Distribution rules of systematic absence and generalized de Wolff figure of merit applied to EBSD ab-initio indexing

R. Oishi-Tomiyasu\textsuperscript{a,b}, T. Tanaka\textsuperscript{c}, and J. Nakagawa\textsuperscript{c}

\textsuperscript{a}Yamagata University, Yamagata, Japan\textsuperscript{*}
\textsuperscript{b}JST PRESTO, Kawaguchi, Japan
\textsuperscript{c}Nippon Steel Corporation, Japan

April 1, 2020

Abstract

For EBSD ab-initio indexing, a new method that adopts several methods originally invented for powder indexing, is reported. Distribution rules of systematic absence and error-stable Bravais lattice determination are used to eliminate negative influence of non-visible bands and erroneous information from visible bands. In addition, generalized versions of the de Wolff figures of merit are proposed as a new sorting criterion for the obtained unit-cell parameters that can be used in both orientation determination and ab-initio indexing from Kikuchi patterns. Computational results show that the new figures of merit work well, similarly to the original de Wolff $M_n$. Ambiguity of indexing solutions is also pointed out, which happens in particular for low-symmetric cells, and may generate multiple distinct solutions even if very accurate positions of band center lines and the projection center are given.

1 Introduction

Electron backscatter diffraction (EBSD) is a characterization technique for the microstructure of crystalline or polycrystalline materials, developed by Venable & Harland (1973), and later refined by Dingley & Baba-Kishi (1986) with an aid of computers. This technique can be applied for the determination of crystal orientation, texture analysis, and phase identification.

In orientation determination, the unit-cell parameters are priorly given. Only the center lines of the Kikuchi bands are utilized for acquisition of the unit-cell orientation [Wright & Adams, 1992; Kogure, 2003]. However, in EBSD ab-initio indexing, the unit-cell parameters and its symmetry are also determined. Since all the bands with the indices $m(hk\ell)$ ($m \neq 0$: integer) completely overlap in EBSD patterns, all the derivative lattices (the sublattices and superlattices) of the true solution have identical positions of Kikuchi center lines. Therefore, information about $d$-spacings contained e.g., in the widths of Kikuchi bands is indispensable for uniquely determining the solution.

Each bandwidth of Kikuchi band corresponds to a Miller index $hk\ell$. It is approximately proportional to the inverse of the interplanar spacing (i.e., $d$-spacings) of the diffracting plane. This information has been also used for EBSD ab-initio indexing [Michael, 2000; Dingley & Wright, 2009; Li & Han, 2015], including the recent software EBSDL [Li et al., 2014]. However, due to the complex profile of band edges [Nolze et. al., 2015; Nolze & Winkelmann, 2017], and also due to the small Bragg angle caused by short wavelengths of the incident electron beam, the error in the band-width measurement is 5-20\% [Dingley and Wright, 2009].

Better accuracy could be obtained by analyzing the higher-order Laue Zone (HOLZ) rings [Langer & Däbritz, 2007; Dingley & Wright, 2009]. However, it is not straightforward to analyze the $d$-spacings from the HOLZ rings [Nolze et al., 2015], and it depends on the crystal structure whether the HOLZ rings are clearly visible.

\textsuperscript{*}E-mail: tomiyasu@imi.kyushu-u.ac.jp, Current affiliation: Institute of Mathematics for Industry (IMI), Kyushu University
Therefore, the \(d\)-spacings extracted from the bandwidths are also used in this article, although our improvements can be similarly applied to the analyses based on Holz rings.

As a feature of our software, it allows errors in the projection center, band positions and band edges to some degree, as seen from the input file (the first parameter in Table 1 of Appendix A in the supplementary materials). In particular, the necessity of a precise projection center has been pointed out \(\text{e.g., in Nolze & Winkelmann (2017)}.\)

Technically, the novel points of our indexing method are as follows:

1) Non-visible band edges in EBSD patterns, are frequently caused by reflections with relatively small structure factors [Nolze and Winkelmann, 2017]. We propose a method that works for all the types of systematic absence (SA), as a result of the theorems in Oishi-Tomiyasu (2013) that are available without any knowledge on the Bravais types and the space groups.

2) a method for error-stable Bravais lattice determination [Oishi-Tomiyasu, 2012], and new figures of merit with a definition similar to the de Wolff figure of merit [de Wolff, 1968], are applied to EBSD indexing for the first time.

As for (1), the method of Dingley & Wright (2009) needs a reciprocal-lattice basis \(l_1^*, l_2^*, l_3^*\) such that all of \(l_1^*, -l_1^*, l_1^* + l_i^*\) are not extinct for both \(i = 2, 3\), although such a basis does not exist for some space groups and settings (\(\text{e.g., No.70 c, d, No.88 c, d}\)). It is explained in Section 2 how to extract information about non-visible bands from visible bands, without being adversely affected by extinct reflections. Our idea is basically same as that proposed in Ito (1949) for powder indexing, and later developed for CONO-GRAPH [Esmaeili et al., 2017]. However, Day (2008) reported that in the experimental pattern of Si, the bandwidth of the forbidden \(\{222\}\) was the most visible among all of \(\{hhh\}\). Therefore, reflection rules might be violated to some degree in case of Kikuchi patterns, owing to dynamical scattering of the electron beam.

With regard to the Bravais lattice determination, even very small errors in the parameters such as rounding errors can cause failure in Bravais lattice determination [Grosse-Kunstleve et al., 2004]. Owing to this, the first author provided a method for error-stable Bravais lattice determination with rigorous proofs in Oishi-Tomiyasu (2012), as explained in Section 3.2.

We also propose figures of merit for orientation determination and ab-initio indexing (Section 4), by extending the definition of the de Wolff figure of merit for the 2D and 3D (= 2D image + bandwidths) experimental data. The de Wolff \(M_n\) has been used as the most efficient indicator in powder indexing, and the generalized ones are presenting very similar properties to those of the original one.

In this article, two types of ambiguities in EBSD ab-initio indexing, which sometimes allow multiple distinct solutions, are also explained. The first one happens when the observed bandwidths are not the narrowest ones. In this case, the above uniqueness problem occurs again, because all the sublattices of the true crystal lattice can have identical bandwidths, in addition to the band positions. Thus, it is necessary to assume that at least some of the narrowest band edges are the most visible. Another ambiguity is caused by inaccuracy of the projection center (in particular, by the shift \(\Delta_z\) perpendicular to the phosphor screen).

For definition, the 3-dimensional (3D) lattice \(M\) is a derivative lattice of another 3D lattice \(L\), if they have a common 3D sublattice \(M \cap L\).

2 Background and formulas for EBSD indexing

A method to gain the unit-cell length-ratios and angles from the center-line positions of the Kikuchi bands (in particular without bandwidths) is explained. The notation used here is basically the same as that in Kogure (2003).

The general situation of electron backscattering is depicted in Figure 1. In Figures 1 and 2, it may be thought that the positions of the projection centers are exact, even if they may be unknown. In the study of EBSD indexing, the relationship between Kikuchi bands has been explained, mainly by using the coordinates of Kikuchi bands. However, the same thing is more easily understood by using the coordinates of the projection of the reciprocal lattice points, which are computable from the band information and the projection center.

The reciprocal space is now embedded into the real space, so that any of three orthogonal unit-vectors of the reciprocal space are mapped to those of the real space. As shown in Figure 1(b), each
Kikuchi line corresponds to the gnomonic projection of a reciprocal-lattice vector on the screen. If \((\tan \sigma \cos \varphi, \tan \sigma \sin \varphi)\) is the \((x, y)\)-coordinate of the foot of the perpendicular line from the pattern center \(O\) to the center line of a Kikuchi band as shown in Figure 1(a), it is understood that the corresponding reciprocal lattice vector \(a^*\) has the direction in the real space:

\[
a^* \propto \begin{pmatrix} -\cos \varphi \cos \sigma, -\cos \varphi \sin \sigma, \sin \sigma \end{pmatrix},
\]

where the \(z\)-axis is vertical to the screen.

The second \(\propto\) implies that \(a^*\) intersects the screen at the coordinate \((-\cos \varphi/\tan \sigma, -\sin \varphi/\tan \sigma, 1)\), if the scale is adjusted so that the camera length equals 1. The first two entries \((-\cos \varphi/\tan \sigma, -\sin \varphi/\tan \sigma)\) are the coordinate of the projection of \(a^*\) on the screen.

![Figure 1: Relationship between the center line of a Kikuchi band and the direction of the corresponding reciprocal lattice vector. In (a), the phosphor screen is parallel to the sheet. In both (a) and (b), the projection \(P\) of \(Q\) corresponds to the gnomonic projection of the reciprocal-lattice vector. In (a), the phosphor screen is parallel to the sheet. In both (a) and (b), the scale is adjusted so that the camera length (= distance from the electron probe to the screen) equals 1.](image)

This interpretation is useful for obtaining a geometric intuition of Kikuchi patterns. In particular, it is easily seen that all the bands corresponding to \(ma^*\) \((m \neq 0; \text{integer})\) have an identical center line. Furthermore, the three projected lattice points \(P_1, P_2, P_3\) are aligned on the phosphor screen, if and only if their reciprocal lattice vectors \(a_1^*, a_2^*, \text{and } a_3^*\) are coplanar. As is well known, this happens if and only if the corresponding Kikuchi lines intersect at one point (Figure 2).

![Figure 2: Kikuchi bands intersecting at one point: if all the Kikuchi lines \(\ell_i\) \((i = 1, 2, 3)\) intersect at \((x_0, y_0)\), then their corresponding projected lattice points \(P_i\) will all be on the line \(x_0X + y_0Y = -1\). Even if the projection center is not exact, the bands intersect at one point, as long as of the lattice points are aligned as in (2). This property can be used when one wants to improve results of automatic band detection. Under the projection-center shift \((\Delta_x, \Delta_y, \Delta_z)\), the intersection is varied from \((x_0, y_0)\) to \((x_0(1 + \Delta_z) - \Delta_x, y_0(1 + \Delta_z) - \Delta_y)\) (see Eq.(5), (6) for the used formulas).

The ratios of the lattice-vector lengths can be determined from the positions of the Kikuchi lines, by using linear equations satisfied by the coplanar lattice vectors. If \(a_1^*\) is the lattice vector that provides the projection \(P_1\) in Figure 2, there are rational numbers \(p, q\) such that \(a_1^* = pa_2^* + qa_3^*, \text{because they are coplanar. In the 3D coordinate system of Figure 1 (b), } P_i\text{ is given by } (-a_i/(a_ix_0 + b_iy_0), -b_i/(a_ix_0 + b_iy_0), 1). \text{Therefore, for some } c_i > 0, \]

\[
a_i^* = c_i(-a_i, -b_i, a_ix_0 + b_iy_0).\]
From the equation \( a_3^* = pa_1^* + qa_2^* \), the ratio \( c_1 : c_2 : c_3 \) satisfies the following:

\[
\begin{pmatrix}
-a_1 & -a_2 & -a_3 \\
-p & -q & -p_c \\
a_1x_0 + b_1y_0 & a_2x_0 + b_2y_0 & a_3x_0 + b_3y_0
\end{pmatrix} \begin{pmatrix}
p c_1 \\
q c_2 \\
-c_3
\end{pmatrix} = 0. \tag{1}
\]

if the values of \( p, q \) are known, the ratio \( c_1 : c_2 : c_3 \) can be computed from the inner products \( \alpha_{ij} = (a_i, b_i) \cdot (a_j, b_j) \), because \( a_1^2 + b_1^2 = 1 \) is now assumed:

\[
c_1 : c_2 : c_3 = \frac{1}{p} \begin{vmatrix}
a_2 & a_3 \\
b_2 & b_3
\end{vmatrix} : \frac{1}{q} \begin{vmatrix}
a_3 & a_1 \\
b_3 & b_1
\end{vmatrix} : \begin{vmatrix}
a_1 & a_2 \\
b_1 & b_2
\end{vmatrix} = \frac{\sqrt{1 - \alpha_{12}^2}}{|p|} : \frac{\sqrt{1 - \alpha_{13}^2}}{|q|} : \sqrt{1 - \alpha_{12}^2}.
\]

The slope \((a_i, b_i)\) of the corresponding Kikuchi line on the screen is independently determined from the position of the pattern center. This explains why the obtained \( c_1 : c_2 : c_3 \) are not affected by the shift of the projection center. Only the third entry of \( a_i^* \) is affected by the shift \((\Delta_x, \Delta_y)\).

Furthermore, the influence of the shift \( \Delta_z \) can be eliminated just by adjusting the scale of the \( z \)-axis, which is seen from Eqs. (5), (6) in Section 3.3. Namely, if \((\Delta_x, \Delta_y)\) are well refined, even if \( \Delta_z \) is imprecise, it is possible to index the band center lines, although the \( z \)-scale of the obtained unit-cell parameters might contain large errors. This problem is also discussed in the following sections. In ab-initio indexing, although \( p \) and \( q \) in In Eq. (1) are unknown, the ratio \( c_1 : c_2 : c_3 \) can be computed by setting \((p, q)\) to specific values e.g., \((1, 1), (2, 1), \) or \((1, 2)\) as in Section 3.

Owing to the gnomonic distortion, the Kikuchi bandwidth \( \beta \) on the screen is related to the Bragg angle \( \theta \), as follows:

\[
\beta = \tan(\sigma + \theta) - \tan(\sigma - \theta), \tag{2}
\]

The information about the \( d \)-spacing \((= 1/|ma^*|)\) of \( ma^* \) and its inverse \( d^* \) can be obtained from this \( \theta \) by using the Bragg equation:

\[
d^* = \frac{1}{d} = \frac{2\sin \theta}{\lambda}, \tag{3}
\]

where \( \lambda \) is the wavelength of the electron beam.

In practice, it is difficult to obtain accurate values of the bandwidths, especially from experimental patterns, and to judge whether the obtained value corresponds to the narrowest band, although the same thing happens when HOLZ rings are used [Michael and Eades, 2000].

In order to handle this situation in a better way, the following theorems are used in Section 3, as general properties of SA. In the statements of the theorems, \( L^* \) is the reciprocal lattice of the crystal lattice. For simplicity, the primitive lattice is always considered here. \( \{l_1^*, l_2^*\} \) is called a primitive set, if it is a subset of some basis \( l_1^*, l_2^*, l_3^* \) of \( L^* \).

**Theorem 1** (Theorem 2 in Oishi-Tomiyasu (2013)). Regardless of the type of systematic absence (SA), there are infinitely many primitive sets \( \{l_1^*, l_2^*\} \) of \( L^* \) such that none of \( l_1^*, l_2^*, l_1^*+2l_2^*, 2l_1^*+l_2^* \) corresponds to an extinct reflection due to the systematic absence. Furthermore, there exist infinitely many 2D sublattices \( L_2^* \) of \( L^* \) such that \( L_2^* \) is expanded by such \( l_1^*, l_2^* \).

The reciprocal lattices \( l_1^*, l_2^*, l_1^*+2l_2^*, 2l_1^*+l_2^*, \) and \( l_1^*+l_2^* \) are coplanar. In the method of CONOGRAPH, their relationship is illustrated as in Figure 3, by using a graph:

**Theorem 2** (Theorem 4 in Oishi-Tomiyasu (2013)). Regardless of the type of systematic absence (SA), there are infinitely many bases \( \{l_1^*, l_2^*, l_3^*\} \) of \( L^* \) such that the following hold:

(a) the reflections of \( \pm l_1^* + l_2^* + l_3^* \) are not forbidden.

(b) For both \( i = 2, 3 \), (i) none of the reflections of \( ml_1^* + (m-1)(-l_2^* + l_1^*) \) are forbidden for any integer \( m \), or otherwise, (ii) none of the reflections of \( ml_1^* + (m-1)(l_1^* - l_3^*) \) are forbidden for any integer \( m \geq 0 \).

That is, none of the underlined lattice vectors in Figure 4 corresponds to a forbidden reflection.

4
Figure 3: A subgraph of a topograph corresponding to the reflections $l_1^*, l_2^*, l_1^* + 2l_2^*, 2l_1^* + l_2^*$ that are not forbidden, and $l_1^* + l_2^*$ that might be forbidden owing to SA. This graph was originally used in Conway (1997), where the term “topograph” was first coined.

Figure 4: Outline of Theorem 2, which ensures that none of the underlined lattice vectors are forbidden.
3 New methods for ab-initio EBSD indexing and scale determination of the unit-cell

3.1 Acquisition of candidates for the primitive lattice

Ab-initio indexing methods can be classified into two categories, depending on their strategy. In the first category, various \(hkl\)s are assigned to a few of selected reflections, in order to generate multiple candidate solutions. In the second category, various combinations of reflections are assumed to satisfy some fixed relationship. In both the categories, it is checked whether there is a candidate that can well predict all the observed reflections. The advantage of the latter is that the true solution is normally generated multiple times from distinct observed reflections, hence it can be more robust against errors in the input than the former.

Our method also belongs to the latter. The basic algorithm, which uses only the positions of the Kikuchi center-lines, is provided in Table 1.

Unlike step (3-a) in which the direction of \(\mathbf{a}_1^* + \mathbf{a}_2^*\) is observed, in steps (3-b) and (3-c), \(\mathbf{a}_1^* + \mathbf{a}_2^*\) predicted from the other input bands is stored in \(\mathcal{A}\). The algorithm is simplified by this use of virtual bands. Figure 5 shows which combinations of visible bands and non-visible bands are used to construct a unit cell.

![Figure 5: Combinations of bands and their reciprocal lattice vectors assigned for indexing. If 5 or 6 bands intersect as in either of (i)–(iii), the lattice basis \((\mathbf{b}_1^*, \mathbf{b}_2^*, \mathbf{b}_3^*) = (-l_1^*, l_2^*, l_3^*)\) is saved in step (4) of Table 1. All of their length-ratios and inner products are determined from the band positions. Every observed band (gray) and computed band (white) are assigned either of \(+l_1^*, +l_2^*, +l_3^*\), or their linear sums \(\pm(l_1^* \pm l_2^*), \pm(2l_1^* + l_3^*), \pm(l_1^* + 2l_2^*)\) \((i = 2, 3)\). The vectors assigned to the observed bands are chosen from the underlined not-forbidden reflections in Figure 4.

So far, it has been unnecessary to use Theorems 1, 2, because influence of SA (and the breakdown of Friedels law [Marthinsen and Høier, 1988]) can be ignored, because of exactly overlaid band positions. However, if information about bandwidths are used in step (3) in order to obtain the lengths of \(\mathbf{a}_1^*, \mathbf{a}_1^* + \mathbf{a}_2^*\), it is necessary to assume that the visible band edges at the positions of \(\mathbf{a}_1^*, \mathbf{a}_2^*\) are the narrowest ones.

In the execution of the algorithm, the identical lattice should be found multiple times by using different combinations of observed bands. The unit-cell scale is properly computed for each combination, if the above assumption on \(\mathbf{a}_1^*, \mathbf{a}_2^*\) is true.
Table 1: Indexing algorithm in which only the positions of Kikuchi center lines are used.

(Input)
Inp: array of unit vectors \((-\cos \sigma \cos \varphi, -\cos \sigma \sin \varphi, \sin \sigma\) obtained from the Kikuchi band center positions \((\tan \sigma \cos \varphi, \tan \sigma \sin \varphi)\) on the screen (the positions may be affected by the error of the projection center).

(Output)
Ans: array of candidates for the reciprocal lattice basis (here, the basis vectors are the edges of the primitive cell).

(Algorithm)
Detection of zones (2D sublattices):
(1) for any distinct vectors \(u_1 \neq u_2\) in Inp, search for all the \(u_3\) in Inp that may be considered to be linearly dependent on \(u_1, u_2\). All of such \(u_3\) are saved in a new array \(\text{Inp}_{u_1, u_2}\).
(2) (Computation of \(\lambda_1, \lambda_2\) with \(u_3 = \lambda_1 u_1 + \lambda_2 u_2\) for each \(u_3 \in \text{Inp}_{u_1, u_2}\), the following equation is solved:
\[
\begin{pmatrix}
  u_{11} & u_{21} & u_{31} \\
  u_{12} & u_{22} & u_{32}
\end{pmatrix}
\begin{pmatrix}
  \lambda_1 \\
  \lambda_2 \\
  -1
\end{pmatrix}
= 0,
\]
where \((u_{1i}, u_{2i}, u_{3i})\) are the entries of \(u_i\) \((i = 1, 2, 3)\). If \(\lambda_1 \leq 0\) or \(\lambda_2 \leq 0\), go to the next \(u_3 \in \text{Inp}_{u_1, u_2}\). Otherwise, carry out step (3) and store a pair of vectors \(\{a_1^*, a_2^*, a_3^*\}\) \((i = 1, 2)\) in a common array \(A\), before going to the next \(u_3\).

(3) In what follows, \(a_i^*\) is the reciprocal lattice vector with the direction \(u_i^*(i = 1, 2, 3)\). The assumption \(a_i^* = pa_i^* + qa_2^*\) is tested for each of \((p, q) = (1, 1), (2, 1), (1, 2)\) in the following (3-a)–(3-c):
(3-a) (Case of \((p, q) = (1, 1)\), i.e., \(a_i^* = a_1^* + a_2^*\)) in this case, \(|a_1^*| : |a_2^*| : |a_3^*| = \lambda_1 : \lambda_2 : 1\) holds (cf. Eq.(1)). Hence, \(\{\lambda_1 u_1, u_3\}\), \(\{\lambda_2 u_2, u_3\}\), are stored in \(A\).
(3-b) (Case of \((p, q) = (2, 1)\), i.e., \(a_i^* = 2a_1^* + a_2^*\)) Similarly, \(|a_1^*| : |a_2^*| : |a_3^*| = \lambda_1/2 : \lambda_2 : 1\) is obtained. Hence \(a_1^*, a_2^*, a_3^*\) are constant multiples of \((\lambda_1/2)u_1, \lambda_2 u_2, \lambda_1/2u_1 + \lambda_2 u_2\). If the direction of \(a_1^* + a_2^*\) is not observed (i.e., not in \(\text{Inp}_{u_1, u_3}\)), \(\{\lambda_1/2 u_1, (\lambda_1/2)u_1 + \lambda_2 u_2\}\), \(\lambda_2 u_2, (\lambda_1/2)u_1 + \lambda_2 u_2\) are stored in \(A\).
(3-c) (Case of \((p, q) = (1, 2)\), i.e., \(a_i^* = a_1^* + 2a_2^*\)) in this case, \(|a_1^*| : |a_2^*| : |a_3^*| = \lambda_1 : \lambda_2/2 : 1\). Hence \(a_1^*, a_2^*, a_3^*\) are proportional to \(\lambda_1 u_1, (\lambda_2/2)u_2, \lambda_1 u_1 + (\lambda_2/2)u_2\). If the direction of \(a_1^* + a_2^*\) is not in \(\text{Inp}_{u_1, u_2}\), \(\{\lambda_1 u_1, \lambda_1 u_1 + (\lambda_2/2)u_2\}\), \(\{\lambda_2/2 u_2, \lambda_1 u_1 + (\lambda_2/2)u_2\}\) are stored in \(A\).

Construction of candidates for the lattice basis:
(4) for any \(\{b_1^*, b_2^*, b_3^*\}\) \((c_1^*, c_2^*, c_3^*\) \(\in A\) including vectors \(b_1^*, c_1^*\) with the same direction, if \(b_1^*, b_2^*, b_3^*\) are linearly independent and pass the following check (\(^*\))\(^\dagger\), the basis \(\{b_1^*, b_2^*, b_3^*\}\) is stored in Ans as a candidate solution.
\(^*\)\(^\dagger\) The check (\(^*\)) is imposed to reduce the number of solutions and computation time. By removing (\(^*\)), it is possible to carry out a more exhaustive search.
3.2 Bravais lattice determination & refinement of the projection center, unit-cell parameters and orientation

In Bravais lattice determination, its corresponding Bravais-type and the parameters of the conventional cell are determined from the parameters of the primitive cell. This process needs to be error-stable, if the unit-cell parameters contains at least rounding-off errors.

As seen from the notation of the lattice characters in the International Tables Vol.A, this determination can be regarded as the process to find lattice vectors that intersect at a specific angle such as 90° or 120°. Even for erroneous unit-cell parameters, it can be also carried out by using the lattice-basis reduction theory. In addition, e.g., if the unit-cell is almost orthorhombic, triclinic and monoclinic cells should be also candidates, under consideration of errors. The latter part will not be so confusing for crystallographers. The main problem for software developers is how to find a reliable program.

As pointed out in Oishi-Tomiyasu (2012), rigorous mathematics can handle well this determination. The method proposed by the first author is fast in the sense that it requires the same number of steps as in the case when the unit-cell parameters are exact. Nevertheless, it is error-stable in the sense that the program is guaranteed to output the correct cell, except for the case when the input cells contain huge errors (the precise condition is also provided in the article). In addition, the program has been used both for determinations under observation errors (e.g., CONOGRAPH for powder indexing) and rounding-off errors [Oishi-Tomiyasu, 2016].

Therefore, it may be thought that the determination of Bravais types and conventional cell can be carried out without failure. In our indexing software, it is executed between acquisition of the primitive-cell and the calculation of figures of merit (Figure 6).

Figure 6: Flowchart of the software; sections 3.1, 3.2, and 4 explain the respective parts.

In the refinement stage, the following parameters are fit to the band positions (and widths, according to the user’s choice):

- \( s \): scale of the unit cell.
- \((\Delta_x, \Delta_y, \Delta_z)\): projection-center shift.
- \( \theta', \sigma', \psi' \): Euler angles to represent an orthogonal matrix:

\[
g(\theta', \sigma', \varphi') = \begin{pmatrix}
\cos \theta' & \sin \theta' & 0 \\
- \sin \theta' & \cos \theta' & 0 \\
0 & 0 & 1
\end{pmatrix}
\times
\begin{pmatrix}
1 & 0 & 0 \\
0 & \cos \sigma' & \sin \sigma' \\
0 & - \sin \sigma' & \cos \sigma'
\end{pmatrix}
\begin{pmatrix}
\cos \varphi' & \sin \varphi' & 0 \\
- \sin \varphi' & \cos \varphi' & 0 \\
0 & 0 & 1
\end{pmatrix}.
\]

- unit-cell parameters represented by the five entries of the lower triangle matrix:

\[
A := \begin{pmatrix}
1 & 0 & 0 \\
a_{21} & a_{22} & 0 \\
a_{31} & a_{32} & a_{33}
\end{pmatrix}.
\]
The above $A$ is obtained by applying the Cholesky decomposition to the following symmetric matrix, and setting the reciprocal unit-cell parameter $a^*$ to 1 in order to normalize the scale:

$$AA^T = \begin{pmatrix}
(a^*)^2 & a^*b^* \cos \gamma^* & a^*c^* \cos \beta^* \\
 a^*b^* \cos \gamma^* & (b^*)^2 & b^*c^* \cos \alpha^* \\
 a^*c^* \cos \beta^* & b^*c^* \cos \alpha^* & (c^*)^2
\end{pmatrix}
$$

$$= \begin{pmatrix}
 a^2 & ab \cos \gamma & a c \cos \beta \\
 ab \cos \gamma & b^2 & bc \cos \alpha \\
a c \cos \beta & bc \cos \alpha & c^2
\end{pmatrix}^{-1}.$$

The refinement process is carried out by non-linear least squares method (more precisely, Levenberg-Marquardt algorithm). The indexing results and $(\Delta_x, \Delta_y, \Delta_z) = 0$ are used as initial parameters, and improved.

When a Kikuchi band corresponds to the Miller index $m(hk\ell)$, the intersection $(X^{cal}, Y^{cal})$ of the Kikuchi line and its perpendicular line from the pattern center can be computed by:

$$X^{cal} = \frac{-xz}{x^2 + y^2}(1 - \Delta_z) + \Delta_x,$$

$$Y^{cal} = \frac{-yz}{x^2 + y^2}(1 - \Delta_z) + \Delta_y,$$

$$(x \ y \ z) = m( h \ k \ \ell) \ Ag(\theta', \sigma', \varphi').$$

$X^{cal}, Y^{cal}$ are independent of the choice of $m$. In addition, $\Delta_z$ and the scale of the $z$-axis cannot be simultaneously determined only from $(X^{cal}, Y^{cal})$. Hence, in order to obtain both, it is necessary to use the bandwidths represented as:

$$\beta^{cal} = (\tan(\sigma^{cal} + \theta^{cal}) - \tan(\sigma^{cal} + \theta^{cal}))(1 + \Delta_z).$$

The values of $\sigma^{cal}$ and the Bragg angle $\theta^{cal}$ are computed by

$$\sigma^{cal} = \arctan(z/\sqrt{x^2 + y^2}),$$

$$\theta^{cal} = \arcsin(s\lambda/\sqrt{x^2 + y^2 + z^2/2}),$$

where $s$ is the scale of the unit cell (required by the above scaling of $A$), and $\lambda$ is the wavelength of the electron beam.

By using the method described in the last paragraph of Section 3.1, the initial value of the $s$ can be obtained before the refinement. The integer $m$ in Eq.(7) can be reassigned in every iteration of the fitting process by checking which $m$ gives the $\beta^{cal}$ closest to the observed $\beta^{obs}$.

4 De Wolff figures of merit for EBSD indexing

After obtaining the candidate solutions, some sorting system is required for finding the most plausible one in a short time. For determining the orientation from the EBSD patterns, the Confidence Index (CI, Field (1997)) based on the number of “votes” (Wright(1992)), the Fit based on the difference between the computed bands and the detected bands), and the Image Quality (IQ) are used.

In what follows, in order to obtain an efficient sorting criterion for EBSD indexing, use of the de Wolff figure of merit $M_n$ generalized to data of dimensions $> 1$ is proposed. Although a number of new figures of merit have been proposed for powder indexing, $M_n$ is still the most efficient, and has been in use for long. In particular, it is possible to judge whether or not a plausible solution is included in the output, just by checking the largest value of $M_n$. In Section 5, we shall see that its generalizations also have this property.

The de Wolff figure of merit $M_n$ evaluates the similarity between the set of observed $q$-values ($= 1/d^2$, $d$: d-spacing) $0 < Q^{obs}_1 < \ldots < Q^{obs}_n$ and the set of computed $0 < q_1 < \ldots < q_N$ by:

$$M_n = \epsilon/\delta,$$

where $\epsilon$ and $\delta$ are the average discrepancy and the actual discrepancy, respectively, which are defined by:
\[ \bar{\epsilon} := \frac{Q_{\text{obs}}}{2N}, \]
\[ \delta := \frac{1}{n} \sum_{i=1}^{n} |Q_{\text{obs}} - Q_{\text{cal}}^i|, \]

\( Q_{\text{cal}}^i \): computed q-value closest to the observed \( Q_{\text{obs}}^i \).

The above \( \bar{\epsilon} \) is an approximation of the mean value of \( \delta \), when it is assumed that \( Q_{\text{obs}}^i (i = 1, \ldots, n-1) \) and \( q_i \) are uniformly distributed in the interval \([0, Q_{\text{obs}}^n]\), without changing the order \( 0 < q_1 < \ldots < q_N \) (also see Wu (1988)). Namely, the following equality holds:

\[ \bar{\epsilon} \approx E \left[ \min_{i=1, \ldots, N} \{|Q - q_i|\} \right]. \]

In Appendix C, this idea is extended to general dimensions. In particular, the obtained figures of merit is scale-free, similarly to the original \( M_n \).

If a set of computed points \( x_1, \ldots, x_N \) and \( X \) are uniformly distributed in an \( s \)-dimensional hypersphere of radius \( R \), the mean value can be approximated by the following asymptotic formula:

\[ E \left[ \min_{i=1, \ldots, N} \{|X - x_i|\} \right] \sim \frac{\Gamma(1/s)}{s N^{1/s}}, \]

where \( \Gamma(z) \) is the Gamma function \( \int_0^{\infty} t^{z-1} e^{-t} \, dt \). By using the volume \( V = (\sqrt{s}R^s) / (\Gamma(s/2 + 1)N) \) of the \( s \)-dimensional hypersphere, the following is obtained:

\[ E \left[ \min_{i=1, \ldots, N} \{|X - x_i|\} \right] \sim \frac{\Gamma(s/2 + 1/3)}{s \sqrt{\pi} N} \left( V / N \right)^{1/s}. \] (10)

For any point configuration in a convex body of volume \( V \), Eq.(10) holds, because the influence of the boundary can be ignored for sufficiently large \( N \).

In particular, the formulas for the dimensions \( s = 2, 3 \) are:

**Case of 2D objects of volume \( V \)**

\[ E \left[ \min_{i=1, \ldots, N} \{|X - x_i|\} \right] \sim \frac{1}{2} (V/N)^{1/2}. \] (11)

**Case of 3D objects of volume \( V \)**

\[ E \left[ \min_{i=1, \ldots, N} \{|X - x_i|\} \right] \sim \frac{\Gamma(1/3)}{\sqrt{\pi} N} \left( \frac{V}{36 \pi N} \right)^{1/3} \]
\[ \approx 2.6789 \left( \frac{V}{36 \pi N} \right)^{1/3}. \] (12)

In the following examples, it is explained how Eqs.(11) and (12) are used for orientation determination and ab-initio indexing.

**Example 1** (For comparison of band positions)

The intersections \( P_{\text{obs}}^i = (X_{\text{obs}}^i, Y_{\text{obs}}^i) (i = 1, \ldots, n) \) of the Kikuchi lines and their perpendiculars through the pattern center can be regarded as a set of coordinates distributed in a 2D circle of radius \( R := \max_{i=1, \ldots, n} \left( (X_{\text{obs}}^i)^2 + (Y_{\text{obs}}^i)^2 \right) \). If \( (X_{\text{cal}}^j, Y_{\text{cal}}^j) (j = 1, \ldots, N) \) are the computed band positions in the circle, these two sets can be compared by setting \( V \) in Eq.(11) to the area \( V = \pi R^2 \) of the circle; the ratio \( M_{n,N} = \bar{\epsilon} / \delta \) is computed by:

\[ \bar{\epsilon} := \frac{R}{2} \sqrt{\frac{\pi}{N}}, \]
\[ \delta := \frac{1}{n} \sum_{i=1}^{n} |P_{\text{obs}}^i - P_{\text{cal}}^i|, \]

where \( P_{\text{cal}}^i = (X_{\text{cal}}^i, Y_{\text{cal}}^i) \) is the computed point closest to \( P_{\text{obs}}^i \).
Example 2 (For comparison of band positions and widths)

The 3D coordinates representing the Kikuchi bands are obtained by considering the bandwidths $\beta_i^{\text{obs}}$ as the third coordinates:

- $P_{i}^{\text{obs}} = (X_i^{\text{obs}}, Y_i^{\text{obs}}, \beta_i^{\text{obs}})$ ($i = 1, \ldots, n$),
- $(X_j^{\text{cal}}, Y_j^{\text{cal}}, \beta_j^{\text{cal}})$ ($j = 1, \ldots, N$), where $\beta_j^{\text{cal}}$ is approximated by using Eq. (8) and the following:

$$\sigma^{\text{cal}} = \arctan\left(\frac{z}{\sqrt{x^2 + y^2}}\right),$$

$$2\theta^{\text{cal}} \approx 2\sin\theta^{\text{cal}} = s\lambda\sqrt{x^2 + y^2 + z^2}.$$

The above $P_{i}^{\text{obs}}$ are distributed in the cylinder with the radius $R = \max_{i=1,\ldots,n}\{(X_i^{\text{obs}})^2 + (Y_i^{\text{obs}})^2\}$ and the height $h = \max_{i=1,\ldots,n}\{\beta_i^{\text{obs}}\}$. Therefore, $V$ in Eq. (12) is set to $\pi R^2 h$. The figure of merit $M_{n,N} = \bar{\epsilon} / \delta$ is computed by:

$$\bar{\epsilon} := \frac{\Gamma(1/3)}{(R^2 h/36N)^{1/3}},$$

$$\delta := \frac{1}{n} \sum_{i=1}^{n} |P_{i}^{\text{obs}} - P_{i}^{\text{cal}}|,$$

where $P_{i}^{\text{cal}} = (X_i^{\text{cal}}, Y_i^{\text{cal}}, \beta_i^{\text{cal}})$ is the computed point closest to $P_{i}^{\text{obs}}$.

In the definition, the number of computed points $N$ is also a parameter, because infinitely many non-visible bands are theoretically included in the observed range. This $N$ can be automatically determined, by setting a lower threshold for the $d$-values as follows.

The software does the following to maximize the value of $M_{n,N}$ and rank the true solutions higher than their derivative lattices:

- for each indexing solution, the lower threshold for the $d$-values of computed bands is set to the largest values necessary for indexing all the bands.
- The position $(X_i^{\text{cal}}, Y_i^{\text{cal}})$ of completely overlapping bands $m(hk\ell)$ ($m \neq 0$: integer) is counted only once in the number $N$. The computed bandwidths $\beta_i^{\text{cal}}$ is set to the narrowest one (i.e., $m = 1$ in Eq. (7)).

As a result, solutions that assign $hk\ell$ with smaller $|h|, |k|, |\ell|$ to the bands, are more likely to obtain the largest $M_{n,N}$. Very flat or thin unit cells, are less likely to be selected, because $hk\ell$s are generated in an isotropic manner, depending on only their $d$-values.

Although these properties of $M_{n,N}$ are considered to be empirically reasonable, the preference does not have firm grounds, especially in the case of low-symmetric cells.

5 Computational results and discussion

The proposed method was implemented using C++ code, and applied to the analysis of dynamically simulated Kikuchi patterns (Figures 1–3 in Appendix B) and experimental patterns (Figures 7–8). The program was run on an Intel Core i7-5930k CPU (3.50 GHz) without parallel computation. The results are presented in Tables 2 and 3.

The quick search and exhaustive search were carried out by using the algorithm shown in Table 1 with/without the check (*), respectively. The search parameters commonly used for the test, are listed in Table 1 of Appendix A.

The simulated patterns were created by using the Bruker’s commercial software DynamicS [Winkelmann et al., 2007]. The parameters used for the simulation are presented in Table 1 of Appendix B. In particular, the used coordinates of the projection centers were exact.

The experimental EBSD patterns by using an SEM-EBSD system (JEOL 7001F-EDAX DigiView camera) with the $20kV$ electron-accelerating voltage and the beam current up to $14nA$. However, with regard to these experimental data prepared prepared this time, precise projection centers were also
available. (the second author used the pattern matching technique to obtain them (see Nolze et. al. (2017); this technique requires pattern simulation based on the phase information.) This is mainly because our study aims to develop a reliable method and software for ab-initio indexing. Improvement of automatic band detection and projection-center determination are also studied by different research groups for their own demands.

The part of band detection and ab-initio indexing was carried out by the first author’s laboratory without prior information. In particular, the positions and widths of the Kikuchi bands were extracted manually, because satisfactory results could not be obtained by automatic detection.

In what follows, after showing the results for the precise projection centers, the results for imprecise projection centers are also presented. The software also refines the projection center, although there is ambiguity in determination of $\Delta_z$, as mentioned in the paragraph following Eq.(7).

Firstly, if comparatively precise projection centers are used, except for the triclinic case (among 5 test data), the program succeeded in acquisition of the correct cell and ranking it to the top among the solutions of the same Bravais-type. However, the obtained unit-cell parameters has 5-10 % propagated errors, when they are compared to the literature values.

In particular, owing to the inaccuracy of the band widths, the $M_{n,N}$ values in Table 3 are smaller than those in Table 2. As for the cementite sample, a rather small value $M_{n,N} = 4.33$ was gained when the bandwidths were used, although the solution was almost identical to the result obtained without band widths. With regard to the Silico-ferrite sample, the difference from the literature values was increased by using the bandwidths.

Figure 7: Band positions and widths extracted from an experimental pattern of Cementite ($1040 \times 1040$ px$^2$)

Firstly, if comparatively precise projection centers are used, except for the triclinic case (among 5 test data), the program succeeded in acquisition of the correct cell and ranking it to the top among the solutions of the same Bravais-type. However, the obtained unit-cell parameters has 5-10 % propagated errors, when they are compared to the literature values.

In particular, owing to the inaccuracy of the band widths, the $M_{n,N}$ values in Table 3 are smaller than those in Table 2. As for the cementite sample, a rather small value $M_{n,N} = 4.33$ was gained when the bandwidths were used, although the solution was almost identical to the result obtained without band widths. With regard to the Silico-ferrite sample, the difference from the literature values was increased by using the bandwidths.
Figure 8: Band positions and widths extracted from an experimental pattern of Silico Ferrite of Ca & Al (1040 × 1040 px²)
Table 2: Comparison of the optimum solutions and the unit cells in the literature when bandwidths are not used†

| Number of used bands | Success/Failure in quick search (time (sec.), \(M_{n,N}\)) | Success/Failure in exhaustive search (time (sec.), \(M_{n,N}\)) | \(a/c\) | \(b/c\) | \(\alpha\) (deg.) | \(\beta\) (deg.) | \(\gamma\) (deg.) |
|----------------------|--------------------------------------------------------|---------------------------------------------------------------|------|------|-----------------|----------------|-----------------|
| Ni (simulated, cubic(F), \(a/c = b/c = 1\)): 20 | S (7.40, \(M_{20,41} = 46.60\)) | (243.83, \(M_{20,41} = 47.56\)) | 1 | 1 | 90 | 90 | 90 |
| Fe (simulated, cubic(I), \(a/c = b/c = 1\)): 23 | S (11.78, \(M_{23,57} = 57.73\)) | (426.62, \(M_{23,58} = 59.03\)) | 1 | 1 | 90 | 90 | 90 |
| Zn (simulated, hexagonal, \(a/c = b/c = 0.5387\)): 23 | S (3.17, \(M_{23,107} = 39.47\)) | (144.68, \(M_{23,105} = 39.85\)) | 0.540 | 0.540 | 90 | 90 | 90 |
| Spheroidal cementite (experimental, orthorhombic (F), \(a/c = 0.6711, b/c = 0.7546\)): 19 | F (0.91) | (20.03, \(M_{19,27} = 16.73\)) | 0.663 | 0.745 | 90 | 90 | 90 |
| Silico-ferrite of calcium and aluminum (experimental, triclinic, \(a/c = 0.881, b/c = 0.897, \alpha = 94.11, \beta = 111.4, \gamma = 110.3\)): 21 | F (0.77) | (53.67, \(M_{21,347} = 7.75\)) | 0.764 | 0.829 | 91.37 | 100.36 | 102.29 |

†The pattern-center shift \(\Delta_n\) in the direction perpendicular to the screen was fixed to 0 for low-symmetric cells, considering the ambiguity mentioned in the paragraph following Eqs. (5) and (6).

Table 3: Comparison of the optimum solutions and the unit cells in the literature when bandwidths are used

| Number of used bands | Success/Failure in quick search (time (sec.), \(M_{n,N}\)) | Success/Failure in exhaustive search (time (sec.), \(M_{n,N}\)) | \(a\) (\(\AA\)) | \(b\) (\(\AA\)) | \(c\) (\(\AA\)) | \(\alpha\) (deg.) | \(\beta\) (deg.) | \(\gamma\) (deg.) |
|----------------------|--------------------------------------------------------|---------------------------------------------------------------|------|------|----------------|----------------|----------------|----------------|
| Ni (simulated, cubic(F), \(a = b = c = 3.516(\AA)\)): 20 | S (6.15, \(M_{20,41} = 22.62\)) | (243.13, \(M_{20,39} = 23.00\)) | 3.397 | 3.397 | 3.397 | 90 | 90 | 90 |
| Fe (simulated, cubic(I), \(a = b = c = 2.866(\AA)\)): 23 | S (13.82, \(M_{23,57} = 37.78\)) | (426.15, \(M_{23,57} = 38.13\)) | 2.805 | 2.805 | 2.805 | 90 | 90 | 90 |
| Zn (simulated, hexagonal, \(a = b = 2.065, c = 4.947(\AA)\)): 23 | S (2.63, \(M_{23,103} = 24.29\)) | (144.41, \(M_{23,103} = 24.29\)) | 2.567 | 2.567 | 4.706 | 90 | 90 | 120 |
| Spheroidal cementite (experimental, orthorhombic (F), \(a = 4.526, b = 5.089, c = 6.744(\AA)\); Gardin (1962)): 19 | F (0.25) | (20.14, \(M_{19,262} = 4.33\)) | 4.122 | 4.659 | 6.245 | 90 | 90 | 90 |
| Silico-ferrite of calcium and aluminum (experimental, triclinic, \(a = 10.40, b = 10.59, c = 11.81(\AA)\), \(\alpha = 94.11, \beta = 111.4, \gamma = 110.3(\text{deg.)}\)): 21 | F (1.01) | F (48.49) |

§The unit-cell parameters obtained by the Rietveld refinement of X-ray diffraction data (a little distinct from the literature values in Takayama et. al. (2018), owing to the different composition).

It is known that the de Wolff \(M_n\) attains a large value (e.g., > 10) only for very plausible solutions, and does not exceed 3 for invalid solutions. From the values in Tables 2–4, it can be seen that the generalized \(M_{n,N}\) also has this property. The following are the other well-known properties of the de Wolff \(M_n\):

(a) \(M_n\) is sensitive to the existence of reflections observed but not computed from the model, and insensitive to the reflections computed but not observed in the pattern, because of the asymmetric definition of \(M_n\) with regard to the observed and calculated reflection sets.

(b) If unit cells with almost identical parameters but distinct Bravais types are compared, the higher-symmetric cell attains a larger \(M_n\), because the peak overlap caused by the symmetry make the number of computed reflections smaller.

According to its definition, \(M_{n,N}\) also has the property (a), which is desirable for use in EBSD indexing, because there are a number of computed but not observed bands in EBSD patterns. However, \(M_{n,N}\) does not possess the property (b), as seen from the values in Table 4, because band overlapping of \(m(hk\ell)\) (\(m \neq 0\): integer) occurs regardless of the symmetry, although it is straightforward to change the definition in an ad-hoc way so that \(M_{n,N}\) also has the property (b).

Although this is a kind of heuristic, the cell with the highest symmetry frequently provides the true solution. In addition, even if the current \(M_{n,N}\) is used, the users can easily find plausible solutions with the highest-symmetry, just by checking the output of the software as in Table 4. This is the reason we did not modify the definition of \(M_{n,N}\) this time, in order to gain the property (b).
As seen from the output in Table 4, the figures of merit also work well to judge which Bravais type is the true. For example, in the case of Ni, the cubic (P, I) solutions obtained much smaller $M_{n,N}$ values, compared to those provided to the cubic(F) solutions. The same thing is observed, when the cubic(P, F) and cubic(I) solutions for the Fe pattern, and the hexagonal and trigonal solutions for the Ni pattern are compared. Considering that all the derivative lattices can index the same band positions and widths, this seems to be owing to the heuristics adopted to generate computed reflections, as described in the last two paragraphs of Section 4.
Table 4: The maximum $M_{n,N}$ values of each Bravais type attained in the exhaustive search; in all the cases, the maximum $M_{n,N}$ value is provided to relaxed parameters of the correct unit cell, because more degrees of freedom is given to lower-symmetric cells in the fitting process. However, the correct Bravais type can be estimated even so, just by checking which Bravais type is provided to a comparatively large $M_{n,N}$ value and has the highest symmetry.

| Bravais type      | Ni (cubic (F)) | Fe (cubic (I)) | Zn (hexagonal) | Cementite (orthorhombic (P)) |
|-------------------|---------------|----------------|----------------|-------------------------------|
|                   | Ni Not used   | Ni Used        | Fe Not used    | Fe Used                      | Zn Not used | Zn Used | Zn Not used | Zn Used | Cementite Not used | Cementite Used |
| Triclinic         | 51.54         | 23.61          | 54.36          | 34.58                        | 42.93       | 23.95   | 20.75       | 4.33    | 20.75           | 4.33            |
| Monoclinic(P)     | < 36          | < 12           | < 44           | < 15                         | 39.70       | 24.00   | 20.10       | 4.33    | 16.73           | 4.33            |
| Monoclinic(C)     | 53.90         | 23.27          | 57.35          | 37.31                        | 42.49       | 24.47   | < 10        | < 3     | < 10            | < 3             |
| Orthorhombic(P)   | < 36          | < 12           | < 44           | < 15                         | < 30        | < 12    | 16.73       | 4.33    | < 10            | < 3             |
| Orthorhombic(C)   | < 36          | < 12           | < 44           | < 15                         | 40.08       | 24.14   | < 10        | < 3     | < 10            | < 3             |
| Orthorhombic(I)   | 53.01         | 23.26          | 52.03          | 33.99                        | < 30        | < 3     | < 10        | < 3     | < 10            | < 3             |
| Orthorhombic(F)   | 50.63         | 22.96          | 57.15          | 36.74                        | < 3         | < 3     | < 3         | < 3     | < 3             | < 3             |
| Tetragonal(P)     | < 36          | < 12           | < 44           | < 15                         | < 30        | < 12    | < 3         | < 3     | < 3             | < 3             |
| Tetragonal(I)     | 52.33         | 23.38          | 54.55          | 37.85                        | < 30        | < 3     | < 3         | < 3     | < 3             | < 3             |
| Trigonal          | 46.31         | 22.81          | 59.14          | 37.96                        | < 30        | < 12    | < 3         | < 3     | < 3             | < 3             |
| Hexagonal         | < 3           | < 3            | < 44           | < 15                         | 39.85       | 24.29   | < 3         | < 3     | < 3             | < 3             |
| Cubic(P)          | < 36          | < 12           | < 44           | < 15                         | < 3         | < 3     | < 3         | < 3     | < 3             | < 3             |
| Cubic(I)          | < 3           | < 3            | 59.03          | 38.13                        | < 3         | < 3     | < 3         | < 3     | < 3             | < 3             |
| Cubic(F)          | 47.56         | 23.00          | < 44           | < 3                          | < 3         | < 3     | < 3         | < 3     | < 3             | < 3             |

Table 5 is the indexing result for the Cementite sample. The bandwidths assigned the Miller indices (022) and (022) were probably due to systematic absence, considering that (011) and (011) are forbidden by the rules (0kl with an odd $k$) of $P\ b\ n\ m$ (No.62). However, influence of non-visible narrowest bandwidths (and also underestimation of the unit-cell scale) is also observed from the assigned Miller indices (e.g., (006) and (330)). In the simulated patterns, similar cases were not detected, and all the input bandwidths were the narrowest.
Table 5: Indexing result for the cementite pattern; the observed/calculated band positions \((X,Y)\) and bandwidths \(\beta\) are compared (the following values have no unit, since the camera lengths is set to 1).

| Miller index | \((X^{\text{cal}}, Y^{\text{cal}})\) | \((X^{\text{obs}}, Y^{\text{obs}})\) | distance between \((X^{\text{cal}}, Y^{\text{cal}})\) and \((X^{\text{obs}}, Y^{\text{obs}})\) | \(\beta^{\text{cal}}\) | \(\beta^{\text{obs}}\) |
|--------------|-------------------------------|-------------------------------|---------------------------------|----------------|----------------|
| -2 -3 3      | -0.3951                       | -0.0993                       | 0.3948                          | 0.0008         | 0.9937         | 0.8992         |
| † 0 0 6      | 0.0655                        | -0.2868                       | -0.0649                         | 0.2885         | 0.9869         | 0.8973         |
| -3 -3 0      | -0.4663                       | 0.1836                        | -0.4653                         | 0.1876         | 0.9942         | 0.9894         |
| -1 0 -3      | 0.0090                        | 0.1526                        | 0.0088                          | 0.1499         | 0.9942         | 0.9401         |
| -2 -3 -3     | -0.0600                       | 0.0654                        | -0.0615                         | 0.0665         | 0.9919         | 0.8083         |
| -1 0 3      | -0.3670                       | -0.7893                       | -0.3626                         | -0.7930        | 0.0058         | 0.8130         | 0.7579         |
| 0 4 -3       | 0.1443                        | 0.0584                        | 0.1477                          | 0.0591         | 0.0034         | 0.8481         | 0.8000         |
| -1 4 0       | 0.0775                        | -0.0141                       | 0.0790                          | -0.0154        | 0.0021         | 0.0767         | 0.0739         |
| -1 4 0       | 0.6853                        | -0.0091                       | 0.6879                          | -0.0069        | 0.0034         | 0.1126         | 0.1118         |
| 1 2 5       | 0.1794                        | -0.7803                       | 0.1845                          | -0.7818        | 0.0053         | 0.1328         | 0.1331         |
| 1 2 5       | 0.0536                        | -0.1553                       | 0.0532                          | -0.1549        | 0.0006         | 0.0826         | 0.0816         |
| 0 4 3       | 0.3773                        | -0.2902                       | 0.3747                          | -0.2907        | 0.0026         | 0.1035         | 0.0990         |
| -1 -2 5     | -0.2421                       | -0.2327                       | -0.2415                         | -0.2352        | 0.0026         | 0.0896         | 0.1095         |
| -1 -2 1     | -0.2346                       | -0.0082                       | -0.2334                         | -0.0072        | 0.0016         | 0.0467         | 0.0526         |
| †0 2 -2    | 0.0870                        | 0.0486                        | 0.0894                          | 0.0482         | 0.0025         | 0.0462         | 0.0440         |
| -1 2 -5     | 0.0770                        | 0.1222                        | 0.0758                          | 0.1206         | 0.0020         | 0.0820         | 0.0872         |
| 0 2 5       | 0.1253                        | -0.4276                       | 0.1219                          | -0.4292        | 0.0038         | 0.0933         | 0.0874         |
| -1 -1 3     | -0.4078                       | -0.3758                       | -0.4045                         | -0.3765        | 0.0033         | 0.0649         | 0.0741         |
| †0 2 2    | 0.3381                        | -0.3489                       | 0.3388                          | -0.3445        | 0.0045         | 0.0567         | 0.0617         |

\(^{†}\)From the reflection rules of \(Pb\) \(n\) \(m\) (No. 62), \{00\} \((t: \text{odd})\) may be excluded.

\(^{c}\)\{01\̅\} and \{011\} were forbidden by the reflection rules \((0kl\text{ with an odd }k)\) of \(Pb\) \(n\) \(m\) (No. 62).

Lastly, since two types of ambiguities are pointed out so far, use of the whole information in a single EBSD pattern is mentioned for future studies; the ambiguity caused by the derivative lattice, is mainly caused by sublattices \(M\) of the true crystal lattice \(L\) with a small index \([L : M]\) when bandwidths are used.

Conversely, if HOLZ rings are used, the roles of sublattices and superlattices are exchanged, because each radius of a HOLZ ring provides an integer multiple of the shortest length of the lattice-vector with the same direction as the corresponding zone axis [Michael and Eades, 2000]. In this case, ambiguity is mainly caused by superlattices \(M\) of \(L\) with small \([M : L]\).

This situation indicates that combinatorial use of the bandwidths and the Holz rings may work as a simple solution for the ambiguity due to the derivative lattices.

6 Conclusion

For ab-initio indexing, a new method based on distribution rules of systematic absence and error-stable Bravais lattice determination was proposed. In addition, the de Wolff figures of merit for 2D images and data of multiple dimension were defined for use in orientation determination and ab-initio indexing of Kikuchi patterns. However, erroneous band widths can cause ambiguity of solutions in particular in case of low-symmetric cells.

The new figures of merit have properties similar to those of the original de Wolff \(M_n\), except for the preference for higher-symmetric cells. From the software’s output, users can efficiently find the optimal solution and Bravais type.

Acknowledgments This study was financially supported by the PREST (JPMJPR14E6). We would like to extend our gratitude to Dr. A. Esmaeili, Ms. T. Ueno of Yamagata University and Mr. G. Sebastian of Kyushu University, who helped us in coding the software, preparing the input files, and performing the computation.
References

[Conway, 2006] Conway, J. H. (2006). *The sensual (quadratic) form*. Carus Mathematical Monographs 26, Mathematical Association of America.

[Day, 2008] Day, A. P. (2008). Spherical ebsd. *Journal of Microscopy*, 230:472–486.

[de Wolff, 1968] de Wolff, P. M. (1968). A simplified criterion for the reliability of a powder pattern indexing. *J. Appl. Cryst.*, 1:108–113.

[Dingley and Baba-Kishi, 1986] Dingley, D. J. and Baba-Kishi, K. (1986). *Scanning Electron Microscopy*, 2:383–391.

[Dingley and Wright, 2009] Dingley, D. J. and Wright, S. I. (2009). Determination of crystal phase from an electron backscatter diffraction pattern. *J. Appl. Cryst.*, 42:234–241.

[Esmaeili et al., 2017] Esmaeili, A., Kamiyama, T., and Oishi-Tomiyasu, R. (2017). New functions and graphical user interface attached to powder indexing software conograph. *J. Appl. Cryst.*, 50:651–659.

[Field, 1997] Field, D. P. (1997). Recent advances in the application of orientation imaging. *Ultramicroscopy*, 67:1–9.

[Gardin, 1962] Gardin, A. I. (1962). An electron-diffraction study of the structure of cementite. *Soviet Physics-Crystallography (Kristallografiya)*, 7:694–700.

[Grosse-Kunstleve et al., 2004] Grosse-Kunstleve, R. W., Sauter, N. K., and Adams, P. D. (2004). Numerically stable algorithms for the computation of reduced unit cells. *Acta Cryst.*, A60:1–6.

[Ito, 1949] Ito, T. (1949). A general powder X-ray photography. *Nature*, 164:755–756.

[Kogure, 2003] Kogure, T. (2003). Computer program for beginners to assist interpretation of electron diffraction patterns (japanese). *Short Review & Scientific Communication*, 32:96–101.

[Langer and Dähritz, 2007] Langer, E. and Dähritz, S. (2007). Investigation of holz rings in ebsd patterns. *Physica Status Solidi (c)*, 6:1867–1872.

[Li and Han, 2015] Li, L. and Han, M. (2015). Determining the bravais lattice using a single electron backscatter diffraction pattern. *J. Appl. Cryst.*, 48:107–115.

[Li et al., 2014] Li, L., Ouyang, S., Yang, Y., and Han, M. (2014). Ebsdll: a computer program for determining an unknown bravais lattice using a single electron backscatter diffraction pattern. *J. Appl. Cryst.*, 47:1466–1468.

[Marthinsen and Høier, 1988] Marthinsen, K. and Høier, R. (1988). On the breakdown of friedel’s law in electron backscattering channelling patterns. *Acta Cryst.*, A44:700–707.

[Michael, 2000] Michael, R. (2000). *Electron Backscatter Diffraction in Materials Science*. New York: Kluwer Academic/Plenum Publishers.

[Michael and Eades, 2000] Michael, R. and Eades, J. A. (2000). Use of reciprocal lattice layer spacing in electron backscatter diffraction pattern analysis. *Ultramicroscopy*, 81:67–81.

[Nolze et al., 2015] Nolze, G., Grosse, C., and Winkelmann, A. (2015). Kikuchi pattern analysis of noncentrosymmetric crystals. *J. Appl. Cryst.*, 48:1405–1419.

[Nolze et al., 2017] Nolze, G., Hielscher, R., and Winkelmann, A. (2017). Electron backscatter diffraction beyond the mainstream. *Crystal Research & Technology*, 52:1600252.

[Nolze and Winkelmann, 2017] Nolze, G. and Winkelmann, A. (2017). Crystallographic and projective properties of kikuchi diffraction patterns. *J. Appl. Cryst.*, 50:102–119.

[Oishi-Tomiyasu, 2012] Oishi-Tomiyasu, R. (2012). Rapid bravais-lattice determination algorithm for lattice parameters containing large observation errors. *Acta Cryst.*, A68:525–535.
[Oishi-Tomiyasu, 2013] Oishi-Tomiyasu, R. (2013). Distribution rules of systematic absences on the conway topograph and their application to powder auto-indexing. *Acta Cryst.*, A69:603–610.

[Oishi-Tomiyasu, 2016] Oishi-Tomiyasu, R. (2016). A table of geometrical ambiguities in powder indexing obtained by exhaustive search. *Acta Cryst.*, A72:73–80.

[Takayama et al., 2018] Takayama, T., Murao, R., and M., K. (2018). Quantitative analysis of mineral phases in iron-ore sinter by the rietveld method of x-ray diffraction patterns. *ISIJ Int.*, 58:1069–1078.

[Venables and Harland, 1973] Venables, J. A. and Harland, C. J. (1973). *Phil. Mag.*, 27:1193–1200.

[Winkelmann et al., 2007] Winkelmann, A., Trager-Cowan, C., Sweeney, F., Day, A. P., and P., P. (2007). Many-beam dynamical simulation of electron backscatter diffraction patterns. *Ultramicroscopy*, 107:414–421.

[Wright and Adams, 1992] Wright, S. I. and Adams, B. L. (1992). Automatic analysis of electron backscatter diffraction patterns. *Metallurgical and Materials Transactions A*, 23:759–767.

[Wu, 1988] Wu, E. (1988). A modification of the de wolff figure of merit for reliability of powder pattern indexing. *J. Appl. Cryst.*, 21:530–535.
A  Search parameters and experimental patterns used in testing the software

| Table 1: Search parameters |
|-----------------------------|
| # The input phi, sigma are assumed to contain errors within this range (degree). | 1.0 |
| # The unit-cell scales s1, s2 computed from the bandwidths are supposed to be equal, if both s1<=s2*(this.value) and s2<=s1*(this.value) hold. | 3.0 |
| # (Used only for Bravais lattice determination) squared lattice-vector lengths | 0.05 |
| # | |
| # | |
| | |
| | |
| | |
| # (Used only for Bravais lattice determination) squared lattice-vector lengths | 1 |
| # | 1 1 0 |
| # Only the solutions with the figure of merit larger than this value is output. | 3.0 |

a The Miller indices are generated in descending order of d-values.

b These flag was set to 1 1 0 in the analyses without using the bandwidths.

B  Simulated patterns and the parameters used for the simulation

Figures 1–3 present the band positions and widths extracted from simulated EBSD patterns and used for testing the software. The parameters used for the simulation are as follows:
Table 2: Calculation settings used by DynamicS software

| Sample | AV (kV) | $d_{min}$ (Å) | $I_{min}$ (%) | Absorption length (Å)$^d$ | Exitation depth (Å)$^e$ | Debye-Waller B: Crystal ($Å^2$)$^f$ | Debye-Waller B: Source ($Å^2$)$^g$ |
|--------|---------|---------------|---------------|---------------------------|-------------------------|--------------------------------------|--------------------------------------|
| Ni     | 20      | 0.5           | 15            | 54                        | 42                      | 0.74                                 | 0.3                                  |
| Fe     | 20      | 0.5           | 15            | 61                        | 48                      | 0.79                                 | 0.3                                  |
| Zn     | 20      | 0.5           | 15            | 68                        | 54                      | 0.83                                 | 0.3                                  |

$^a$ acceleration voltage of electron beam,
$^b$ minimum lattice spacing of the reflecting lattice planes,
$^c$ minimum intensity of a reflector relative to the strongest reflector,
$^d$ IMFP (Inelastic Mean Free Path) of the electrons extrapolated from the NIST Standard Reference Database,
$^e$ mean depth of backscattering,
$^f$ Debye-Waller factor of the crystal,
$^g$ Debye-Waller factor of the source.

Figure 1: Band positions and widths extracted from a simulated pattern of Ni (956 × 956 px$^2$)
Figure 2: Band positions and widths extracted from a simulated pattern of Fe (1600 × 1152 px²).

Figure 3: Band positions and widths extracted from a simulated pattern of Zn (956 × 956 px²).
C De Wolff figure of merit and its generalization

As a figure of merit used for powder indexing, the de Wolff $M$ is defined as follows:

$$M_n = \bar{\epsilon} / \delta,$$

where $\bar{\epsilon}$ and $\delta$ are the average discrepancy and the actual discrepancy, respectively, defined by:

$$\bar{\epsilon} := Q_{n}^{obs} / 2N, \quad (A.1)$$

$$\delta := \frac{1}{n} \sum_{i=1}^{n} |Q_i^{obs} - Q_i^{cal}|, \quad (A.2)$$

$Q_i^{cal}$: computed line closest to the observed line $Q_i^{obs}$.

In order to generalize the definition of $M$ to the sets of points in $\mathbb{R}^s$, it is important to understand why the average discrepancy is defined as (A.1), when the actual discrepancy is defined as (A.2); as explained in Wu (1988), if $q_0 < q_1 < \cdots < q_N$ is specified, the average discrepancy with regard to the $q_i$’s, is defined as:

$$\epsilon_W := \frac{1}{4q_N} \sum_{k=1}^{N} (q_k - q_{k-1})^2, \quad (A.3)$$

This $\epsilon_W$ equals to the mean value of the distance of $Q$ from its nearest $q_k$, where $Q$ is assumed to be uniformly distributed in the interval $[0, q_N]$.

If the computed lines $q_1, \ldots, q_{N-1}$ are also assumed to be uniformly distributed under the constraint $0 < q_1 < \cdots < q_{N-1} < q_N$, the mean value of $\epsilon_W$ is given by:

$$\frac{1}{4q_N} \int_0^{q_2} \cdots \int_0^{q_{N-1}} \sum_{k=1}^{N} (q_k - q_{k-1})^2 dq_1 \cdots dq_{N-1} = \frac{q_N}{2(N+1)}. \quad (A.4)$$

The de Wolff average discrepancy $\bar{\epsilon} := Q_{n}^{obs} / 2N$ gives a good approximation of the formula (A.4).

In order to generalize this to a general space of dimension $s$, we consider $x_1, \ldots, x_N$ and $X_{1}^{obs}, \ldots, X_{n}^{obs}$ as points in some fixed domain $\Omega \subset \mathbb{R}^s$. In this case, it is difficult to provide a formula for the average discrepancy for the specified $x_i$’s, as in (A.3). However, it is possible to estimate the mean value as in (A.4), if $x_1, \ldots, x_N$ also run over $\Omega$ with uniform probability.

Suppose that $X$ is a coordinate in $\Omega$. If $x_1, \ldots, x_{N-1}$ are uniformly distributed in $\Omega$, and $x_N$ is uniformly distributed on the boundary of $\Omega$, we have:

$$\text{Prob} \left( \min_{1 \leq i \leq N} |X - x_i| > r \right) = \left( 1 - \frac{\text{Vol} (\Omega \cap B_r(X))}{\text{Vol}(\Omega)} \right)^{N-1} \left( 1 - \frac{\text{Area} (\partial \Omega \cap B_r(X))}{\text{Area}(\partial \Omega)} \right),$$

where $B_r(X)$ is the ball with radius $r$ and center at $X$ of dimension $s$. $\partial \Omega$ is the surface of $\Omega$. The probability density function of $\min_{1 \leq i \leq N} |X - x_i|$ is then given by:

$$\text{Prob} \left( \min_{1 \leq i \leq N} |X - x_i| = r \right) = -\frac{d}{dr} \text{Prob} \left( \min_{1 \leq i \leq N} |X - x_i| > r \right).$$


Hence, if $X$ is also uniformly distributed in $\Omega$, the mean value of $\min_{1 \leq i \leq N} \{|X - x_i|\}$ is given by:

\[
E \left[ \min_{1 \leq i \leq N} \{|X - x_i|\} \right] = \int_0^\infty \frac{r}{\text{Vol}(B_R(0))} \left( \int_{|X| \leq R} \text{Prob} \left( \min_{1 \leq i \leq N} \{|X - x_i|\} = r \right) dX \right) dr
\]

\[
= -\frac{1}{\text{Vol}(B_R(0))} \int_0^\infty \frac{d}{dr} \left( \int_{|X| \leq R} \text{Prob} \left( \min_{1 \leq i \leq N} \{|X - x_i|\} > r \right) dX \right) dr
\]

\[
= \frac{1}{\text{Vol}(B_R(0))} \int_{|X| \leq R} \left( \int_0^\infty \text{Prob} \left( \min_{1 \leq i \leq N} \{|X - x_i|\} > r \right) dr \right) dX.
\]

In what follows, we shall lead the formulas for the ball $\Omega = B_R(0)$ of dimension $s$.

(Case of $s = 1$)

\[
\text{Vol} \left( B_R(0) \cap B_r(X) \right) = \begin{cases} 
2R & \text{if } r \geq R + |X|, \\
2r & \text{if } r \leq R - |X|, \\
r + R - |X| & \text{if } R - |X| \leq r \leq R + |X|.
\end{cases}
\]

\[
\text{Area} \left( \partial B_R(0) \cap B_r(X) \right) = \begin{cases} 
2 & \text{if } r > R + |X|, \\
0 & \text{if } r < R - |X|, \\
1 & \text{if } R - |X| < r < R + |X|.
\end{cases}
\]

Hence,

\[
\int_0^\infty \text{Prob} \left( \min_{1 \leq i \leq N} \{|X - x_i|\} > r \right) dr = \int_0^{R - |X|} \left( 1 - \frac{r}{R} \right)^{N-1} dr + \frac{1}{2} \int_{R - |X|}^{R + |X|} \left( \frac{|X| - r + R}{2R} \right)^{N-1} dr
\]

\[
= \frac{R}{N} \left[ \left( 1 - \frac{R}{R} \right)^N \right]_0^{R - |X|} - \frac{R}{N} \left[ \left( \frac{|X| - r + R}{2R} \right)^N \right]_{R - |X|}^{R + |X|}
\]

\[
= \frac{R}{N} \left( \left( \frac{|X|}{R} \right)^N - 1 \right) + \frac{R}{N} \left( \frac{|X|}{R} \right)^N = \frac{R}{N}.
\]

Therefore,

\[
E \left[ \min_{1 \leq i \leq N} \{|X - x_i|\} \right] = \frac{R}{N} = \frac{\max_{1 \leq i \leq N} \{|x_i|\}}{N}.
\]
(Case of \( s \neq 1 \))

\[
\text{Vol} \left( \overline{B_R(0)} \cap B_r(X) \right) = \begin{cases} \frac{\pi^{s/2} R^s}{\Gamma \left( \frac{s}{2} + 1 \right)} \frac{\pi^{(s-1)/2} R^s}{\Gamma \left( \frac{s}{2} + 1 \right)} 
& \text{if } r \geq R + |X|, \\
\frac{\pi^{s/2} R^s}{\Gamma \left( \frac{s}{2} + 1 \right)} \frac{\pi^{(s-1)/2} R^s}{\Gamma \left( \frac{s}{2} + 1 \right)} 
& \text{if } r \leq R - |X|, \\
\int_0^1 \int_0^{1-r^2} (1 - t^2)^{(s-1)/2} dt & \text{if } \text{Vol} \left( \overline{B_R(0)} \cap B_r(X) \right) \end{cases}
\]

\[
\text{Area} \left( \partial B_R(0) \cap B_r(X) \right) = \begin{cases} \frac{2\pi^{s/2} R^{s-1}}{\Gamma \left( \frac{s}{2} + 1 \right)} 
& \text{if } r > R + |X|, \\
0 & \text{if } r < R - |X|, \\
\frac{2\pi^{(s-1)/2} R^{s-1}}{\Gamma \left( \frac{s}{2} + 1 \right)} \int_0^1 \int_0^{1-r^2} (1 - t^2)^{(s-3)/2} dt & \text{if } R - |X| \leq r \leq R + |X|. \end{cases}
\]

From geometrical considerations, we have:

\[
\\text{Vol} \left( \overline{B_{\frac{R-R+|X|}{2}}} \right) < \text{Vol} \left( \overline{B_R(0)} \cap B_r(X) \right) < \min \{ \text{Vol} \left( \overline{B_R(0)} \right), \text{Vol} \left( \overline{B_r(X)} \right) \},
\]

\[
0 < \text{Area} \left( \partial B_R(0) \cap B_r(X) \right) < \min \{ \text{Area} \left( \partial B_R(0) \right), \text{Area} \left( \partial B_r(X) \right) \}.
\]

Hence, we have:

\[
\int_0^\infty \text{Prob}_X \left( \min_{1 \leq i \leq N} \{|X - x_i| > r\} \right) dr > \int_0^{|X|} \left( 1 - \frac{r^s}{R^s} \right)^{N-1} dr \\
+ \int_{|X|-|X|}^R \left( 1 - \frac{r^s}{R^s} \right)^{N-1} \left( 1 - \frac{r^{s-1}}{R^{s-1}} \right) dr,
\]

\[
\int_0^\infty \text{Prob}_X \left( \min_{1 \leq i \leq N} \{|X - x_i| > r\} \right) dr < \int_0^{|X|} \left( 1 - \frac{r^s}{R^s} \right)^{N-1} dr \\
+ \int_{|X|-|X|}^{|X|} \left( 1 - \frac{R - |X| + r^s}{2s R^s} \right)^{N-1} dr.
\]

Using the variable transformations \( r = R y \) and \( r_2 = R y_2 \),

\[
\frac{1}{\text{Vol}(B_R(0))} \int_{|X| \leq R} \left( \int_0^{|X|} \left( 1 - \frac{r^s}{R^s} \right)^{N-1} dr \right) dX = \frac{s}{R^s} \int_0^R \int_0^{r_2} \left( 1 - \frac{r^{s-1}}{R^{s-1}} \right)^{N-1} dr_2 \\
= s R \int_0^1 \int_0^{y_2} (1 - y^s)^{N-1} (1 - y_2^s)^{N-1} dy dy_2 \\
= s R \int_0^1 (1 - y^s)^{N-1} (1 - y_2^s) dy_2 \\
= R \int_0^1 (1 - y^s)^{N-1} dy.
\]
E \left[ \min_{1 \leq i \leq N} |X - x_i| \right]

As a result, the lower and upper bounds of \( E [\min_{1 \leq i \leq N} |X - x_i|] \) are obtained:

\[
\begin{align*}
\frac{1}{\text{Vol}(B_R(0))} & \int_{|X| \leq R} \left( \int_{R-|X|}^R \left( 1 - \frac{r^s}{R^s} \right)^{N-1} \left( 1 - \frac{r^{s-1}}{R^{s-1}} \right) dr \right) dX \\
& = \frac{s}{R^s} \int_0^R r_2^{-1} \left( \int_{R-r_2}^R \left( 1 - \frac{r^s}{R^s} \right)^{N-1} \left( 1 - \frac{r^{s-1}}{R^{s-1}} \right) dr \right) dr_2 \\
& = sR \int_0^1 y_2^{s-1} \left( \int_{1-y_2}^1 (1 - y^s)^{N-1} (1 - y^{s-1}) dy \right) dy_2 \\
& = sR \int_0^1 (1 - y^s)^{N-1} (1 - y^{s-1}) \left( \int_{1-y}^1 y_2^{s-1} dy_2 \right) dy \\
& = R \int_0^1 (1 - y^s)^{N-1} (1 - y^{s-1}) \left( 1 - (1 - y)^s \right) dy.
\end{align*}
\]

\[
\begin{align*}
\frac{1}{\text{Vol}(B_R(0))} & \int_{|X| \leq R} \left( \int_{R-|X|}^{R+|X|} \left( 1 - \frac{(R - |X| + r)^s}{2^s R^s} \right)^{N-1} dr \right) dX \\
& = \frac{s}{R^s} \int_0^R r_2^{-1} \left( \int_{R-r_2}^{R+r_2} \left( 1 - \frac{(R - r_2 + r)^s}{2^s R^s} \right)^{N-1} dr \right) dr_2 \\
& = sR \int_0^1 y_2^{s-1} \left( \int_{1-y_2}^1 (1 - (1 - y_2 + y)^s)^{N-1} dy \right) dy_2 \\
& = 2sR \int_0^1 (1 - y_3^{s-1}) \left( \int_{1-y_3}^1 y_2^{s-1} dy_2 \right) dy_3 \quad (y_3 = \frac{1 - y_2 + y}{2}) \\
& = 2sR \int_0^1 (1 - y_3^{s-1}) \left( 1 - (1 - y_3)^s \right) dy_3.
\end{align*}
\]
\[
E \left[ \min_{1 \leq i \leq N} \{|X - x_i|\} \right]
\]
\[
< R \int_0^1 (1 - y)^{N-1} (2 - (1 - y)^s) dy
\]
\[
= R \int_0^1 (1 - y)^{N-1} dy - R \sum_{i=1}^s \binom{s}{i} (-1)^i \int_0^1 (1 - y)^{N-1} y^i dy
\]
\[
= \frac{R}{s} \int_0^1 (1 - z)^{N-1} z^{(1-s)/s} dz - \frac{R}{s} \sum_{k=1}^s \binom{s}{k} (-1)^k \int_0^1 (1 - z)^{N-1} z^{(1-s+k)/s} dz \quad (z = y^s)
\]
\[
= \frac{R}{s} B(N,1/s) - \frac{R}{s} \sum_{k=1}^s \binom{s}{k} (-1)^k B(N,(k+1)/s).
\]

From \( \lim_{n \to \infty} \Gamma(n)n^z/\Gamma(n + z) = 1 \), the beta function \( B(x,y) = \Gamma(x)\Gamma(y)/\Gamma(x + y) \) has the asymptotic formula \( B(x,y) \sim \Gamma(y)x^{-y} \). As a result, we have:

\[
E \left[ \min_{1 \leq i \leq N} \{|X - x_i|\} \right] \sim \frac{\Gamma(1/s)}{s} \frac{R}{N^{1/s}} \quad (N \to \infty).
\]

In particular, the formulas for \( s = 2, 3 \) are as follows:

(Case of point configurations in a 2D ball of radius \( R \))

\[
E \left[ \min_{1 \leq i \leq N} \{|X - x_i|\} \right] \sim \frac{\sqrt{\pi}}{2} \frac{R}{\sqrt{N}} \quad (N \to \infty).
\]

(Case of point configurations in a 3D ball of radius \( R \))

\[
E \left[ \min_{1 \leq i \leq N} \{|X - x_i|\} \right] \sim \frac{\Gamma(1/3)}{3} \frac{R}{N^{1/3}} \quad (N \to \infty).
\]