A continuum particle model for micro-scratch simulations of crystalline silicon

Citation for published version (APA):
A continuum particle model for micro-scratch simulations of crystalline silicon

S.O. Sperling, J.P.M. Hoefnagels, K. van den Broek, M.G.D. Geers

Abstract

In order to suffice the stringent surface requirements imposed on silicon wafers, thorough investigation of the fabrication methods is necessary. Typically, the processing conditions for slicing and grinding operations are determined via scratch experiments, where the information is mostly derived from the wear scars at the scratched surface. In order to cultivate an improved understanding of the subsurface deformations caused by silicon micro-scratching, numerical simulations are instrumental. For this purpose, a continuum particle-based micro-scale formulation is here proposed that addresses two challenges: (i) the need for a numerical methodology enabling mechanical continuum-discontinuum transitions and (ii) a constitutive model that accounts for the mechanics resulting from the underlying phase transformations. In this paper, both aspects are covered in order to simulate the basics of silicon micro-scratching.

1. Introduction

Over the past decades, Integrated Circuits (ICs) have become an essential part of modern society and industry and are embedded in, e.g., smartphones, personal computers and automation. ICs are typically build from silicon-based semiconductors, which are fabricated from silicon wafers containing particularly stringent surface requirements. Accordingly, fabrication methods that convert silicon ingots into wafers are widely investigated in order to optimize surface quality while minimizing processing time and costs. The manufacturing routine for silicon wafers consists of several steps: (i) slicing, (ii) grinding, (iii) etching, (iv) polishing and (v) cleaning (Liu et al., 2007). Steps (i) and (ii) involve subsurface phase transformations, damage and fracture that have to
be removed in the subsequent steps (iii)–(v) to achieve a defect free, smooth and clean wafer surface. Hence, investigating the surface and subsurface deformation and damage behavior during slicing and grinding is industrially relevant since it determines the etching/polishing depth, which should be minimal since it is a time consuming and expensive process (Brinksmeier et al., 2010). Not only manufacturing, but also wafer handling for transport and storage involves contact that affects the surface quality, introducing possible fabrication risks that are unfavorable. To investigate suitable process conditions related to ingot slicing (Ge et al., 2018; Gassilloud et al., 2005) and wafer grinding (Tang and Zhang, 2014; Gogotsi et al., 2001), micro-scratch experiments are performed. Yet, in-depth analysis of micro-scratching induced subsurface phase transformations, damage and fracture is severely hampered by experimental constraints. As a solution, computational modeling can assist in uncovering subsurface damage mechanisms.

Numerical scratch simulations on silicon are often performed using Molecular Dynamics (MD) simulations where, for instance, the nano-scale removal characteristics of Chemical Mechanical Polishing (CMP) have been investigated (Nguyen and Fang, 2020). MD is widely recognized as a powerful tool for the simulation of silicon machining and allows to investigate nano-scale mechanisms of various mechanical, physical and chemical processes that are combined with machining to optimize post-machining surface quality (Goel et al., 2015). However, expanding MD simulations to length-scales relevant for grinding and slicing operations is unfeasible with current computational resources. Alternatively, continuum models for silicon micro-scratching have been proposed. A silicon micro-scratch was modeled using the Finite Element Method (FEM) in Budnitzki and Kuna (2019), who reported an adequate agreement with experimental values using a dedicated constitutive model. Also, FEM models including fracture have been proposed. The eXtended Finite Element Method (XFEM) model of Wang et al. (2020) demonstrated the ability to simulate a single radial crack induced by silicon scratching. Wallburg et al. (2020) showed that their finite element model with cohesive zone elements can reproduce the same global subsurface damage zone as found in the scratch experiments.

Even though these continuum models look promising, their ability to include complex scratch-induced fracture beneath the surface and chip formation at the surface is limited. To remedy this, particle methods such as the Smoothed Particle Hydrodynamics (SPH) method have recently been applied to model chip formation (Mir et al., 2017; Klippel et al., 2021) and subsurface defects (Li et al., 2022) for silicon micro-scratching. The growing interest for scratch modeling with particle methods is due to their ability to straightforwardly capture large deformations, introduce continuum-discontinuum transitions and handle contact boundary conditions. A prominent example, besides SPH, is the Material Point Method (MPM) which has been used to investigate the scratching of steel (Mishra et al., 2019) and high purity copper (de Vaucorbeil et al., 2022). On the other hand, it has been shown that both SPH and MPM suffer from computational artifacts originating from their numerical description, which compromise the continuum consistency of the modeled response (Ganzenmüller, 2015a; Steffen et al., 2010). In this study, a recently developed particle method called the Continuum Bond Method (CBM) is employed. CBM was introduced by the authors in Sperling et al. (2022) where its FEM-like continuum consistency and ability to account for complex fracture events was demonstrated. The CBM framework was developed and demonstrated on simple two-dimensional examples. In the present paper, the method is extended to the full three-dimensional setting, enabling its application to scratch simulations.

In order to convert the CBM setup to a silicon micro-scratch model, a suitable continuum constitutive law is required. Cross-sectional Transmission Electron Microscopy (XTEM) micrographs of indented mono-crystalline silicon from, e.g., Zhang and Zarudi (2001) and Vandeperre et al. (2007) clearly reveals the presence of phase transformed material at the indented surface and subsurface. A strong relationship between the indentation response, loading conditions and phase transformed material was found by Jang et al. (2005), indicating that the phase transformations need to be considered in the model. During an indentation cycle, a sequence of phase transformations occurs. During the loading stage, the initial diamond cubic structure (Si-I) transforms to a metallic tetragonal (Si-II) arrangement, involving a densification of ∼ 20%. The phase transformation from Si-I to Si-II under compressive loading has been extensively investigated using diamond anvil cell tests (Hu et al., 1986), resistivity measurements during indentation (Haberl et al., 2006), MD simulations (Mizushima et al., 1994) and Density Functional Theory (DFT) calculations (Zarkevich et al., 2018). Upon unloading, either a mixture of crystalline phases (Si-III/Si-XII) or amorphous silicon (Si-a) is formed where the unloading rate has a significant influence on the resulting product phase. Slow unloading is associated with the formation of the Si-III/Si-XII phases, whereas fast unloading stimulates the formation of the Si-a phase (Jang et al., 2005). Also, it was observed by Gassilloud et al. (2005) that a fast scratch contains a higher degree of Si-a in the wear scar than a slow scratch. Since fast loading rates better represent the conditions for wafer slicing and grinding, only the formation of Si-a upon unloading is considered in this study and the formation of other silicon phases is omitted.

Modeling the mechanical effects due to phase transformations with continuum constitutive models requires clear information about the material properties of the individual phases and the transformation criteria. This is particularly difficult for silicon, where the plastic material behavior of the Si-II phase is difficult to characterize since it is a phase that is only stable under high pressure. Moreover, a precise criterion governing the Si-II to Si-a transformation is still unclear since it depends on several loading conditions. An elaborate continuum model on Si-I → Si-II transformations is discussed in Levitas (2018) and Babaei and Levitas (2018). A single transformation variant following an interpolation function was included in the kinematic description. The parameters were fitted according to molecular dynamics simulations done in Levitas et al. (2017) and the observed band structure in the intermediate microstructural state during phase transformation could be reproduced. The model was extended to larger scales and multiple variants in Babaei and Levitas (2020), achieving a closer resemblance to experimental observations found in Popov et al. (2015) on the formation of poly-crystals. The aim of these models is to uncover the mechanisms and physics related to Si-I → Si-II transformations in silicon crystals by incorporating various nano-scale mechanisms. However, the formation of Si-a and its mechanical implications at the microscopic scale are outside the scope of these models, which is essential for micro-scale scratch simulations of silicon.
For silicon scratching, the literature provides phenomenological constitutive models exploiting well developed tools originally designed for continuum plasticity to capture the essence of the underlying phase transformations. Typically, a Drucker–Prager type model is employed to model the densification corresponding to the Si-I to Si-II transformation (Wang et al., 2020; Wallburg et al., 2020; Mir et al., 2017). However, the volumetric expansion corresponding to the Si-II to Si-a transition had to be realized with a separate numerical algorithm. A thermodynamically consistent model was proposed by Budnitzki and Kuna (2016), which is arguably the most advanced continuum constitutive model for silicon that accounts for phase transformations under contact conditions to date. Nevertheless, a broad range of phenomenological fitting parameters is introduced, which complicates the exploitation of this model. So, to reduce the amount of phenomenological parameters, while keeping the essence and the thermodynamic consistency of the constitutive model, in this study their preceding model presented in Budnitzki and Kuna (2012) will be adopted. Note that, this constitutive model has not been applied to scratching problems yet and was originally formulated in an infinitesimal strain setting. Hence, the model is here extended to the finite strain inelastic regime using the mathematical principles introduced by Bilby et al. (1955), Kröner (1959), Lee (1969) and Simo (1992). The extended model is incorporated in the above mentioned three-dimensional CBM framework, allowing to investigate its response under scratch conditions.

The goal of this paper is to present an integrated CBM model for silicon micro-scratch simulations, emphasizing the implementation details, model parameters and simulation characteristics. To this end, the novelties of this work can be summarized as follows:

- A three-dimensional extension of the recently-introduced particle method CBM (Sperling et al., 2022) with implementation details for the particle simulator LAMMPS.
- A finite strain generalization of the infinitesimal framework of Budnitzki and Kuna (2012) according to Bilby et al. (1955), Kröner (1959), Lee (1969) and Simo (1992) complemented with the numerical details for an explicit dynamic scheme.
- Investigation of the simulated global and local material behavior in silicon under scratch conditions.

This contribution paves the way to intricate numerical-experimental comparisons and the extension to bond-fracture routines to model surface and subsurface fracture, which will be the subject of future studies.

The paper is organized as follows, in Section 2 the three-dimensional CBM formulation is discussed, followed by a brief overview of numerical methods for scratch simulations and closing with the LAMMPS implementation details including a specific indenter-surface contact description. Then, in Section 3, the derivation and implementation details of the constitutive model are given. Afterwards, a discussion on the different physical, fitting and numerical parameters is given in Section 4 where the parameters for the scratch simulation are identified. Finally, a scratch simulation is discussed in Section 5 and conclusions are given in Section 6.

Throughout this paper, the following notation conventions will be employed (using Einstein summation convention and Cartesian tensors):

- scalars $a$
- time derivative $\dot{a} = \frac{\partial a}{\partial t}$
- basis vectors $\vec{e}_i$
- Kronecker delta $\delta_{ij}$
- vectors $\vec{a} = a_i \vec{e}_i$
- vector norm $\| \vec{a} \| = \sqrt{\vec{a} \cdot \vec{a}}$
- second order tensors $\mathbf{A} = A_{ij} \vec{e}_i \vec{e}_j$
- dyadic product $\vec{a} \vec{b} = a_i b_j \vec{e}_i \vec{e}_j$
- inner product $\mathbf{A} : \vec{b} = A_{ij} b_j \vec{e}_i$
- double product $\mathbf{A} : \mathbf{B} = A_{ij} B_{ij}$
- second order tensor norm $\| \mathbf{A} \| = \sqrt{\mathbf{A} : \mathbf{A}}$
- fourth order tensors $\mathbf{A}^d = A_{ijkl} \vec{e}_i \vec{e}_j \vec{e}_k \vec{e}_l$
- fourth order identity tensor $\mathbf{I}^d = \delta_{ij} \delta_{kl} \vec{e}_i \vec{e}_j \vec{e}_k \vec{e}_l$
- deviatoric part of second order tensor $\mathbf{A}^d = \mathbf{A} - \frac{1}{3} \mathbf{I}^d$
- trace of second order tensor $\text{tr}(\mathbf{A}) = \mathbf{A} : \mathbf{I} = A_{ii}$

2. Numerical framework

The CBM framework in a two-dimensional setting has been derived in Sperling et al. (2022). Here, the extension to the three-dimensional framework is given, illustrating two main differences: (i) threefold vector sets are employed instead of twofold vector sets and (ii) a single bond between two particles is no longer part of either one or two twofold vector sets, but can be part of an arbitrary amount of threefold vector sets.

2.1. Three-dimensional CBM formulation

The continuum body is discretized into a set of inter-connected particles, where each particle carries a fraction of the body’s mass. The connectivity between the particles is determined via a triangulation procedure and consists of threefold vectors spanning
tetrahedrons. In this work, the open-source program Gmsh (Geuzaine and Remacle, 2009) is employed to facilitate the particle discretization, bond connectivity and tetrahedral vector sets. In Fig. 1(a) a schematic representation is given of particle \( i \) including its connected neighbors in the reference position \( \vec{X}_i \). The deformed position \( \vec{x}_i \) according to the deformation map \( \phi(\vec{X}, t) \) is shown as well. The deformation map \( \phi(\vec{X}, t) \) is a function of the reference position \( \vec{X} \) and time \( t \). For clarity, one tetrahedral vector set is colored in red. This single red tetrahedral set, denoted with \( \Lambda \), is extracted from the cloud in Fig. 1(b), whereby the transformation of the relative position vectors from the reference configuration \( \vec{x}_{ia} = \vec{X}_a - \vec{X}_i \) to the deformed configuration \( \vec{x}_{ia} = \vec{x}_a - \vec{x}_i \) according to the locally affine deformation gradient \( \mathbf{F}_\Lambda \) is illustrated. Note that, the locally affine deformation gradient \( \mathbf{F}_\Lambda \) can be obtained exactly according to

\[
\mathbf{F}_\Lambda \cdot \vec{X}_i = \vec{x}_i, \\
\mathbf{F}_\Lambda \cdot \vec{x}_{i\alpha} = \vec{x}_{i\alpha}, \\
\mathbf{F}_\Lambda \cdot \vec{x}_{i\beta} = \vec{x}_{i\beta}, \\
\mathbf{F}_\Lambda \cdot \vec{x}_{i\gamma} = \vec{x}_{i\gamma}.
\]

Multiplying with the corresponding relative position vector in the reference configuration and summing gives

\[
\mathbf{F}_\Lambda \cdot \vec{X}_{A,i} = \sum_{I=a,\beta,\gamma} \vec{x}_{iI} \vec{x}_{iI}.
\]

with

\[
\vec{X}_{A,i} = \sum_{I=a,\beta,\gamma} \vec{x}_{iI} \vec{x}_{iI}.
\]

The locally affine deformation gradient for tetrahedral set \( \Lambda \) is then given by

\[
\mathbf{F}_\Lambda = \sum_{I=a,\beta,\gamma} \vec{x}_{iI} \left( \vec{X}_{A,i}^{-1} \cdot \vec{x}_{iI} \right),
\]

since \( \vec{X}_{A,i} \) is a symmetric tensor by construction. The approximate deformation gradient for the particle can now be formulated according to a volume weighted average over a fraction of all tetrahedral sets in the particle’s cloud

\[
\bar{\mathbf{F}}_i = \frac{1}{\Omega_i} \sum_{\Lambda \in \mathcal{T}_i} \frac{\Omega_{\Lambda}}{4} \mathbf{F}_\Lambda,
\]

where \( \Omega_\Lambda \) is the volume spanned by the threefold set \( \Lambda \) and \( \Omega_i \) is the total volume corresponding to particle \( i \) such that

\[
\Omega_i = \sum_{\Lambda \in \mathcal{T}_i} \frac{\Omega_{\Lambda}}{4}.
\]

with \( \mathcal{T}_i \) the list of all tetrahedra attached to particle \( i \). Substituting Eq. (6) into Eq. (7) yields

\[
\bar{\mathbf{F}}_i = \frac{1}{\Omega_i} \sum_{\Lambda \in \mathcal{T}_i} \sum_{I=a,\beta,\gamma} \vec{x}_{iI} \vec{\psi}_{A,iI},
\]

with

\[
\vec{\psi}_{A,iI} = \frac{\Omega_{\Lambda}}{4} \left( \vec{X}_{A,i}^{-1} \cdot \vec{x}_{iI} \right).
\]

This formulation allows to convert the summation over the attached tetrahedra \( \mathcal{T}_i \) into a summation over the bonds \( P_i \) attached to particle \( i \)

\[
\bar{\mathbf{F}}_i = \frac{1}{\Omega_i} \sum_{j \in P_i} \vec{x}_{ij} \vec{\zeta}_{ij},
\]
where
\[
\tilde{e}_{ij} = \sum_{A \in \Gamma_{ij}} \tilde{\psi}_{A,ij},
\] (12)
is referred to as the bond vector and \( \Gamma_{ij} \) is the list of tetrahedra containing the bond between particles \( i \) and \( j \). From the definition of the approximate deformation gradient, i.e. Eq. (11), the particle interaction forces can be recovered from the discrete formulation of the internal stress power formulated as
\[
\delta W^{\text{int}} = \sum_{i=1}^{N} \mathbf{P}_i : \delta \mathbf{F}_i^T \Omega_i,
\] (13)
where \( \mathbf{P}_i \) denotes the first Piola–Kirchhoff stress acting on particle \( i \) and \( N \) denotes the amount of particles constituting the discrete body. Note that any constitutive relation \( \overline{\sigma} = C(F) \) may be employed. Following the derivation given in Sperling et al. (2022) for the discrete balance equation, which is largely equivalent for both two- and three-dimensional kinematic formulations, the interaction force is given as
\[
\mathbf{f}_{ij} = \mathbf{P}_j \cdot \mathbf{\tilde{e}}_{ij} - \mathbf{P}_i \cdot \mathbf{\tilde{e}}_{ij}.
\] (14)
Finally, for completeness, an schematic overview of the explicit time integration scheme solving the particle evolution is given in Algorithm 1 listed in the Appendix. Note that, for clarity, the particle subscript \( i \) is omitted for all quantities and only the time step subscript \( n \) is included in the algorithm. Within CBM, particles and material points are considered interchangeable, as commonly done in continuum particle methods such as SPH, Peridynamics and MPM. Hence, all continuum quantities reside in the material points, as a result of Eq. (13), which is represented as a particle of finite size with visualization software OVITO (Stukowski, 2010). The particle size depends on the average bond length, which is assumed to be much larger than that of an atom. Visualization of local continuum quantities is done by color coding of the particles. Unlike MD or DEM, in CBM the particle stacking does not affect the material behavior, which here follows the applied continuum constitutive formulations. For discrete damage and fracture events, the particle configuration has a topological influence. For an isotropic model at the considered length scale, a random particle distribution is preferred for CBM simulations since the continuum behavior is left unaffected and the random distribution reduces the numerical bias in fracture.

2.2. Numerical methods for scratch simulations

Modeling of scratching is numerically challenging. A broad range of numerical methods exists in the literature and many have been applied to scratching already. When considering silicon micro-scratching, from a purely mechanical perspective, several aspects need to be considered: (i) contact boundary conditions, (ii) large local deformations, (iii) complex fracture events, (iv) material detachment (v) constitutive accuracy and (vi) resolution of different length scales. Below, a concise comparison of CBM against: MD, FEM, explicit 

- **MD** So far, MD seems to be the most suitable method to simulate silicon scratching (Goel et al., 2015). However, since atoms are explicitly modeled, it is impossible to upscale the computational domain to several micrometers (vi). The usage of continuum constitutive laws in CBM alleviates this restriction, but at the cost of requirements (i)–(v) since a mixed continuum-discrete formulation is employed rather than a purely discrete method. Additionally, the widely used Tersoff potential cannot be adopted to model material behavior anymore.

- **FEM** For FEM the biggest challenges are requirements (iii) and (iv) since it employs a continuous interpolated domain. Using a phase-field description can help to account for (iii), but since fracture is modeled by an additional continuum variable (Miehe et al., 2010), strong discontinuities cannot be trivially introduced and thus (iv) remains a challenge. Less challenging but still cumbersome are requirements (i) and (ii) for continuum based methods, because the mesh is highly sensitive to distortions. The discrete nature of CBM allows to account for (iii) and (iv) via bond-breakage, eliminating the interaction between particles to introduce discontinuities. The averaging procedure allows for higher mesh distortions than FEM, improving (ii) and the point-wise description enables simple contact formulations (i).

- **SPH** Both SPH and CBM are continuum particle methods, where a particle represents a material point that incorporates all information. The main similarity between SPH and CBM is that they are more suitable for (i)–(iv) than FEM because of their point-wise description (Leroch et al., 2016). The main difference lies in the definition of the averaging procedure over neighboring particles. While SPH employs a manually defined interaction region, CBM has local tetrahedral connections only. Consequently, CBM is much more accurate on a local level (v) and is independent of numerical tuning parameters, as proven in Sperling et al. (2022). However, by including explicit connectivity between particles, CBM might be more sensitive to distortions (ii) than SPH.

- **Peridynamics** Peridynamics exist in a several variants but the method is rarely used for scratch simulations. The so-called Ordinary State-Based (OSB) Peridynamics formulation is used by Xu and Zhu (2022) to model the scratch behavior of cubic silicon carbide, but the kinematics of OSB Peridynamics is based on bond elongations instead of continuum kinematic quantities. On the other hand, the Non-Ordinary State-Based (NOSB) Peridynamics formulation does employ continuum kinematics allowing to include continuum constitutive laws. There exist significant similarities between the NOSB Peridynamics and the SPH formulations (Ganzenmüller et al., 2015), but different approaches are considered to alleviate NOSB Peridynamics from numerical deficiencies (v), e.g. Chen (2018). However, this has not been applied to scratching yet.
• MPM: The use of an independent computational background grid to find the gradient quantities of the particles gives MPM its strong Eulerian character which allows for high degrees of plasticity and deformation. This characteristic makes the method promising for scratch modeling (Mishra et al., 2019). However, stability issues have been observed which challenge the continuum consistency (v) (Steffen et al., 2010). Also, since it does not entail explicit connectivity between particles, modeling fracture (iii) is less straightforward than for CBM and the detaching debris (iv) cannot displace outside the specified computational background grid.

2.3. LAMMPS integration

While LAMMPS is commonly used as an MD simulator, over the past few years LAMMPS has been generalized to an all-round particle simulator with the ability to operate at multiple scales (Thompson et al., 2022). Nowadays, packages that contain subroutines for Discrete Element Method (DEM), Peridynamics and SPH have been incorporated into the code. Methods such as MPM are not included in the standard LAMMPS code but user implementations have been realized (Leroch et al., 2018). Continuum scale particle implementations for LAMMPS are nowadays possible due to its modular code structure and its open-source character. This subsection is divided into two parts: (i) an overview of the implementation details for the three-dimensional CBM formulation in LAMMPS and (ii) an appropriate indenter formulation suitable for CBM.

2.3.1. Implementation details

As previously mentioned, the Gmsh output facilitates the particle discretization and the tetrahedral threefold vector sets. From this, the bond topology within the discretized domain is established and the particle volumes, neighbor lists and bond vectors can be determined. Since standard LAMMPS input files have a specific layout and sequence, an alternative interface between Gmsh and LAMMPS to import the Gmsh output file structure is implemented as an input command. The particles are identified as atoms in the LAMMPS code and the other data structures such as the tetrahedral sets and bond connectivity are stored in a separate class, a so-called fix class, which is tailored to CBM and instantiated during initialization. Since the particle neighbor lists are set through the bonds given by the Gmsh output file, the standard LAMMPS neighbor list functionality can be bypassed and simply handled by the CBM fix. Note that, using a specifically designed fix to store non-atomic quantities related to the method was inspired by the Total Lagrangian SPH subroutine in LAMMPS, which can be found in the MACHDYN package (Ganzenmüller, 2015b).

The LAMMPS code is well-known for its thoroughly developed domain decomposition features that efficiently minimize computation times. In domain decomposition, each processor handles a part of the domain’s atoms, called the owned atoms. Additionally, each processor contains an overlapping fraction of the neighboring processor’s atoms, referred to as ghost atoms. If a bond crosses the processor’s subdomain boundary, it means that the bond consists of an owned atom and a ghost atom. In order to obtain the correct continuum quantities and interaction forces for atoms of crossing bonds, information needs to be transferred from owned atoms to ghost atoms between the processors. The constitutive evaluation sequence (lines 5 to 8 of Algorithm 1 in the Appendix) for the CBM method as implemented in LAMMPS is given below

(i) Communication of atom positions $\vec{x}_i$ between processors (between lines 3 and 5).
(ii) Calculation of the approximate deformation gradient $\vec{F}_i$ according to Eq. (11) of atom $i$ where the neighbor list $P_i$ and bond vectors $\vec{c}_{ij}$ are supplied by the CBM fix (lines 5 and 6).
(iii) Evaluation of the constitutive law $C(\vec{F}_i)$ to calculate the first Piola–Kirchhoff stress $\vec{P}_i$ acting on each atom (line 7).
(iv) After the stresses $\vec{P}_i$ for all atoms are calculated, a second processor communication step is performed to update the ghost atom stresses (between lines 7 and 8).
(v) Execution of a loop over all the bonds in the processor’s subdomain to calculate the interaction forces according to Eq. (14) (line 8).

While step (i) is a standard step in every LAMMPS routine, steps (ii)–(v) are specific for the CBM scheme and are implemented in a dedicated pair style. This specific pair style together with the CBM fix constitute the core of the CBM subroutine in LAMMPS.

2.3.2. Indenter formulation

LAMMPS has a built-in indenter formulation exploiting a penalty algorithm if overlap is detected over a predefined region (line 10 of Algorithm 1 in the Appendix). For a spherical indenter of radius $r_c$ where the center is located at $\vec{x}_c$, the closest distance is given by

$$\delta_i = r_c - \|\Delta\vec{x}_i\|,$$

where $\Delta\vec{x}_i = \vec{x}_i - \vec{x}_c$. Under the condition $\delta_i \geq 0$ a penalty force is applied to particle $i$ taking the form

$$f_i^{pen} = K\delta_i^2\vec{n}_i,$$

where $K$ is the contact stiffness in GPa and the radial direction $\vec{n}_i$ is determined by $\Delta\vec{x}_i/\|\Delta\vec{x}_i\|$. The penalty force represents the repulsive force of the indenter surface acting on the particle. Even though this method is straightforward and effective for DEM and MD, for CBM this contact routine proves to be a source of instabilities at the contact surface. In Fig. 2(a) an example of the contact surface of an elastically loaded specimen is shown, obtained with the build-in indenter formulation as described by Eq. (16). The hydrostatic pressure, defined as $p_h = -\frac{1}{3}\text{tr}(\sigma)$ where $\sigma$ is the Cauchy stress, is plotted on the surface to illustrate the spurious...
fluctuations that arise during indentation. To resolve these fluctuations, additional information on the particle’s equivalent surface orientation is required.

In continuum mechanics, the relation between the reference and deformed surface is described according to Nanson’s formula

\[
d\overrightarrow{A} \overrightarrow{n} = \det (\overrightarrow{F}) \overrightarrow{A}_0 \overrightarrow{n}_0, \tag{17}
\]

where \(d\overrightarrow{A}_0\) and \(d\overrightarrow{A}\) denote the infinitesimal surface areas in the reference and deformed configuration, respectively. Additionally, vectors \(\overrightarrow{n}_0\) and \(\overrightarrow{n}\) denote the corresponding outward surface normal in the reference and deformed configuration, respectively. Using the left polar decomposition \(\overrightarrow{F} = \overrightarrow{V} \cdot \overrightarrow{R}\) gives the following relation for the surface orientation vector in the deformed configuration

\[
\overrightarrow{n} = \det (\overrightarrow{F}) \frac{d\overrightarrow{A}_0}{d\overrightarrow{A}} \overrightarrow{V}^{-1} \cdot \overrightarrow{R} \cdot \overrightarrow{n}_0, \tag{18}
\]

where \(\overrightarrow{R}\) represents the orthogonal rotation tensor and \(\overrightarrow{V}\) the left stretch tensor. Implementing the left polar decomposition reveals that the deformed surface orientation given in Eq. (18) consists of a deformation and a rotation component. The exact orientation of the deformed surface normal would require the full evaluation of Eq. (18). Yet, solely including the surface rotation, i.e. neglecting \(\det(\overrightarrow{F})\frac{d\overrightarrow{A}_0}{d\overrightarrow{A}} \overrightarrow{V}^{-1}\), is a good first-order approximation for \(\overrightarrow{n}\) and is assumed to be adequate for the current application. Note that \(\|\overrightarrow{n}\| = 1\) is satisfied automatically.

In CBM, just like in other particle methods, the domain’s surface is combined with the first particle layer of the body. Hence, for each overlapping particle, an approximate deformation gradient \(\overrightarrow{F}_i\) is already pre-determined from the constitutive update. If the particle \(i\) overlaps with the indenter surface, the rotation tensor \(\overrightarrow{R}_i\) can be found through the left polar decomposition of the particle deformation gradient \(\overrightarrow{F}_i\). The approximate surface orientation of particle \(i\) is then defined as

\[
\overrightarrow{n}_i = \overrightarrow{R}_i \cdot \overrightarrow{n}_i, \tag{19}
\]

which replaces the radial \(\overrightarrow{n}_i\) used in the indenter force formulation in Eq. (16). With this updated formulation, the same indentation problem yields a more consistent result at the contact surface as demonstrated in Fig. 2(b). Accounting for the surface rotation through the rotation tensor only, i.e. neglecting the surface deformation, suffices to extract the normal contact surface direction for the current application. For the scratch and indentation examples presented in this paper, a sphero-conical indenter surface is employed, for which the same contact formulation is applied.

3. Silicon constitutive law

In this section the physical, mathematical and implementation aspects of the continuum constitutive phase transformation model for crystalline silicon are discussed. First, an overview of the relevant physics is given followed by a kinematical basis and constitutive transformation model for the mechanically induced phase transformations.

3.1. Physical background

The mechanical behavior of silicon under compressive loading conditions such as hydrostatic pressure, indentation and scratching have been extensively investigated in the literature, uncovering different silicon phases (He et al., 2016) and the accompanying inelastic behavior (Zhang et al., 2015). Even though these experimental studies give valuable insights into the material behavior, precise transformation criteria and material properties of the different phases are still not available. Hence, in order to establish a constitutive model, observations from literature are translated to existing continuum modeling principles, allowing to perform qualitative scratch analyses of crystalline silicon. XTEM micrographs of scratched silicon specimens (e.g. Tang and Zhang (2014))
uncover a significant degree of amorphous silicon present at the surface and subsurface, indicating that the mechanism of Si-a formation dominates the material behavior of silicon under microscale scratching conditions. Therefore, to construct a model that reflects the material behavior under microscale scratching conditions, the underlying mechanics belonging to the formation of Si-a needs to be included. As mentioned in the introduction, upon compressive loading the Si-I to Si-II transformation occurs whereas unloading initiates the transformation from Si-II to Si-a (Gaál-Nagy and Strauch, 2006; Zarkevich et al., 2018; Cheong and Zhang, 2000). Some MD simulations, executed by e.g. Levitas et al. (2017), have taken into account the reverse transformation from Si-II to Si-I. However, The backward transformation has not been established in the literature so far. Some continuum models investigating the phase Si-I → Si-II phase transformation have included the backward transformation in their model, but Babaei and Levitas (2020) clearly mention that such transformations are not confirmed by experiments. The phase transformation sequence Si-I → Si-II → Si-a for fast unloading under indentation conditions is widely accepted (Jiapeng et al., 2017) and therefore constitutes the working hypothesis in the presented constitutive model. The transformation sequence is accompanied by significant volumetric compression in the first step and subsequent expansion in the second. Nano-scratch experiments conducted by Wu et al. (2010) reported an alternative phase transformation path Si-I → Si-III/II-XII. On the other hand, the scratch study executed by Ge et al. (2018) found a significant recovery of the scratched surface, which supports the argument that the transformation path includes Si-II. Scratching is a complex loading case where the dominating mechanisms for silicon depend on the considered scale, tip geometry, normal force and scratch speed. Shear deformations play a more active role at the contact surface under scratching conditions. The effect of phase transformations under plastic shearing for silicon have been investigated in the literature and it was found that plasticity stimulates the formation of Si-III (Levitas and Zarechnyy, 2006; Blank and Estrin, 2014), reduce the transformation pressure (Levitas, 2019) and promote the formation of poly-crystals (Levitas et al., 2017). However, so far it is unclear how these mechanisms affect the material response under actual scratching conditions and relations coupling plasticity and phase transformations are unknown for silicon. Hence, for the current constitutive model, the authors restrict the model to experimental scratch observations and established knowledge of silicon material behavior under contact conditions. The purpose of the continuum constitutive model is to capture the underlying mechanics related to phase transformations and corresponding volume changes. In Section 3.2 the volumetric transformation strain is derived within a finite strain framework and in Section 3.3 the transformation kinematics related to the transformation strain will be detailed. Then, after elaborating the elastic behavior in Section 3.4, the transformation criteria are given in Section 3.5 followed by the details regarding its numerical implementation in Sections 3.6 and 3.7.

3.2. Finite strain kinematics

In a broad range of experimental indentation and scratch studies of crystalline silicon, dislocations and slip bands were observed accompanied by phase transformations (Jiapeng et al., 2017), so the presence of plasticity in the silicon crystal is undeniable. It was confirmed by Bradby et al. (2001) that plasticity in the form of dislocation movement in the Si-I phase has a minor effect on the overall indentation response. This suggests that the influence of plasticity in the Si-I phase is relatively small compared to the inelastic effects of phase transformation. Hence, conventional plastic mechanisms in the Si-I phase are neglected and the transformation induced deformation, i.e. the changing structure affected by phase transformation, is considered as the prime source of inelastic deformation to be captured in the current model. In terms of continuum kinematics, the multiplicative decomposition as proposed by Bilby et al. (1955), Kröner (1959), Lee (1969) is employed to split the total deformation gradient into

\[ \mathbf{F} = \mathbf{F}^e \cdot \mathbf{F}^i, \]  

(20)

where \( \mathbf{F}^e \) represents the elastic part and \( \mathbf{F}^i \) the inelastic part. The transformation deformation gradient \( \mathbf{F}^i \) represents the structural changes related to the phase transformations. Both volumetric and deviatoric parts depend on a single transformation criterion defined in stress space and are therefore inherently coupled by a kinematical evolution rule. The advantage of the adopted formulation is that existing mathematical tools and solution algorithms, originally developed for continuum plasticity, can be exploited. This makes the proposed model more accessible to a broader audience interested in modeling silicon micro-scratching. The disadvantage, though, is that the physics underlying the phase transformations is incorporated in a reduced and simplified manner.

The volumetric component of the transformation deformation gradient tensor is related to the phase transformation behavior through the volumetric transformation strain as

\[ \nu^i = -\ln \left( J^i \right), \]  

(21)

where \( J^i = \det (\mathbf{F}^i) \). To relate the phase evolution to the volumetric compression and expansion of the transformation strain, a kinematical law is required. As a point of departure, the inelastic kinematical rule from Simo (1992) (adapted from its original elasto-plasticity context) is adopted

\[ L^i_t (\mathbf{b}) = -2\dot{\gamma} \frac{\partial f (\mathbf{r}, \nu^i)}{\partial \mathbf{r}} \cdot \mathbf{b}^e, \quad \text{with} \quad \mathbf{r} = 2\frac{\partial \psi^e}{\partial \mathbf{b}^e} \cdot \mathbf{b}^e, \]  

(22)

combined with the conditions

\[ \dot{\gamma} \geq 0, \quad f (\mathbf{r}, \nu^i) \leq 0, \quad \dot{\gamma} f (\mathbf{r}, \nu^i) = 0. \]  

(23)
This rule is compatible with the multiplicative decomposition proposed in Eq. (20). In Eq. (22), \( \psi^e \) denotes the elastic strain energy density, \( \tau \) the Kirchhoff stress, \( b^e = F^e \cdot F^{eT} \) the elastic left Cauchy–Green deformation tensor, \( \gamma \) the inelastic multiplier and \( f(\tau, \psi^e) \) the transformation surface which will be further detailed in Section 3.5. \( L^\psi(b^e) \) is the Lie derivative of the elastic left Cauchy–Green deformation tensor expressed as

\[
L^\psi(b^e) = b^e - L \cdot b^e - b^e \cdot L^T,
\]

where \( L \) represents the velocity gradient tensor. Next, consider the material time derivative of the volumetric transformation strain given in Eq. (21)

\[
\dot{\psi}^I = -\frac{d}{dt}(\ln(J^I)).
\]  

(25)

Exploiting the multiplicative decomposition \( J = \det(F) = J^e J^I \) and the evolution of the volumetric deformation \( J = J tr(d) \) where \( d = \frac{1}{2} (L + L^T) \) is the rate of deformation tensor, the volumetric transformation kinematics can be written in a general way as

\[
\dot{\psi}^I = \frac{J^e}{J^I} - tr(d).
\]  

(26)

The evolution of the volumetric elastic deformation is described as

\[
J^e = J^e tr \left( \frac{1}{2} b^e \cdot b^{e-1} \right),
\]

which stems from the straightforward relation

\[
\frac{d}{dt}(\det(b^e)) = \frac{d}{dt}(J^{e^2}) = 2 J^e J^e = J^e 2 b^{e-1} : b^e.
\]  

(28)

Rearranging Eq. (24) for \( b^e \), substituting in Eq. (27), using the equality \( tr(d) = tr(L) \), and substitution into Eq. (26) gives

\[
\dot{\psi}^I = tr \left( \frac{1}{2} L^\psi(b^e) \cdot b^{e-1} \right).
\]  

(29)

Finally, substituting the expression of the Lie derivative of \( b^e \) for the inelastic flow rule given in Eq. (22) results in

\[
\dot{\psi}^I = -\dot{\gamma} tr \left( \frac{\partial f(\tau, \psi^I)}{\partial \tau} \right),
\]

(30)

providing a direct relation for the volumetric transformation evolution. From Eq. (30), it can be deduced that \( \dot{\gamma} \) denotes the transformation evolution since it is directly coupled to the volumetric transformation strain. From a physical perspective, \( \dot{\gamma} \) represents the rate at which the parent structure grows into the product structure.

3.3. Phase evolution relations

Typically, phase transformation models employ a phase mixture parameter that is implemented in the kinematic description, as proposed in Levitas (2018), Babaei and Levitas (2018, 2020). This mixture parameter is then considered as an internal variable, for which additional kinematic laws are established through thermodynamic consistency conditions. The current approach deviates from that procedure by introducing two history parameters that follow the inelastic volumetric transformation strain related to compression and subsequent recovery. The phase mixture residing in a material point is derived from these history parameters, providing an intuitive description of the material’s state. This way, the phase evolution is solely dependent on the inelastic flow governed by the transformation criterion specified later. The material constitutes of a composition of three possible phases represented by fractions \( a_I \) (Si-I), \( a_{II} \) (Si-II) and \( a_a \) (Si-a) where \( a_I + a_{II} + a_a = 1 \) and hence for the evolution

\[
a_I + a_{II} + a_a = 0,
\]

(31)

holds and \( a_I = 1 \) in the pristine state. The goal is construct a clear relation between the phase evolution and inelastic volumetric flow since the silicon phases are related to the volumetric transformation strain \( \dot{\psi}^I \). Recall from Eq. (25) that the flow direction is defined such that for compression \( \dot{\psi}^I > 0 \) and expansion \( \dot{\psi}^I < 0 \). Since volumetric compression is associated with the Si-I to Si-II transformation the following holds

\[
\dot{\psi}^I > 0 \quad \rightarrow \quad \dot{a}_I < 0, \quad \dot{a}_{II} > 0, \quad \dot{a}_a = 0,
\]

(32)

and the volumetric expansion is associated with the Si-II to Si-a transformation so

\[
\dot{\psi}^I < 0 \quad \rightarrow \quad \dot{a}_I = 0, \quad \dot{a}_{II} < 0, \quad \dot{a}_a > 0.
\]

(33)

Now, two history parameters are introduced that track the amount of volumetric transformation deformation that is related to the formation of Si-II and Si-a, presented by \( \dot{\psi}^{II} \) and \( \dot{\psi}^a \) respectively. Both \( \dot{\psi}^{II} = 0 \) and \( \dot{\psi}^a = 0 \) in the undeformed reference situation, since the pristine material consists solely of Si-I. The parameter \( \dot{\psi}^{II} \) is related to the volumetric transformation strain as

\[
\dot{\psi}^I > 0 \quad \text{and} \quad \dot{\psi}^{II} \leq \dot{\psi}^I \rightarrow \quad \dot{\psi}^{II} = \dot{\psi}^I, \quad \dot{\psi}^a = 0.
\]  

(34)
Fig. 3. Schematic representation of the transformation surfaces in stress space where the vertical axis denotes the equivalent von Mises stress $\tau = \sqrt{\frac{2}{3}||\mathbf{\tau}||^2}$ and the horizontal axis the pressure $p = -\frac{1}{3}\text{tr}(\mathbf{\tau})$. The gray part represents the admissible (elastic) region where $f(\tau, \nu) < 0$ and the red part denotes the transformation band where the width is determined by so-called transformation resistance $Q$. The situations for (a) Si-I $\rightarrow$ Si-II and (b) the Si-II $\rightarrow$ Si-a transformation are sketched.

stating that $\nu^{II}$ is a monotonically increasing function that captures the forward transformation from Si-I to Si-II. The parameter $\nu^a$ is related to the volumetric transformation strain as

$$\dot{\nu}^a < 0 \quad \text{and} \quad \nu^{II} \geq -\nu^a \rightarrow \nu^{II} = 0, \quad \nu^a = \nu^v,\quad (35)$$

enforcing that $\nu^a$ is a monotonically decreasing function that only decays if there is non-recovered $\nu^{II}$ present. From the above relations, the evolution of the phases can be defined as

$$\dot{\alpha}_I = -\frac{\nu^{II}}{\nu^{II, C}}, \quad \dot{\alpha}^{II} = \frac{\nu^v + \nu^a}{\nu^{II, C}}, \quad \dot{\alpha}_a = -\frac{\nu^a}{\nu^{II, C}},\quad (36)$$

where $\nu^{II, C}$ denotes the maximal volumetric transformation strain serving as upper bound such that the phase fractions remain between 0 and 1. On a final note, it is clear from the above that the phase fractions are determined via the volumetric part of $\mathbf{F}$ and the transformation bounds are set according to volumetric criteria. The upper bound of the shape change for the Si-I $\rightarrow$ Si-II transformation is coupled to the volumetric bound through the transformation criterion defined in stress space. The transformation criterion, discussed later, includes an angle in stress space that dictate the ratio between volumetric and deviatoric transformation strain. A realistic range for this ratio can be found from studies investigating the transformed structure such as the study performed by Zarkevich et al. (2018).

3.4. Elastic material behavior

The elastic material behavior results from the definition of the elastic part of the strain energy density. A common choice for the elastic free energy is the quadratic logarithmic model because of its suitable properties for moderately large elastic strains (Geers, 2004). The strain energy density is described as

$$\psi^e(\mathbf{b}^e) = \frac{\lambda}{8} \ln(\det(\mathbf{b}^e))^2 + \frac{\mu}{4} \ln(\mathbf{b}^e) : \ln(\mathbf{b}^e),$$

where $\lambda$ and $\mu$ are the Lamé parameters. The relation for the Kirchhoff stress tensor can be found by substituting Eq. (37) into the second equation of Eq. (22) resulting in

$$\mathbf{\tau} = \frac{\lambda}{2} \ln(\det(\mathbf{b}^e)) \mathbf{I} + \mu \ln(\mathbf{b}^e) = \mathbf{^4C}^e : \mathbf{\varepsilon},\quad (38)$$

where the fourth order stiffness tensor is defined as

$$\mathbf{^4C}^e = \lambda \mathbf{I} + 2\mu \mathbf{I},\quad (39)$$

and the elastic logarithmic strain tensor as

$$\mathbf{\varepsilon} = \frac{1}{2} \ln(\mathbf{b}^e).\quad (40)$$
3.5. Transformation surfaces

The description of the transformation surfaces are taken from Budnitzki and Kuna (2012) and are revisited here. Note that, the original constitutive model was constructed within an infinitesimal strain context. Here, the transformation surfaces and corresponding parameters are adopted with two adjustments: (i) the Cauchy stress \( \sigma \) is replaced by the Kirchhoff stress \( \tau \) and (ii) the volumetric transformation strain is now defined according to Eq. (21) instead of the trace of the inelastic infinitesimal strain tensor. The transformation surface dictating the Si-I \( \rightarrow \) Si-II transformation is described as

\[
f_{\text{I-I}}(\tau, \nu^t) = \tau_q - \frac{b}{a} \sqrt{(p_t - p^c(\nu^t) - a)^2 - a^2},
\]

where \( p_t = -\frac{1}{3} \text{tr} (\tau) \) denotes the pressure and \( \tau_q = \sqrt{\frac{2}{3}} \| \tau \| \) the equivalent von Mises stress in terms of the Kirchhoff stress tensor. A schematic representation in two-dimensional stress space is given in Fig. 3(a), where the gray marked region underneath the curve represents the purely elastic region and the red part denotes the transformation band width, which is governed by the critical pressure related to the volumetric transformation strain as

\[
p^c(\nu^t) = p^0 + Q \nu^t.
\]

Here, \( p^0 \) serves as the initial critical pressure for the onset of phase transformation and \( Q \) is identified as the transformation resistance, entailing an increasing critical pressure for increasing volumetric transformation strain. Furthermore, the parameter \( a \) denotes the tip rounding of the transformation surface which introduces smoothness at the hydrostatic axis. Lastly, \( b/a \) represents the pressure-deviatoric angle which governs the balance between the hydrostatic and deviatoric contributions to the phase transformation, incorporating the mechanism of reduced transformation pressure for non-hydrostatic stresses. The physics behind the reduced transformation pressure under non-hydrostatic conditions is investigated in e.g. Gaál-Nagy and Strauch (2006) or Zarkevich et al. (2018), for the bi-axial and uni-axial compression and in Levitas (2019) for shear deformations.

The transformation surface related to the Si-II \( \rightarrow \) Si-a transformation is defined as

\[
f_{\text{II-a}}(\tau, \nu^t) = \frac{b}{a} \sqrt{(p_t - p^c(\nu^t) - a)^2 - a^2 - \tau_q^3}.
\]

While the shape is equivalent to the transformation surface \( f_{\text{I-I}} \), the \( f_{\text{II-a}} \) surface is shifted along the hydrostatic axis by \( d \), which is referred to as the hydrostatic offset. Since the transformation surface \( f_{\text{II-a}} \) is related to unloading, its admissible region is now above the transformation surface as schematically illustrated in Fig. 3(b) by the gray elastic region. The shapes for \( f_{\text{I-I}} \) and \( f_{\text{II-a}} \) are set to be equivalent for simplicity and computational convenience (i.e. no surface intersections), since there is no clear reasoning to justify otherwise for the Si-II \( \rightarrow \) Si-a transformation. The lack of quantitative information on structural changes associated with the Si-II \( \rightarrow \) Si-a transformation, motivates the restriction of purely volumetric recovery of \( F^t \) to simulate the Si-a formation, even though it is probable that some deviatoric strain is associated with the formation of the Si-a phase. The model presented in Budnitzki and Kuna (2016) accounts for deviatoric strain in Si-a by reversing a portion of the deviatoric strain formed in the Si-II formation. The implication is, however, that transformation surfaces need to be introduced that depend on the shape of the inelastic deformation and that the restrictions that are imposed on the flow directions, recovering the inelastic shape, result in complex thermodynamic considerations. This results in an additional unknown parameter, where the deviatoric strain recovery for Si-II \( \rightarrow \) Si-a transformation can be included according to a fitting procedure. Omitting this deviatoric recovery seems justified as long as more precise nano-scale information is still missing. On a final note, the transformation surfaces operate under specific conditions for the volumetric transformation strain, as listed in Figs. 3(a) and 3(b). So when \( \nu^t = \nu^t_c \), states where \( f_{\text{I-I}} > 0 \) are elastic and for \( \nu^t = 0 \) states where \( f_{\text{II-a}} > 0 \) are elastic. These conditions are incorporated in the numerical scheme which is discussed next.

3.6. Numerical implementation

In Algorithm 2, listed in the Appendix, an overview of the constitutive routine (expansion of line 7 in Algorithm 1) is provided for a single particle (material point) during a time step \( \Delta t \) where the deformation gradient tensor is incremented from \( F_n \) to \( F_{n+1} \). First, the Kirchhoff trial stress \( ^* \tau \) is calculated from the strain increment. Then, depending on the current value of the Kirchhoff trial stress \( ^* \tau \) and volumetric transformation strain \( \nu^t \), it is determined whether the transformation surfaces \( f_{\text{I-I}} \) and \( f_{\text{II-a}} \) are exceeded or if the behavior is purely elastic. A general radial return algorithm (as designed for elasto-plasticity, e.g. Zeng et al. (1996)) can be used for the transformation update with the assigned transformation surface, which is rather standard and therefore omitted from this paper. The radial return algorithm calculates the increment of the inelastic multiplier \( \Delta \tau \) and the direction of the inelastic transformation strain defined as

\[
N_{\text{II}} = \left. \frac{\partial f}{\partial \tau} \right|_{\tau_{n+1}, \nu^t_{n+1}}.
\]

Before updating the Kirchhoff stress and the volumetric transformation strain, it is verified that the phase evolution bounds are properly enforced within a step. For the Si-a phase evolution bound, i.e. \( \nu^t = 0 \), this is quite straightforward and realized by simply shortening the inelastic multiplier increment \( \Delta \tau \) by the exceeded amount of volumetric transformation strain. The reason this straightforward correction is applicable is because the Si-II \( \rightarrow \) Si-a transformation only entails volumetric flow, so the flow direction is essentially independent of the location on the transformation surface in stress space. However, for the Si-II phase evolution
The incremental steps are small within the explicit time integration scheme adopted here. Clearly, the left cube violates the transformation cube. The pressure response for the lower and upper transformation bounds are included and it is clear that when the transformation bound, i.e. $\nu = \sqrt{\frac{\tau^e}{\nu^C}}$, a dedicated routine is constructed that calculates the correct kinematical direction within the step $t_n$ and $t_{n+1}$ when exactly the phase evolution bound is reached. More details about the Si-II evolution bound algorithm can be found in Section 3.7. Finally, the stress $\tau_{n+1}$ and volumetric transformation strain $\nu^e_{n+1}$ are updated and the value for $\bar{b}^e_{n+1}$ is extracted from $\nu^e_{n+1} = \frac{1}{2} \ln \left( \bar{b}^e_{n+1} \right)$.

### 3.7. Si-II evolution bound algorithm

In order to employ a classical radial return algorithm for the inelastic update, a routine accounting for the physical bound of the phase transformation evolution within the increment (which cuts off the inelastic deformation within a time step) is necessary because this is not incorporated in standard routines. When the volumetric transformation strain exceeds the threshold marking a phase transformation bound within an increment, it essentially means that the step contains a purely elastic tail so the determined direction and magnitude of the inelastic deformation may be incorrect. This routine calculates the correct location on the transformation surface within this time increment where $\nu^e = \nu^e_C$ such that the correct value for $N_{n+1}$ can be determined. One could argue that for small time increments, which are typically used in explicit dynamics analyses, a bound enforcement routine would be redundant because the overshoot beyond $\nu^e_C$ is expected to be small. However, in explicit dynamic analyses the switch between the inelastic and purely elastic behavior causes local oscillations that may entail large local strain rates, resulting in large strain steps beyond the volumetric transformation bounds.

The numerical routine is described in the Appendix in Algorithm 3 (expansion of line 23 in Algorithm 2). When the volumetric stress $\tau_{n+1}$ in the Si-II phase is exceeded at $n+1$ (i.e. transformation is complete), a fictional intermediate step $n+\delta$ is introduced at which the Kirchhoff pressure $\rho_{n+\delta}$ at $\nu^e_{n+\delta} = \nu^e_{n+1} = \nu^e_C$ is calculated. To find $\rho_{n+\delta}$ (step 10 in Algorithm 3), the intersection point is calculated between the volumetric response found by the incremental step from $n$ to $n+1$ in the radial return algorithm and the elastic pressure response of the Si-II phase defined as

$$p^e_{n+\delta} = \kappa (\nu - \nu^e_C),$$

where $\nu = -\ln(J)$ and $\kappa$ denotes the bulk modulus. The excess volumetric strain is attributed to the elastic strain of the Si-II phase. To find the correct inelastic deformation direction, the full Kirchhoff stress tensor $\tau_{n+\delta}$ at $n+\delta$, at which the completion bound is reached, needs to be calculated. Note that, $N_{n+\delta} = N_{n+1}$ because the response between $n+\delta$ and $n+1$ is purely elastic. To find the Kirchhoff stress $\tau_{n+\delta}$, the previously found pressure $\rho_{n+\delta}$ and the relation $\int_{I_{n+\delta}} = 0$ from Eq. (41) is employed to calculate the equivalent Kirchhoff stress $\epsilon_{n+\delta}$. With the value for $\tau_{n+\delta}$, a purely deviatoric back projection, similar to non-hardening von Mises plasticity, on the trial stress $^t\tau$ allows to find $^t\tau_{n+\delta}$. Finally, by adding $^t\tau_{n+\delta}$ and $-\rho_{n+\delta}1$ one arrives at $\tau_{n+\delta}$, which gives $N_{n+\delta}$ since the value $\nu^e_{n+\delta} = \nu^e_C$ is given.

In Fig. 4(a) the simulated effective Kirchhoff pressure response is given for the compression and subsequent decompression of a cube. The pressure response for the lower and upper transformation bounds are included and it is clear that when the transformation bounds are omitted in the numerical scheme of the constitutive routine, the transformation criteria are not properly satisfied. In Fig. 4(b), the discretized cube in compressed (top) and decompressed (bottom) state are given. The color schemes indicates the amount of volumetric transformation strain above and below the imposed bounds. Clearly, the left cube violates the transformation bounds and the right one does not. This illustrates the necessity of the imposed bound enforcement routine for CBM, even though the incremental steps are small within the explicit time integration scheme adopted here.
4. Scratch simulation parameters

In this section the simulation setup and parameter choices for the scratch model are detailed. The previously introduced model parameters are divided into three categories: (i) physical parameters, (ii) fitting parameters and (iii) numerical parameters. The physical interpretation and the adopted values of the parameters are discussed in the corresponding sections below. In order to clearly evaluate the accuracy of the indentation response, an experimental reference is taken from literature and the indentation setup is chosen such that it matches the corresponding experimental setup as close as possible.

4.1. Physical parameters

The physical parameters denote the set of parameters that have a clear physical meaning in the model and can be directly taken from experimental data. The parameters belonging in this section are the bulk modulus $\kappa$, shear modulus $G$, the initial critical pressure $p^0$ and volumetric transformation strain $\nu^{\text{tr,C}}$.

- **Bulk modulus $\kappa$:** The elastic parameters for the cubic crystal structure of mono-crystalline silicon measured by Hall (1967)

  \[ c_{11} = 165.64 \text{ GPa}, \quad c_{12} = 63.94 \text{ GPa}, \quad c_{44} = 79.51 \text{ GPa}. \]

  where the indices are in Voigt notation. From these parameters the bulk modulus is calculated according to

  \[ \kappa = \frac{c_{11} + 2c_{12}}{3} = 97.84 \text{ GPa}. \]  

- **Shear modulus $G$:** Since the crystal structure is cubic, but the material model is isotropic, a precise choice for the shear modulus is not evident. However, from the elastic constants the minimum and maximum values are known

  \[ G_{\text{min}} = \frac{c_{11} - c_{12}}{2} = 50.85 \text{ GPa}, \quad G_{\text{max}} = c_{44} = 79.51 \text{ GPa}. \]

  Both moduli will be evaluated against experimental indentation data in Section 4.2 to identify the most suitable value.

- **Initial critical pressure $p^0$:** The phase transformation pressure calculated by MD in Mizushima et al. (1994) equals $p^0 = 12.7 \text{ GPa}$, which falls within the experimentally observed range. Since the transformation surfaces are defined in the Kirchhoff stress space while the given critical pressure is defined in terms of the Cauchy stress, a translation step is required. The Kirchhoff transition pressure is found by solving the relation

  \[ p^0 = J^e p^0_e = \kappa \ln(J^e). \]  

  Filling in the bulk modulus found in Eq. (47) and $p^0_e$ gives $J^e = \det(F^e) = 0.891$, which in turn gives $p^0 = 11.31 \text{ GPa}$. Note that the value found for $J^e$, which corresponds to the volume change at the start of phase transformation for the Si-I crystal, only deviates $\sim 2\%$ from the value measured by Hu et al. (1986).

- **Volumetric transformation strain $\nu^{\text{tr,C}}$:** The volume change found for the transformed crystals from Hu et al. (1986) is $V_{II}/V_I = 0.796$. From the multiplicative split that is employed in this model, the volume corresponding to the Si-I and Si-II crystals can be calculated according to $V_{II} = J^{e,II} J^{\text{tr,II}} V_0$ and $V_I = J^{e,I} J^{\text{tr,I}} V_0$, respectively. Here, $J^{e,I}$, $J^{e,II}$ denote the elastic and $J^{\text{tr,I}}$, $J^{\text{tr,II}}$ the inelastic volume change corresponding to Si-I and Si-II phases. The reference volume is given by $V_0$. These definitions provide the straightforward relation $V_{II}/V_I = (J^{e,II} J^{\text{tr,II}}) / (J^{e,I} J^{\text{tr,I}})$. Since Si-I is the parent phase, $J^{\text{tr,I}} = 1$ because no inelastic volumetric changes can occur. By assuming the elastic volumetric change to be equal in the Si-I and Si-II crystal after complete transformation, i.e. $J^{e,I} = J^{e,II}$, one obtains $V_{II}/V_I = J^{\text{tr,II}}$. From this, the upper limit for the volumetric transformation strain can be defined as

  \[ \nu^{\text{tr,C}} = -\ln(J^{\text{tr,II}}) = 0.228. \]

  As far as the authors know, there exists no clear experimental values for $V_{II}/V_I$, so the lower limit for volumetric recovery is assigned as $J^{\text{tr,I}} = 1$ which simply gives $\nu^{\text{tr,I}} = -\ln(J^{\text{tr,I}}) = 0$.

4.2. Fitting parameters

The fitting parameters denote the set of parameters that have no precise physical meaning and typically result from phenomenological model assumptions. The parameters concerned are the pressure-deviatoric angle $b/a$ and the hydrostatic offset $d$. For the fitting parameter evaluation, the indentation response serving as reference is taken from Budnitzki and Kuna (2017) (also in Budnitzki (2014)). To reflect the Berkovich indenter setup used for the experiments, an equivalent radius and angle for the spherico-conical indenter in the model needs to be obtained. For this, the measurements for the tip area function of the Berkovich indenter which was used for the experiments was extracted from Budnitzki (2014) and compared to Eq. (8) in Čech et al. (2016). Fig. 5 presents the comparison between the analytical expression and the experimental data for the tip area function. For the spherico-conical indenter, the tip radius is estimated to be 330 nm and the center-to-face angle 70.3°.
Fig. 5. Comparison between the tip area function for the Berkovich indenter obtained from Budnitzki (2014) and the spherico-conical approximation according to Eq. (8) from Čech et al. (2016).

Fig. 6. (a) Effective indentation response for two shear moduli and corresponding values for $b/a$ compared to the averaged experimental response from Budnitzki and Kuna (2017). (b) Unloaded subsurface cross-section for different shear moduli values indicating the local unrecoverable deviatoric transformation strain (top) and recovered volumetric strain (bottom).

• **Pressure-deviatoric angle $b/a$:** In Fig. 6(a) the indentation response of the minimal and maximum shear moduli are given compared to an average of experimental indentation curves for different crystallographic orientations of silicon, taken from Budnitzki and Kuna (2017). The values for $b/a$ are adjusted for $G_{\min}$ and $G_{\max}$ such that it coincides with the loading portion of the experimental reference. It is clear that the unloading recovery is more accurately captured for $G_{\min}$, and significantly deviates from $G_{\max}$. To better understand this distinction, an unloaded cross-section of the indented subsurface, shown in Fig. 6(b), is investigated. The domain is divided into two parts where the left side is set to the minimal shear modulus $G_{\min}$ and $b/a = 0.98$ and the right side contains the settings $G_{\max}$ and $b/a = 0.66$. The top image shows the equivalent transformation strain, defined as $\varepsilon^t_{\text{eq}} = \sqrt{\frac{2}{3}} || \varepsilon^t ||$ with $\varepsilon^t = \frac{1}{2} \ln (b^t)$, indicating the permanent deviatoric transformation strain. The bottom image plots the fraction of recovered volumetric transformation strain from the $f_{11\rightarrow a}$ transformation surface, relative to the volumetric bound $\psi_{\text{C}}$. It can be observed that for the settings for $G_{\max}$, a higher amount of deviatoric transformation strain is present in the subsurface. This traps the transformed material and limits the ability for volumetric recovery corresponding to the Si-II $\rightarrow$ Si-a transformation. This mechanism of volumetric recovery, which represents the physics, is more active in the settings for $G_{\min}$, which explains the closer resemblance to the experimental response. Therefore, it is concluded that the settings $G = G_{\min} = 50.85$ GPa and $b/a = 0.98$ is most suitable in this case.
Fig. 7. Effective indentation response for different values of $d$ compared to the range of responses spanning different crystallographic orientations given by Budnitzki and Kuna (2017).

- **Hydrostatic offset $d$:** In Fig. 7, the indentation response for different values of $d$ are given. As a reference, the range of responses for different crystallographic orientations taken from Budnitzki and Kuna (2017) are included to indicate the accuracy of the current set of parameters. From Fig. 7 it can be observed that $d = 6$ GPa is an adequate value.

On a final note, it can be concluded from the results of Budnitzki and Kuna (2017) presented in Fig. 7 that the crystallographic orientation has a small influence on the indentation response. This observation justifies the assumption of isotropic material behavior for nearly all silicon models including (Budnitzki and Kuna, 2012, 2016) and Wang et al. (2020). However, this assumption is challenged by the results found in Gerbig et al. (2009), where strong differences were found between the indentation responses of (001) (110) and (111) surface orientations for spherical indentations. From a crystallographic perspective, Zarkevich et al. (2018) showed that the Si-I $\rightarrow$ Si-II transformation favors a specific deformation related to the unit cell, which would indicate that the crystallographic orientation does influence the overall indentation response. However, the surface and subsurface stresses are very different for sharp and blunt indenters. The high local contact pressure caused by sharp indenters result in immediate phase transformation on the silicon surface at the first point of contact, therefore undetectable on the indentation curve and active throughout the complete loading cycle.

4.3. **Numerical parameters**

The numerical parameters denote the set of parameters that have no physical substantiation and are introduced purely for algorithmic purposes. The parameters belonging to this section are the transformation surface tip radius $a$ and the transformation resistance $Q$.

- **Transformation surface tip radius $a$:** The parameter $a$ indicates the tip radius of the conical transformation surface in stress space as illustrated in Fig. 3. A perfect cone, i.e. $a = 0$ GPa, would introduce a non-smooth transformation surface on the intersection with the hydrostatic axis. Consequently, particles subjected to nearly hydrostatic deformations could experience abrupt flow changes causing the simulation to become unstable. This is especially important for explicit dynamic routines. The value is set to $a = 1$ GPa, resulting in the shape of the conical transformation surface shown in Fig. 3. This value is determined in a qualitative sense, introducing sufficient smoothness at the hydrostatic axis without drastically changing the conical shape of the transformation surface.

- **Transformation resistance $Q$:** The parameter $Q$ controls the magnitude of hydrostatic shift of the transformation surfaces related to the volumetric transformation strain. From a physical perspective, a displacive phase transformation occurs instantaneously, although this may occur at slightly different pressures depending on local microstructural variations. Still, a sudden change in volume is expected when the transformation pressure is reached. Such an event will introduce a discontinuity in the model response, which may lead to numerical instabilities. Hence, the transformation resistance is set to $Q = 0.01$ GPa to ensure a continuous transition between the phases without affecting the response significantly.

5. **Scratch example**

In this section, the scratch setup and silicon constitutive model are evaluated for a scratch test. First, an overview of all the parameters and the simulated geometry is given. Afterwards, the simulated scratch results are discussed with the aid of experimental observations from literature.
Fig. 8. Sketch of the scratch setup including the domain size with the coarsely discretized ‘bulk’ domain and the refined ‘contact’ domain for which the purple surface indicates the pre-defined surface within which all contact with the indenter takes place. The arrows $F_N$ and $F_T$ represent the reaction forces experienced by the indenter. The scratched distance is assigned with $u$. The geometrical parameters for the ‘bulk’ and ‘contact’ domain, respectively, are $W \times L \times H = 10 \times 13 \times 12 \, \mu m$ and $w \times l \times h = 2 \times 5 \times 1 \, \mu m$.

Fig. 9. Particle discretization of the scratch domain including the coarsely discretized ‘bulk’ domain, the refined ‘contact’ domain and a surface plot of the modeled indenter surface. For the indenter surface the same geometrical parameters are employed as for the parameter identification: Center-to-face angle of $70.3^\circ$ and a tip radius of 330 nm.

Table 1
Model Parameters for the scratch simulation.

<table>
<thead>
<tr>
<th>Description</th>
<th>Symbol</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bulk modulus</td>
<td>$\kappa$</td>
<td>97.84</td>
<td>GPa</td>
</tr>
<tr>
<td>Shear modulus</td>
<td>$G$</td>
<td>50.85</td>
<td>GPa</td>
</tr>
<tr>
<td>Initial critical pressure</td>
<td>$p^0$</td>
<td>11.31</td>
<td>GPa</td>
</tr>
<tr>
<td>Volumetric transformation strain</td>
<td>$\nu^{\mu\ell}$</td>
<td>0.228</td>
<td></td>
</tr>
<tr>
<td>Volumetric recovery strain</td>
<td>$\nu^{\mu\ell}$</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Transformation surface tip radius</td>
<td>$a$</td>
<td>1</td>
<td>GPa</td>
</tr>
<tr>
<td>Transformation resistance</td>
<td>$Q$</td>
<td>0.01</td>
<td>GPa</td>
</tr>
<tr>
<td>Pressure-deviatoric angle</td>
<td>$b/a$</td>
<td>0.98</td>
<td>–</td>
</tr>
<tr>
<td>Hydrostatic offset</td>
<td>$d$</td>
<td>6</td>
<td>GPa</td>
</tr>
<tr>
<td>Indenter stiffness</td>
<td>$K$</td>
<td>1140</td>
<td>GPa</td>
</tr>
<tr>
<td>Bulk domain particle distance</td>
<td>$\ell_C$</td>
<td>0.7</td>
<td>$\mu m$</td>
</tr>
<tr>
<td>Contact domain particle distance</td>
<td>$\ell_s$</td>
<td>0.04</td>
<td>$\mu m$</td>
</tr>
<tr>
<td>Numerical damping coefficient</td>
<td>$\beta$</td>
<td>0.01</td>
<td>[-]</td>
</tr>
</tbody>
</table>
Fig. 10. Reaction forces experienced by the indenter during a load-controlled scratch at a depth of 250 nm over a scratching distance of 3 μm. (a) Normal force in z direction and (b) Tangential force in y direction. (c) The ratio between the normal and tangential force for a scratch depth of 250 nm and 125 nm.

5.1. Simulation setup

The setup consists of a perfectly flat substrate and an indenter surface which is initially moved in the -z direction for indentation and subsequently in the y direction for scratching. The purpose here is to simulate the scratch in a sufficiently large domain, where boundary effects do not contribute. The simulation domain is split into a ‘bulk’ domain, which has a coarse discretization, and a ‘contact’ domain which has a fine discretization for computational efficiency. For clarity, a schematic illustration is provided in Fig. 8, where the purple region represents the surface on which contact with the indenter will occur. The scratched distance from left to right is denoted by \( u \), which indicates the current y position for the indenter relative to the initial indentation position (the covered scratch track is marked by black arrow). The red arrows at the top of the indenter symbolize the normal \( F_N \) and tangential \( F_T \) reaction forces from the displacement-controlled scratch acting on the indenter. An image, generated in OVITO (Stukowski, 2010), of the complete simulation domain and indenter surface is shown in Fig. 9. The corresponding domain size is described in the caption of Fig. 8, the indenter geometry is maintained from the parameter identification and is repeated in the caption of Fig. 9 and a summary of the material and simulation parameters is provided in Table 1. Note that, in the following images, the indenter surface and ‘bulk’ domain are omitted and only the refined ‘contact’ domain is plotted for clarity. The scratch simulations are performed under quasi-static conditions, neglecting any dynamic effects. This is ensured by enforcing the indenter velocity to be a factor 1000 lower than the speed of sound of the pristine Si-I material. Recall that, friction and frictional heating are omitted from the current analysis since the main objective is to evaluate capability of the CBM framework and extended constitutive model to capture phase transitions under scratching conditions. To expand the scratch model to construct a physically more representative situation, the incorporation of friction and frictional heating is recommended.

5.2. Results

A displacement controlled scratch is performed at a depth of 250 nm and a distance of 3 μm. The evolution of the effective normal and tangential forces experienced by the indenter over the scratch distance is shown in Fig. 10. When scratching is initiated, a significant drop in the normal force occurs, as can be observed in Fig. 10(a). The head of the scratch contains the pre-deformed material from the initial indentation stage of the scratch cycle. Scratching over indentation-deformed material required less effort from the indenter than scratching over undeformed material, hence a lower normal reaction force is observed at this stage while the scratch depth is maintained. In Fig. 10(b) the tangential force over the scratch depth is plotted. It can be concluded soon after indenter is translated in the y direction, i.e. initiation of the scratching stage, the vertical reaction force \( F_T \) reaches a steady state response. The observation of quick steady state for \( F_T \) was confirmed by the displacement controlled scratch experiment of Budnitzki and Kuna (2019).

In Fig. 11 a cross-sectional snapshot of the contact region along the scratch (yz-plane) at \( u = 2 \) μm is given where the indenter location is assigned by the green arrow. The top image shows the current phases predicted by the volumetric transformation deformation and its history, the bottom indicates the deviatoric (permanent) transformation induced deformation in the product...
Fig. 11. Cross-sections along the $yz$-plane of the ‘contact’ domain along the scratch direction where the indenter, omitted for clarity but indicated by the green arrows, is located at $u = 2 \, \mu m$. The top image indicates the phase distribution where blue represents Si-I ($\alpha_I > 0.9$), white Si-II ($\alpha_{II} > 0.1$) and red Si-a ($\alpha_a > 0.1$) giving an upper bound for the region affected by phase transformations. The bottom image shows the non-recoverable deviatoric deformation at the surface and subsurface. The yellow dashed lines in both top and bottom image designate the location behind which the deformation is no longer affected by the indenter. A cross-section in the $xz$-plane at the yellow dashed line is provided on the right.

phases. A steady-state scratching regime has been achieved when the indenter is located at $u = 2 \, \mu m$, the separation between permanent deformation (deformation that is no longer affected by the indenter) and the material still strained by the indenter is given by the yellow dashed line. On the right side a cross-section in the $xz$-plane is given at the yellow dashed line to provide a clear view on the phase and strain distribution found by the simulation. From the phase distribution plot (Fig. 11 top) it can be concluded that model is able to qualitatively capture the mechanism of subsurface phase transformations, which is widely reported in experimental indention and scratch studies. Also, it can be observed that behind the permanent deformation line (yellow dashed line), residual Si-II remains at the bottom of phase transformed region. From Fig. 11 (bottom), it can be observed that along the scratch, close to the surface, the amount of deviatoric transformation strain is higher and it fades out deeper into the subsurface. The amount of deviatoric transformation deformation close to the surface, restricts the volumetric recovery mechanism for Si-a formation, basically trapping the Si-II phase. This mechanism of Si-II entrapment under scratching conditions was postulated by Mylvaganam and Zhang (2011) who performed MD simulations investigating the scratch induced deformation of silicon. Experimentally confirming this phenomenon via direct XTEM observations is challenging since the volumetric strain gets easily relaxed in the sliced specimens. Note also, that fracture is not included in the present model and XTEM micrographs of scratch experiments by e.g. Tang and Zhang (2014) showed that median crack formation typically occurs in the subsurface below the phase transformed material. The omission of fracture in the current model might suppress the mechanism of subsurface release at the bottom of the phase transformed region via crack formation, trapping the Si-II and restricting volumetric expansion.

A second observation from the phase distribution plot (Fig. 11 top), is the remaining Si-II at the top, which seems to contradict experimental observations done with XTEM and Raman spectroscopy from e.g. Gassilloud et al. (2005), Tang and Zhang (2014), Zhang et al. (2015). A post-scratch snapshot of the phase distribution and equivalent von Mises stress in terms of Cauchy stress at the surface is given in Fig. 12. From the phase distribution plot it can be observed that on both sides of the center of the scratch, wide but thin laths of Si-II remain. Further away from the center to the side of the scratch, at the interface between the Si-II and Si-a at the top view, very thin bands of Si-II are formed and from the stress plot it can be concluded that the volumetric mismatches have a significant influence on the surface stresses. It is important to realize that when fracture is incorporated, the
surface stresses will have a large influence on the crack formation. The mechanical behavior of silicon, along with the occurring phase transformations, is significantly more complex than that of most materials and it is thus not surprising that only the most prominent and well investigated mechanisms have been included in the present constitutive model. However, one cannot deny the existence of other transformation mechanisms creating possible distinctions between simulations and experiments. For instance, He et al. (2016) discusses a shear-driven pathway to transform Si-I to Si-a, circumventing the transformation through the metastable Si-II phase as incorporated in the current model. Since high shear deformations occur at the scratched surface, this mechanism can play a significant role in the formation of Si-a at the surface in scratch experiments and can explain the discrepancy with the modeled results close to the surface.

The established model can be employed to perform more extensive parameter investigations on silicon micro-scratching. For instance, the effect on the indenter forces for different scratch depths as demonstrated in Fig. 10(c). Here, the ratio between the normal and tangential force is plotted over the scratch length for a scratch depths of 250 nm and 125 nm. It can be concluded that the ploughing resistance remain similar for both depths in the considered case. Experimental comparisons can be performed using the scratch track topology and indenter forces for different indenter geometries for validation. The corresponding subsurface phase evolution and resulting stresses captured by the model can contribute to the understanding of silicon scratching behavior. However, such studies and the extension to fracture are outside the scope of this paper and subject of future research.

6. Conclusions

In this paper, a novel discrete particle based silicon scratching model is presented. First, the three-dimensional extension of the Continuum Bond Method (CBM) introduced in two dimensions by Sperling et al. (2022). The implementation details in open-source particle simulator LAMMPS was discussed. Subsequently, the continuum transformation model based on volumetric deformation
S.O. Sperling et al.

20

S.O. Sperling et al. to designate silicon phase evolution described by Budnitzki and Kuna (2012) was extended to finite strain formulation. For this extension, the mathematical principles originally developed for large strain plasticity from Bilby et al. (1955), Kröner (1959), Lee (1969) and Simo (1992) were exploited. Then, the numerical implementation and phase completion bound routine suitable for explicit dynamic analysis was discussed in detail. The conclusions regarding the established model can be summarized as follows:

1. For the set of parameters and considered indentation geometry, the simulated indentation curve falls within the experimentally found crystallographic range from Budnitzki and Kuna (2017). It was shown that the loading portion of the indentation curve can be obtained for a range of shear moduli, but the unloading recovery is represented more accurately with the smallest shear modulus, limiting the deviatoric transformation strain and thereby enabling the volumetric expansion corresponding to the Si-II → Si-a transformation.

2. CBM is a suitable method for scratch simulations. An alternative, but simple, penalty contact algorithm was introduced to enable a stable mechanical interaction behavior with the indenter. The specimen was subjected to large local deformations. The resulting stress and strain fields obtained at the surface and subsurface are smooth, indicating stable progression throughout the simulation.

3. The scratch simulation reveals characteristics in line with experimental observations for silicon scratching. Yet, the interpretation of the surface stresses with the current constitutive model must be done with care.

The detailed model description, parameter identification and numerical scratch evaluation presented in this paper forms a thorough foundation for further experimental comparisons and extension to scratch induced fracture simulations. Both will be the subject of future research.

CRediT authorship contribution statement

S.O. Sperling: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration. J.P.M. Hoefnagels: Conceptualization, Methodology, Writing – review & editing, Supervision. K. van den Broek: Supervision, Funding acquisition. M.G.D. Geers: Conceptualization, Methodology, Resources, Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Marc Geers reports financial support was provided by VDL ETG Technology & Development B.V.

Data availability

Data will be made available on request.

Acknowledgments

This research is co-financed by Holland High Tech, The Netherlands through a public–private partnership in research and development within the Dutch top sector of High-Tech Systems and Materials.

Appendix

Algorithm 1: Explicit time integration scheme

```
for n = 1 to M steps do
    \vec{v}_{n+1/2} = \vec{v}_n + \frac{1}{2} \vec{f}_n \Delta t
    \vec{x}_{n+1} = \vec{x}_n + \vec{v}_{n+1/2} \Delta t
    \vec{F}_{n+1} = \sum \vec{f}_i
    \vec{F}_{n+1} = \sum \vec{f}_i
    \vec{P}_{n+1} = c \left( \vec{F}_{n+1}, \vec{F}_{n+1}, \ldots \right) ← Silicon constitutive law
    \vec{f}_{n+1} = \sum \vec{f}_i
    \vec{f}_{n+1} = \vec{f}_{n+1} + \frac{1}{2} \vec{f}_{n+1}
    \vec{v}_{n+1} = \vec{v}_{n+1/2} + \frac{1}{2} \vec{f}_{n+1}
end
```
Algorithm 2: Silicon constitutive law

1. \textbf{Input:} \( F_{s+1}, F_{e+1}, \mathbf{b}, \mathbf{v}^u, \tau, \alpha \)
2. \( F = F_{s+1}, F_{e+1} \)
3. \( \tau = \frac{1}{2} \ln(\mathbf{b}) \) \quad \text{Eigenvalue decomposition} \quad \text{of} \quad \mathbf{b} \)
4. \( \mathbf{v}^u = \mathbf{b} \cdot F^\top \)
5. \( \tau = \mathbf{C} : \tau \)
6. \( f_{s+1} = f_{s+1}(\tau, \mathbf{v}^u) \)
7. \( f_{e+1} = f_{e+1}(\tau, \mathbf{v}^u) \)
8. \( \text{if} \quad \mathbf{v}^u < \nu^C \quad \text{and} \quad f_{s+1} > 0 \quad \text{then} \)
9. \( \quad \text{[\( N_{s+1}, \Delta \)]} \Rightarrow \text{Radial return algorithm} \)
10. \( \text{else if} \quad \mathbf{v}^u > 0 \quad \text{and} \quad f_{s+1} > 0 \quad \text{then} \)
11. \( \quad \text{[\( N_{s+1}, \Delta \)]} \Rightarrow \text{Radial return algorithm} \)
12. \( \text{else} \)
13. \( \quad N_{s+1} = 0 \)
14. \( \quad \Delta = 0 \)
15. \( \Delta V = -\Delta^{tr} (N_{s+1}) \)
16. \( \quad \text{if} \quad \mathbf{v}^u + \Delta V > \nu^C \quad \text{then} \)
17. \( \quad \tau_{\text{new}} = \tau - \Delta^{tr} \mathbf{C} : N_{s+1} \)
18. \( \quad I_{s+1} = I_{s+1}, I_{e+1} \)
19. \( \quad \text{if} \quad \mathbf{v}^u + \Delta < 0 \quad \text{then} \)
20. \( \quad \Delta V = -\Delta^{tr} (N_{s+1}) \)
21. \( \quad \mathbf{v}^u_{s+1} = \mathbf{v}^u_{s} - \Delta V (N_{s+1}) \)
22. \( \quad \mathbf{b}_{s+1} = \mathbf{b}_{s+1} - \Delta V (N_{s+1}) \)
23. \( \quad \mathbf{b}_{s+1} \Rightarrow \text{Eigenvalue decomposition} \quad \tau_{s+1} = \frac{1}{2} \ln(\mathbf{b}_{s+1}) \)
24. \( \quad F_{s+1} = I_{s+1} : F_{s+1}^\top \)

Algorithm 3: Si-II evolution bound enforcement

1. \textbf{Input:} \( I_{s+1}, F_{s+1}, \tau_{\text{new}}, \tau, \star \tau \)
2. \( \Delta \nu = -\tau_{\text{new}} \cdot \nu_{\text{new}} \)
3. \( \nu_{\text{new}} = -\ln(\det(F_{s+1})) \)
4. \( \nu_{\text{new}} = -\frac{1}{2} \nu(\tau_{s+1}) \)
5. \( \nu_{\text{new}} = -\frac{1}{2} \nu(\tau_{\text{new}}) \)
6. \( \quad A = (\nu_{\text{new}} - \nu_{s+1}) \Delta \nu \)
7. \( \quad B = \nu_{\text{new}} - \nu_{s+1} \)
8. \( \quad \nu_{\text{new}} = (B + k^C_\nu) / (k - A) \)
9. \( \quad \nu_{\text{new}} = \nu_{\text{new}} \nu_{s+1} \)
10. \( \quad \nu_{\text{new}} = \nu_{\text{new}} \nu_{s+1} \)
11. \( \quad \nu_{\text{new}} = \nu_{\text{new}} \nu_{s+1} \)
12. \( \quad \nu_{\text{new}} = \nu_{\text{new}} \nu_{s+1} \)
13. \( \quad \nu_{\text{new}} = \nu_{\text{new}} \nu_{s+1} \)
14. \( \quad \nu_{\text{new}} = \nu_{\text{new}} \nu_{s+1} \)
15. \( \quad \nu_{\text{new}} = \nu_{\text{new}} \nu_{s+1} \)
16. \( \quad \nu_{\text{new}} = \nu_{\text{new}} \nu_{s+1} \)
17. \( \quad \nu_{\text{new}} = \nu_{\text{new}} \nu_{s+1} \)
18. \( \quad \nu_{\text{new}} = \nu_{\text{new}} \nu_{s+1} \)
19. \( \quad \nu_{\text{new}} = \nu_{\text{new}} \nu_{s+1} \)
20. \( \quad \nu_{\text{new}} = \nu_{\text{new}} \nu_{s+1} \)
21. \( \quad \nu_{\text{new}} = \nu_{\text{new}} \nu_{s+1} \)

References


