A general analytical method for interface normal determination in TEM A general analytical method for interface normal determination in TEM

Rui-Xun Xie, Melvyn Larranaga, Frédéric Mompiou, Nicolas Combe, Wen-Zheng Zhang

HAL Id: hal-02910108
https://hal.science/hal-02910108

Submitted on 31 Jul 2020

Highlights

A general analytical method for interface normal determination in TEM

XIE Rui-Xun,LARRANAGA Melvyn,MOMPIOU Frédéric,COMBE Nicolas,ZHANG Wen-Zheng

- We present a general method to determine the interface normal from arbitrary orientations.

- We developed a general model of interface projection, including existing methods as special cases.

- The method is proved to be robust, efficient, and scalable in different situations.

A general analytical method for interface normal determination in TEM

XIE Rui-Xun[a], LARRANAGA Melvyn[b], MOMPIOU Frédéric[b], COMBE Nicolas[b] and ZHANG Wen-Zheng[a]

[a]*Key Laboratory of Advanced Materials (MOE), School of Materials Science and Engineering, Tsinghua University, Beijing 100084, PR China*
[b]*CEMES-CNRS, Université de Toulouse, 29 rue J. Marvig, 31055 Toulouse, France*

ARTICLE INFO

ABSTRACT

This paper presents a new analytical method to determine interface normals from a series of bright/dark field images taken from arbitrary orientations. This approach, based on a general geometrical model of interface projection, provides a generalized formulation of existing methods. It can treat an excessive number of inputs, i.e. orientation conditions. Given 6 or more sets of inputs, even with considerable experimental errors, we prove that this method is still very likely to yield results with satisfactory trueness. The robustness of the method can thus allow its implementation in problems dealing with a large amount of data. We show that this method can also be applied to determine 1D features or to check the planarity of microstructural features.

1. Introduction

Transmission electron microscopy (TEM) is a powerful tool to examine a variety of defects in materials, including point defects (zero dimension, 0D), dislocation lines (1D), interfaces (2D) and inclusions (3D). Since TEM images are 2D projections of the 3D space, the geometrical features of non-0D defects, e.g., their shapes, must be reconstructed from their projections on the observation plane. The identification of 3D inclusions, curved lines or curved planes is a cumbersome work [1][2]. However, many 1D or 2D defects, e.g., interfacial dislocations or faceted interfaces, often have preferred orientations. They can be considered as straight lines or flat planes. Accurately determining their crystallographic orientations is fundamental to unravel the mechanism of associated microstructure evolution [3]. In this context, the problem is restricted to the line direction or the interface plane normal determination, using a technique called trace analysis [4][9].

The concept and methods of trace analysis using TEM were clearly introduced in the well-known book by Hirsch et al. [4]. A line feature is always on the plane defined by its projection and the electron beam direction. Therefore it can be determined if observed from two directions. However, the projection of a planar feature usually contains the projection of two traces, the intersections of the interface with two foil surfaces, separated by a certain width. An interface can be directly determined at its edge-on condition, where the projection width is zero and the interface

ORCID(s):

normal lies in the viewing screen [4]. Since an edge-on condition often carries non-negligible uncertainty, several modifications were later proposed to improve the accuracy, such as the single edge-on method [9][5], the double edge-on method [10][11], and the trace & edge-on method [12]. Despite their good accuracy, these methods are not easy to use, as finding an exact edge-on condition is usually time consuming or even impossible if it is out of the tilting range of the sample holder.

An edge-on condition is not always necessary, if the interface contains an additional sharp line feature, such as a straight dislocation line [5][6][7] or an intersection with another plane [8], using the so-called double-trace method. The orientations of both trace and line feature in the plane can be measured by the trace analysis method of 1D features. Then one can obtain the interface normal by making a cross product of the line and the trace direction. However, additional line features are not always present, which limits the application of this method.

When the condition for the double-trace method is not satisfied, the projection width can be used to calculate interface normals at arbitrary orientations. In this approach, the trace direction is assumed to lie on the screen plane at zero tilt. The interface normal can be determined, once the inclination angle between the foil surface and the interface is determined. Traditionally, this was done by measuring the foil thickness and the projection width of the interface at an orientation near to zero tilt [9]. Since this method contains considerable uncertainty, Zhang and Kelly [13] made an improvement by tilting the sample along the trace direction. But it is usually difficult to ensure the tilting axis exactly parallel with the trace direction. Qiu an Zhang [14] solved this problem on a single tilt holder by taking the angle between the trace and the tilting axis into consideration. In order to avoid ambiguous solutions, this approach still needs to track the trend of the projection width change during sample tilting. All these methods are based on the assumption that the upper and lower foil surfaces are both perpendicular to the electron beam direction at zero tilt. This may introduce systematic error when the foil has thickness variation or not flat.

Above methods are highly sensitive to experimental errors, since they rely on limited information — usually one or two sets of inputs (beam and projected trace directions, projection width). Particularly, the error could be greatly amplified by the cross product operation in double trace and double edge-on methods [12]. To improve the reliability, one may use excessive experimental data to calculate several solutions of the interface normal. The scattering of results can be plotted in a pole figure [12][14], but the selection of the final result and the estimation of its uncertainty are largely based on the operator's experience.

In order to improve the accuracy of interface normal determination and simplify the TEM operation, we present here a close-form algorithm to optimize the result with multiple sets of inputs from arbitrary orientations. In this method, there is no specific requirements on the beam direction, nor conditions on the interface and the foil.

In Section 2, we will describe the methodology used to compute the interface normal. Section 3 will present experimental details, followed by Section 4 illustrating the applications of the present method. The results and method will finally be discussed in Section 5. The accuracy of the results will be adressed by an error analysis, and the method will be compared with other methods proposed in the literature. A generalization of our approach will be proposedat the end.

2. Geometrical model of interface projection

The approach presented below is a purely geometrical analysis of the orthogonal projection of planar features on the observation screen, without any requirement on the beam direction (projection direction) or the foil surface. It is based on the measurement of the interface width and apparent trace direction for different beam direction. The symbols used in the derivation and their meaning are listed in Table 1.

Table 1
Definition of symbols [1]

Symbol	Definition
b_e	Reversed electron beam direction, unit vector
n	Interface normal, unit vector
r	Arbitrary vector in the screen plane, unit vector, $b_e \cdot r = 0$
s	Foil surface normal, unit vector
t	Trace direction, i.e., direction of the intersection between the interface and the foil surface, unit vector $t = n \times s / \|n \times s\|$
t_p	Projected direction of t
h	Foil thickness
η	Azimuth angle from x axis of the screen to t_p
d	Real interface width, i.e., real distance between traces
w	Projected interface width, i.e., distance between projected traces
i	Input number, index, $1 \leq i \leq m$
m	Total number of inputs
$Rx(\theta)$, $Ry(\theta)$, $Rz(\theta)$	Right-handed rotation matrix with rotation angle θ about x, y, z axis respectively
$sgn(i)$	A function that gives a value of ± 1 respect to the sign of $n \cdot b_{ei}$

A flat interface has two traces, the intersections with upper and lower foil surfaces. Firstly, let us assume the two foil surfaces to be parallel planes, resulting in two parallel traces separated by a certain width d. This case is

---

[1]If not pointed out, all the vector symbols noted in bold and italics are column vectors.
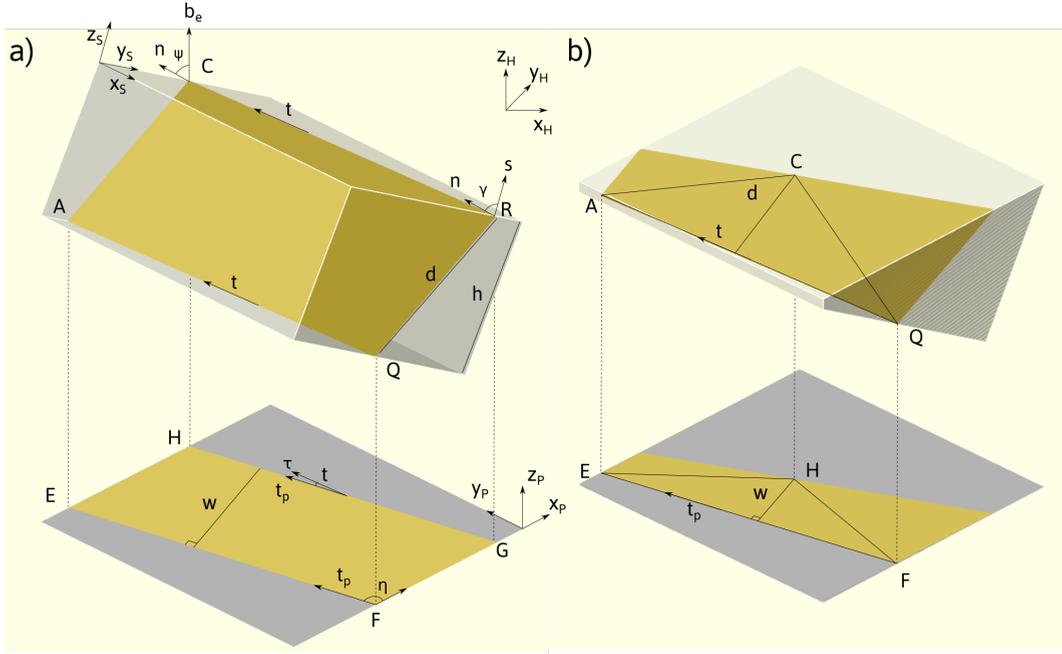
Figure 1: Geometrical description of interface projection. (a) Parallel traces case: The interface $AQRC$ is projected on the viewing screen, giving the projection $EFGH$. The inset is the cross section view along the trace direction. (b) Non-parallel traces case: A traceable point C on the interface and a trace AQ are projected to the viewing screen, giving the projection EFH.

shown in Figure 1a, where the interface is enclosed by the red parallelogram $AQRC$. Line CR and AQ are upper

and lower traces of the interface, intersections between the foil surfaces (orange planes) and the interface. The blue

parallelogram $EFGH$ is the projected interface on the screen. Areas of these parallelograms, $S_{AQRC} = d|AQ|$ and

$S_{EFGH} = w|EF|$, are related by $S_{EFGH} = S_{AQRC} \cos \psi$, where ψ is the angle between reversed electron beam

direction $\boldsymbol{b}_e$ and interface normal $\boldsymbol{n}$. Since $|EF|$ is the projection of $|AQ|$, they are related by $|EF| = |AQ| \cos \tau$,

where τ is the angle between trace $\boldsymbol{t}$ and its projection $\boldsymbol{t}_p$.

When the foil is bent or has a wedge shape, the assumption of parallel foil surfaces is no longer valid. Because of

this, trace AQ and trace CR are not parallel, $AQRC$ in Figure 1 is no longer a parallelogram (Figure 1b). However,

if a point on the interface can be tracked, it is still possible to resolve the relationship between d and w. As shown in

Figure 1b, by tracking point C on the interface AQC, one can redefine d as the distance between C and the trace AQ,

intersection with one foil surface (orange plane), while the distance between the projected point H and projected trace

EF is measured as w. In this case, it is crucial to measure w at the same position each time. The area of interface

triangle $S_{AQC} = d|AQ|/2$ and that of projection triangle $S_{EFH} = w|EF|/2$ are still related by $S_{EFH} = S_{AQC} \cos \psi$.

The relationship of $|EF| = |AQ| \cos \tau$ remains valid.

Therefore, no matter two foil surfaces are parallel or not, the interface projection width w is always given by

$$w = \frac{d \cos \psi}{\cos \tau} \tag{1}$$

Or alternatively using dot product:

$$w = \frac{|d\boldsymbol{n} \cdot \boldsymbol{b}_e|}{\sqrt{1 - \left(\boldsymbol{b}_e \cdot \boldsymbol{t}\right)^2}} \tag{2}$$

In Eq. (2), w and $\boldsymbol{b}_e$ can be measured, while $d\boldsymbol{n}$ and $\boldsymbol{t}$ are unknown.

Vector $\boldsymbol{t}$ can be determined from its measured projections $\boldsymbol{t}_p$'s at different $\boldsymbol{b}_e$'s, using:

$$\begin{bmatrix} \boldsymbol{b}_{e1} \times \boldsymbol{t}_{p1} & \boldsymbol{b}_{e2} \times \boldsymbol{t}_{p2} & ... & \boldsymbol{b}_{em} \times \boldsymbol{t}_{pm} \end{bmatrix}^{\mathrm{T}} \boldsymbol{t} = \boldsymbol{0} \tag{3}$$

When $m = 2$, Eq. (3) is solved by using the cross product $\boldsymbol{t} = (\boldsymbol{b}_{e1} \times \boldsymbol{t}_{p1}) \times (\boldsymbol{b}_{e2} \times \boldsymbol{t}_{p2})$. When the number of input sets $m \geq 3$, it is an overdetermined homogeneous linear equation, and can to be solved by the least square method (LSM, see Appendix A).

The determination of $\boldsymbol{t}$ allows to compute the non-linear part $\sqrt{1 - \left(\boldsymbol{b}_e \cdot \boldsymbol{t}\right)^2}$ in Eq. (2), turning Eq. (2) into a linear problem. However, the width of interface projection, w_i, is always positive, while $\boldsymbol{n} \cdot \boldsymbol{b}_{ei}$ could be negative. Thus sgn(i), the sign of $\boldsymbol{n} \cdot \boldsymbol{b}_{ei}$, is needed to remove the absolute sign in Eq. (2). Using the fact that $d\boldsymbol{n} \cdot \boldsymbol{t} = 0$, a synthetic formula incorporating the input measurements (w_i, $\boldsymbol{b}_{ei}$) can be derived:

$$\begin{bmatrix} \boldsymbol{b}_{e1}^{\mathrm{T}} \\ \boldsymbol{b}_{e2}^{\mathrm{T}} \\ ... \\ \boldsymbol{b}_{en}^{\mathrm{T}} \\ \boldsymbol{t}^{\mathrm{T}} \end{bmatrix} (d\boldsymbol{n}) = \begin{bmatrix} \mathrm{sgn}(1)w_1\sqrt{1 - \left(\boldsymbol{b}_{e1} \cdot \boldsymbol{t}\right)^2} \\ \mathrm{sgn}(2)w_2\sqrt{1 - \left(\boldsymbol{b}_{e2} \cdot \boldsymbol{t}\right)^2} \\ ... \\ \mathrm{sgn}(m)w_m\sqrt{1 - \left(\boldsymbol{b}_{em} \cdot \boldsymbol{t}\right)^2} \\ 0 \end{bmatrix} \tag{4}$$

This equation can be solved by LSM, but the uncertainty of sgn(i) would result in multiple solutions of $d\boldsymbol{n}$. When

$m = 2$, Eq. (4) is determined and two different solutions are expected. It is impossible to assess the actual one without additional information. When $m \geq 3$, Eq. (4) is overdetermined, and its residual error can be computed (see Appendix A). The uncertainty of sgn(i) will give 2^{m-1} solutions of $\boldsymbol{n}$. The one with the minimum residual error is the best solution, because an improper sgn(i) would result in significantly large residual error. The contribution to the residual error from each set of input ($w_i, \boldsymbol{b}_{ei}, \boldsymbol{t}_{pi}$) can be quantified by the respective element in the residual vector of Eq. (4) (see Appendix A). The element with the maximum absolute value and different sign from the others may indicate an abnormal input with significantly large deviation. One may drop this input and recalculate the result, which can reduce the residual error and improve the input consistency. Thus, with at least 3 different sets of experimental data ($w_i, \boldsymbol{b}_{ei}, \boldsymbol{t}_{pi}$), LSM will give the optimized solutions of $\boldsymbol{t}$, d and $\boldsymbol{n}$.

The accuracy of the interface normal, which is the deviation between the determined value and the true value is an important concern. Unfortunately, the true value of $d\boldsymbol{n}$ is usually unknown, and needs to be estimated. The estimation of the true value is a range with a certain confidence, i.e., a confidence interval, which can represent the accuracy of the result, based on the internal consistency of inputs. In the present work, the 95% confidence interval of d and that of $\boldsymbol{n}$ can be calculated by the bootstrap method (see Appendix B) with Eq. (4).

A C++ implementation of the present algorithm can be found in Supplementary Materials.

3. Experimental details

The present method was tested in two cases. In both cases, bright field images were recorded at different orientation conditions in an FEI Tecnai G2 20 TEM operated at 200 kV. Observations were performed using a double tilt sample holder with the coordinate conventions as shown in Figure 2: the x-axis is parallel to the holder axis, and the y-axis is perpendicular to the x-axis and the beam direction. In the following, the subscripts $_H$, $_P$ and $_S$ denote the indices in holder, projection screen, and sample coordinate system, respectively . In the holder coordinate system, the sample is firstly tilted about the $-\boldsymbol{y}$ axis by angle β, and then tilted about the $-\boldsymbol{x}$ axis by angle α.

The rotation angle from holder $\boldsymbol{x}_H$ axis to the projection screen $\boldsymbol{x}_P$ axis is ϕ about $\boldsymbol{z}_H = \boldsymbol{z}_P$ axis. This accounts for the rotation of the image in the microscope column. Hence, the beam direction $\boldsymbol{b}_e$ perpendicular to the screen and the direction $\boldsymbol{t}_p$ in the observation plane can be calculated using right-handed rotation matrices Rx(θ), Ry(θ), and Rz(θ), as shown below:
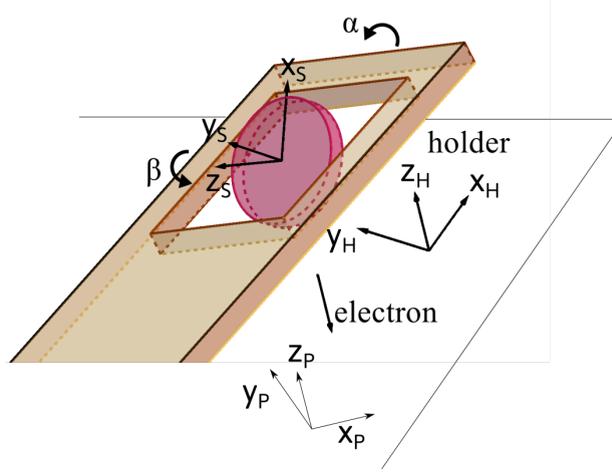
Figure 2: Illustration of the holder coordinate system when the sample is mounted in a double-tilt holder.

$$\boldsymbol{b}_{\mathrm{e}} = \mathrm{Ry}(\beta)\mathrm{Rx}(\alpha)[0,\ 0,\ 1]_{\mathrm{H}}^{\mathrm{T}}$$
$$= [\sin\beta\cos\alpha,\ -\sin\alpha,\ \cos\beta\cos\alpha]_{\mathrm{S}}^{\mathrm{T}} \tag{5}$$

$$\boldsymbol{t}_{\mathrm{p}} = \mathrm{Ry}(\beta)\mathrm{Rx}(\alpha)\mathrm{Rz}(\phi)[\cos\eta,\ \sin\eta,\ 0]_{\mathrm{P}}^{\mathrm{T}} \tag{6}$$

where η is the azimuth angle between the projected trace and the x_P-axis on the viewing screen.

In the first experiment, a grain with two interfaces was investigated in a duplex stainless steel Fe-24.9Cr-7.0Ni-3.1Mo (wt%) sample prepared by the same procedure as reported in [15]. The images were taken in conditions where $\phi = 90°$.

In the second case, a grain boundary with non-parallel traces was observed in an aluminum bicrystal with a misorientation close to a coincident $\Sigma 41\ 12.68° < 0\ 0\ 1 > \{5\ 4\ 0\}$. The sample was first ground to 50 microns using SiC grain disks and then electro-polished to obtain electron transparency. The TEM foil was strained in-situ at c.a. 400°C, as reported earlier in [?] (ref missing). Plastic deformation leads to a complex microstructure of dislocations, resulting in significant bending of the wedge foil and possible deviation from the original orientation. Eq. (5) and Eq. (6) are still valid in this case, using here $\phi = 157°$ for Eq. (6).

4. Application examples

4.1. Measuring two interfaces with a double-tilt holder

Figure 3 shows a series of bright field (BF) images of a faceted austenite grain (A) in a ferrite matrix (F) taken at different orientations. Two FCC/BCC facets, IntA and IntB, intersect and form a well defined corner. In order to show the ability to determine multiple interfaces, normals of these two facets are determined simultaneously, using 6 sets of inputs among 8 different imaging conditions. The determination was performed both in the holder coordinate system and in the lattice coordinate system. The former is more convenient to use, as it only needs tilt angles of the sample holder, while the latter, which requires the analysis of Kikuchi patterns, usually has better accuracy. The normal of IntA is determined by Figures 3a~f, and that of IntB by Figures 3a ~ e and 3g. Figure 3h, a near edge-on condition of both IntA and IntB, was only used to check the results.

The first step of the method is to determine t. The b_e's in holder coordinates were calculated with data in Table 2 and Figure 3, using Eq. (5) and Eq. (6), while the t_p's of IntA and IntB are given by their azimuth angle with respect to the x axis of Figures 3a~g. The lattice coordinates of b_e's and t_p's were defined in austenite using the Kikuchi patterns shown in Figure 3, as listed in Table 2. Hence t was determined by solving Eq. 3 in either coordinate system.

The second step is to substitute b_e's, t, and w_i's, into Eq. 4 to solve dn. By enumerating all the possibilities of $\text{sgn}_A(i)$, and $\text{sgn}_B(i)$, dn and its residual error in both conventions are solved by LSM. Among them, the one with the minimum residual error was chosen as the final result, whose $\text{sgn}(i)$ is shown in Table 2. Then, the optimized results were calculated, as shown by Table 3. The n's in holder coordinates were transformed to lattice coordinates using the transformation matrix determined from Figure 3e2 (see Supplementary materials).

The t's of IntA and IntB in holder coordinates are almost in the xOy plane, indicating that the foil was almost perpendicular to the beam direction at zero tilt. For each interface, the real width (d) and the normal (n) calculated from different coordinate systems show a good consistency. This means the holder coordinate system can be used instead of the lattice coordinate system to simplify the process while reserving the accuracy. The correctness of n is verified by tilting both interfaces to a near edge-on condition (Figure 3h), where the interface normal n can be determined by $b_e \times t_p$. Using the data of Figure 3h in Table 2, the normal of IntA is [0.562 0.592 − 0.579], and that of IntB is [−0.594 − 0.237 − 0.766]. These results are inside the confidence range of the n's determined in Table 3.
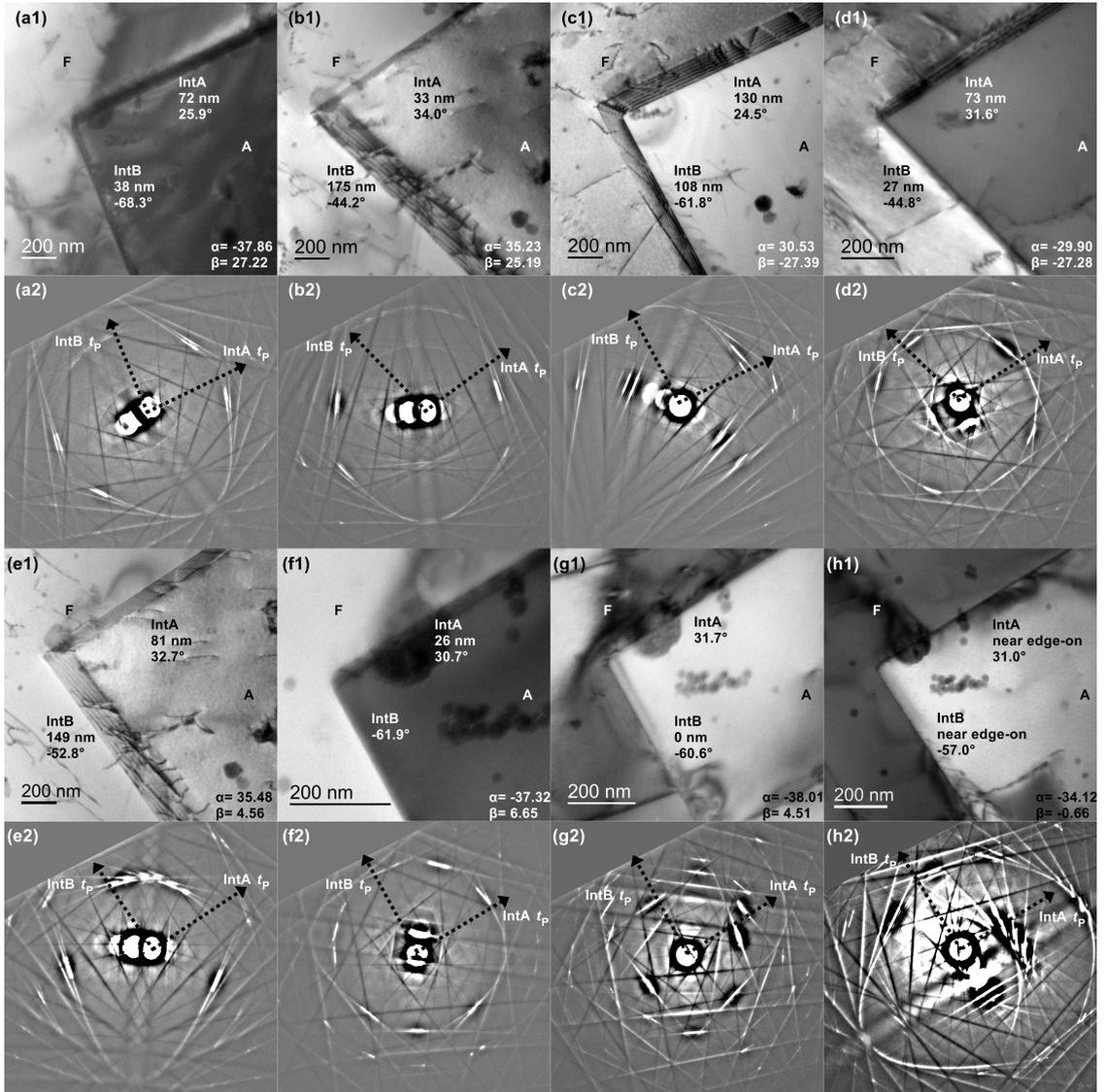
Figure 3: BF images of the interfaces and the corresponding Kikuchi patterns taken in austenite at different tilt angles. The α and β tilt angles are reported in the bottom right. The projection widths and the azimuthal angle of both IntA and IntB interfaces are also indicated. Arrows in the Kikuchi patterns indicate the $\boldsymbol{b}_e \times \boldsymbol{t}_p$ directions of IntA and IntB.

Hence, the present method is able to accurately determine multiple interfaces simultaneously, either in the holder or lattice coordinate system.

4.2. Measuring a grain boundary with non-parallel traces

Figure 4 shows a series of BF images of an interface with non-parallel traces in an Al bicrystal sample taken at different tilt angles. The tracked point (pointed out by arrows) is defined by the position where the lower trace (dotted line) starts to deviate. Figures 4a~e are used to solve the interface normal, while Figure 4f is a near edge-on condition

Table 2
Directions in lattice coordinates from the Kikuchi patterns in Figure 3 and sgn(i) for solving Eq. (4)

Figure	$\boldsymbol{b}_e$	IntA $\boldsymbol{t}_p$ IntB $\boldsymbol{t}_p$	sgn$_A(i)$ sgn$_B(i)$	w_{Ai} w_{Bi}	$\alpha(°)$	$\beta(°)$
3a	[-0.811 0.485 0.327]	[0.546 0.427 0.721] [0.249 0.792 -0.558]	-1 +1	72 38	-37.86	27.22
3b	[-0.629 0.147 -0.763]	[-0.250 0.891 0.378] [0.772 0.239 -0.590]	+1 +1	33 175	35.23	25.19
3c	[-0.107 0.628 -0.771]	[0.186 0.774 0.605] [0.963 -0.128 -0.238]	+1 +1	130 108	30.53	-27.39
3d	[-0.281 0.951 0.130]	[0.485 0.023 0.874] [0.691 0.295 -0.660]	+1 -1	73 27	-29.90	-27.28
3e	[-0.465 0.341 -0.817]	[-0.048 0.912 0.408] [0.885 0.157 -0.438]	+1 +1	81 149	35.48	4.56
3f	[-0.677 0.684 0.272]	[0.597 0.295 0.746] [0.457 0.680 -0.574]	-1 not used	26 -	-37.32	6.65
3g	[-0.654 0.704 0.277]	[0.612 0.277 0.741] [0.469 0.665 -0.582]	not used +1	- 0	-38.01	4.51
3h	[-0.592 0.776 0.219]	[0.578 0.220 0.786] [0.541 0.584 -0.601]	not used not used	- -	-34.12	-0.66

Table 3
Results of trace directions and interface normals

Coordinates	Interface	$\boldsymbol{t}$	$\boldsymbol{n}$	Confidence range of $\boldsymbol{n}$ (°)	d (nm)
holder	IntA	$[0.453, -0.891, 0.003]_H$	$[-0.863, -0.432, 0.262]_H$ [0.567 0.588 $-$ 0.577]	1.0	166 ± 5
	IntB	$[0.776, 0.630, 0.037]_H$	$[0.549, -0.670, 0.500]_H$ [-0.605 -0.208 -0.769]	1.8	178 ± 6
lattice	IntA	$[-0.215 \ -0.581 \ -0.785]$	[0.581 0.585 $-$ 0.567]	1.6	172 ± 8
	IntB	$[-0.722 \ -0.253 \ 0.644]$	[$-0.610 - 0.240 - 0.755$]	1.6	182 ± 6

for verification. The trace $\boldsymbol{t}$ of the interface in holder coordinates deviates from the xOy plane, confirming that the foil is bent. Through a similar process as the previous example, the optimized result in holder coordinates is calculated as following: $\boldsymbol{t} = [0.883, -0.348, 0.315]_H$, $\boldsymbol{n} = [-0.473, -0.646, 0.600]_H \sim 2.4°$, and $d = 224 \pm 12$ nm. (what is 2.4° ?) Using the transformation matrix determined from a Kikuchi pattern (Figure S1), the lattice coordinates of the interface normal is $\boldsymbol{n} = [4.000 \ -4.176 \ 0.144]$ with 5.3° deviation from $[4 \ -5 \ 0]$. Noting that the 95% confidence range is 2.4°, this deviation is considered to be the result of plastic deformation, rather than experimental errors. Using Eq. 5, the $\boldsymbol{b}_e$ at the near edge-on condition in Figure 4f is $[0.278, 0.536, 0.797]_H$, which is almost perpendicular to the interface normal with 0.1° deviation. This edge-on condition confirms the result of $\boldsymbol{n}$. Therefore, the present method remains effective even on interfaces with non-parallel traces.
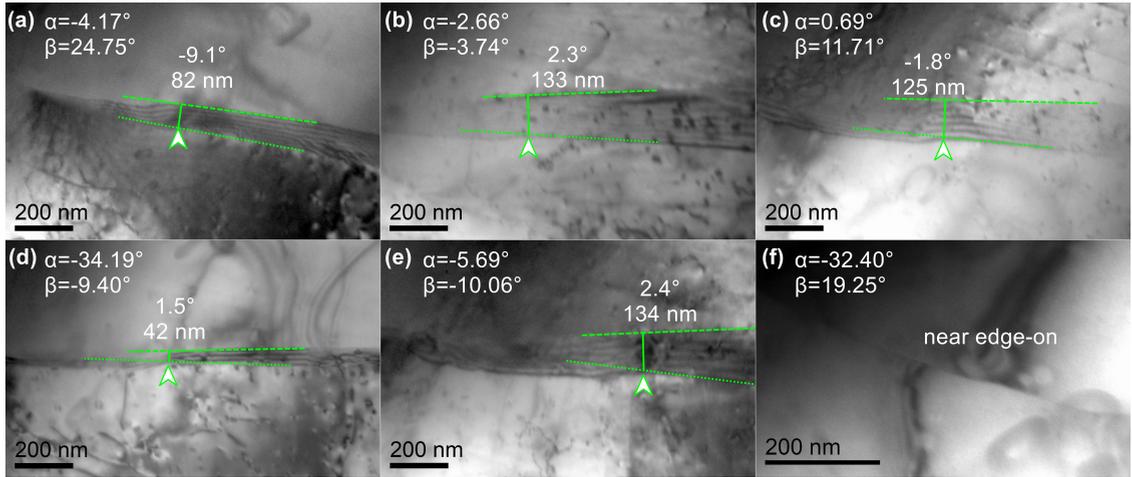
Figure 4: BF images of the interface with non-parallel traces at different tilt angles. The α and β tilt angles are reported in the top left. The widths between the tracked point (indicated by arrows) and the trace (dash line), and the azimuth angles of the trace (dash line) are also indicated.

5. Discussion

5.1. Accuracy of n

The use of the method presented here raises three questions concerning the accuracy of n: (i) is the true value of the interface normal inside the bootstrap confidence interval; (ii) what is the main error source on interface normal determination; (iii) what is the optimal number of input datasets for an acceptable accuracy. To answer these questions, statistical analysis was performed using simulated erroneous data.

The simulated datasets were generated in holder coordinates, using a 100-nm-thick virtual sample with two parallel surfaces normal to z. As the prerequisite, the true value of the interface normal n must be known. Then, the α and β angles were randomly chosen in $\pm 40°$ range, and the true values of w and η were calculated. The dataset (α, β, η, w) was randomly varied to simulated the experimental error. The α, β and η angles were randomly varied in $\pm 1°$ range, while the projection widths in ± 5 nm range. With several erroneous datasets, the interface normal was determined, and the deviation from the true value, which is called trueness for clarity, was calculated.

In the first analysis, the interface normal was randomly chosen and determined with 6 sets of simulated data. This process was repeated 10000 times, and the true value was inside the bootstrap confidence interval for 9369 times. This agrees to the confidence level (95%) of the interval. The true value could be outside the confidence interval when the datasets (α, β, η, w) have relatively large systematic error, as this problem cannot be figured out from the consistency
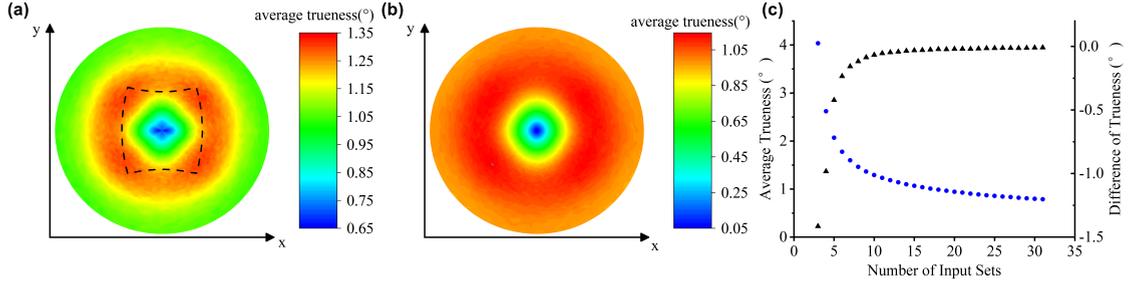
Figure 5: Statistical analysis of the trueness of the results.
(a) Distribution of the average trueness for interface normals chosen within the pole figure viewed from z axis. α, β, η and w have random errors. The beam directions reached by sample tilting are limited by dashed lines. (b) Same as (a) when the error is restricted to w. (c) The change of trueness with respect to the number of inputs for a fixed interface normal.

of the datasets.

In the second analysis, 4949 unit vectors evenly distributed on a hemisphere were used as known interface normals. For each normal, 6 arbitrary conditions (α, β, η, w) were used to compute back an erroneous normal, and its trueness. Because the trueness is dependent on the orientation conditions, it was averaged over 10000 arbitrary condition sets for each known normal.

The results are plotted in a pole figure in Figure 5a. It shows an *mm* symmetry due to the angular area covered by the sample tilt, limited by the dashed lines in Figure 5a. The average trueness of interface normals ranges between $0.65°$ and $1.35°$, but most of the value is due to the error on w as shown in Figure 5b, where the average trueness is computed when varying only w. The ring-like shape of trueness in Figure 5a can be understood by inspecting how errors affect the value of $\cos \psi$ in Eq. 4. Using $h = d \sin \gamma = d \sqrt{1 - (\boldsymbol{n} \cdot \boldsymbol{s})^2}$ (Figure 1a inset) and Eq. 2, we can derive

$$\cos \psi = \boldsymbol{n} \cdot \boldsymbol{b}_{\mathrm{e}} = \frac{w}{h} \sqrt{1 - \left(\boldsymbol{b}_{\mathrm{e}} \cdot \boldsymbol{t}\right)^2} \sin \gamma \tag{7}$$

When the interface normal $\boldsymbol{n}$ moves from the center to the edge in the pole figure, statistically, $\sin \gamma = \sqrt{1 - (\boldsymbol{n} \cdot \boldsymbol{s})^2}$ will increase, making the value of $\cos \psi$ more sensitive to the error of w. In other words, largely inclined planes in the foil are less precisely determined, because they statistically present orientation conditions where the projection width is small. Meanwhile, the increase in ψ makes the direction of $\boldsymbol{n}$ less sensitive to the error of $\cos \psi$, as $\mathrm{d}\psi = -\mathrm{d}(\cos \psi)/\sin \psi$. The two effects having opposite variations, this leads to a trueness distribution with a maximum value in a ring area inclined about $54°$ with the foil normal (Figure 5b). In common cases, the projection of the interested interface is not very wide, which means that its normal is not located at the blue part in Figure 5a. Therefore,

the error of $\boldsymbol{n}$ mainly comes from the error of w.

In the third analysis, the interface normal was fixed to $[-0.547,\ 0.514,\ 0.661]_{\mathrm{H}}$, an orientation with the worst trueness in Figure 5a, while the number of input datasets varied from from 3 to 15. The average of trueness were computed, again by randomly varying the condition sets over 10000 configurations. As shown in Figure 5c, the trueness of $\boldsymbol{n}$ would reach 1.7° with 6 sets of inputs. It would reach 1.5° with 8 sets of inputs, but barely improves afterwards. Based on the above analysis, 6 to 8 sets of inputs would be the optimal choice.

In practice, it is trivial to reach the data quality ($\pm 1°$, ± 5nm) in this analysis, and hence to get the results with the corresponding accuracy. Moreover, data quality can even be improved by image measurements at higher magnification, and by using Kikuchi patterns.

5.2. Comparison with existing methods

The method proposed here relies on the measurements of the direction and width of interface projections, and the beam directions. It works both in lattice or holder coordinate systems. The double trace and edge-on methods can be considered as special cases of the present method in the lattice coordinate system. The geometry model used here offers a framework to understand the existing projection width methods, which, as shown below, used specific simplifications of the non-linear part $\sqrt{1 - \left(\boldsymbol{b}_{\mathrm{e}} \cdot \boldsymbol{t}\right)^{2}}$ in Eq. 2. The summary of the comparison is listed in Table 4.

Table 4
A comparison between this work and previous methods

Method	Coordinates	Formulas	Special requirements	Number of inputs
double-trace [5]	lattice	Eq. (3)	another line feature	2
single edge-on [9]	lattice	Eq. (4)	edge-on condition	1
double edge-on [10]	lattice	Eq. (4)	edge-on condition	2
trace & edge-on [12]	lattice	Eq. (3) and Eq. (4)	edge-on condition	3
projection width [13]	holder	Eq. (4)	tilting about the trace	2
projection width [14]	holder	Eq. (12)	tilting axis in screen	2
this work	either	Eq. (3) and Eq. (4)	none	≥ 3

In the double-trace method and edge-on methods, the interface normal is always derived by the cross product of two vectors. These two vectors can be two traces (double-trace), a projected trace and a beam direction (single edge-on), a trace and a beam direction (trace & edge-on), or two beam directions (double edge-on). Since the cross product is equivalent to LSM when solving a determined equation, these methods can all be treated as special cases of Eq. (4), with the traces determined by Eq. (3).

Existing projection width methods [13][14] treated the non-linear part $\sqrt{1 - \left(\boldsymbol{b}_e \cdot \boldsymbol{t}\right)^2}$ in Eq. 2 with different strategies. Zhang and Kelly [13] tilted the sample about the trace direction, resulting in $\boldsymbol{b}_e \cdot \boldsymbol{t} = 0$ (?) and $w = |d\boldsymbol{n} \cdot \boldsymbol{b}_e|$. In their work, the sample thickness h was used rather than d. These two variables have a relationship, $h = d \cos \gamma$ (Figure 1a inset), which is valid only if the two foil surfaces are both perpendicular to the beam direction at zero tilt. Qiu and Zhang [14] extended this work using the latter assumption, but they substituted $\sqrt{1 - \left(\boldsymbol{b}_e \cdot \boldsymbol{t}\right)^2}$ by taking the tilting axis of the sample holder as a reference vector. The derivation of this method using the present geometry model is given in Appendix C.

The limitation of the reference vector method is that this vector $\boldsymbol{r}$ must be parallel to the tilting axis of the sample holder, otherwise the $\boldsymbol{n} \times \boldsymbol{s} \cdot \boldsymbol{r}$ in Eq. 11 would not have a simple result like x, making it too hard to be solved. Therefore, the reference vector method cannot be applied to double-tilt holders, because the resulting tilting axis is neither constant, nor in the screen plane.

In addition, these two methods cannot determine the sign of w_i's without external help, as they only use two sets of inputs. As shown in this work, the use of at least 3 input sets enables the determination of the sign of w_i's straightforwardly.

5.3. Generalization of the present method

The present method can also be used as an enhanced method to determine the length and the direction of 1D features. Let l be the projected length of the 1D feature and d be the original length of the 1D feature. Each measurement of a 1D feature projection ($\boldsymbol{b}_{ei}$, $\boldsymbol{l}_{pi}$, l_i) will give

$$\left[\begin{array}{cc} \boldsymbol{b}_{ei} \times \boldsymbol{l}_{pi} & \boldsymbol{b}_{ei} \end{array}\right]^T d\boldsymbol{l} = \left[\begin{array}{cc} 0 & l_i \end{array}\right]^T \tag{8}$$

Eq. 8 can be solved the same way as described above. The length vector of the 1D feature can thus be determined.

This approach may be used, for instance, to track accurately the distance between pining points on a dislocation line or the size of dislocation loops, as long as, 2 traceable features, such as intersection points between dislocation segments, can be detected. If 3 traceable points can be detected, for instance, to determine dislocation habit plane if at least 3 points (or 2 lines) can be tracked [16], or to assess the planarity of 4 points (or 3 lines). (i'm not sure to

understand this sentence)

Since all algorithms proposed here employed LSM, they are equally robust against random errors and scalable on large datasets. Efficient LSM functions can be found in plenty of math libraries, such as Eigen, LAPACK and Intel®MKL. Hence the present algorithms can be implemented as a real-time solver (as shown in Supplementary Material). They also have the potential to be integrated with automatic feature tracking in a tilting series, which could be a fast and efficient way to automatically measure microstructural features in a foil.

6. Conclusions

A new analytical method to determine interface normals with excessive inputs has been proposed. It is a robust algorithm based on a generalized geometrical model of interface projection, and it can also automatically deal with a large amount of data. The validity of the method was verified using experimental observations of interfaces in TEM. This study proves the reliability and effectiveness of this method, regardless of the input orientation or foil surface configuration. Given 6 or a few more sets of inputs, even with considerable experimental errors, this method is still capable of yielding reliable results. It can also be extended to determine 1D features or to check the planarity of a set of features. The present method is compared with existing approaches, showing that many existing approaches can be treated as special cases of the present method, while the present method has only few constraints.

A. LSM for solving overdetermined linear equations

For an overdetermined $Ax = b$ equation, e.g., Eq. 4 with $m \geq 3$, the least square solution of vector x is $x^+ = A^+b$, where A^+ is Moore-Penrose inverse, or pseudo-inverse, of matrix A. The solution has the residual vector, $e = AA^+b - b$, and the residual error, $\|AA^+b - b\|_2$. This method is applicable for most circumstances. However, when $b = 0$, e.g., Eq. 3, this method will only give a trivial solution of $x = 0$. $Ax = 0$ is a typical problem of overdetermined homogeneous linear system, whose solution is the eigen-vector of A^TA with the smallest eigen-value. The residual error of the solution is the smallest eigen-value.

B. Bootstrap method for estimating confidence intervals

Bootstrap is a statistical technique to estimate the variation of statistics that are computed from a set of data [?]. Here we use the bootstrap method to estimate the variation, i.e., confidence intervals, of d and $\boldsymbol{n}$ determined by Eq. 4. In Eq. 4, $\boldsymbol{t}$ and $\text{sgn}(i)$ are considered as constant parameters in the equation, as they have already been determined. Thus, the data is composed of m data points of $(\boldsymbol{b}_{ei}, w_i)$. Then, the data is resampled with replacement to generate a resampled data set of size m, with which the statistics d^* and $\boldsymbol{n}^*$ (star denoting resampled data) can be calculated. This procedure is repeated 10000 times, the deviations between d^* and d and the angle between $\boldsymbol{n}^*$ and $\boldsymbol{n}$ can also be calculated. The 95% confidence interval of d is given by the 95th percentile of the deviation of d^*. The 95% confidence interval of $\boldsymbol{n}$ is given by the 95th percentile of the deviation angle of $\boldsymbol{n}^*$, denoting a confidence cone around $\boldsymbol{n}$ [?].

C. The projection width method reported in [14]

Here we used the symbols and equations in present work to derive the equations reported in [14].

By replacing $\boldsymbol{t}$ in Eq. 2 with $\boldsymbol{n} \times \boldsymbol{s}/\|\boldsymbol{n} \times \boldsymbol{s}\|$, and temporarily ignoring the absolute sign on $\boldsymbol{b}_e \cdot \boldsymbol{n}$, we get:

$$w = \frac{h\boldsymbol{b}_e \cdot \boldsymbol{n}}{\sqrt{1 - (\boldsymbol{n} \cdot \boldsymbol{s})^2 - (\boldsymbol{n} \times \boldsymbol{s} \cdot \boldsymbol{b}_e)^2}} \tag{9}$$

By introducing $\boldsymbol{r}$, the angle η between $\boldsymbol{t}_p$ and $\boldsymbol{r}$ is expressed by:

$$\cos \eta = \frac{\boldsymbol{t}_p \cdot \boldsymbol{r}}{\|\boldsymbol{t}_p\|} = \frac{\boldsymbol{n} \times \boldsymbol{s} \cdot \boldsymbol{r}}{\sqrt{1 - (\boldsymbol{n} \cdot \boldsymbol{s})^2 - (\boldsymbol{n} \times \boldsymbol{s} \cdot \boldsymbol{b}_e)^2}} \tag{10}$$

By substituting Eq. 10 into Eq. 9, we get Eq. 11. h and $\boldsymbol{n}$ have to be combined, for they both need solving.

$$\boldsymbol{b}_e \cdot (\frac{h}{\boldsymbol{n} \times \boldsymbol{s} \cdot \boldsymbol{r}}\boldsymbol{n}) = \frac{w}{\cos \eta} \tag{11}$$

In the work of Qiu and Zhang, the sample is tilted about $\boldsymbol{y}$ axis with the screen remains unchanged. The $\boldsymbol{x}^r, \boldsymbol{y}^r, \boldsymbol{z}^r$ vectors at zero tilt in their work are the basis vectors in the holder coordinates system. Therefore, the parameters should be set as $\boldsymbol{b}_e = [0,\ 0,\ 1]_H^T, \boldsymbol{r} = [0,\ -1,\ 0]_H^T, \boldsymbol{n} = \text{Ry}(\alpha)[x,\ y,\ z]_H^T$, and $\boldsymbol{s} = \text{Ry}(\alpha)[0,\ 0,\ 1]_H^T$. In addition, the trace

direction was measured at zero tilt angle by η_0, which would directly give t. Similar as Eq. (4), Eq. 12 is constructed using two sets of Eq. 11 and $t \cdot n = 0$. The $\pm$ sign is caused by the uncertainty of w_i signs, and should be eliminated by inspecting the trend of projection width change.

$$
\begin{bmatrix}
-\sin\alpha_1 & 0 & \cos\alpha_1 \\
-\sin\alpha_2 & 0 & \cos\alpha_2 \\
\sin\eta_0 & -\cos\eta_0 & 0
\end{bmatrix}
\frac{h}{x}
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix}
=
\begin{bmatrix}
w_1/\cos\eta_1 \\
\pm w_2/\cos\eta_2 \\
0
\end{bmatrix}
\tag{12}
$$

The solution of Eq. 12 is Eq. 13, the same as reported in their work.

$$
\begin{bmatrix}
h \\
hy/x \\
hz/x
\end{bmatrix}
=
\frac{1}{\sin(\alpha_2-\alpha_1)}
\begin{bmatrix}
\frac{w_1\cos\alpha_2}{\cos\eta_1} \mp \frac{w_2\cos\alpha_1}{\cos\eta_2} \\
(\frac{w_1\cos\alpha_2}{\cos\eta_1} \mp \frac{w_2\cos\alpha_1}{\cos\eta_2})\tan\eta_0 \\
\frac{w_1\sin\alpha_2}{\cos\eta_1} \mp \frac{w_2\sin\alpha_1}{\cos\eta_2}
\end{bmatrix}
\tag{13}
$$

References

[1] P. A. Midgley, R. E. Dunin-Borkowski, Electron tomography and holography in materials science, Nature Materials 8 (2009) 271.

[2] J. S. Barnard, J. Sharp, J. R. Tong, P. A. Midgley, High-resolution three-dimensional imaging of dislocations, Science 313 (2006) 319–319.

[3] W.-Z. Zhang, X.-F. Gu, F.-Z. Dai, Faceted interfaces: a key feature to quantitative understanding of transformation morphology, npj Computational Materials 2 (2016) 16021.

[4] P. Hirsch, A. HOWIE, R. B. NICHOLSON, D. PASHLEY, M. WHELAN, Electron Microscopy of Thin Crystals, Butterworths, London, 1965. URL: https://books.google.fr/books?id=eNIeAQAAIAAJ.

[5] B. Sandvik, C. Wayman, Characteristics of lath martensite: Part i. crystallographic and substructural features, Metallurgical Transactions A 14A (1983) 809–822.

[6] C. T. Young, J. H. Steele, J. L. Lytton, Characterization of bicrystals using kikuchi patterns, Metallurgical Transactions 4 (1973) 2081–2089.

[7] Q. Liu, A new method for determining the normals to planar structures and their trace directions in transmission electron-microscopy, Journal of Applied Crystallography 27 (1994) 762–766.

[8] S. Li, Y. Zhang, C. Esling, J. Muller, J.-S. Lecomte, G. W. Qin, X. Zhao, L. Zuo, Determination of surface crystallography of faceted nanoparticles using transmission electron microscopy imaging and diffraction modes, Journal of Applied Crystallography 42 (2009) 519–524.

[9] J. W. Edington, K. T. Russell, Practical electron microscopy in materials science, Macmillan International Higher Education, 1977.

[10] C. P. Luo, X. L. Xiao, D. X. Wu, A tem method for accurate measurement of habit plane (interface): double edge-on trace analysis, Progress

in Natural Science 7 (1997) 742–748.

[11] D. Qiu, W.-Z. Zhang, A tem study of the crystallography of austenite precipitates in a duplex stainless steel, Acta Materialia 55 (2007) 6754–6764.

[12] Y. Meng, L. Gu, W. Zhang, Precise determination of the irrational preferred interface orientation by tem, Acta Metall Sin 46 (2010) 411.

[13] M. X. Zhang, P. M. Kelly, J. D. Gates, Determination of habit planes using trace widths in tem, Materials Characterization 43 (1999) 11–20.

[14] D. Qiu, M. Zhang, A simple and inclusive method to determine the habit plane in transmission electron microscope based on accurate measurement of foil thickness, Materials Characterization 94 (2014) 1–6.

[15] J. Du, F. Mompiou, W.-Z. Zhang, In-situ tem study of dislocation emission associated with austenite growth, Scripta Materialia 145 (2018) 62–66.

[16] S. Naanani, J.-P. Monchoux, C. Mabru, A. Couret, Pure climb of [001] dislocations in tial at 850Âř c, Scripta Materialia 149 (2018) 53–57.