 439–448.
- Brown, I. D. & Poeppelmeier, K. R. (2014). *Bond Valences*. Berlin, Heidelberg: Springer.
- Brunner, G. O. (1977). *Acta Cryst.* **A33**, 226–227.
- Brunner, G. O. & Schwarzenbach, D. (1971). *Z. Kristallogr.* **133**, 127–133.
- Carter, F. L. (1978). *Acta Cryst.* **B34**, 2962–2966.
- Christy, A. G. & Mills, S. J. (2013). *Acta Cryst.* **B69**, 446–456.
- Eon, J.-G. & Nespolo, M. (2015). *Acta Cryst.* **B71**, 34–47.
- Frank, F. C. & Kasper, J. S. (1958). *Acta Cryst.* **11**, 184–190.
- Gagné, O. C. & Hawthorne, F. C. (2015). *Acta Cryst.* **B71**, 562–578.
- Gagné, O. C. & Hawthorne, F. C. (2018a). *Acta Cryst.* **B74**, 79–96.
- Gagné, O. C. & Hawthorne, F. C. (2018b). *Acta Cryst.* **B74**, 63–78.
- Giacovazzo, C., Monaco, H. L., Artioli, G., Viterbo, D., Milanesio, M., Gilli, G., Gilli, P., Zanotti, G. & Ferraris, G. (2011). *Fundamentals of Crystallography*. Oxford, New York: Oxford University Press.
- Guńka, P. A., Dranka, M., Hanfland, M., Dziubek, K. F., Katrusiak, A. & Zachara, J. (2015). *Cryst. Growth Des.* **15**, 3950–3954.
- Guńka, P. A., Dziubek, K. F., Gładysiak, A., Dranka, M., Piechota, J., Hanfland, M., Katrusiak, A. & Zachara, J. (2015). *Cryst. Growth Des.* **15**, 3740–3745.
- Guńka, P. A., Hapka, M., Hanfland, M., Dranka, M., Chałasiński, G. & Zachara, J. (2018). *ChemPhysChem*, **19**, 857–864.
- Hellenbrandt, M. (2004). *Crystallogr. Rev.* **10**, 17–22.
- Hill, M. O. (1973). *Ecology*, **54**, 427–432.
- Hoppe, R. (1970). *Angew. Chem. Int. Ed. Engl.* **9**, 25–34.
- Hoppe, R. (1979). *Z. Kristallogr.* **150**, 23–52.
- Hoppe, R., Voigt, S., Glaum, H., Kissel, J., Müller, H. P. & Bernet, K. (1989). *J. Less-Common Met.* **156**, 105–122.
- IUCr (1992). *Acta Cryst.* **A48**, 922–946.
- IUPAC (1997). *Compendium of Chemical Terminology: IUPAC Recommendations*. Oxford, UK; Malden, MA, USA: Blackwell Science.
- Jost, L. (2006). *Oikos*, **113**, 363–375.
- Lalik, E. (2005). *J. Appl. Cryst.* **38**, 152–157.
- Lima-de-Faria, J., Hellner, E., Liebau, F., Makovicky, E. & Parthé, E. (1990). *Acta Cryst.* **A46**, 1–11.
- Nespolo, M. (2016). *Acta Cryst.* **B72**, 51–66.
- Nespolo, M., Ferraris, G., Ivaldi, G. & Hoppe, R. (2001). *Acta Cryst.* **B57**, 652–664.
- Nespolo, M., Ferraris, G. & Ohashi, H. (1999). *Acta Cryst.* **B55**, 902–916.
- O’Keeffe, M. (1979). *Acta Cryst.* **A35**, 772–775.
- Rahm, M. (2015). *J. Chem. Theory Comput.* **11**, 3617–3628.
- Rahm, M. & Christe, K. O. (2013). *ChemPhysChem*, **14**, 3714–3725.
- Rutherford, J. S. (1998). *Acta Cryst.* **B54**, 204–210.
- Sidey, V. (2008). *Acta Cryst.* **B64**, 515–518.
- Spek, A. L. (2009). *Acta Cryst.* **D65**, 148–155.
- Varn, D. P. & Crutchfield, J. P. (2016). *Philos. Trans. R. Soc. A*, **374**, 20150067.
- Wang, X. & Liebau, F. (1996). *Acta Cryst.* **B52**, 7–15.
- Zachara, J. (2007). *Inorg. Chem.* **46**, 9760–9767.