Full length article

Automated identification and tracking of deformation twin structures in molecular dynamics simulations

H.J. Ehrich [a], A. Dollmann [b,c], P.G. Grützmacher [a], C. Gachot [a], S.J. Eder [a,d,*]

[a] *Institute for Engineering Design and Product Development, TU Wien, Lehárgasse 6 – Objekt 7, 1060 Vienna, Austria*
[b] *Institute for Applied Materials (IAM), Karlsruhe Institute of Technology (KIT), Kaiserstraße 12, 76131 Karlsruhe, Germany*
[c] *MicroTribology Center (µTC), Straße am Forum 5, 76131 Karlsruhe, Germany*
[d] *AC2T research GmbH, Viktor-Kaplan-Straße 2/C, 2700 Wiener Neustadt, Austria*

ARTICLE INFO

ABSTRACT

Deformation twinning significantly influences the microstructure, texture, and mechanical properties of metals, necessitating comprehensive studies of twin formation and interactions. While experimental methods excel at analyzing individual samples, they often lack the capability for temporal analysis of twinned structures. Molecular dynamics simulations offer a temporal dimension, yet the absence of suitable tools for automated crystal twin identification has been a significant limitation. In this article, we introduce a novel computational tool integrated into the visualization and analysis software OVITO. Our tool automates the identification of coherent twin boundaries, links related twin boundaries, validates twin structures through orientation analysis, and tracks twins over time, providing quantifiable data and enabling in-depth investigations. Validation on a copper single crystal under shear loading demonstrates successful tracking of various twins, revealing their genesis and growth over multiple timesteps. This innovative approach promises to advance the computational materials science domain by facilitating the study of deformation twinning, offering profound insights into the behavior and mechanical performance of materials.

1. Introduction

In the realm of materials science and engineering, the mechanical properties of metals and alloys play a pivotal role in determining the structural stability, wear resistance, and longevity of components [1]. Responsible for the mechanical properties are the underlying deformation mechanisms, with deformation twinning being one of the prominent avenues for plastic deformation [2–4]. Deformation twinning involves the homogeneous shear of atomic layers within crystal lattices, resulting in the creation of distinct twin boundaries that separate regions of mirror-image symmetry. While slip is the predominant deformation mechanism for metals, twinning constitutes the second of the two primary mechanisms through which metals (particularly those with few independent slip systems) accommodate external stresses, thereby shaping their behavior under various loading conditions [5].

Twinning can be significantly pronounced in materials with low stacking fault energy [6,7], at high strain rates [4,8,9] or at low temperatures [7,10–12]. These conditions promote an increase in stacking fault width and therefore the occurrence of deformation twinning. Deformation twinning, in turn, enhances a material's strength and ductility [13] by increasing the work hardening behavior [6,14–16].

Deformation twins act as obstacles for dislocation glide, reducing the mean free path of dislocations [17], a phenomenon often referred to as the "dynamic Hall-Petch effect" [14,18,19]. In the context of steels, in the ductility-tensile strength or global formability diagram, twinning-induced plasticity (TWIP) steels lie above the range classically referred to as "material banana" [20]. In contrast to conventional steels, which often see a decrease in ductility as they become harder, TWIP steels, instead of losing ductility, actually exhibit an increase in the same [21].

Understanding and characterizing deformation twinning have emerged as vital endeavors for both fundamental research and practical applications [22]. The prevalence and evolution of twin boundaries dictate the material's response to stress, influencing aspects ranging from work hardening to crack initiation and propagation [23].

With the exception of some brave attempts, traditional experimental methods are often limited in their ability to capture the dynamic temporal evolution of twin structures during deformation processes [24,25]. These methods usually provide static snapshots of twinning events in isolated samples but struggle to reveal the intricate details of twin growth and interaction over time. To bridge this gap and gain a comprehensive understanding of twinning mechanisms, computational

tools have become indispensable. Molecular dynamics (MD) simulations present a powerful avenue for investigating the dynamic evolution of materials at atomic scales [26–28]. These simulations offer the unique advantage of observing the temporal progression of materials under external stresses [29,30]. However, despite their potential, MD simulations have long lacked a suitable tool to systematically identify and analyze twins automatically or at least semi-automatically. Such a tool would enable researchers to track twin formation, evolution, and interactions, facilitating a deeper understanding of twinning as a deformation mechanism.

This paper introduces a novel algorithm, built upon the OVITO platform [31], that addresses this crucial gap in the field. The algorithm serves a multifaceted purpose, automating the identification of coherent twin boundaries, establishing connections between related twin boundaries, and ultimately tracking the temporal evolution of the identified twin structures. By seamlessly integrating these functionalities, the algorithm provides quantifiable data with an option of detailed investigations into individual twins. To demonstrate the algorithm's efficacy, we applied it to study the deformation behavior of a copper single crystal subjected to shear. The algorithm successfully tracked and analyzed various twin boundaries, revealing insights into their genesis and growth over multiple timesteps. Furthermore, the orientation analysis reliably validated twins, showcasing the algorithm's accuracy and utility in elucidating the intricate mechanisms that underlie twinning.

This paper addresses the critical need for a powerful tool to study deformation twinning through MD simulations. By introducing an innovative algorithm capable of identifying and tracking twins, this study empowers researchers with a robust platform to explore the fundamental aspects of twinning, fostering deeper insights into material behavior and its implications for mechanical performance.

2. Used software

The basis for any visualization and analysis of MD simulations is the per-atom data of a successfully conducted simulation run. This atom or geometry data contains a snapshot of information such as position, velocity, stresses, order parameters, or other relevant properties of each atom in the simulated system. The obtained data is then usually displayed and evaluated with a software different from the one that produced the raw data.

The required model systems should be characterized by homogeneous composition and microstructure (single-crystal) and only subjected to uniaxial load. Based on established theories [15], it is possible to predict their response to the ruling conditions and subsequently compare these predictions to the obtained results. These simple custom samples were constructed to systematically test and validate the featured code. AtomSK [32] was used to generate realistic single-crystalline face-centered cubic (fcc) specimens as an input for the MD simulation. Simple normal-load cases were applied to single-crystalline structures using LAMMPS [33], an open-source molecular dynamics simulation software, in order to analyze unambiguous stress fields. Shear simulations were then conducted for the final goal of being able to successfully analyze more complex stress fields.

The algorithm developed in the course of this work is based on the powerful visualization and analysis software OVITO [31]. Important functions include the reliable recognition of crystal structures and their spatial position and rotation, determination of property relationships between neighboring atoms, and tracking of atomic motion over multiple time steps. Unfortunately, the predefined so-called modifiers in OVITO do not allow the user to identify twinned structures by default. However, the possibility to include custom Python scripts expands OVITO's capabilities considerably. As the stacking of the atoms is reversed at the twin boundary, twin boundaries in fcc crystals appear as single layers of atoms with hexagonal-close-packed (hcp) structure in OVITO. In contrast, stacking faults are detected as two adjacent layers of hcp-like atoms [34], which form the next-most similar structure.

This implies a difference between both patterns concerning an atom's number of nearest neighbors with the same structure type, which can be exploited to find planar structures. One of the important features of the presented tool is to robustly disambiguate between stacking faults and proper twins. Parameterized from the acquired data, planes can be compared to find crystallographically meaningful connections.

3. Implementation

The utilization of this tool is divided into two parts. First, it includes an automated script that acts on a range of given files containing atom or geometry data, e.g., LAMMPS dump files (see Fig. 1, steps 1–5). It builds on the problem of distinguishing between single-layer and all other hcp structures, in particular double hcp layers mentioned in Section 2. These formations are identified, parameterized, and tracked over time. The obtained data is visualized or stored for further analysis.

The second part is an interactive OVITO session in which detected twin boundary pairs can be examined in detail (see Fig. 1, step 6). It serves as validation for all detected potential twins, since the occurrence of structures, erroneously identified as twins, cannot be entirely excluded. For twin boundary pairs verified as true twins, the relative shift between the outside fcc regions can be computed, based on the atomic motion of the environment relative to a previous time step. In situations involving non-uniaxial loading, this may offer insight into the dominant stress direction that initiated its formation.

3.1. Automated Python script

The script makes use of some of OVITO's predefined modifiers and is complemented by custom Python functions. Algorithm 1 shows how clusters of possible twin boundaries are found utilizing basic OVITO modifiers and a short custom selection script. It receives a series of files and applies modifiers as follows.

Algorithm 1: Find possible twin boundary atoms

  input : OVITO compatible file with atom data
  output: List of clusters with planar structure

  Apply: *polyhedral template matching*
  if *structure type is hcp or unidentified* **then**
     Apply: *coordination Analysis*
     if *unidentified atom connects two well-defined planes* **then**
        append to selection
     else if *hcp atom is part of well-defined plane* **then**
        append to selection
     if *selected* **then**
        Apply: *cluster analysis*
        if *cluster size > threshold* **then**
           add to list of planar structures

The structure type property is determined using polyhedral template matching (PTM) [35], currently being the most reliable (and robust at high temperature [36] and strain [28]) structure identification algorithm, and the only one that has the capability to determine lattice orientation. Information about the orientation is needed to analyze the lattice environment of a suspected twin to confirm or reject its designation. Next, the coordination number for every atom of hcp-like or unidentified structure is computed, ignoring all other atom types (they are not considered as neighbors). The coordination number is determined by counting the neighboring atoms within a given cutoff
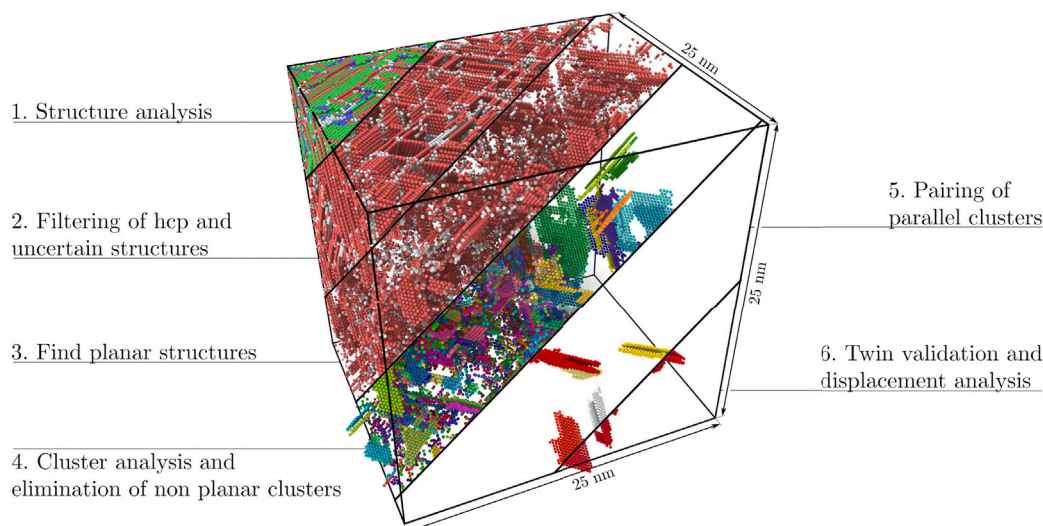
Fig. 1. Schematic step-wise break-down of the tool processes. Color schemes: Steps 1–3: according to structure type: fcc (green), hcp (red), bcc (blue), unidentified (white), Step 4: according to cluster size (arbitrary rainbow colors), Step 5: according to pair size (arbitrary colors from linear color gradient), Step 6: (if validated) according to relative shift between twin fcc matrix regions: narrowing (red), widening (blue), not validated (white).

distance. This cutoff value, which also applies to the cluster analysis, can be determined by evaluating the atoms' radial distribution function (RDF). The RDF is a measure of the probability density of finding two particles at a given distance in an atomic system. The RDF curve shows sharp peaks at certain radial distances, indicating the presence of well-defined interatomic distances for the first, second, and nth neighbors [37]. With increasing temperature, the peaks become blurred, but can still be automatically analyzed to find the cutoff distance, which is the RDF minimum between the peaks of the first and second nearest neighbors.

In fcc lattices, the densely packed planes are of type {111}, and these also become twinning planes in twins. Any atom located within such a layer has six nearest neighbors in that (perfect) plane. Large portions of the structures we are looking for are defined by the above criterion. However, because of the deformation twins' nature and distortion of the matrix due to strain, the formations often exhibit steps in otherwise planar structures, and they may also become warped. This partly results in atoms not being identified as twin boundary atoms and, e.g., in the case of steps, two separate planes are identified instead of one continuous twin boundary. To be able to successfully include the twin boundary in full, atoms of unidentified structure type are used to connect clusters of well-defined layers in close proximity. These linking atoms are defined by having at least two nearest neighbors of structure type hcp, and at least one with a coordination number of 6. If even one nearest neighbor exceeds a coordination number of 6, the center atom is discarded. Atoms of well-defined layers are associated with having a coordination number of 6 and structure type hcp as well as having a minimum of 3 nearest neighbors with the same characteristics.

Fig. 2 shows a step between two well-defined planes with unidentified atoms and holes (i.e., out-of-plane or non-hcp atoms) in between, where the former act as links between both planes. In cases where other formations such as double- or multi-layer stacking faults, undefined hcp-like structures, and other crystal defects intersect or border ideal plane configurations, the resulting selection may not be perfectly planar. This lack of perfect planarity can have a negative impact on the quality of the parameterization process' results. Specifically, if holes within the selection grow too large, it is possible for coherent twin boundaries to become disconnected. OVITO's cluster analysis modifier finds disconnected groups (so-called clusters) based on a distance cutoff between single atoms. The result is disturbed by the noise from single-atom, non-planar, or one-dimensional (string of atoms) clusters. This disturbance can be reduced by discarding all clusters with a size of less

than a threshold number of atoms (default value: 50). If one intends to use the temporal tracking feature, the threshold should be kept high, discarding the numerous small twins, as their spatial density heavily affects the tracking accuracy. The removal of unwanted structures not covered by this will be performed at a later stage. Thereby, every single layer and its corresponding atoms are tagged and sorted into a list according to size. In case of periodic boundary conditions, clusters straddling the periodic boundaries must be unwrapped, meaning their atoms are partly re-positioned. This is necessary to ensure a correct parameterization of the plane. However, clusters outside the simulation cell will no longer have a connection to any atoms discarded earlier. Thus, twin validation and displacement analysis are not available for these objects, since the re-positioned atoms have no adjacent fcc neighbors whose orientation could be analyzed.

Algorithm 2: Parameterize planes

   input : List of clusters with supposed planar structure
   output: Data table containing plane-defining vectors

   for *all clusters* **do**
      principal component analysis
      if *shape not planar* **then**
         delete cluster from planar structures list
      else
         save normal vector and one orthogonal eigenvector to
         a list

In the following, regression planes are calculated based on each cluster's atom positions. Algorithm 2 summarizes how this helps to eliminate any further shapeless clusters and to allow further processing of mathematically described planes. Principal component analysis (PCA) is an unsupervised method with all independent but still correlated variables [38], which results in the desired orthogonal regression (see Fig. 3).

PCA essentially causes a principal axis transformation, finding a more reasonable basis of a vector space for a certain dataset. With the origin at the center of the point cloud, the axes of the new coordinate system are rotated successively in the direction of the next largest variance [40]. Generally, the first few principal components will account for most of the variation in the original variables if a set of more than two variables shows substantial correlations among them [38]. Transferred to the spatial coordinates at hand, this allows a rough estimate of the shape of the point cloud. Since the studied
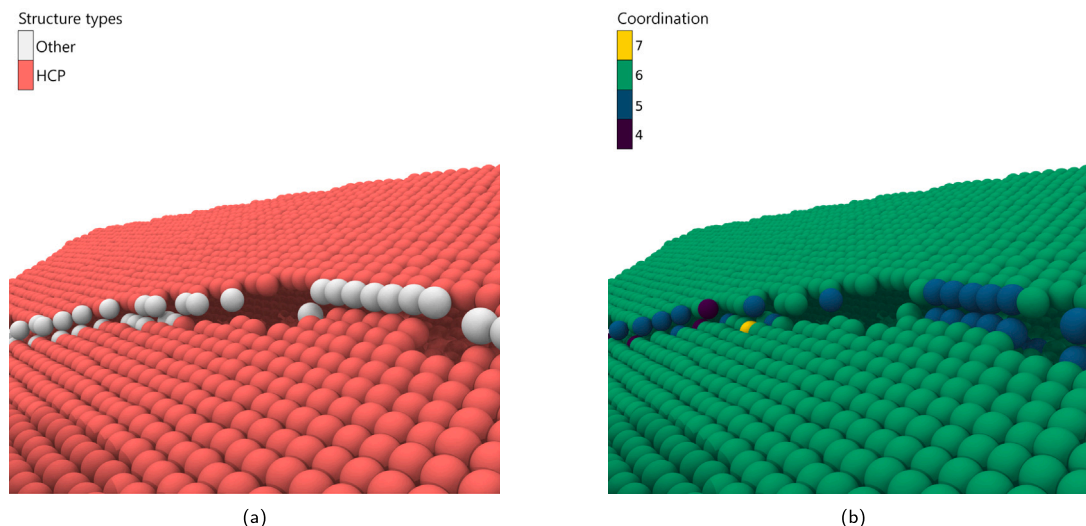
Fig. 2. Visualization with focus on the linking atoms between two adjacent planes. (a) the planes consist only of hcp type atoms, the linking atoms are primarily of unidentified ("other") structure type. (b) in-plane atoms have a coordination of 6; that of the linking atoms varies from 4 to 7. The combination of this information allows to identify the linking atoms, so that the cluster analysis correctly connects both planes.
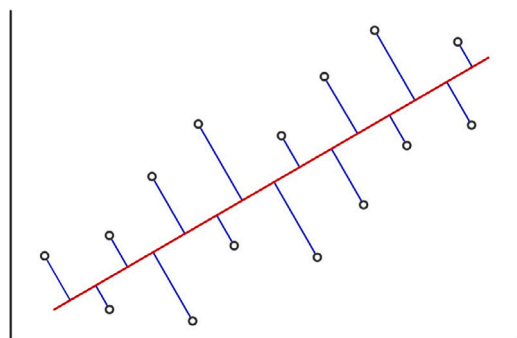

Fig. 3. Regression line in two dimensions computed with principal component analysis. *Source:* Adapted from [39]

twin boundary atoms exhibit a strong linear correlation regarding their positions in the direction of the plane-spanning vectors, PCA can be successfully applied to parameterize planes.

In practice, the application part is limited to setting up the covariance matrix [41] and computing its eigenvalues and eigenvectors. These resemble the variances in atom positions and their directions, respectively. From the above stochastic and mathematical relationships, the approximate geometry of a parameterized cluster can be obtained. Since a perfect plane has the minimal extent, thus variance, in the direction of its normal vector, this plane-defining vector is the eigenvector corresponding to the smallest eigenvalue. The remaining two eigenvalue-eigenvector-pairs allow the plane's aspect ratio to be determined, which facilitates the elimination of non-planar structures.

The obtained list of possible twin boundary forming clusters is now unfolded into a 2-D array, comparing the clusters individually. The following geometrical relations (for abbreviations, see Table 1) are calculated for each comparison (see Fig. 4 for a visualization):

Table 2 displays the empirically obtained conditions used in the plane pairing algorithm 3. The tolerance in condition C_α is necessary because of unavoidable deviations from a perfect plane configuration (single atoms) and the consideration of stress-related twin boundary curvature. The remaining conditions reduce the number of crystallographically unreasonably paired planes. C_β prevents that two distant planes, which in reality could never form a twin, are paired (if D_N is not small against D_{max}). C_{COM} forbids that two different clusters in the

Table 1

Abbreviations for geometrical relations computed between two plane-shaped clusters.

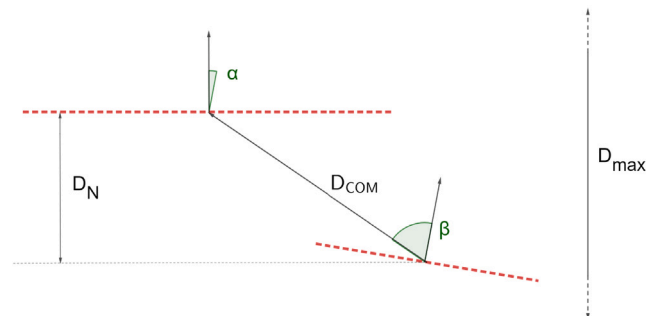|	
α	angle between normal vectors
β	angle between one normal vector and each cluster's center of mass (COM) connecting vector
D_N	distance in normal vector direction between COMs
D_{COM}	distance between COMs
S_{rel}	size relation (ratio between number of atoms per cluster)
D_{max}	largest dimension of the simulation cell

Fig. 4. 2-dimensional visualization of geometrical relations computed for pair finding.

Table 2

Empirically acquired conditions for cluster pairing algorithm 3.

C_α	$\alpha < 5°$
C_β	$\beta < 90° - 90° \times 20 D_N / D_{max}$
C_{COM}	$D_{COM} < 3 \times D_N$
C_{Size}	$0.2 < S_{rel} < 5$

same plane are assigned to each other. Lastly, C_{Size} provides a balanced size relation between twin boundaries.

All important values of the found pairs are exported in case the user decides to do an interactive OVITO session afterwards. If multiple matches are identified, a comparison of the values of D_N is performed for each of the matches. The two matches with the closest values of D_N are then selected as the final result.

The tracking algorithm 4 relies on comparison of position and orientation between timesteps. The mean values of the COM and normal vector properties of the two twin boundaries are compared between

---

Algorithm 3: Find parallel planes

    input : List of normal vectors of clusters with planar structure
    output: LAMMPS dump file and data tables

    for *all permutations of the computed planes* **do**
        Compute: α, β, D_N, D_{COM}, S_{rel}
        see `Table 1`

        if *plane combination satisfies all conditions*
        see `Table 2`
        then
            if *plane already paired* **then**
                if D_{COM} *(selected pair)* $<$ *than* D_{COM} *(already paired)* **then**
                    replace previously found pair with selected pair
            else
                append plane pair to list of possible twin boundary pairs

    Assign: *"possible twin" marker to atoms of selected pairs*
    Export: *related data tables*

---

the currently detected structures and the information stored from the previous timestep. At the first provided frame, the list of detected twin structures is initiated and will serve as reference for the next frame. Starting at frame 2, the data table containing information from the previous timestep is read. A twin is tracked if a structure from the referenced timestep is of the same orientation (5° tolerance), less than 15 Å (based on a sampling interval of 2 ps), away in normal direction, and if the combined COM lies within a range of 50 Å. The previous twin's ID will be assigned to the successfully tracked twin. If no match is found, the selected twin receives its own unique ID. After iterating through all structures of the current timestep, the timestep of origin and of disappearance of non-tracked twins are transferred to the updated data table. In case of multiple matches, the one closest to the reference position is selected. This tends to be the most computationally demanding aspect, depending on the number of found clusters.

Algorithm 4: Track twins over provided timesteps

    input : List of positional data of twins detected in previous timestep
    output: Altered atom data file and data table of tracked, lost, and new twins

    for *all twins found at current timestep* **do**
        if *orientation and position similar to twin in previous timestep* **then**
            Tracking

        else if *no match found* **then**
            New twin

    carry over data of lost twins
    Export: *LAMMPS dump file and updated tracking data table*

3.2. Interactive OVITO session

While the successfully paired planes have passed the first requirement for twin identification, they have not yet undergone validation as genuine twins. The results may still contain erroneously paired twin boundaries or structures that do not qualify as twin boundaries at all. Therefore, further identifiers such as the associated changes in lattice orientation and other related characteristics must be confirmed. To address this, each twin boundary pair is systematically re-evaluated. For identification of the crystallographic orientations in the twins' immediate vicinities, the results of the first script can be passed into the OVITO data pipeline. OVITO can load either a single file or an entire sequence. In both cases, all required data tables are automatically imported for each timestep. The evaluation of previously found potential twin boundary pairs is first done by examining the orientation of the areas between and to some extent also outside the recognized planes. The difference in crystallographic orientation between the twin and the matrix atoms can be seen in Fig. 5(a) (in this case the twin is formed during compression). This well known schematic representation of a deformation twin, served as inspiration to apply analogous segmentation to the detected twin boundary pairs (see Fig. 5(b)). For this purpose, the orientation of all fcc-type and twin boundary atoms inside a cylinder aligned with the vector connecting the centers of mass are saved together with their normal distance to one of the planes. The cylinder's radius is determined by the size of the cluster, but has a minimum of 10 Å to compensate for possible holes and small defects between planes.

Because of the twin forming in parallel planes, clusters of similar crystallographic orientation appear at regular intervals when the collected data is plotted. Every cluster represents one parallel layer that the cylindrical mask has intersected. Only if steps of a twin boundary are intersected, the cluster will scatter along the vertical axis, as this layer shares atoms of matrix or inner twin and twin boundary. The PTM stores the orientation of any analyzed atom as a quaternion, a four-component vector whose entries can be a suitable choice for a scalar identifier of orientation.

A normalization of the resulting scattered orientation data to a range between zero and one helps spread the (possibly small) deviations in orientation between atoms in close proximity. This enables a more meaningful, high-contrast display of plots, enabling automatic identification of special features (in this context sudden changes in orientation), but it also improves floating point precision. The aim is to achieve clear demarcation between spatial regions of different orientation. To determine which component of the orientation quaternion provides the highest contrast between matrix and twin, all four components are compared to the pattern that would be expected for a perfect twin. For better comparability, these orientation components are then separately replaced with their median values over all atoms in the previously mentioned clusters. Grouping the data, this time by orientation, makes it possible to autonomously identify the graphs' shape. Large jumps in the graph come from a significant change in orientation, which ideally indicates the position of the first twin layer, the twin boundary. In OVITO, the respective atoms are identified as hcp-like and are not connected to the surrounding fcc matrix, which is why in some cases the twin boundary is followed again by a jump in the graph (see Fig. 6).

The two patterns shown in this figure can be divided into five sections of similar orientation in the case of Fig. 6(a) and three for Fig. 6(b). All detected structures sharing one of the patterns are reliably identified as true twins. In the last step, the pairing algorithm's results are reviewed. Due to the division of the areas, they can also be considered individually so that knowledge about which atoms are located outside the twin boundaries is gained.

This segmentation is performed according to the quaternion component that has the highest contrast in orientation between matrix and twin regions. Then, from the median values of the quaternions representing the orientations of these newly defined regions, we can calculate both the rotation angle θ and the associated axis **n** between the twinned fcc volume and the surrounding fcc (matrix) region. If the rotation axis comes out normal to the twinning plane and the rotation angle is close to 60° ($\Sigma 3$ boundary) [43], the twin has been successfully confirmed.

Both quantities θ and **n** can be determined, according to Ref. [44], by first forming the (non-commutative) quaternion product of one
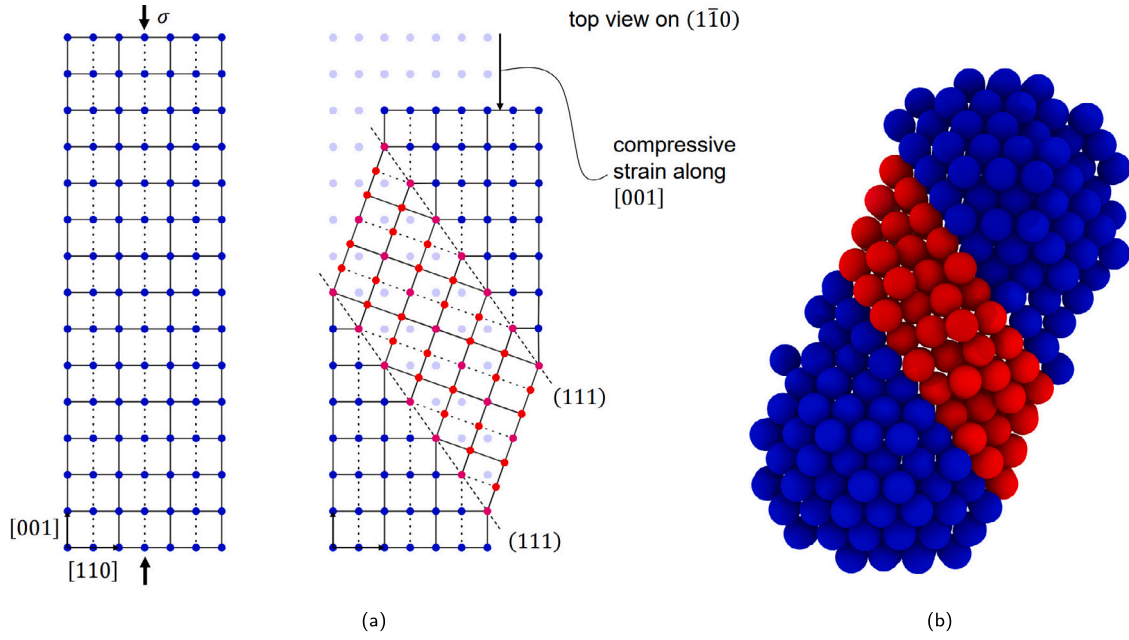
Fig. 5. Visualized concept of dividing the selection into distinct regions, which facilitates the orientation and displacement analysis: relative displacements of the blue regions after twinning as an indication of the type of formation. (a) 2-dimensional scheme of compression twin formation in fcc lattices [42, p. 4]. (b) Atom selection according to algorithm 5: 3-dimensional equivalent to figure (a): twin atoms (red) and matrix atoms (blue).
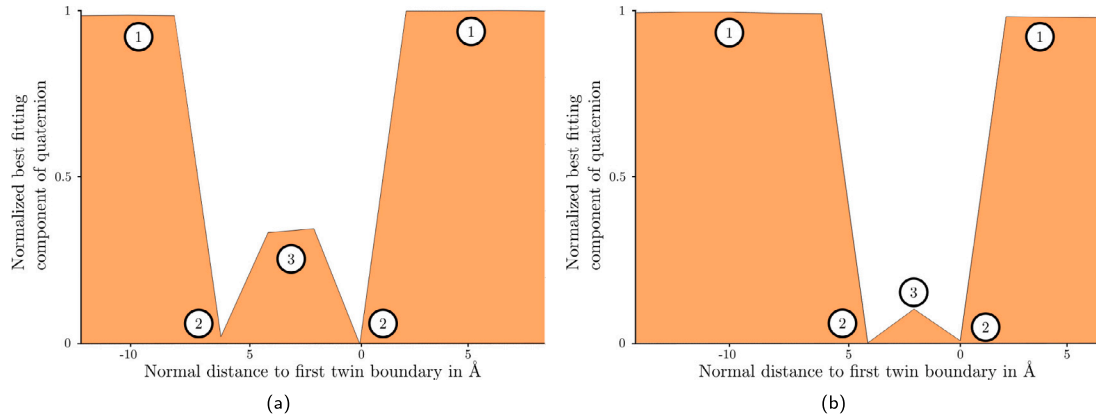

Fig. 6. Average crystallographic orientation against normal distance of atom layers parallel to twin boundaries: (a) twin fcc structure with distinct difference in orientation to twin boundaries. (b) twin fcc structure with little to no difference in orientation to twin boundaries. 1: surrounding matrix, 2: twin boundaries of hcp-like structure, 3: twin fcc structure enclosed by twin boundaries.

quaternion with the inverse of a second one, yielding a single quaternion that represents the rotation from one crystal region/orientation into the other:

$$
\begin{aligned}
x\,y =&(x_0 y_0 - x_1 y_1 - x_2 y_2 - x_3 y_3)\\
&+(x_0 y_1 + x_1 y_0 + x_2 y_3 - x_3 y_2)\mathbf{i}\\
&+(x_0 y_2 - x_1 y_3 + x_2 y_0 + x_3 y_1)\mathbf{j}\\
&+(x_0 y_3 + x_1 y_2 - x_2 y_1 + x_3 y_0)\mathbf{k}
\end{aligned}
\tag{1}
$$

This quaternion is of the form $q = q_0 + \mathbf{i}q_1 + \mathbf{j}q_2 + \mathbf{k}q_3$. The corresponding rotation matrix can then be obtained as

$$
\mathbf{Q} = \begin{bmatrix}
1 - 2q_2^2 - 2q_3^2 & 2q_1 q_2 + 2q_0 q_3 & 2q_1 q_3 - 2q_0 q_2\\
2q_1 q_2 - 2q_0 q_3 & 1 - 2q_1^2 - 2q_3^2 & 2q_2 q_3 + 2q_0 q_1\\
2q_1 q_3 + 2q_0 q_2 & 2q_2 q_3 - 2q_0 q_1 & 1 - 2q_1^2 - 2q_2^2
\end{bmatrix}.
\tag{2}
$$

From this, the rotation axis $\mathbf{n}$ is calculated component-wise using the non-diagonal elements a_{ij} of $\mathbf{Q}$.

$$
\mathbf{n} = \frac{(a_{23} - a_{32}), (a_{31} - a_{13}), (a_{12} - a_{21})}{\sqrt{(a_{23} - a_{32})^2 + (a_{31} - a_{13})^2 + (a_{12} - a_{21})^2}}
\tag{3}
$$

The diagonal elements of $\mathbf{Q}$ are used to compute the rotation angle θ.

$$
\theta = \arccos 0.5(\mathrm{tr}(\mathbf{Q}) - 1)
\tag{4}
$$

It is obvious that the success of the orientation analysis depends significantly on the selection of atoms. Modifying the cutoff radius of the sphero-cylindrical mask used for the selection brings only visual improvements. Fig. 7(a) suggests an advantage of smaller cutoffs, as the drop in orientation on the left is avoided (compared to the other plots). Since this is highly individual though, and the tolerance for outliers is lowered, a minimum cutoff of 10 Å was selected. Increasing the value to 15 Å shows no remarkable improvement (compare Figs. 7(b) and 7(c)). However, it is certain that an overly large radius will exacerbate the discontinuity (see in particular the region between the twin boundaries in Fig. 7(d) and previous plots). In highly dynamic systems, increasing the cutoff will always deface the plots, as the density of hcp-like structures is high, which mostly account for the arbitrary changes in orientation.

The issue of intersecting steps within twin boundaries during orientation analysis can potentially be addressed by combining information
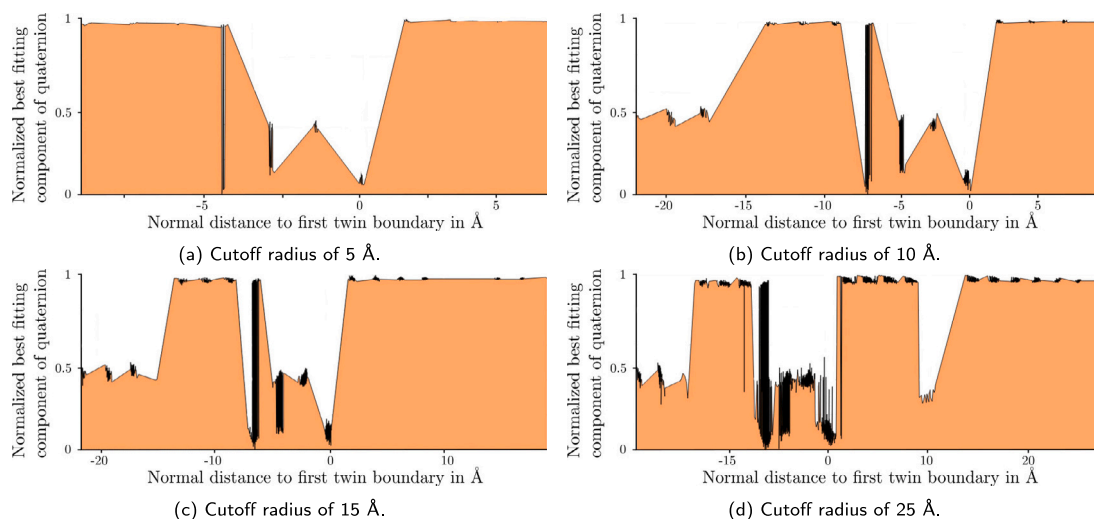
(a) Cutoff radius of 5 Å.

(b) Cutoff radius of 10 Å.

(c) Cutoff radius of 15 Å.

(d) Cutoff radius of 25 Å.

Fig. 7. Various cutoff radii for the selection mask during orientation analysis.

Algorithm 5: Validation and displacement analysis

    input : Atom data files of previous algorithm

    output: Orientation-vs.-normal-distance plots

    for *every twin boundary pair* **do**

        select atoms within cylindrical mask

        find best fitting quaternion component

        plot orientation against normal distance to one plane

        if *graph matches pattern* **then**

            compute rotation angle and axis for validation

            get displacement of regions outside the twin

            if *distance between region's COMs decreased* **then**

                mark as narrowing twin

            else

                mark as widening twin

---

from the entire preceding analysis sequence. This approach can help consolidate twin validation and determine its relative shift in each system snapshot. Since the steps resemble the moving Shockley partial dislocations, there will typically be at least one frame where the plot is unambiguous.

Comparing the displacement of the sections surrounding the twin (see Fig. 6 Section 1 and Fig. 5 blue atoms), their relative movement can be computed. This relative shift, either narrowing or widening, is assigned to every successfully detected and validated twin. In non-uniaxial load cases, this might provide the possibility to deduce the predominant sign of stress that led to that twin's formation. This approach is based on the observable relative movement of the areas outside the twin (see Fig. 5(a)). This concept is transferred to the masked region (see Fig. 5(b)) by computing the average position (equivalent to COM) and displacement of both atom groups (see blue atoms) and comparing the resulting average positions of both frames. If, after subtracting the displacement from the COM for both regions, the distance between the new positions has decreased in comparison to the distance only between the centers of mass, it suggests the outside fcc regions have moved closer together. Analogous thoughts apply in the opposite case. Algorithm 5 summarizes the key points of the interactive session.

4. Representative case study

In this section, we present a short case study to showcase the capabilities and benefits of the presented tool. The object of the study was a copper single crystal with initial dimensions of $25 \times 25 \times 25$ nm^3 under shear, and an embedded atom method (EAM) potential was used to model the atomic interactions [45]. The NVT (canonical) ensemble was simulated using a Nosé–Hoover thermostat [46–48] with a damping parameter of 0.2 ps.

The system was equilibrated at a temperature of 200 K with the x, y, and z axes corresponding to the [100], [010] and [001] directions, respectively. A rigid top section consisting of two monolayers of Cu is sheared in positive x direction at a velocity of 360 m/s (resulting in an effective strain rate of $\simeq 0.0113$ ps^{-1}). As a result of the applied strain rate, perturbations near the rigid regions emerged, acting as zones that absorbed and mitigated additional artifacts, thus preventing their propagation into the region of interest.

A velocity gradient was applied to the region of atoms enclosed between the two rigid layers that impose the kinematics, uniformly inducing shear to quickly reach dynamic equilibrium. This approach allows imposing high shear rates without risking artifacts resulting from acceleration of the boundary regions that may not be perfectly smooth.

The timestep size was set to 2 fs, and the first twin structures were detected after 12 ps. Fig. 8 gives an overview of which structures are identified as possible twins. The center left structure found in both Fig. 8(a) and 8(b) (indicated by the white arrow) is noteworthy as it is bordered by a stacking fault, which disfigures the resulting structure. The presence of atom accumulations within twin boundaries, particularly when situated centrally, is observed to have a limited influence on the direct detection of coherent twin boundaries. However, these accumulations can significantly perturb the correct pairing (owing to contorted plane parameterization) and hinder the twin validation through atom orientation analysis.

The algorithm for twin boundary atom identification 1 results in considerable noise of monatomic and amorphous clusters (see Fig. 9(a)). As larger structures are expected and of more interest, a generous cluster size threshold of 100 atoms was set to de-clutter the picture. Non-planar accumulations were removed after being identified via PCA (Alg. 2). Of over 6000 found clusters, only 40 were of the desired shape, and 20 were successfully paired (see Fig. 9(b)).

Fig. 10 shows a twin that was successfully tracked over several frames. At $t = 18-22$ ps, we see the moving Shockley partial dislocation that sets the twin apart from a regular stacking fault, thus contributing
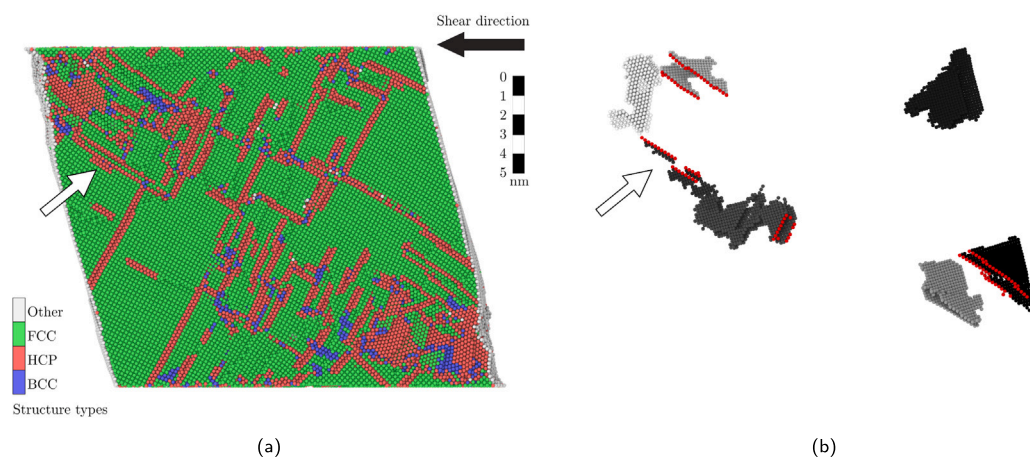
Fig. 8. Same cross-section of the simulation showing (a) the structure identified with PTM and (b) the identified twin-like structures with atoms in the slicing plane highlighted in red. The system has been cropped to emphasize the region under analysis.
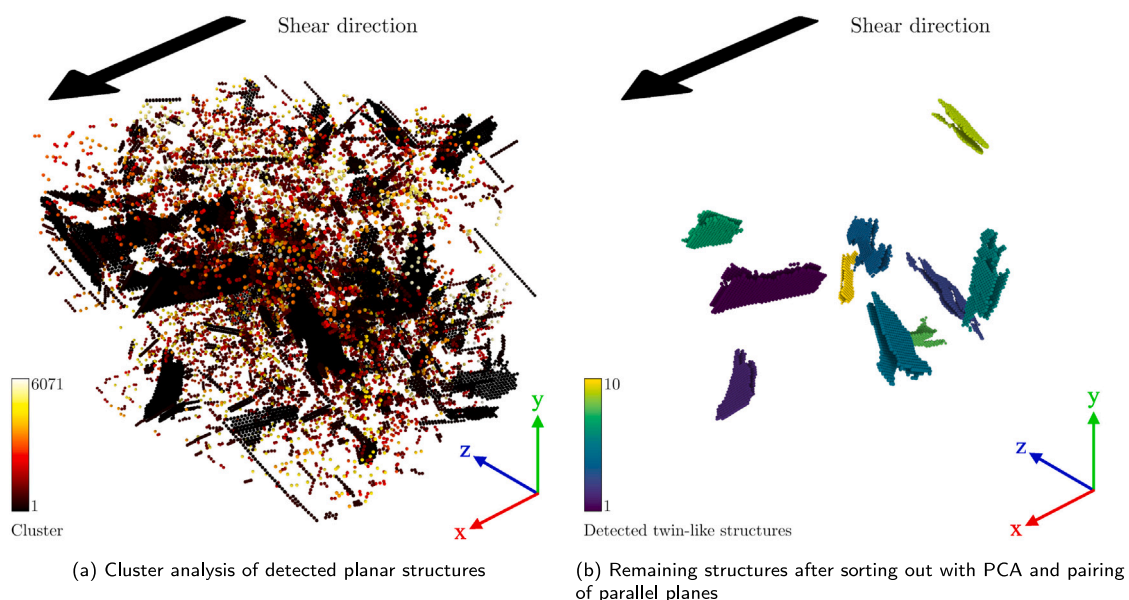

(a) Cluster analysis of detected planar structures

(b) Remaining structures after sorting out with PCA and pairing of parallel planes

Fig. 9. Interim results of twin finding algorithm after 28 ps, corresponding to 2.8 nm of shearing.
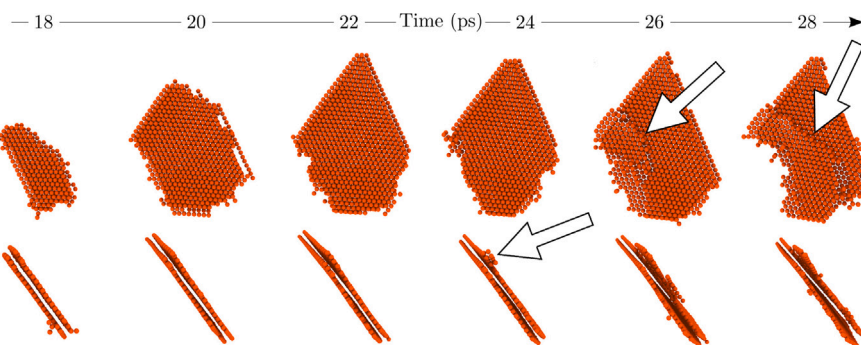

Fig. 10. Tracking of single twin, movement of Shockley partials clearly visible.

to the twin's growth. Results after 24 ps demonstrate that because of crystal rotation, distortion, and other external influencing factors, the selection of atoms belonging to the twin boundaries cannot be perfect. Nevertheless, the rough shape remains constant, allowing the twin to be tracked. The time range $t = 24 - 28$ ps illustrates the next twin layer building up.

Fig. 11 offers a detailed look on a clean and readily identifiable example of a four-layer twin. No step or moving Shockley partial is captured within the selection and no disfiguring holes or crystal defects interfere with the twin boundaries. Thus, the plot of the quaternion component describing the crystallographic orientation, used to automatically identify the separated outside matrix regions from the
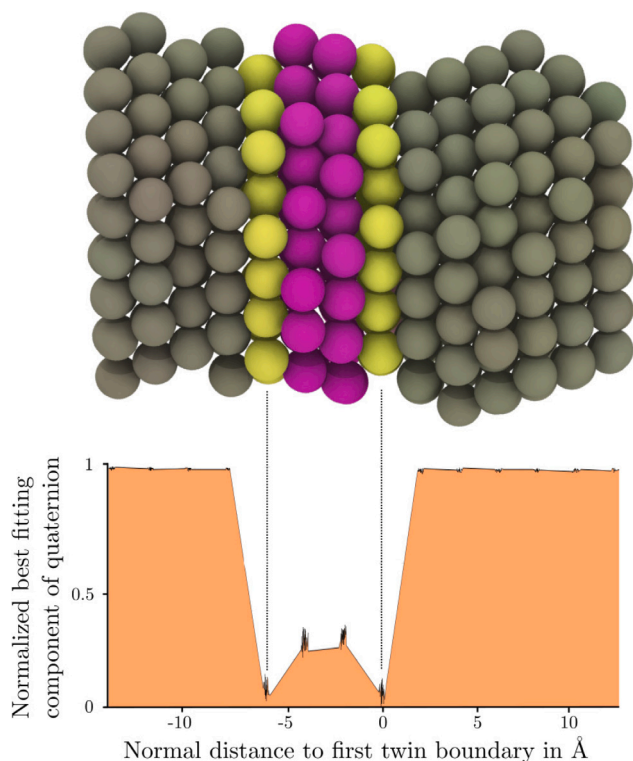
Fig. 11. Selection of atoms used to capture and compare the orientation of the outside fcc regions (dark gray) and the twinned volume (violet), separated by the twin boundaries (yellow). The color scheme was derived after translating each per-particle quaternion into an RGB color. The best fitting quaternion-component found in algorithm 5 is plotted against the normal distance to one twin boundary.

twinned regions, exhibits a clear shape that allows for a distinction of twinned and regular matrix volume for the selection.

Calculating the rotation axis and angle twice, once using quaternions from the left region and once from the right region, confirms the twin characteristics. This yields an angle of 59.89° with a 0.28° deviation from the normal vector of the twin boundary plane for the left-side approach, and 60.40° with a 2.67° deviation for the right-side approach, respectively.

However, not all twins can be identified so easily. Further inspecting the same twin introduced in Fig. 10 at $t = 28$ ps reveals interesting as well as challenging details.

Fig. 12(b) demonstrates that, in contrast to the right twin boundary B, its left counterpart is divided into two. The segment labeled A2 lies within the twinned region and shares a vertical {111} plane with it, which leads to a broader scattering of orientation values along the vertical axis. Segment A1 forms the remainder of the left twin boundary and shares its layer with some of the outside fcc lattice, exhibiting a significant difference in orientation that results in an even larger spread among the values. While the reason for the split twin boundary can be a step (a consequence of the deformation twin's shape), Fig. 10 suggests that it corresponds to a forming Shockley partial. In this case, as the cluster atom furthest from the right twin boundary happens to be a left twin boundary atom, the graph is again characterized by a jump, making rapid analysis by eye difficult. The gap between the light red and white region belongs to a different planar structure that was excluded from the evaluation in order to avoid jumps in the data. Despite these challenges, computing the median for each cluster effectively smooths the plot, making automatic validation possible. Due to the ambiguous position of the left twin boundary, the calculated angle and axis deviation from the left-side approach deviate significantly from the expected values of 60° and 0°, respectively. The orientation

difference is 61.56°, which is within an acceptable range. However, the axis deviation measures 12.61°. This underscores the importance of cross-verification, as the right-side approach, considering the planar twin boundary, yields an angle of 59.97° and a deviation between rotation axis and normal vector of only 0.23°.

5. Challenges and improvements

This section discusses the issues and other challenges encountered during twin identification and tracking. The first attempt to identify planes showed, on closer examination of individual twins, that structures belonging together were not correctly identified as such if the selection criterion was kept too strict. Finding single layers by the coordination number of hcp-like atoms does not account for possible steps within twin boundaries that form due to the mechanism of deformation twinning. Studying the structural environment of single layers revealed that atoms of unidentified lattice type could act as the connecting links between well-defined planes. Adapting algorithm 1 to this knowledge not only dramatically improved the pairing process, but finally allowed to detect twin boundaries in their actual lenticular shape (see Fig. 13). However, by relaxing the conditions, more atoms are selected, which may establish unwanted connections to other structures. The identification of linking atoms is a substantial element of the algorithm that is currently being optimized towards more accurately detecting the twin boundary's entire shape on one hand, and reducing the number of artificial connections on the other.

Several of the empirically derived values for the conditions for pairing twin boundaries and tracking twins listed in Table 1 remain debatable. Unlike indisputable parameters such as average neighbor distance obtained from an RDF-analysis or coordination numbers of atoms in a perfect fcc lattice, the most convenient values for pairing and tracking can slightly change at every timestep. The pairing algorithm 3 contains a design decision that may have a noticeable negative impact on the outcome in some cases. Whenever multiple parallel matches for one plane are found, the two closest clusters are paired. Thus, the adjacent twin boundaries of two parallel and close twins might be falsely paired, and since the tracking algorithm is built into the autonomous script, it is possible for non-twins to be erroneously tracked. This can be resolved by performing the orientation analysis parallel to the pairing process. First, the conditions must be mitigated so that a broader selection of possible matches is made. That selection can be filtered by considering the orientation between supposed twin and outside region. The implementation of this approach would significantly improve the validity of the pairing. The final pairing would no longer rely on empirically determined values, but on values based on clear references in the literature. The current separation of automatic pairing and tracking and interactive OVITO session was chosen for a better insight and reduced computational costs during testing.

While the case study focused on showcasing the robust capabilities of our tool in various aspects, it is worth noting that we did not present the results of the displacement analysis introduced in 3.2. At the time of this research, development and testing of this particular feature were still in their initial phases. Linking the results of the displacement analysis to concrete local stress fields required more extensive testing and validation than was possible within the scope of this study. However, we emphasize the inherent potential of this analysis. As a tool in continuous development, it holds promise for future research and applications. Our future goals include further testing and refinement to provide a meaningful link between displacement and local twin formation stress.

At this point, it should be noted that OVITO has recently introduced an implementation for identifying planar defects in fcc lattices [49]. Our tests revealed that the modifier robustly identifies coherent twin boundaries, which is suitable for finding annealing twins. However, it falls short when it comes to detecting deformation twins, due to their out-of-plane shape. By contrast, our tool can detect twin boundaries
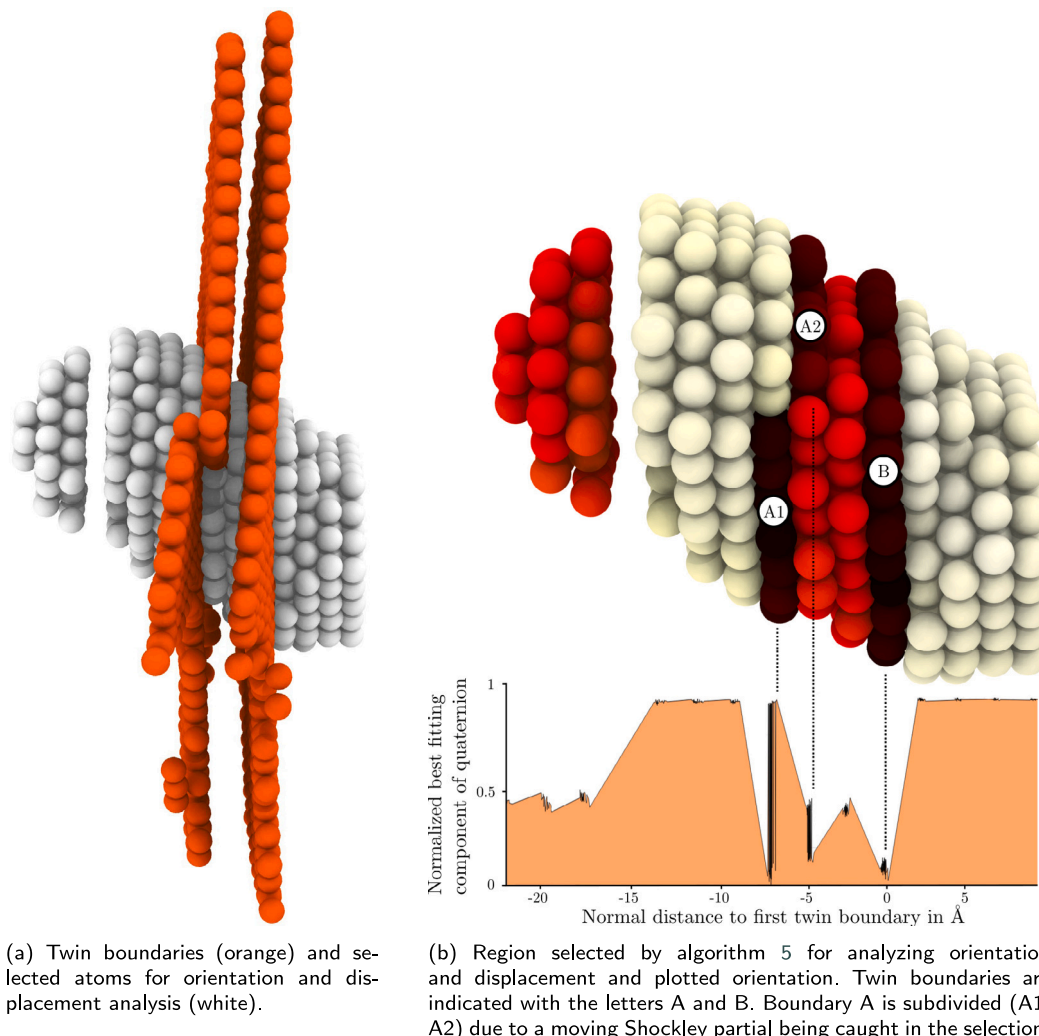
(a) Twin boundaries (orange) and se-
lected atoms for orientation and dis-
placement analysis (white).

(b) Region selected by algorithm 5 for analyzing orientation
and displacement and plotted orientation. Twin boundaries are
indicated with the letters A and B. Boundary A is subdivided (A1,
A2) due to a moving Shockley partial being caught in the selection.

Fig. 12. Case of twin containing a step that complicates automatic validation.
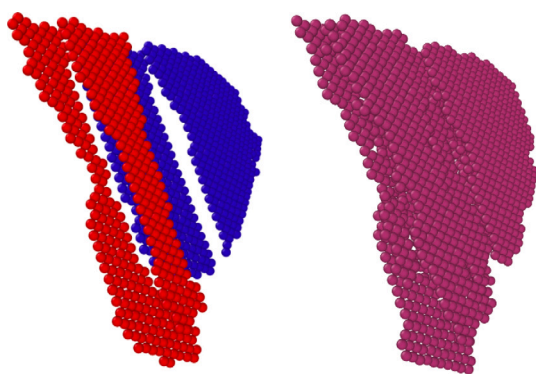

Fig. 13. Comparison of results of algorithm 1 before (left) and after (right) revision
of the same. The large gap between two formerly individual pairs (red and blue) is
closed by including atoms outside a well-defined plane leading to the detection of the
entire structure (purple).

that extend over multiple planes, offering a more comprehensive analy-
sis. Most notably, our tool distinguishes itself through the capacity to
validate its own findings, employing an orientation analysis between
twinned volume and regular fcc region, thereby ensuring a higher
degree of accuracy and reliability.

6. Conclusion and outlook

The objective of this work was to develop a post-processing analysis
tool for molecular dynamics simulations that allows a comprehensive
examination of deformation twins in fcc lattice materials. In detail,
the goals were (1) identification of coherent twin boundaries, (2)
connection of boundary pairs, (3) displacement analysis of validated
twins and (4) their temporal tracking. These goals were pursued by (1)
implementing an algorithm taking advantage of OVITO's capabilities
to detect twin boundaries as hcp-like single layers, (2) plane parame-
terization by PCA and pairing based on simple geometric relations, (3)
orientation and displacement analysis of sensibly selected regions, and
(4) cross-timestep spatial analysis.

The results show that twin structures can be detected, tracked,
and even examined individually. We can clearly distinguish between
stacking faults and twins, while a temporal analysis allows detailed
views and an understanding of the formation mechanism of individual
twins.

However, some challenges appearing during the methods develop-
ment could not yet be solved universally. The first selection of possibly
planar structures is crucial, and inaccuracies at this stage degrade
the results of all later computations. Sub-graph-matching could be a
promising alternative for the empirically determined conditions, as
successfully proven by Stukowski [50]. In the course of this work, it
became apparent that most of the difficulties stem from empirically
determined numerical values used as hard-coded conditions.

Nonetheless, this tool can already provide valuable insights into twin formation by detecting the majority of twin-like structures and allowing reliable validation of the same. Especially in sufficiently large systems, where twinning is the preferred deformation mechanism over dislocation slip, good results can be obtained. Materials scientists may use the tool to collect quantifiable data and examine individual structures, enabling them to better study the formation mechanisms of deformation twins in complex systems.

CRediT authorship contribution statement

H.J. Ehrich: Data curation, Methodology, Software, Visualization, Writing – original draft. **A. Dollmann:** Validation, Writing – review & editing. **P.G. Grützmacher:** Validation, Writing – review & editing. **C. Gachot:** Supervision, Writing – review & editing. **S.J. Eder:** Conceptualization, Data curation, Project administration, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The code used in this study is available on GitHub at https://github.com/hehrich/automated-twin-tracking.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used GPT-3.5 in order to shorten and improve the readability of the introduction. After using this tool, the authors reviewed and edited the section as needed and take full responsibility for the content of the publication.

Acknowledgments

References

[1] W. Zhai, L. Bai, R. Zhou, X. Fan, G. Kang, Y. Liu, K. Zhou, Recent progress on wear-resistant materials: Designs, properties, and applications, Adv. Sci. 8 (11) (2021) 2003739.

[2] J.A. Venables, Deformation twinning in face-centred cubic metals, Philos. Mag. 6 (63) (1961) 379–396.

[3] S. Mahajan, D.F. Williams, Deformation twinning in metals and alloys, Int. Metall. Rev. 18 (2) (1973) 43–61.

[4] J.W. Christian, S. Mahajan, Deformation twinning, Prog. Mater. Sci. 39 (1–2) (1995) 1–157.

[5] M.A. Meyers, O. Vöhringer, V.A. Lubarda, The onset of twinning in metals: A constitutive description, Acta Mater. 49 (19) (2001) 4025–4039.

[6] S. Martin, S. Wolf, U. Martin, L. Krüger, D. Rafaja, Deformation mechanisms in austenitic trip/twip steel as a function of temperature, Metall. Mater. Trans. A 47 (2016) 49–58.

[7] E. Kwon, S. Fujieda, K. Shinoda, S. Suzuki, Texture evolution and FCC/HCP transformation in Fe–Mn–Si–Cr alloys by tensile deformation, Mater. Sci. Eng. A 527 (2010) 6524–6532.

[8] O. Grässel, L. Krüger, G. Frommeyer, High strength Fe–Mn–(Al, Si) trip/twip steels development—properties—application, Int. J. Plast. 16 (10–11) (2000) 1391–1409.

[9] G. Frommeyer, U. Brüx, P. Neumann, Supra-ductile and high-strength manganese-trip/twip steels for high energy absorption purposes, ISIJ Int. 43 (3) (2003) 438–446.

[10] T.H. Blewitt, R.R. Coltman, J.K. Redman, Low-temperature deformation of copper single crystals, J. Appl. Phys. 28 (6) (1957) 651–660.

[11] T. Mori, H. Fujita, Twinning deformation in single crystals of cu-8 at, Trans. Japan Inst. Metals 18 (1) (1977) 17–24.

[12] S. Allain, J.-P. Chateau, O. Bouaziz, S. Migot, N. Guelton, Correlations between the calculated stacking fault energy and the plasticity mechanisms in Fe–Mn–C alloys, Mater. Sci. Eng.: A 387–389 (2004) 158–162, 13th International Conference on the Strength of Materials.

[13] B.C. De Cooman, O. Kwon, K.-G. Chin, State-of-the-knowledge on twip steel, Mater. Sci. Technol. 28 (5) (2012) 513–527.

[14] I. Gutierrez-Urrutia, D. Raabe, Dislocation and twin substructure evolution during strain hardening of an Fe–22Wt, contrast imaging, Acta Mater. 59 (16) (2011) 6449–6462.

[15] I. Karaman, H. Sehitoglu, K. Gall, Y.I. Chumlyakov, H.J. Maier, Deformation of single crystal hadfield steel by twinning and slip, Acta Mater. 48 (6) (2000) 1345–1359.

[16] H. Idrissi, K. Renard, D. Schryvers, P.J. Jacques, On the relationship between the twin internal structure and the work-hardening rate of twip steels, Scr. Mater. 63 (10) (2010) 961–964.

[17] W.S. Choi, S. Sandlöbes, N.V. Malyar, C. Kirchlechner, S. Korte-Kerzel, G. Dehm, P. Choi, D. Raabe, On the nature of twin boundary-associated strengthening in Fe-Mn-C steel, Scr. Mater. 156 (2018) 27–31.

[18] D. Barbier, N. Gey, S. Allain, N. Bozzolo, M. Humbert, Analysis of the tensile behavior of a twip steel based on the texture and microstructure evolutions, Mater. Sci. Eng. A 500 (2009) 196–206.

[19] O. Bouaziz, S. Allain, C. Scott, Effect of grain and twin boundaries on the hardening mechanisms of twinning-induced plasticity steels, Scr. Mater. 58 (6) (2008) 484–487.

[20] World Auto Steel, Advanced high-strength steels application guidelines, 2021, https://ahssinsights.org/. (Accessed: 10 October 2023).

[21] L. Remy, A. Pineau, Twinning and strain-induced F.C.C. → H.C.P. transformation in the Fe-Mn-Cr-C system, Mater. Sci. Eng. 28 (1) (1977) 99–107.

[22] Q. Yu, L. Qi, K. Chen, R.K. Mishra, J. Li, A.M. Minor, The nanostructured origin of deformation twinning, Nano Lett. 12 (2) (2012) 887–892.

[23] J.R. Greer, J.T.M. De Hosson, Plasticity in small-sized metallic systems: Intrinsic versus extrinsic size effect, Prog. Mater. Sci. 56 (6) (2011) 654–724.

[24] A. Dollmann, A. Kauffmann, M. Heilmaier, A. Srinivasan Tirunilai, L.S. Mantha, C. Kübel, S.J. Eder, J. Schneider, C. Greiner, Dislocation-mediated and twinning-induced plasticity of CoCrFeMnNi in varying tribological loading scenarios, J. Mater. Sci. 57 (36) (2022) 17448–17461.

[25] A. Dollmann, J.S. Rau, B. Bieber, L. Mantha, C. Kübel, A. Kauffmann, A.S. Tirunilai, M. Heilmaier, C. Greiner, Temporal sequence of deformation twinning in cocrni under tribological load, Scr. Mater. 229 (2023) 115378.

[26] N. Argibay, M. Chandross, S. Cheng, J.R. Michael, Linking microstructural evolution and macro-scale friction behavior in metals, J. Mater. Sci. 52 (5) (2017) 2780–2799.

[27] S.J. Eder, P.G. Grützmacher, M. Rodríguez Ripoll, J.F. Belak, Elucidating the onset of plasticity in sliding contacts using differential computational orientation tomography, Tribol. Lett. 69 (3) (2021) 79.

[28] S.J. Eder, P.G. Grützmacher, M. Rodríguez Ripoll, C. Gachot, D. Dini, Does speed kill or make friction better?—designing materials for high velocity sliding, Appl. Mater. Today 29 (2022) 101588.

[29] P. Grützmacher, C. Gachot, S.J. Eder, Visualization of microstructural mechanisms in nanocrystalline ferrite during grinding, Mater. Des. 195 (2020) 109053.

[30] S.J. Eder, U. Cihak-Bayr, C. Gachot, M. Rodríguez Ripoll, Interfacial microstructure evolution due to strain path changes in sliding contacts, ACS Appl. Mater. Interfaces 10 (28) (2018) 24288–24301.

[31] A. Stukowski, Visualization and analysis of atomistic simulation data with OVITO–The open visualization tool, Model. Simul. Mater. Sci. Eng. 18 (1) (2009) 015012.

[32] P. Hirel, Atomsk: A tool for manipulating and converting atomic data files, Comput. Phys. Comm. 197 (2015) 212–219.

[33] A.P. Thompson, H.M. Aktulga, R. Berger, D.S. Bolintineanu, W.M. Brown, P.S. Crozier, P.J. in 't Veld, A. Kohlmeyer, S.G. Moore, T.D. Nguyen, R. Shan, M.J. Stevens, J. Tranchida, C. Trott, S.J. Plimpton, LAMMPS - A flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, Comp. Phys. Comm. 271 (2022) 108171.

[34] A. Stukowski, How to identify the twin structure in FCC and BCC crystals using OVITO?, 2020, https://www.ovito.org/forum/topic/how-to-identify-the-twin-structure-in-fcc-and-bcc-crystals-using-ovito/. (Accessed: 05 July 2023).

[35] P.M. Larsen, S. Schmidt, J. Schiøtz, Robust structural identification via polyhedral template matching, Modelling Simul. Mater. Sci. Eng. 24 (5) (2016) 055007.

[36] S.J. Eder, P.G. Grützmacher, M. Rodríguez Ripoll, D. Dini, C. Gachot, Effect of temperature on the deformation behavior of copper nickel alloys under sliding, Materials 14 (1) (2021).

[37] S.V. Sukhomlinov, M.H. Müser, Determination of accurate, mean bond lengths from radial distribution functions, J. Chem. Phys. 146 (2) (2017) 024506.

[38] I.T. Jolliffe, Principal Component Analysis, Springer New York, New York, NY, 2002, pp. 1–7.

[39] RainerGewalt, PCA vs linear regression – therefore you should know the differences, 2021, https://starship-knowledge.com/pca-vs-linear-regression. (Accessed: 20 August 2023).

[40] J. Grieser, Hauptkomponentenanalyse, 2023, http://www.juergen-grieser.de/germanoldcentury/pdfs/EOF.pdf. (Accessed: 21 August 2023).

[41] G.G. Judge, Introduction to the Theory and Practice of Econometrics, second ed., Wiley, New York, 1988.

[42] A. Kauffmann, M. Heilmaier, Plasticity - lecture for mechanical engineering and materials science and engineering, 2014, https://www.iam.kit.edu/wk/downloads/PoMI/7-Other_deformation_mechanisms_web.pdf. (Accessed: 26 January 2023).

[43] H.M. Otte, Faulting in austenite, Acta Metall. 2 (2) (1954) 349–352.

[44] A. Rollett, T. Ozturk, 27-750 Texture, microstructure & anisotropy, 2016, http://pajarito.materials.cmu.edu/lectures/27750-Microstructure-Syllabus-Spg16.pdf. (Accessed: 15 September 2023).

[45] S.M. Foiles, M.I. Baskes, M.S. Daw, Embedded-atom-method functions for the fcc metals Cu, Ag, Au, Ni, Pd, Pt, and their alloys, Phys. Rev. B 33 (1986) 7983–7991.

[46] S. Nosé, A molecular dynamics method for simulations in the canonical ensemble, Mol. Phys. 52 (2) (1984) 255–268.

[47] S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, J. Chem. Phys. 81 (1) (1984) 511–519.

[48] W.G. Hoover, Canonical dynamics: Equilibrium phase-space distributions, Phys. Rev. A 31 (1985) 1695–1697.

[49] C. Kalcher A. Stukowski, Identify Fcc Planar Faults, OVITO GmbH, Germany, 2023, Available at https://docs.ovito.org/reference/pipelines/modifiers/identify_fcc_planar_faults.html. (Accessed: 15 September 2023).

[50] A. Stukowski, Crystal Analysis Tool User's Manual, Technische Universität Darmstadt, Darmstadt, Germany, 2017.