Primary radiation damage

Nordlund, Kai

2018-12-15


http://hdl.handle.net/10138/268058
https://doi.org/10.1016/j.jnucmat.2018.10.027

Downloaded from Helda, University of Helsinki institutional repository.
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.
Please cite the original version.
Primary radiation damage: A review of current understanding and models

Kai Nordlund a,*, Steven J. Zinkle b,c, Andrea E. Sand a, Fredric Granberg a, Robert S. Averback d, Roger E. Stoller c, Tomoaki Suzudo c, Lorenzo Malerba f, Florian Banhart g, William J. Weber h,c, Francois Willaime i, Sergei L. Dudarev j, David Simeone k

a University of Helsinki, Department of Physics, P.O.Box 43, 00014, Finland
b University of Tennessee, Department of Nuclear Engineering, Knoxville, TN, 37996, United States
c Oak Ridge National Laboratory, Materials Science & Technology Division, P.O. Box 2008, Oak Ridge, TN, 37831, United States
d University of Illinois, Department of Materials Science & Engineering, Urbana, IL, 61801, United States
e Japan Atomic Energy Agency, Center for Computational Science and e-Systems, Tokai, Ibaraki, 319-1195, Japan
f SCK-CEN, Institute for Nuclear Materials Science, B-2400, Mol, Belgium
g University of Strasbourg, CNRS, Institut de Physique et Chimie des Matériaux, UMR 7504, F-67034, Strasbourg, France
h University of Tennessee, Department of Materials Science and Engineering, Knoxville, TN, 37996, United States
i DEN-Département des Matériaux pour le Nucléaire, CEA, Université Paris-Saclay, F-91191, Gif-sur-Yvette, France
j Culham Centre for Fusion Energy, UK Atomic Energy Authority, Abingdon, Oxfordshire, OX14 3DB, United Kingdom
k DEN/DMN/LA2M-LRC CARMEN, CEA, Université Paris-Saclay, F-91191, Gif-sur-Yvette, France

A R T I C L E   I N F O

Article history:
Received 23 July 2018
Received in revised form 16 October 2018
Accepted 16 October 2018
Available online 19 October 2018

Keywords:
dpa
Displacement cascades
Defect production
Thermal spike

A B S T R A C T

Scientific understanding of any kind of radiation effects starts from the primary damage, i.e. the defects that are produced right after an initial atomic displacement event initiated by a high-energy particle. In this Review, we consider the extensive experimental and computer simulation studies that have been performed over the past several decades on what the nature of the primary damage is. We review both the production of crystallographic or topological defects in materials as well as radiation mixing, i.e. the process where atoms in perfect crystallographic positions exchange positions with other ones in non-defective positions. All classes of materials except biological materials are considered. We also consider the recent effort to provide alternatives to the current international standard for quantifying this energetic particle damage, the Norgett-Robinson-Torrens displacements per atom (NRT-dpa) model for metals. We present in detail new complementary displacement production estimators ("athermal recombination corrected dpa", arc-dpa) and atomic mixing ("replacements per atom", rpa) functions that extend the NRT-dpa, and discuss their advantages and limitations.

© 2018 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

* Corresponding author.
E-mail address: kai.nordlund@helsinki.fi (K. Nordlund).

https://doi.org/10.1016/j.jnucmat.2018.10.027
0022-3115/© 2018 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
1. Introduction

Particles with kinetic energies clearly above conventional thermal energies, i.e. with $E_{kin} > 1$ eV, exist in nature due to cosmic radiation and radiation decay, but are nowadays produced in a wide range of man-made devices for basic research and practical applications. For instance, the great accelerators at CERN and other particle physics laboratories in the world attempt to unravel the fundamental nature of the universe [1,2], and numerous smaller devices are widely used for equally exciting research in physics [3], chemistry [4], medicine [5] and nanoscience [6]. On the application side, ion implantation is one of the key technologies in silicon chip manufacturing [7], and electron accelerators are one of the key ways to treat cancer [8]. All of these activities make it interesting and important to understand what are the fundamental effects of high-energy particles on matter. Moreover, in nuclear fission reactors, which currently provide about 13% of the world’s electricity, materials degradation associated with neutron irradiation damage is a key factor [9].

One of the main consequences of the interaction of high energy particles (electrons, ions, neutrons and photons, or more exotic particles such as neutrinos [10] or possibly even dark matter particles [11]) with materials is the formation of lattice defects resulting from the energy transfer to the atoms. Other consequences include production of non-damage-producing phonons, excitons and plasmons, secondary electrons and photons, and heating of the material. The effects on the material can be either detrimental and beneficial. The defects can take many forms: in a crystal it is easy to understand that an atom can be kicked out from its initial lattice site, leaving an empty site (a vacancy) behind and creating an atom at an interstitial site in front [12]. It is, however, important to realize that the crystal defects formed can also be much more complicated in structure [13]: they can for instance be defect clusters [14], amorphous zones [15,16], dislocation loops [17] or three-dimensional defects [18,19]. On surfaces the damage can also take the form of adatoms [20], craters [21,22] and ripples [23]. In amorphous materials the generated defects can be over- or under-coordinated atoms [24] or empty regions [25]. Photon irradiation creates damage largely by electronic excitation processes causing bond breaking [26], although very high-energy gamma photons can also produce damage by atomic recoil processes [27].

The damage production mechanisms can in most cases be well divided into two categories by time scale. The primary damage is formed immediately after the particle impact by atomic collision processes and strong material heating caused by the colliding atoms far from thermodynamic equilibrium. Numerous computer simulation and experimental studies have shown that the time scale for the ballistic atom collision processes is of the order of 0.1 – 1 ps and the time scale for subsequent thermalization of the collisions about 1 – 10 ps [13,28,29], see Fig. 1 and e.g. Ref. [13]. After this athermal (in the sense that thermally activated processes at equilibrium are not significant) stage, long-time scale (nanoseconds to years) damage evolution caused by thermally activated processes can occur. The precise duration of the primary damage stage depends on primary knock-on atom (PKA) energy [13] as well as temperature, becoming shorter at high temperature due to a higher final thermalization temperature (smaller energy change to become thermalized) and exponentially increasing atomic mobility which accelerates the follow-on thermally activated defect evolution stage.

In this review, we discuss the current understanding of primary radiation damage from neutrons, ions and electrons on inorganic materials. We leave out many other important and interesting aspects of radiation damage. Specifically, we exclude from our consideration:

- Damage by photons, molecular and nanocluster projectiles
- More exotic elementary particles (such as muons, positrons, neutrinos, possible dark matter, etc.)
- The properties (stability and mobility) of the defects produced
- Damage at elevated temperatures and its consequences (solute segregation and precipitation, void swelling, irradiation creep, embrittlement in all its forms, etc.)
- Response functions, reaction to external driving forces (such as stress, strain or electromagnetic fields)
- Electronic defect production, including swift heavy ions and ion track production by them
- Organic materials

2. Understanding of primary damage in materials

2.1. Threshold displacement energy

The threshold displacement energy (TDE) (sometimes just called the displacement energy or the threshold energy) is in a
Fig. 1. a) Collision cascade induced by a 10 keV recoil in Fe at an ambient temperature of 0 K. The circles indicate atom positions in a 2 unit cell (4 atom layer) thick cross section through the center of the simulation cell, and the color scale the kinetic energy of the atoms. Note how a very large number of atoms is initially displaced, but when the cascade cools down, almost all of them return to perfect crystal positions. This is the athermal recombination effect discussed extensively in this report. b) Exactly same collision cascade, but with the color scale being potential energy (upper row) or atom movement distance from original positions (lower row). Comparison of these two shows that even though almost all atoms in the cascade do return to some perfect crystalline position, many of them do not return to the same position they started in. Hence the number of atom replacements is much larger than the number of defects produced. Original figure for this Article. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
sense the most fundamental quantity for describing radiation damage in materials. It is often defined in a very simple form, as the minimum recoil energy given to an atom in a material needed to create a stable defect \[13,30,31\]. The displacement process is traditionally described as a step function, with zero probability of generating a stable displacement below the threshold energy and full certainty of producing a stable displacement above the threshold energy. With this definition, it can be given as a well defined quantity. However, the true nature of the threshold displacement energy is more subtle, and hence it is important to understand in detail its nature and limitations.

The velocity of a recoil produced by an incident high-energy particle is much larger than thermal velocity. Hence as a rough first approximation, one can imagine that the energy required to form a defect can be estimated by assuming that only the recoil atom moves, and determining the potential energy barrier along a straight line in the initial recoil direction, see Fig. 2. This so called “sudden approximation” has been shown to work well at least for the low-index crystal directions in simple crystals [32]. This model and illustration makes it rather obvious that the threshold energy must depend on the direction of the recoil in relation to the crystal lattice.

One has to recognize that this defect production mechanism is different from the thermally activated formation of Frenkel pairs (pairs of vacancies and interstitials [33]). As the latter typically involves numerous \(10^2 - 10^{15}\) lattice vibrations in the potential well before a defect is formed [34], in metals this can occur only close to the melting point [35]. Indeed, the TDEs are typically of the order of \(10 - 50\) eV [36–41], roughly a factor of 5 – 10 higher than the Frenkel pair formation energies [33]. This difference between the Frenkel pair formation energy (referring to the creation of a monovacancy and a self-interstitial at infinite separation) and the threshold displacement energy is due to spontaneous recombination effects. The lattice strain associated with the production of a Frenkel pair can induce spontaneous recombination of the vacancy and self-interstitial if they are located within several lattice periods from each other, even at cryogenic temperatures where in the classical approximation point defect motion does not normally occur (although it does occur if quantum zero point vibrations of atoms are taken into account [42–45]). The experimental [46–48] and computational [49,50] model-based estimates of this spontaneous recombination distance are on the order of 0.5 – 2 nm. Since typical ion stopping powers in solids are \(-20 – 100\) eV/nm at low energies, additional kinetic energy of the order of \(10 – 60\) eV must be imparted to the PKA to allow it to become separated from its paired vacancy by greater than the spontaneous recombination distance.

It is also important to recognize that on the atomic level, there is no unique value of the threshold displacement energy, but each crystal direction has its own TDE due to the anisotropy of the crystal structure [39]. This is due to the crystallographic dependence of both kinetic energy transfers (initial outward movement of the self-interstitial) as well as the subsequent spontaneous recombination probability [51]. In Fig. 3 the damage production factor is plotted as a function of recoil energy for polycrystalline copper (experimental average of \(40\) collision directions) [52]. From the figure it is clear that the displacement probability increases gradually above the minimum recoil energy \((-19\) eV), which can cause displacement in the lowest energy crystal direction. An apparent plateau in the

![Image](https://via.placeholder.com/150)

**Fig. 2.** Illustration of the threshold displacement energy concept. If the bottom left atom 0 at \(R_0\) receives a recoil e.g. in directions A or B, it will cause a complex many-body collisional process. In case the atom would move in a straight path along path B and the other atoms would not have time to move at all (which of course in reality will not be the case), the potential energy in the system could be illustrated as on the right side in the “Sudden approximation” curve. At some point there is a maximum in the potential energy curve, before the atom goes into an interstitial position in one of the minima to the right of the original position. This maximum is an estimate of the threshold displacement energy. In reality the atoms respond somewhat to the motion of the recoiling atom 0, which lowers the threshold energy. In equilibrium, the threshold is much lower since all atoms have ample time to relax their position and a defect is formed only when the atom positions and movements happen to be such that a defect can form. Path A illustrates a case where the original atom 0 is directed directly towards another atom 1. In this case the original atom most likely replaces the atom at 1, while atom 1 receives a secondary recoil and becomes an interstitial atom. However, after this initial defect production, recombination effects may still affect the threshold, see discussion in main text. Figure from OECD review article [41], reprinted with permission TO BE OBTAINED.

![Image](https://via.placeholder.com/150)

**Fig. 3.** Illustration of the damage function as a function of energy for Cu. The graph clearly shows that there is not a single TDE value for damage production. Overlaid on the original figure is a step function illustrating the NRT-dpa equation with a threshold of 30 eV. Figure adapted from Ref. [52], reprinted with permission from Elsevier.
defect production probability occurs above 30 eV (the accepted crystallographic-averaged TDE for Cu), in qualitative agreement with the Kinchin-Pease (KP) or Norgett-Robinson-Torrens (NRT) damage functions (these models are described in detail in Section 4.1). However, the displacement probability for energies above the average TDE is less than unity and does not approach 1 until the recoil energy is 2–3 times the average TDE (or approximately 4 times the minimum TDE). Therefore, it is not possible to use a single TDE value to describe near-threshold displacements in crystalline materials.

At a finite temperature the situation is even more complex. Each atom in a lattice has some kinetic energy, \( E_{\text{kin}} \), distributed according to the Maxwell-Boltzmann statistical law. If the thermal velocity vector happens to be in the direction of the recoil energy received, this increases the damage production probability, and vice versa. Moreover, also thermal fluctuations in atom positions and velocities can strongly affect the recombination probability.

The effect of temperature on the TDE in several metals and ceramics has been observed in experiments [53–57] and analytical model [58] studies to gradually decrease with increasing temperature; this has been attributed to a combination of reduced elastic stiffness, lattice thermal vibration contributions to the kinetic energy of the PKA, and lattice vibration contributions to the reduction of the spontaneous recombination distance [55,57]. On the other hand, recent molecular dynamics simulations (MD) study of rutile TiO\(_2\) reported an increase in the TDE value with increasing temperature [59]. However, in this case the effect appears to be associated with enhanced thermally-induced correlated recombination, which is inappropriate to consider for quantification of primary damage parameters such as TDE.

The stochastic nature of primary damage events is an additional important consideration that requires a large numbers of primary knock-on atom collisions to be evaluated to achieve accurate statistics of computational simulations. For example, molecular dynamics simulations in Fe at 36 K [59] showed that even for recoil energies as high as 400 eV, the net amount of defects after all recombinations can sometimes even be zero (even though the energy is high enough in principle to produce tens of defects by formation energy considerations alone). This effect is illustrated in Fig. 4, where the lines in different shades of grey show damage production as a function of energy for recoils starting in exactly the same direction from the exact same position in a lattice initially at 0 K. The large fluctuations in the lines and the fact that they sometimes go back to zero even at energies approaching 400 eV illustrate the major degree of chaoticity and role of recombination in near-threshold damage production. All these complications mean that even for a single crystallographic direction, there is an extended energy range where the defect production probability gradually rises from 0 to 1 rather than exhibiting a sharp integer step [59]. In other words, defect production is stochastic in nature, rather than deterministic.

We note that the above subtleties do not, however, impede the use of the threshold displacement energy concept. This is because it is possible to calculate an average over the threshold displacement functions in different directions for use in analytical equations such as the KP or NRT models of displacements per atom (dpa). Even though there is some non-uniqueness in how the average should be calculated, but this does not amount to more than a couple of eV (less than 10% of the total) [59], which is a relatively small uncertainty compared to the other ones related to damage production (described later in this paper). We further note that to account for the non-uniquenesses, in principle it should be possible to extend the analytical NRT formalism and higher-level models to include various stochastic and fluctuating terms.

Although the threshold energy, \( E_d \), is usually considered to be independent of a crystal direction, we do note that since \( E_d \) is a user-defined parameter, the NRT equation could actually be used to calculate the dpa values for different crystal directions. This calculation would only be strictly relevant for near-threshold displacement events (since multiple directions are involved in energetic displacement cascades).

The TDE has been studied extensively both by experiments and simulations (see, e.g., Refs. [32,36,37,39,56,60–72and38]), there are still uncertainties for several technologically important materials. In particular, for Fe there are insufficient experimental values (only 3 directions [39]) to determine the average threshold, and there are big variations in the threshold displacement energies predicted by classical interatomic potentials [59]. Similarly in Si, there is a large variation in both experimental and classical molecular dynamics simulations. Recently, quantum mechanical density-functional theory (DFT) simulations have given more reliable theoretical values for the threshold displacement energies for Si, Ge and Fe [73–75].

On the other hand, the DFT work in Si also showed that below the threshold energy for the formation of a Frenkel pair, it may be possible to form a topological bond order defect [76,77]. This defect is metastable, but has been suggested to be an important factor for the amorphization of silicon under typical ion beam fluxes [78]. Contrarily, at low flux levels and for light ions or electrons, it is likely the bond defects would all recombine before damage overlap occurs. In this case, the relevant threshold energy would be that for Frenkel pair formation. In other words, in systems where metastable topological defects can form, it may make sense to define two different TDEs, one for low and another for high flux conditions. Somewhat surprisingly, one of the materials where the threshold displacement energy is known quite accurately is in carbon nanotubes and graphene [63,79,80]. This is because in these materials, aberration-corrected transmission electron microscopy
(TEM) experiments enable following the displacement of individual atoms. The graphene works are particularly interesting in that they showed explicitly that on a very precise level, the threshold is not an exact value even in a single crystal direction due to thermal vibrations of atoms [79].

2.2. Metals

Radiation damage in metals has been studied very extensively over many decades. This has been motivated primarily by the importance of controlling neutron-induced degradation of fission reactor materials. Until recently, the level of understanding of metal damage was also at the highest level among all classes of materials - now two-dimensional nanomaterials may actually be at a comparable level, as will be reviewed in Section 2.5. As many of the concepts of radiation effects originally introduced for metals have been transferred to other materials as well, it is natural to start the review for metals.

Classical MD simulations have been crucial in giving insight into the underlying physics for primary damage processes. Although in general MD simulations in materials science suffer from uncertainty of the reliability of the results due to the choice of interaction potentials [80], remarkably, MD simulations of cascades in metals very consistently give a similar behavior reproducing the experimental damage recombination [28]. We emphasize that as in the case of the thresholds, the cascades are highly chaotic, i.e. a small perturbation can lead to a large difference in the outcome. In the keV energy regime, the variability of outcomes is further strengthened by recoil channeling effects [81,82]. Hence any quantitative deductions from cascade results should always be obtained from a statistically representative set of simulations. The further discussion in this Section is based on studies where large data sets have indeed been obtained to gain a reliable view of typical behavior.

A typical example of the MD simulation results is given in Fig. 1, which shows how initially (first about 200 fs) a set of ballistic collisions lead to a major production of displaced atoms. It is noteworthy that this number of displaced atoms roughly corresponds to the NRT prediction, or what would be obtained from a binary collision approximation (BCA) simulation, that does not consider thermal effects. However, after this ballistic phase, the cascade becomes a heat spike (thermal spike), i.e. a dense region of many-body atomic collisions that is low density (“depleted zone”) in the middle and high density at the outskirts of the cascade [83]. After about 1 ps, the atoms in the spike (if it is dense enough) get a Maxwell-Boltzmann-like distribution of kinetic energies, i.e. behave like a thermodynamic system [28]. This justifies the use of the word “thermal” or “heat” spike (both terms are used with the same meaning, roughly as commonly). Calculating the average energy of the atoms in the heat spike and converting this into units of Kelvin gives a “temperature” of the order of 10 000 K for the heat spike [28]. Using the word “temperature” in this context is somewhat controversial, as the high-kinetic-energy phase only lasts some tens of picoseconds, and hence this is naturally not a normal equilibrium temperature. However, the textbook definitions of temperature [84] state that local thermodynamic equilibrium can be used if the time scale of the system is much larger than the relaxation time in the same system. As the heat spike relaxation time is about 3 lattice vibrations [28], roughly 1 ps, a heat spike lasting about 10 ps can be considered a thermodynamic system.

The cooling down of the heat spike can be considered a very rapid (ps time scales) recrystallization process of the hot liquid. Since a recrystallization process tends to produce perfect crystal, it is natural that much of the displaced atoms generated in the thermodynamic phase of the cascade regenerate (in other words, interstitials and vacancies recombine) and the produced damage is less than the initial number of displaced atoms (the NRT prediction). This explains the damage efficiencies clearly below 1 for energetic displacement cascades. Since this recombination process does not require any thermally activated defect migration (atom motion is caused primarily by the high kinetic energy introduced by the recoil atom), this recombination is called “athermal” (i.e. it can also happen if the ambient temperature of the sample would be 0 K), as explicitly shown in Fig. 4. The time scales of the three stages of cascade behavior are also illustrated schematically in Fig. 5.

By the argumentation given above, one could argue that all damage in metals would recombine, and the resulting $\xi = 0$, where $\xi$ is the surviving fraction of defects. This is, however, not observed either experimentally or computationally. A systematic comparison of MD cascade results in Fe [85] showed that all potentials used to simulate this material gave a similar trend of $\xi$ as a function of energy, and always saturation values for damage energies above about 10 keV between 0.2 – 0.5, with the dominant values being about 0.3, in agreement with experiments (see Fig. 6). Instead of total recombination, what typically happens is that a few isolated interstitials and interstitial clusters remain at the outskirts of a heat spike region, some vacancies are randomly distributed, and the center of the cascade remains a vacancy cluster. Most parts of this behavior are well understood. Isolated interstitials can be created by ballistic displacements of high-energy recoils and replacement collision sequences [51]. If these processes transport the interstitial clearly outside the heat spike region, the interstitial is not likely to re-enter the heat spike region and be recombined. Isolated vacancies can be created if the recrystallization front moves so rapidly that there is not enough time to create perfect lattice (there is always a finite probability of vacancy creation proportional to exp$(-E_v/k_B T)$, where $E_v$ is the vacancy formation energy and $T$ the temperature at the recrystallization front, which is close to the melting temperature). The formation of vacancy clusters in the middle of the volume of a cascade can be understood as follows: since a few interstitials are formed at the outskirts of the cascades, vacancies are left in the heat spike. These are pushed towards the center by the recrystallization front, and hence form a cluster there when recrystallization is complete. This mechanism can be observed explicitly by MD simulations of a bilayer system of elements with different melting points [86].

The formation mechanisms of interstitial clusters are somewhat less well established, and it seems several mechanisms can be active. The reported mechanisms include having an asymmetric recrystallization front that isolates a liquid pocket with an excess of atoms, leaving it behind as an interstitial cluster [87]. It has also been shown that the interaction of two supersonic shock fronts from nearby centers of heat spikes can lead to the stronger front injecting atoms into the underdense core of the weaker one, thus leading to the formation of an interstitial-rich region [88]. These two mechanisms are likely related to each other, i.e. the latter may explain the formation of a liquid pocket with an excess of atoms, that becomes isolated. Also dislocation loop punching [89,90] has been suggested to be possible in cascades [91], but the same lead author later found that this observation was due to an insufficiently short simulation time [92]. It seems that direct interstitial loop punching is not possible except near to surfaces [93,94].

The formation mechanisms of defects and defect clusters in cascades are thus reasonably well understood, and the total number of defects has been measured with resistivity recovery experiments and agrees well with MD simulation using modern interatomic potentials at least in Fe [95,96]. However, what fraction of defects is in clusters cannot be fully reliably predicted by MD or easily measured experimentally. Even the most modern MD potentials give significant variation of the fraction of damage in
clusters [95,97], and this fraction is also affected by uncertainties in how to treat the low-energy limit of electronic stopping power and the electron-phonon (e-ph) coupling [98]. Furthermore, recent studies for 50 and 150 keV PKAs in Ni and NiFe have demonstrated that the electronic energy loss by the PKA and energetic recoils, transferred back to the lattice via e-ph coupling, can contribute to athermal recombination of defects, further decreasing defect survival [99,100].

Traditional transmission electron microscopy experiments can readily observe large (>2 nm in diameter) defect clusters [101]. Modern microscopes can also detect smaller clusters, but there is limited work currently on the topic. In contrast, MD simulations mainly capture the formation of small defect clusters. Larger clusters appear infrequently, and then mainly in energetic cascades which demand significant computer capacity to simulate fully. Limited computational resources hence often restrict the feasible number of individual cascade simulations, lowering the likelihood of ever observing such rare events. One way of reconciling the size scales of primary radiation damage in MD simulations, with the observable damage in TEM experiments, is through the statistical distribution of defect sizes. These have been found to closely follow power laws [102,103]. The extrapolated distribution in W agrees well with in situ observations of the dislocation loops in W foils, irradiated at cryogenic temperatures [104] (see Fig. 8). The
combined MD and experimental data provide the basis for a model that describes the size-frequency distribution of defect clusters, extending from the smallest - experimentally invisible - to the largest occurring individual clusters observed in TEM irradiated foils [105]. This method also allows identifying which interatomic potential(s) give predictions in line with experimental results.

In Fe, however, a similar approach is hindered by the upper limit of the defect size distribution in the primary damage, which, judging from ion irradiation experiments at low temperature and low dose, lies mostly [106,107], if not exclusively [108], below the TEM observable limit. Nevertheless, simulations combining MD predictions with kinetic Monte Carlo methods indicate that large interstitial clusters in the primary damage may significantly affect the long-term microstructural evolution [109,110]. Hence clearly more work is needed to establish the fraction of damage in clusters, the structure of the clusters and their effect on the macroscopic damage evolution. Modern aberration-corrected TEM’s could give significant advances on the topic.

2.3. Semiconductors

Radiation damage production mechanisms in semiconductors differ clearly from those in metals [111]. For the primary damage production, this has been attributed to the open crystal structure [87] and the much slower recrystallization [112]. These effects lead to the possibility to form amorphous pockets directly in single primary damage events [15,113,114]. Defects formed in a cascade in Si are illustrated in Fig. 7, which shows how, contrary to metals, a large fraction of the cascading atoms remain in a disordered, “amorphous”, region after the cascade has cooled down. These pockets are also stable at room temperature over macroscopic times in most tetrahedral semiconductor materials [15,114–116], although at least in GaAs the amorphous pockets can recrystallize already at room temperature [117].

During prolonged irradiation at low temperatures, the semiconductors become completely amorphized [118–122]. In other words, the damage saturation effect described in Section 2.2 for metals does not necessarily exist in semiconductors. At elevated temperatures, however, semiconductors do not amorphize at any fluence [123–126]. Particularly, there is a critical temperature for amorphization which is typically a few hundred degrees centigrade above room temperature. For instance, for Si it is about 300°C and for SiC about 200°C (the precise value depends on ion species, energy and flux). The reason to this is attributed to defect migration that allows defects to recombine at elevated temperatures as well as thermal recrystallization of amorphous pockets, that prevents amorphization [124–126]. However, this mechanism is different from the one in metals, where (as discussed in Section 2.2) the recombination can occur without any thermal activation.

The formation of amorphous pockets (disordered zones) makes quantifying the amount of damage challenging, mainly because it is not obvious which atoms to define as a “defect” in such a zone. In the sense of atoms not being in a perfect crystalline environment, all atoms in an amorphous zone could be considered “defect atoms”. On the other hand, such a definition does not scale well with an analysis of Frenkel pairs in less disordered regions (note that even in a single vacancy, 4 atoms around the empty site could be considered “strongly disordered”). Analysis of atom coordinates obtained from MD have shown that the “defect number” can indeed vary as much as an order of magnitude for the exact same configuration depending on criterion used [87]. However, the different numbers were found to be fairly well proportional to each other, which means that any of the criteria can be used as long as the analysis method is clearly described.

In the elemental semiconductors Si and Ge, MD simulations consistently show a linear increase of damage production with ion energy, i.e. there is either no athermal damage recombination effect similar to that in metals, or it does not increase with recoil energy [87,113]. This is consistent with experiments which show, at least as a first approximation, that damage production scales fairly well with energy deposition; a dose of $6 \times 10^{23}$ eV/cm$^2$ at 80 K has been reported to predict well the amorphization dose for a fairly wide range of ions with different masses [127]. If there would be an athermal recombination effect like that in metals, one would not
expect a simple scaling with different ion masses. Some examples of defect production as a function of bulk recoil energy in Si and Ge are given in Table 1. For all potentials, the data scales fairly well with a simple KP/NRT equation (for the “Ge SW mod pot.”, the scaling between 0.4 and 2 keV is not good because in this potential large amorphous zones start to be formed around 2 keV). For instance, for the “Si SW pot.” data set an average threshold energy of 34 eV reproduces with a KP equation all the data points within the statistical uncertainty after the electronic energy loss of about 20% is counted out from the recoil energy.

Comparison of the values for the same element, with different potentials, show that there is a serious potential reliability issue, however. There are hopes that computers are soon powerful enough such that DFT MD simulations of full collision cascades can be used to reduce this uncertainty.

### 2.4. Ionic materials

A moderate amount of experimental information is available on threshold displacement energy surfaces for ionic materials, with MgO and Al₂O₃ being the most extensively studied materials [38,128]. Typical crystallographic-averaged values of the displacement energies for the anion and cation sublattices range from ~40 to ~100 eV. In recent years, improvements in molecular dynamics simulations in nonmetals have led to a significant number of studies on displacement energies in oxides including MgO [129,130], MgAl₂O₄ [131], TiO₂ [132,133], UO₂ [134,135], and ZrSiO₄ [136,137]. Because of the difficulty in developing suitable interatomic potentials and the advancements in computational resources, ab initio MD methods are more frequently employed in determining TDEs in simple and complex oxides, including MgO [138], ZrO₂, CeO₂, ThO₂ [139], Y₂Ti₂O₇ [140], Gd₂TiO₅, Gd₂ZrO₂ [141], La₂Zr₂O₇, Nd₂Zr₂O₇, Sm₂Zr₂O₇ [142], and SrTiO₃ [143]. Limited recent experimental studies of TDEs have also been performed on CeO₂ [144], ZrO₂ [145], ZnO [146], TiO₂ [147], and more complex oxides such as perovskite, titanates and zirconates [128,148].

There are significantly fewer defect production studies that have been performed on ionic materials compared to metals. This is in part due to the lack of a simple in situ measurement technique for ionic materials analogous to the electrical resistivity technique that has been a workhorse for defect production measurements in irradiated metals at cryogenic temperatures. Many of the experimental defect production studies on ionic materials have been performed at room temperature, where significant point defect mobility can lead to undertimates of the defect production efficiency [38]. In addition, many of the experimental measurements have utilized tools such as optical spectroscopy that are valuable for monitoring specific defects such as F-center monovacancies, but are less successful at quantifying the full spectrum of isolated and clustered radiation defects. Typical measured values of defect production efficiency in MgO and Al₂O₃ are 20 – 50% of the NRT calculated displacement value over a broad range of PKA energies, 0.1 – 100 keV [38,149], see Fig. 9, although defect production efficiencies near 100% have been measured in recent ion beam experiments on Al₂O₃ [150] and ZnO [151]. Clearly further work is needed to resolve the source of these discrepancies.

#### 2.5. Carbon-based materials

Graphite is still of importance as a moderator or reflector in reactor technology. Irradiation of graphite at low temperatures leads to a rise in energy due to the accumulation of defects. This is known as the Wigner energy [165], which is an important safety issue in nuclear technology. The sudden release of the Wigner

<table>
<thead>
<tr>
<th>Recoil energy (keV)</th>
<th>Si, SW pot.</th>
<th>Si, Tersoff pot.</th>
<th>Ge, SW pot.</th>
<th>Ge SW mod pot.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>4.1±0.5</td>
<td>8.3±0.2</td>
<td>2.4±0.1</td>
<td>4.6±0.3</td>
</tr>
<tr>
<td>2.0</td>
<td>17±1</td>
<td>39±2</td>
<td>12±1</td>
<td>47±2</td>
</tr>
<tr>
<td>5.0</td>
<td>43±1</td>
<td>84±2</td>
<td>27±1</td>
<td>141±5</td>
</tr>
</tbody>
</table>

*Fig. 9. Surviving defect fractions in irradiated Al₂O₃ and Al.*
energy above 250°C may cause disastrous events such as the Windscale fire in 1957. For this reason, radiation effects and defects in graphite have been subject of numerous studies [166–168]. However, the nature of the defects remained unclear until new nanomaterials on a graphitic basis, such as carbon nanotubes or graphene, became available. These materials are now of major technological interest, and radiation effects have to be taken into account in applications in space and zones of high radiation level. On the other hand, the encapsulation of other layered materials with graphene could prevent the former from radiation damage [169]. The controlled generation and reconstruction of radiation defects can be used for tailoring the properties of graphene and creating new morphologies of graphenic nanomaterials.

Detailed electron microscopy studies of the 2D materials, where electron irradiation and structural characterization can be carried out in situ at the same time, have clarified the picture also with respect to defects in bulk graphite [170–172]. Since the displacement of even single carbon atoms can be observed in situ in the electron microscopes, the situation in graphene is now very well investigated and understood. Graphene is therefore an ideal system to study atom displacements in detail, to test the theoretical concepts, and to determine threshold energies with high precision [173]. On the other hand, precise knowledge of radiation effects is indispensable to distinguish native defects from beam-induced artefacts. Electron irradiation is unavoidable in the electron microscope, but radiation defects have to be avoided nevertheless. This requires detailed knowledge about defect formation and annealing.

Two-dimensional materials, like graphene, consisting of only one coherent atomic layer, have allowed for the first time, to see individual atomic defects in modern aberration-corrected electron microscopes. By varying the acceleration voltage of the microscope and observing the appearance of lattice defects, the displacement threshold can be precisely determined. From these studies it is known that the creation of visible structural defects in graphene or graphite requires a minimum electron energy of approximately 80 keV if the momentum transfer occurs normal to the basal plane of graphite. This corresponds to a displacement threshold of approximately 17 eV (energy of the displaced carbon atom) [170]. Due to the presence of conduction electrons, electronic excitations are quenched and do not lead to structural changes under irradiation. The threshold for ballistic displacements shows a large anisotropy in graphite, and the value for in-plane displacements is higher by almost a factor of two. The McKinley-Feshbach formalism [174] describes the observed ballistic displacement rate rather well. However, lattice vibrations have to be taken into account [79], leading to a slightly higher displacement rate close to the threshold than predicted by the McKinley-Feshbach theory. Displacement thresholds for different carbon isotopes (such as $^{12}$C and $^{13}$C) can be distinguished and compared with theoretical predictions [175]. It has been observed that the TDE of atoms at the edges of graphene layers or around existing structural defects is clearly below the bulk threshold. This may lead to an ongoing degradation of graphitic materials starting from edges or defects at electron energies as low as 20 keV [79].

The configuration and behavior of atomic defects in graphite and graphene is different from other materials [176]. Due to the high structural flexibility of the graphenic lattice, different ways of reconstructing the hexagonal network and locally changing the hybridization of carbon atoms are possible. This allows a relaxation of the lattice via transformations of the Stone-Wales type after the displacement of atoms so that all dangling bonds around defects are saturated. Examples of vacancies in graphene are shown in Fig. 10. Monovacancies are stable against restructuring of the lattice but have a low migration energy of 0.9 – 1.4 eV [177]. They may therefore migrate already at or slightly above room temperature, and coalesce to form stable and immobile divacancies. This is the most important annealing mechanism in graphitic materials. The formation energy of a divacancy (7.2 – 7.9 eV for the removal of two atoms) is of the same order as of a monovacancy (7.3 – 7.5 eV for the removal of one atom), resulting in a defective graphitic lattice at elevated temperatures that consists mainly of arrangements of divacancies. Divacancies may occur in different configurations of pentagonal and heptagonal rings and are immobile up to very high temperatures. The reconstruction leads to a variety of new morphologies of graphitic materials, that can be induced by careful irradiation at elevated temperatures [178]. In contrast to vacancies, carbon ad-atoms are more difficult to study by electron microscopy, in particular since their migration energy is as low as 0.4 eV [179] and their thermal migration accordingly fast, even at room temperature. Therefore, much less is known about interstitials in graphitic materials, but the major features of defect production and annealing can be well understood by assuming that vacancies govern the production and dynamics of radiation defects. At temperatures below 250°C, the agglomeration of defects is seen as an ongoing rupture of basal planes, leading eventually to an amorphization of graphite. This has been confirmed in ion irradiation studies of carbon nanotubes [180].

Diamond has other applications than graphitic materials, and radiation effects in diamond have not been studied in similar detail until now. However, since the atom species are the same as in graphite but the bonding type and lattice structure are completely different, information about radiation effects in diamond is highly important. TDEs of 30 – 48 eV have been reported [38,181], which are higher than in graphitic materials. This is due to the dense packing of carbon atoms in the diamond lattice [182]. This makes diamond more stable under irradiation and therefore a “radiation-hard” material. The anisotropy of radiation damage is much less than in graphitic materials, due to the lack of open space in the diamond lattice. Although diamond is an insulator with a band gap of 5.5 eV, electronic excitations do not lead to visible structural defects under irradiation. Thus, we have to assume that ballistic displacements are the main source of radiation damage. The knowledge about radiation damage in diamond is also limited due to the difficulty of studying the behavior of individual carbon atoms in a densely packed three-dimensional crystal. Nevertheless, the understanding of irradiation of diamond is of importance for space applications and for use in environments with high radiation level. It has been shown that the higher radiation hardness of diamond can be made use of by irradiating a graphite-diamond interface with energetic electrons and so transforming graphite to diamond, even at low pressure [183,184]. An irradiation effect in diamond of current interest is the nitrogen-vacancy center [185] that can be made by irradiating nitrogen-doped diamond with energetic electrons, protons, ions, neutrons or photons, followed by annealing. Interesting applications of the nitrogen-vacancy centers in magnetic field detection and quantum computing are coming in sight [186].

Most recently, chains of sp$_2$-hybridized carbon atoms as a perfectly one-dimensional material attracted renewed interest due to their interesting electrical and mechanical properties [187]. Electron irradiation of carbon chains leads to immediate rupture after the ballistic displacement of an atom which can be followed in situ in electron microscopy studies [188]. Since vacancies, interstitials and annealing effects can be excluded in this case (a single vacancy cuts the chain), fundamental radiation studies with the precise determination of displacement cross-sections in a unique material appear to be feasible.
2.6. Effect of defects, dislocations, and surfaces on primary damage

In applications, it is infrequent that materials, bombarded by energetic ions or neutrons, have an ideal crystal structure. In the majority of cases the crystal structure is not perfectly periodic, as is often assumed in simulations. Real materials contain dislocations, grain boundaries, and various other defects. Surfaces also influence the rate of radiation defect production. Most significantly, even if the initial crystal structure of a materials were ideal, the generation of damage itself results in the gradual defect accumulation and dislocations. Hence for the interpretation of observations it is necessary to know how the microstructure of a material already containing defects changes as a result of a new collision cascade event initiated by an energetic projectile.

Early studies examined the effect of pre-existing cascade debris as a way to model cascade overlap, and reported that the sum of the damage produced by sequential cascades was less than twice that in single cascades [189]. A model study of cascades overlapping with a homogeneous point defect distribution also found that cascades in the heat spike regime can anneal damage [190]. A significant more recent finding is that cascades can give rise to dislocation climb [191]. For example, dislocations may absorb self-interstitial defects produced in a cascade, stimulating the formation of large vacancy clusters and vacancy loops [192,193]. The latter can be interpreted as a “production bias” effect [194], where the formation of large vacancy clusters results from the rapid absorption of interstitial defects by dislocations already on the collision cascade timescale. This is contrary to the conventional interpretation of vacancy clustering due to supersaturation resulting from the stronger elastic interaction between dislocations and self-interstitials than between dislocations and vacancies [195]. Cascades may also alter the morphology of existing defects completely, causing vacancy clusters to collapse into vacancy dislocation loops, changing the Burgers vector of small dislocation loops, or producing complex dislocation networks [196,197].

Recently, this work has extended to studies of multiple overlapping cascades, with the aim to obtain a picture of high-fluence effects [198–206]. The high-fluence effects are discussed further in Section 6.4. We note that, in principle, sequential MD simulations of overlapping cascades, while formally corresponding to the limit of a very high dose rate, in fact can be interpreted as examples of individual cascade events occurring in a material already containing radiation defects, irrespectively of the rate at which such events occur. This assumption will be comparable with experiments when the temperature is low and/or the defect migration barriers are high, and hence the thermal recombination of defects is low.

In the crystallographic sense, a surface can also be treated as a defect, since it breaks the symmetry of the Bravais lattice, and forming a surface requires a certain amount of energy. Surface energy can even be interpreted using the same equations as those defining defect formation energies [207]. Collision cascades occurring in the close proximity to a surface produce large vacancy clusters [103,208], similarly to cascades near dislocations producing large vacancy clusters and loops [192,193], mentioned above. Surface then acts as a sink for mobile point defects [209] and interstitial-like dislocation loops formed at the periphery of a heat spike [93]. If a heat spike intersects a surface, the hot material can experience viscous flow to the surface, leading to cratering [21,22,210,211]. Both effects can alter strongly the production of radiation damage in the bulk, even down to depths of 5–10 nm [93,103].

Naturally, any primary damage event near a surface can also lead to sputtering, a topic very extensively studied since its discovery in 1852 both experimentally and theoretically [212–214]. In recent decades, MD simulations have given detailed insights in the nanoscale origin of sputtering, most of them stemming directly from the primary damage state [21,215–220]. A more detailed discussion of sputtering effects is beyond the aims of this Review.

2.7. Primary damage in nanostructured materials

The presence of grain boundaries may also alter defect production rates [191,221–223]. This effect is particularly strongly pronounced in nanocrystalline metals, where the proximity of grain boundaries to almost any collision cascade event can have a major effect on the statistics of radiation defect production. This is particularly interesting from the point of view that nanocrystalline metals can, at least when prepared in a fully dense state, have much higher tensile strength than the same material in a normal polycrystalline state with grain sizes in the micron range [224–226].
Hence if nanocrystalline metals exhibit high radiation tolerance, they may prove attractive for a variety of applications in radiation environments. Also, since radiation embrittlement is a phenomenon controlled by the spatial scale of microstructure (i.e., not the strength of obstacles but their density) [227], nanostructural materials naturally provide a route through which brittle behavior can be controlled through an alternative mechanism, for example, grain boundary sliding, which circumvents dislocation-controlled plasticity and the fundamental limits associated with it.

There has been a range of studies of radiation hardness of nanocrystalline metals. A grain boundary can act as a sink for vacancies or interstitials, depending on the local atomic density of the boundary, see Fig. 11. For instance in Ni, MD simulations showed that interstitial annihilation at grain boundaries leads to a defect morphology dominated by vacancy-type defects, along with a complex partial dislocation network [221,222]. Such effects have also experimentally been shown to lead to enhanced irradiation resistance [223]. For instance, it was noted that during MeV Ar irradiation of TiNi, a nanocrystalline alloy retained the long-range order while its coarser-grained counterpart was amorphized [228]. The strengthening is likely related to grain boundary segregation, which is known to strengthen metals under some conditions [229,230]. From these observations, it seems that nanocrystalline microstructure could indeed be a major benefit to the radiation hardness of metals. However, one has to keep in mind that since grain boundaries have an interface energy, thermodynamics drives any nanocrystalline system towards grain growth. Irradiation can either work against this (by introducing dislocations or segregation that stabilizes small grains) or enhance it by allowing normally inert grain boundaries to move [231]. Hence, while it appears clear that nanocrystallinity can enhance radiation hardness at low doses and temperatures [232,233], long-term stability in a high-dose high-temperature environment is much less certain and should be ascertained through further work.

A related concept for finding high-radiation hardness materials is the idea to use nanostructured thin film materials as defect sinks. This concept has also been studied in good detail (see Ref. [234] and references therein). In addition to the direct annihilation of cascade defects in the primary damage production state, it was shown that interstitials can also be emitted from interfaces and thus enhance interstitial-vacancy recombination [223]. However, the same caveat as for nanocrystalline metals regarding long-term stability at high doses applies also for these systems.

For realizing the high-strength benefits of both nanocrystalline and nanostructured thin film materials, having a full-density starting material is an important starting point. However, intentionally highly underdense foam-like materials such as aerogels, are very interesting for a range of applications such as heat isolation [235] and capturing interstellar dust grains [236]. Due to the possible use of such nanofoams in space or nuclear reactors, it is also of interest to examine their radiation tolerance. In a nanofoam, where the base material is surrounded by vacuum or a gas, one can analogously to the case of interfaces imagine that interstitials and vacancies could easily migrate to the surface. Experiments and computer simulations of gold nanofoams indeed show that this is the case [237]. Moreover, detailed analysis in that work indicated that there is a window of the diameter of foam ligaments and dose rate where an enhanced radiation endurance is possible. The special geometry of a nanofoam has also been shown lead to markedly reduced sputtering yields due to the high possibility of redeposition of sputtered material to another part of the foam [238]. This simulation result is well in line with experiments on the sputtering of so called W fuzz (a foamlike morphology of tungsten formed after very high dose He bombardment [239–241]), which showed that the sputtering yield can be reduced by about an order of magnitude in a fuzz-like morphology [242].

Radiation damage has also been examined in 0-dimensional materials such as nanoclusters and nanorods (and carbon nanomaterials, which are discussed separately in Section 2.5). As there are extensive reviews of radiation damage in nanomaterials [6,243,244], we here only mention one aspect of primary damage in nanoclusters that relate to the issues discussed above with respect to nanocrystalline metals. One common class of nanocluster morphology is the multiply twinned icosahdron [245], which can be considered a polycrystal composed of exactly 20 individual single-crystalline tetrahedral grains. Experimental studies of the irradiation of these clusters showed that a significant fraction of the clusters can be driven into a single-crystalline morphology [246]. MD simulations showed that this can be explained by heat spike induced grain boundary movement [247]. Taken together, the two studies confirm that pure irradiation effects can induce grain growth in a nanocrystalline system.
2.8. Amorphous and quasicrystalline materials

The simplest definition of an “amorphous” material is that of a material being fully disordered. However, all real materials must have some short-range order at least at a first nearest neighbor distance scale. This is because the size of the core electron shells prevents atoms from being arbitrarily close to each other, which leads to the distance pair correlation function being zero at very small interatomic separations and having a first nearest neighbor peak. In practice, measurements of pair correlation functions show that most amorphous materials have order at least to the second nearest neighbor peak. Hence it is also possible to create and detect radiation damage in amorphous materials, e.g. as disturbances of the pair correlation function. Moreover, in amorphous materials with fully or partly covalent bonding (such as semiconductors and most ionic materials) the coordination number of atoms can also be used as a measure of defects: since e.g. Si chemically prefers a coordination of four, Si atoms with any other coordination number can be considered defects. A region of several nearby under-coordinated Si atoms could be considered a vacancy-like defect. A direct evidence that such a definition makes sense is provided by positron annihilation experiments that have clearly detected open-volume defects in amorphous silicon with characteristics similar to a vacancy in crystalline silicon [25]. Moreover, these open-volume and other possible defects can be annealed by heating the sample [25,248], similar to the well-known characteristics of defects in crystalline materials. The annealing has also been shown to release considerable amounts of heat [248].

In ionic amorphous materials, such as silica-based glasses, it is well established that color centers correspond to dangling bond or impurity defects, similar to those observed in the corresponding crystalline materials (quartz in the case of silica) [249,250].

Computer simulations have given considerable insight into the nature of defects in amorphous materials. They have shown that both vacancy-like open-volume [251–253] and interstitial-like compressed regions [253] can exist in metallic glasses. Such defects can be defined via coordination symmetry of atoms: well optimized amorphous glasses have a large fraction of high-symmetry neighborhoods around atoms, e.g. having icosahedral symmetry. Monitoring the change of such coordination states via Voronoi polyhedron analysis can allow de

The radiation mixing, i.e. permanent displacement of atoms from their initial site, is of clear experimental relevance. For instance in a bilayer sample, it will lead to broadening of the interface [272–274], which can be experimentally measured with several depth profile analysis techniques [275]. In systems with embedded precipitates or nanoclusters, it can lead to partial or full dissolution of the embedded feature [276]. In a fully homogeneous material, the mixing effect is in principle meaningless as all atoms of the same type are equal. However, even a slight concentration gradient of any of the elements involved makes it a meaningful quantity. A few examples that illustrate the critical importance of ion beam mixing to understanding radiation damage are dissolution of precipitates in two-phase alloys [277], re-solution of fission gas in nuclear fuels [278], and chemical disordering and amorphization of intermetallic or ionic compounds [279,280].

A bit of thought makes it clear that the amount of atoms mixed by an irradiation event is likely to be larger than the damage produced. Even for a recoil just above the threshold displacement energy, sequences of atom replacements can lead to several atoms entering new positions. Such a replacement collision sequence has been known from some of the very first MD simulations ever performed [51]. In the heat spike regime, as illustrated in Fig. 1b, it is even visually clear that a very large amount of atoms can be displaced in the molten zone. From this, it is clear that the KP or NRT-dpa equations are not suitable to describe ion beam mixing.

One of the early recognitions of ion beam mixing in fact derives from measurements of disorder in Cu$_x$Au during low temperature, fast-neutron irradiation [281]. In order to explain the large changes in the electrical resistivity that occurred, Siegel estimated already in 1949 that many Cu and Au atoms must exchange locations for each Frenkel pair that was created [281].

Three mechanisms of ion beam mixing have now been identified, each operating in a different phase in the evolution of the cascade.

**Recoil implantation:** As energetic ions slow down in solid targets they undergo a series of collisions with target atoms via a screened Coulomb interaction. Two consequences of the screened Coulomb interaction is: (i) the initial ion only slowly changes its incident direction, and (ii) it transfers large amount of energy to only few atoms. This results in a few target atoms receiving high energies, with their motion close to the direction of the initial ion. For ion irradiations, this results in the forward recoiling of atoms into the sample. For neutron or fast fission fragments, the directions are random, but nevertheless a few atoms recoil very large distances. Recoil implantation can thus play an important role, for example in the resolution of fission gas bubbles in reactor fuels [282] or destabilizing nano-ODS (oxide dispersion strengthened) steel alloys [283] materials presently under consideration for advanced reactors components.

An example of recoil implantation is shown in Fig. 12, where the number of Cu atoms recoiling more than 60 nm into Al is shown as a function dose for 500 keV Xe irradiation [284]. Notice that the number of Cu atoms recoiling across a Cu/Al interface and into
widely used to obtain the magnitude of mixing $NRT$. Scales reasonably well with damage energy density, $\epsilon$, calculations of ballistic mixing are performed in the same spirit as the mean square displacement of a target atom, using computer simulations, such as TRIM. A few examples are Mund and Gras-Marti using Boltzmann transport theory \cite{285} is required for replacement events, $\sim 15$ eV, are much larger than energies for chemical ordering, $\sim 1$ eV. Various models have been developed to calculate ballistic mixing. The result derived by Sigmund and Gras-Marti using Boltzmann transport theory \cite{285} is widely used to obtain the magnitude of mixing $\zeta_{BM}$. They found for the mean square displacement of a target atom, $<R^2>$, normalized by the deposited damage energy density,

$$
\zeta_{BM} = \frac{<R^2>}{\Phi D} = \Gamma_0 \frac{\zeta_{21}^2 A_e^2}{3N_0E_e}
$$

where $\Gamma_0$ is a dimensionless constant ($= 0.608$), $\zeta_{21} = 4(m_1m_2)/(m_1 + m_2)^2 E_e$ is the minimum energy for atomic displacements and $\lambda_c$ is the mean range of a recoil distance of energy $E_e$. Typical values for $\zeta_{BM}$ are $10A^2$/eV or $50A^2$/dpa. Calculations of ballistic mixing are performed in the same spirit as the NRT model of defect production; accordingly, ballistic mixing, like NRT, scales reasonably well with damage energy density, i.e. $\Phi D$. More quantitative calculations of ballistic mixing can be made using computer simulations, such as TRIM. A few examples are provided in Table 2. Notice that the mixing, which is normalized by dpa, is nearly independent of energy and target mass. For lighter targets, especially those with more open structures such as Si, low-energy recoils (i.e., secondaries, terriaries, etc.) appear to increase significantly the cascade mixing.

**Ballistic mixing**: The momentum of displaced atoms in energetic displacement cascades resulting from higher order displacement events is nearly isotropic. Ballistic mixing, therefore, operates similarly to diffusion processes. The mixing is considered ballistic in the sense that alloy components flow down gradients in their concentrations during prolonged irradiation, and not gradients in their chemical potential. This simply reflects the fact that energies required for replacement events, $\sim 15$ eV, are much larger than energies for chemical ordering, $\sim 1$ eV. Various models have been developed to calculate ballistic mixing. The result derived by Sigmund and Gras-Marti using Boltzmann transport theory \cite{285} is widely used to obtain the magnitude of mixing $\zeta_{BM}$. They found for the mean square displacement of a target atom, $<R^2>$, normalized by the deposited damage energy density,

$$
\zeta_{BM} = \frac{<R^2>}{\Phi D} = \Gamma_0 \frac{\zeta_{21}^2 A_e^2}{3N_0E_e}
$$

where $\Gamma_0$ is a dimensionless constant ($= 0.608$), $\zeta_{21} = 4(m_1m_2)/(m_1 + m_2)^2 E_e$ is the minimum energy for atomic displacements and $\lambda_c$ is the mean range of a recoil distance of energy $E_e$. Typical values for $\zeta_{BM}$ are $10A^2$/eV or $50A^2$/dpa. Calculations of ballistic mixing are performed in the same spirit as the NRT model of defect production; accordingly, ballistic mixing, like NRT, scales reasonably well with damage energy density, i.e. $\Phi D$.

More quantitative calculations of ballistic mixing can be made using computer simulations, such as TRIM. A few examples are provided in Table 2. Notice that the mixing, which is normalized by dpa, is nearly independent of energy and target mass. For lighter targets, especially those with more open structures such as Si, low-energy recoils (i.e., secondaries, terriaries, etc.) appear to increase significantly the cascade mixing.

**Thermal spike mixing**: Koehler and Seitz \cite{289} and later Vineyard \cite{290} considered whether atomic jumps could take place during the thermal spike phase of the cascade. The number of jumps is simply calculated by the integral,

$$
\eta = \int d^3r \frac{\epsilon}{k_B T}
$$

where the jump rate per unit volume of material is $R = \exp(-Q/k_B T)$, $Q$ is the activation barrier for atomic jumping. For a cylindrical geometry, assuming the initial cascade energy is deposited as a delta function, $\epsilon \times \delta(r)$ and $\epsilon$ is the damage energy per unit length, the Vineyard model yields,

$$
\eta = \frac{0.02226A_e^{5/3}}{C_3^{1/3}Q^{5/3}}
$$

where $\epsilon = \epsilon_0 \delta(r)$ is the initial energy of the cascade, $C$ is the specific heat, $k_B$ is the thermal conductivity and $Q$ is the activation energy for diffusion. Different models include different assumptions about $Q$, however, MD simulations have shown that unless the local temperature exceeds the melting temperature, diffusion is negligible on the time scale of thermal spikes. As a consequence, thermal spike mixing becomes increasingly important as the melting temperature increases and the energy density increases, i.e., the atomic mass of the material increases (See e.g. Ref. \cite{291}).

While dividing the cascade into different stages is conceptually attractive, accurate calculations of ion beam mixing require MD computer simulations. Fig. 13 illustrates the time evolution of mixing in Si (Fig. 13a) and in several metals (Fig. 13b) for cascade energies up to 10 keV. Clearly noticeable is that in Si the mixing is complete within a few tenths of picoseconds, but it continues up to a few picoseconds in heavier metals. For higher energy cascades, the time required for complete mixing extends even longer. This shows that in Si, which has low atomic number and fairly high melting temperature, the mixing is largely ballistic, but in metals, it derives mostly from the thermal spike.

Only a few MD simulations have been performed at energies typical of recoil events from fast neutrons, $E > 50$ keV. The results from one such study on Ni, Pd, and Pt for energies up to 200 keV is shown in Table 3 and Fig. 14. The mean square displacement of atoms is well described by the expression,

$$
\langle R^2(E) \rangle = a \frac{E^{3/2}}{b^{3/2}} + bT^2
$$

where $a$ and $b$ are constants. At high energy, $\langle R^2(E) \rangle$ scales linear with energy. This is the regime where subcascades form. Values for $a$ and $b$, as well as values of the mixing parameter are listed in Table 2. The table shows two different MD results, one where the $\langle R^2(E) \rangle$ was fit to only five data points \cite{292}, another one into a recent much more extensive data set \cite{41}. The differences show that the $R^2$ fit is sensitive to the quality of the data set. Even then,
both MD data sets agree very well with the experimental values within the uncertainty. A further uncertainty in the analysis comes from a recent argument that the approach of using mean square displacement as measure of mixing may overestimate the quantity \[Q_e\] [293].

Ion beam mixing has been measured on several metals at low temperatures in many systems using tracer impurities. A few results [291] for high energy heavy ion irradiation of metals are listed in Table 4 and an extensive review of values is provided in Ref. [294]. The results illustrate that mixing increases with decreasing melting temperature and increasing atomic number. It is noteworthy that the small mixing parameters for metals like Fe and Ni are comparable to values obtained for ballistic mixing, alone, indicating that thermal spikes are less important in these metals.

4. Analytical models for primary damage

4.1. Kinchin-Pease and NRT models

The first reasonable analytical model for estimating radiation damage in materials was provided by Kinchin and Pease (KP) in 1955 [295]. Their KP model gave the basis for an early model to calculate displacements per atom, by considering kinetic energy transfers above a threshold, the material-specific “displacement energy”. In the original model (still used as is e.g. in the SRIM computer code [31] quick damage calculation mode), the damage is simply calculated as the deposited energy (damage energy) \(T_d\) divided by two times the displacement energy \(E_d\). Originally \(T_d\) was considered to be the total ion or recoil energy, but after the detailed theory of stopping power was developed by Lindhard et al. [296] in the 1960’s, it became clear that one should use the nuclear deposited energy (recoil energy minus energy lost to electronic stopping) for this quantity.

The KP model was later refined by Norgett, Robinson and Torrens about 40 years ago using binary collision computer simulations of ion collisions in solids [49,297]. These simulations showed that even in a random-walk like binary collision sequence, about 20% of the sites from where an atom is kicked out happen to be filled by another atom, which motivated introducing a prefactor of 0.8 to the Kinchin-Pease equation. Note that this “binary collision recombination” is not the same effect as the much more complex thermodynamic recrystallization effect described in Section 2.2 for

### Table 3

Comparison of ion beam mixing of MD simulation with experiments. Two different sets of MD results are given: (1) from Ref. [292], and (2) a fit to a more extensive data set from Ref. [41] Note: \(Q = \frac{1}{2}R^2/6\phi_0\).

<table>
<thead>
<tr>
<th>element</th>
<th>a (Å²/eV)</th>
<th>b (eV)</th>
<th>Q(MD)(1) (Å²/eV)</th>
<th>Q(MD)(2) (Å²/eV)</th>
<th>Ion/energy</th>
<th>Q(exp) (Å²/eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ni</td>
<td>5.638</td>
<td>44400</td>
<td>5.1±0.4</td>
<td>4.7</td>
<td>600 keV Kr</td>
<td>4.8±0.5</td>
</tr>
<tr>
<td>Pt</td>
<td>16.60</td>
<td>5412000</td>
<td>9.5±0.8</td>
<td>12.6</td>
<td>400 keV Kr</td>
<td>9±1</td>
</tr>
<tr>
<td>Pd</td>
<td>4506</td>
<td>7.077×10⁹</td>
<td>14±1</td>
<td>16.9</td>
<td>1 MeV Kr</td>
<td>14±2</td>
</tr>
</tbody>
</table>

### Table 4

Values for the ion beam mixing for several metals [291].

<table>
<thead>
<tr>
<th>element</th>
<th>C</th>
<th>Al</th>
<th>Ti</th>
<th>Fe</th>
<th>Ni</th>
<th>Cu</th>
<th>Mo</th>
<th>Ru</th>
<th>Ag</th>
<th>Hf</th>
<th>Ta</th>
<th>W</th>
<th>Pt</th>
<th>Au</th>
</tr>
</thead>
<tbody>
<tr>
<td>(R^2/\phi_0) (Å²/eV)</td>
<td>14</td>
<td>112</td>
<td>36</td>
<td>27</td>
<td>39</td>
<td>150</td>
<td>28</td>
<td>44</td>
<td>450</td>
<td>90</td>
<td>54</td>
<td>72</td>
<td>115</td>
<td>730</td>
</tr>
</tbody>
</table>
The NRT predicted number of atomic displacements \(N_d(T_d)\) as a function of cascade energy, or the damage function, is given by:

\[
N_d(T_d) = \begin{cases} 
0 & T_d < E_d \\
1 & E_d < T_d < 2E_d / 0.8 \\
0.8T_d & 2E_d / 0.8 < T_d < \infty
\end{cases}
\]

(5)

The damage energy for a single ion or recoil is given by the total ion energy minus the energy lost to electronic interactions (ionization). Typical values of \(E_d\) for different materials range from 20 to 100 eV [40,298]. The dpa number can be obtained by using the damage energy in a given volume, calculating the number of defects in this volume using Eq. (5), and then normalizing with the average volume per atom.

The importance of the calculated dpa parameter is that it is the starting point for calculations of virtually all radiation effects in solids, and thus facilitates quantitative comparisons of different materials irradiated with the same kind of radiation, and also comparing the same material irradiated by different kinds of irradiation such as electron, ion, and neutron irradiation (as long as the main damage production mechanism is due to atomic recoils) [40,49,289,295,297–300]. Estimation of the damage is also important in modern materials processing by focused ion beams, or when irradiating nanomaterials [6,301]. Due to its widespread usage as an international standard for calculating displacement damage levels for the past 43 years, there is great value in retaining the NRT dpa as a standardized parameter.

However, the NRT displacement damage model contains several quantitative shortcomings. At low knock-on energies, near the displacement threshold, molecular dynamics simulations have reported that the defect production probability rises more gradually than the assumed step function in the KP and NRT displacement models, with a typical value near 0.5 at the nominal TDE [59,302] (Figs. 3 and 4). As previously noted in Section 2.1 and Fig. 3, similar behavior has also been reported in a detailed experimental study of the threshold displacement behavior in Cu [52]. In both the MD simulations and experimental study, the damage production typically does not reach a value of 1 until the damage energy is 2–3 times the nominal TDE. For polycrystalline materials, some defects are produced at energies below the crystallographic-averaged TDE; therefore, a damage function that transitions from zero to 1 defect production at a single TDE is not relevant to describe near-threshold displacement events in polycrystals. These observations suggest it is necessary to utilize a continuously varying probability function (rather than the step function at a single TDE followed by a plateau assumed in the KP and NRT models), in order to account for the crystallographic dependence and stochastic nature of near-threshold displacement events.

At higher energies, the existence of heat spikes leads to the major damage recombination effects and atomic mixing effects discussed in detail in Sections 2.2 and 3. These are not accounted for by the NRT equation. Hence Eq. (5) when used for most non-amorphizing metals on the one hand overestimates the number of stable defects by a factor of 3–4 [13,303–305], and on the other hand underestimates the amount of atomic mixing [291,304,306]. Even though the initial effect is on the nanometric scale, it has also been estimated that it can lead to macroscopic consequences such as a five-year underestimation of the lifetime of a nuclear reactor pressure vessel exposed to a very high thermal flux [307]. Similar trends have also been reported for intermetallic and high-entropy alloys [308,309] and ceramics [38,310,311]. This gives a strong motivation for developing improved dpa models.

### 4.2. Athermal recombination corrected (arc)-dpa model for damage production

The binary collision simulations used as the basis of the NRT-dpa model [49] focused on the collisional phase of the displacement cascade and did not consider the dynamics of cascade evolution as atomic velocities fell to the speed of sound (~5 eV) and lower. At this stage, the many-body interactions leading to heat spikes discussed in Section 2.2, become relevant. With increasing PKA energy, the displacement event produces progressively more Frenkel defects that are spatially close to other defects. The ~10–100 jumps occurring per atom during the 1–10 ps cascade cooling phase [13] can induce significant additional recombination events as the cascade atom energies decrease.

Fig. 15 summarizes the defect production as a function of PKA energy as determined from experiments performed in Cu near 4 K (where long range thermally activated defect motion is impossible [33]). The predicted defect production and number of replaced atoms obtained from MD simulations are also shown. The figure shows that the actual defect production is sublinear with respect to damage energy between ~0.1 and 10 keV, becoming about 1/3 of the NRT-dpa prediction. At energies > 10 keV corresponding to the onset of subcascade formation [13,111,312], the defect production increases linearly with damage energy but maintains the factor of ~3 lower defect production compared to the NRT-dpa value.

The physical basis for the reduction in surviving defects, with respect to the NRT model, with increasing knock-on atom energy can be understood by considering the following simplified derivation, first presented by us in Ref. [313]. As discussed in Section 2.2, the ultimate survival of initially created Frenkel defects requires physical separation of the interstitial and vacancy beyond a minimum distance known. In this derivation, we denote this the spontaneous recombination distance \(L\). At low energies (below the subcascade formation regime [13]) the displacement cascades are roughly spherical with radius \(R\), and form a liquid-like zone of dense collisions (the heat spike described above). It is further assumed that only interstitials transported to the cascade outer periphery defined by \(R - 1\) to \(R\) will result in stable defects, whereas Frenkel pairs created in the cascade interior (0 to \(R - L\)) will experience recombination (see Fig. 16). The fraction of initially

![Fig. 15. Experimental and simulation data showing quantitatively the problem with the NRT-dpa equation. In the figure, “exp” stands for experimental data, and MD for simulated molecular dynamics data. The other abbreviations denoted different interatomic potentials. The references are as in Ref. [313]. The figure shows that the NRT-dpa equation does not represent correctly either the actual damage (Frenkel pairs produced) nor the number of replaced atoms. The former is overestimated by roughly a factor of 3, and the latter underestimated by a factor of 30. From Ref. [313], reprinted with open access permission.](image-url)
created NRT-dpa defects that survive is therefore given by the ratio of the outer spherical shell volume to the total cascade volume:

$$\xi_{\text{survive}} = \frac{V_{\text{outer}} - V_{\text{inner}}}{V_{\text{outer}}} = \frac{4\pi R^3 - 4\pi(R-L)^3}{4\pi R^3} = 3 \frac{L}{R} - 3 \frac{L^2}{R^2} = 3 \frac{L}{R}$$ (6)

for \( L < R \). This “surviving defect fraction”, \( \xi_{\text{survive}} \), thus tells which fraction of defects predicted by the NRT-dpa model without any recombination survives. The cascade radius \( R \) can be, within the regime of spherical cascades, estimated from classical theory of nuclear stopping power [13,30,296]. Using the same code [313] we found that low-energy (less than or of the order to 10 keV) recoils of damage energy \( T_d \) have an average movement distance (range) \( R \) that is proportional to \( T_d^{3/4} \), where the exponent \( x \) is \( \sim 0.4 - 0.6 \) for the metals considered in this study. Since \( R \approx T_d^{3/4} \), this further gives:

$$N_d(T_d) = \frac{0.8T_d}{2E_d} \frac{3}{R^3} L \frac{0.8T_d}{2E_d} \frac{3}{R^3} L \times T_d^{-x}.$$ (7)

This simple model thus provides an intuitive explanation for why cascade damage production is sublinear with damage energy in the heat spike regime. Molecular dynamics simulation studies have reported that defect production rates up to the onset of subcascade formation in a variety of metals can be well described by \( N_d \sim T_d^{3/4} \), where \( x \) is between 0.2 and 0.3 [13,314]. These \( x \) values are smaller than the value obtained in our simplified model, because real cascades are not perfectly spherical and some defects form small clusters, reducing the recombination probability. However, since at high energies cascades break up into subcascades [41,87,315], it is reasonable to assume that damage production becomes linear with damage (nuclear deposited) energy. Hence the surviving defect fraction factor \( \xi(T_d) \) that accounts also for subcascade breakdown should have the feature of being a power law at low energies, but becoming a constant \( c \) at high ones. A function which fulfills both criteria is:

$$\xi(T_d) = A T_d^b + c$$ (8)

where \( b < 0 \) is consistent with the damage production efficiency reducing with increasing energy \( T_d \) and the desired limit \( \xi(T_d) \rightarrow c \) when \( T_d \rightarrow \infty \). This thus gives a total damage production:

$$N_d(T_d) = \frac{0.8T_d}{2E_d} \left( A T_d^b + c \right) = \frac{0.8A}{2E_d} T_d^{1+b} + \frac{0.8c}{2E_d}$$ (9)

Note that here the exponent \( b \) is not the same as \( x \), since the latter \( \xi \) function is not a pure power law. The prefactor \( A \) is defined by demanding the function to be continuous, i.e. \( \xi(2E_d/0.8) = 1 \).

Taken together, this derivation lead us to propose [313] a modified defect production model, the athermal recombination corrected displacements per atom.

$$N_d,\text{arc-dpa}(T_d) = \begin{cases} 0 & T_d < E_d \\ 1 & E_d < T_d < \frac{2E_d}{0.8} \\ \frac{0.8d}{2E_d} \xi_{\text{arc-dpa}}(T_d) & \frac{2E_d}{0.8} < T_d < \infty \end{cases}$$ (10)

with the new efficiency function \( \xi_{\text{arc-dpa}}(T_d) \) given by

$$\xi_{\text{arc-dpa}}(T_d) = \frac{1 - c_{\text{arc-dpa}}}{(2E_d/0.8)^2 T_d^{-b_{\text{arc-dpa}}} + c_{\text{arc-dpa}}}$$ (11)

Here \( E_d \) is the average threshold displacement energy [59] which is the same as in the NRT-dpa and \( b_{\text{arc-dpa}} \) and \( c_{\text{arc-dpa}} \) are material constants, that need to be determined for a given material from MD simulations or experiments. The overall form Eq. (5) and the constant 0.8 are retained for direct comparison with the NRT-dpa model; in particular making it easy to modify computer codes that now use the NRT-dpa by simply multiplying with the function \( \xi_{\text{arc-dpa}}(T_d) \).

Fig. 17 compares the derived arc-dpa expression for Fe and W with several recent MD simulation results used for the fitting. We tested that if the fit is limited to energies \( < 10 \text{ keV} \), one also can fit the data well with a power law with an exponent of \( \sim 0.7 - 0.8 \), i.e. the data is consistent with MD reports of power law dependencies. However, the arc-dpa form has the major improvement that it can also describe the saturation. Even though there is some variation in the MD data (due to differences in interatomic potentials), all of the MD results give damage production well below the \( \xi = 1 \) value predicted by the NRT-dpa model for cascade energies \( > 1 \text{ keV} \). The arc-dpa fit to the composite data gives a reasonable averaging description of the decreasing trend in \( \xi \) up to \(-10 \text{ keV} \) and the expected approach to a constant value at higher cascade energies. Values for the constants in the arc-dpa equation for some metals are given in Table 5. Very recently, the arc-dpa form has been used in an analysis of damage under fission reactor conditions, and shown to agree with experimental damage data measured under cryogenic temperatures [316].

It should be noted here, that while the NRT-dpa and arc-dpa models are functions of the damage energy of a projectile, the threshold displacement energy \( E_d \) in Eqs. (5) and (10) nevertheless refers to the total kinetic energy of a recoil. This apparent inconsistency likely stems from the fact that the original Kinchin-Pease model [295] was developed during a time when the inelastic energy losses experienced by energetic ions in matter were not well understood, and hence not accounted for. The NRT model [297] later combined the Kinchin-Pease model with a treatment of the electronic energy losses based on the work by Lindhard et al. [296], without noting the logical mismatch this introduced in the energy parameters. However, this formulation is well in line with the BCA model implemented in the computer code MARLOWE [49], used e.g. to determine the efficiency factor 0.8 in the NRT model. MARLOWE subtracts the inelastic energy losses at the time of collision, so that effectively the criterion within the computer code for defect creation at the threshold level becomes \( T_d = E_d - Q = E_1 + E_2 > E_d + E_d = 2E_d \), where \( E_d \) is the initial kinetic energy of the projectile and \( T_d \) is the corresponding damage energy. \( Q \) is the energy lost to electronic stopping, and \( E_1 \) and \( E_2 \) are the kinetic energies of the projectile and target atom recoil after the collision. The efficiency factor 0.8 must then necessarily be included in the ranges of the piecewise defined NRT model (Eqs. (5) and (10)), in order to yield a
4.3. Other models for damage production

As noted in Section 4.2, the buildup of damage in MD simulations has often been described with power laws also prior to the development of the arc-dpa model, with reported exponents below 1.0 [13,314,317,318]. A single power law clearly cannot describe the transition to a linear damage regime, however, a combination of power laws can. Several such descriptions have been reported. Some lead to a linear or close-to-linear dependence at high energies [13], some to a clearly different limiting behavior [319]. These models, have, however, not been formulated in a way that have a threshold or are consistent with the NRT-dpa equation described earlier.

Although proposed arc-dpa fitting curves described by Eqs. (8)–(11) offer several advantages for improved quantification of the primary damage state in irradiated over a wide range of damage energies, there are also several potential significant shortcomings of these fitting equations. First, the efficiency function for low energy (near-threshold, $T_d = 2E_d < 0.8$) is forced to be equal to one, whereas experimental and modeling studies of near-threshold collisions events indicate the efficiency may often be less than one as discussed in Section 2.1 and illustrated in Figs. 3 and 4.

Second, the transition from the single cascade to the multiple subcascade regime is not defined by the subcascade threshold energy, $E_{sc}$ ($E_{sc}$ can be determined from experiments [312,320–322] or computational simulations [13,323,324]; as a consequence, the approach to a constant efficiency regime that is observed in many materials may not occur until well above $E_{sc}$. Finally, the fitting function assumes that the damage efficiency approaches a constant value at high damage energies. This assumption appears to be valid for many materials, but supralinear defect production has been reported for tungsten at damage energies above 100 keV due to interconnected supersonic shockwaves [324].

An alternative fitting equation to describe the arc-dpa efficiency can be obtained by considering three regimes: near-threshold displacements regime in polycrystalline materials may not occur until well above $E_{sc}$. The lower bound for the near-threshold displacements regime in polycrystalline materials would extend down to the minimum energy that could cause a Na, $E_{d, min}$ $< 0.5E_{d, avg}$. The modified arc-dpa defect production model would be given by the following equations:

$$N_{d,arc-dpa}(T_d) = \begin{cases} 0, & T_d < E_{d, min} \\ N_{NRT} \cdot \xi_{arc-dpa}(T_d), & T_d > E_{d, min} \end{cases}$$

$$\xi_{arc-dpa}(T_d) = \begin{cases} A_1 \cdot (T_d - 2E_d)^{r_1}, & 2E_d < T_d \leq E_{sc} \\ A_2 \cdot (T_d)^{r_3}, & T_d > E_{sc} \end{cases}$$

Typical values for the power law coefficients in the three energy regimes are $r_1 \sim 0.5$, $r_2 \sim 0.7$, and $r_3 \sim 1.0$, respectively [13,314,324]. In the near-threshold region, the alternative arc-dpa fitting function provides a more physically accurate defect...
production based on a gradually changing statistical probability function with non-integer values.

5. Analytical model for radiation mixing

5.1. Replacement-per-atom (rpa) model

As noted in Section 3, since the NRT-dpa model deals with production of defects that are not on perfect lattice sites, it cannot predict the number of atoms that are transported from their initial lattice site to a new lattice site, i.e. replace another atom in a perfect crystal site (right panel in Fig. 18). Typically, the mixing is measured by measuring the broadening of a thin marker layer [291,325]. Via an analogy with random walk atom diffusion, it is possible to relate this measured broadening to the actual number of atom replacements per ion inside the material [28]. Analysis of neutron and ion beam radiation mixing data has shown that the actual number of replaced atoms can be more than an order of magnitude larger than the number of displacements predicted by the NRT-dpa model [292,293,306,326].

The superlinear increase in the number of replaced atoms with increasing knock-on atom energy can be estimated by a model considering the spatial extent of a collision cascade [313]. Consider first low energies (in the keV regime) and dense materials, where cascades are normally compact. As noted in Section 4.2 low-energy cascades in dense materials are roughly spherical. In the liquid-like cascade region all atoms are free to move and hence are almost certain to lead to one or more replacements during the thermal spike phase (as illustrated in Fig. 1, bottom frames). The number of atoms \( N \) in a spherical cascade of radius \( R \) is proportional to the sphere volume, i.e. \( N \propto R^3 \), and (as already noted for the arc-dpa model) \( R \propto T_d^{1/3} \). One thus finds that the number of replaced atoms \( N_{rpa} \propto T_d^{8/3} \). Since \( x > 1/3 \), this simple consideration gives an intuitive explanation for why the number of replaced atoms increases superlinearly with energy at low energies, when cascades are compact. At high energies, when cascades split into subcascades [88,315], the behavior can be expected to change to a linear dependence with damage energy. Similarly to the arc-dpa function, one thus obtains the functional form [313] for the mixing efficiency in relation to the NRT-dpa equation:

\[
\xi_{rpa}(T_d) \propto \frac{r_{cusp}^{\text{bc}}}{b_{rpa}^{\text{bc}} + T_d^{r_{cusp}^{\text{bc}}}}
\]  

(14)

By demanding continuity, \( 2E_d/0.8 = 1 \), one arrives at

\[
\xi_{rpa}(T_d) = \left( \frac{b_{rpa}^{\text{rpa}}}{(2E_d/0.8)^c_{rpa}} + 1 \right) \frac{T_d^{c_{rpa}}}{b_{rpa}^{\text{rpa}} + T_d^{c_{rpa}}}
\]  

(15)

Here \( b_{rpa}^{\text{rpa}} \) and \( c_{rpa}^{\text{rpa}} \) are the new material constants. Values for these constants for several metals are given in Table 5. Since the NRT equation already is proportional to \( T_d \), with this form the prediction is that the number of replaced atoms increases at low energies with \( T_d \) as \( N_{rpa} = T_d^{c_{rpa}} - T_d^{1+c_{rpa}} \), i.e. \( c_{rpa} = 3x - 1 \). At high energies, when \( T_d \gg b_{rpa} \), the form becomes linear with energy, as expected when cascades are split into subcascades. In this functional form, \( b_{rpa}^{\text{rpa}} \) has a physical meaning as the average energy for subcascade breakdown in terms of number of replaced atoms. Moreover, similar to the arc-dpa form, Eq. (15) fulfills the same conditions of continuity and compatibility with the NRT-dpa model. The rpa fits to Fe and W data are shown in Fig. 17.

6. Limitations of simulations and analytical models

6.1. Limitations of binary collision approximation models

Although very well suited to simulate ion ranges and channeling effects at keV and MeV energies [327–329], the binary collision approximation is very limited in its capabilities to predict damage production. It can give a somewhat reasonable description of the spatial distribution of vacancies and interstitials. However, due to the lack of many-body interactions, it cannot tell anything about the atomic structure of these defects. Moreover, the number of vacancies obtained in full cascade BCA simulations depend on the assumptions used in setting up the model. Some codes such as briam are even internally inconsistent in this regard [41,330]. Defect clustering can be simulated, however, doing this requires assumptions on the clustering distance that cannot be obtained without calibration data from MD [331] or other methods.

As discussed in Section 3, the binary collision approach can be used to model ballistic radiation mixing. In materials where heat spikes are not significant (low-density and low-mass ones) the BCA model can give a quite good description of the recoil atom

---

Fig. 18. Schematic illustration of the damage predicted by the three different damage models for the case of \( \approx 1 \) keV damage energy in a typical metal. For illustration purposes, the damage is illustrated as if all damage were produced in the same 2-dimensional plane. Blue circles illustrate atoms in original lattice positions, yellow-brown denotes atoms that are in a different lattice position after the damage event, red atom pairs denote two interstitial atoms sharing the same lattice site, and empty lattice positions denote vacancies. Left: Damage production predicted by the NRT-dpa model. Middle: actual damage production, addressed by the new arc-dpa equation. Right: actual atom replacements, addressed by the new rpa equation, agreeing better with experimental data on number of replaced atoms (ion beam mixing). Note that in real 3 dimensional systems, the difference is even larger than in this 2D schematic. From Ref. [313], reprinted with open access permission. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
displacement distances and hence mixing [332–335]. However, in materials where heat spikes are important, the BCA simulations clearly underestimate the mixing coefficients [336,337].

6.2. Limitations of molecular dynamics models of radiation damage

Provided that MD simulations are carried out with an adaptive time step [338,343] and realistic repulsive potentials [13,30,339–342], the approach can in principle describe the nuclear collisions in a cascade very accurately within the Born-Oppenheimer approximation (fully quantum mechanical scattering calculations have shown that for light projectiles and targets, dynamic change of the polarization state of the electron cloud can modify the trajectory from the Born-Oppenheimer ones [343,344]). Moreover, if one would have a perfectly realistic equilibrium potential as well, MD could also describe the atom motion in heat spikes exactly as they occur in reality.

In reality, there is of course no such thing as a “perfect” interatomic potential. The reliability of the interatomic potential is the key to how reliable the MD simulation result is. Discussing the reliability of the MD potentials systematically is beyond the scope of this Review. We only briefly note that MD simulations with appropriate choice of potentials have demonstrated good agreement with many irradiation quantities, such as range profiles [345,346], sputtering yields [220,347,348], crater shapes and sizes [349,350] and ion beam mixing [292]. On the other hand, such studies have also shown that not all interatomic potentials give a good agreement [348] and that agreement in one energy range does not guarantee agreement in another [347].

However, with the classical MD approach only, electronic excitations effects are completely neglected. The electronic excitations can be divided into two categories (although the limit between the two is not sharp): electronic stopping power and electron-phonon coupling. The former effect, i.e. the slowing down of energetic ions by electrons, has been studied very extensively both theoretically [296,351,352] and experimentally (for a review until 1985, see Refs. [13,30]). It can be very well included in both BCA and MD simulations as a frictional force, and the outcome can be tested by comparing experimental and simulated range profiles [345,346,353–355]. Electronic stopping has been routinely included in MD simulations since the 1990s [87,356–360] (although studies of cascades in Fe for a long time tended not to include it [13,88,361,362]).

The issue of how to include the e-ph coupling is less clear. In general, it means the effect where the atomic ("phonon") and electronic subsystems are out of equilibrium with each other, i.e. have different temperatures. Such a situation is very well known in plasma physics, and one of the key reasons why plasmas are challenging to model. It is also routinely studied in low-temperature physics [363]. In collision cascades, what happens is that the initial high-energy collisions (which are completely outside thermal equilibrium) initially heat up the ionic subsystem on sub-picosecond timescales [28,364]. At the same time, electronic stopping heats up the electronic subsystem. However, these heating mechanisms are not coupled, so the two systems most likely end up at different temperatures. Due to e-ph coupling the two subsystems interact with each other towards equilibrating the two temperatures. In principle, this can be a very important mechanism in metals, since in these electronic heat conductivity is much faster than the ionic one, and thus after cascade energy is transferred to the electrons, the system could cool down very quickly.

The e-ph coupling must also somehow be related to the low-energy limit of electronic stopping power. This statement can be made firmly, since if one implements the traditional electronic stopping powers that are proportional to ion velocity down to zero velocity [296] in MD, any system quenches very quickly down to 0 K. In reality, the electronic stopping must cease to be active when the electronic and ionic subsystems have reached equilibrium, however, it is not obvious how this statement could be quantified. There are many approaches for including e-ph coupling in MD simulations, and many of these also deal with the electronic stopping power [356,358,365,366]. The early ones suffered from having free parameters that could not be directly calibrated with any experimental or simulation data. Test simulations, treating the coupling constant as a free parameter, have shown that the choice of the coupling model and low-energy limit of electronic stopping can affect both damage production and mixing values [98,358]. This makes solving the puzzle of low-energy electronic excitations crucial to achieve a predictive capability of the primary damage production.

In magnetic materials, for example iron, iron alloys and steels, in addition to e-ph coupling it is also necessary to treat energy transfer to magnons. Magnons, and the dynamics of magnetic excitations in general, also ought to be included when simulating the formation of defects in high-energy collision cascade events, since the relaxation timescales for electrons, phonons and magnons differ [36–39]. Radiation defects formed in a magnetic material have a clear magnetic signature [386], e.g. the formation of a self-interstitial defect in iron reduces the magnitude of magnetic moments in the vicinity of a defect whereas the size of magnetic moments increases in the vicinity of a vacancy. On average, the production of defects in iron appears to increase the average magnetic moment of atoms [369,370], also an unusual, fundamentally electronic, effect associated with radiation damage. These experimental observation agree with simulations predicting high rates of production of vacancy clusters in collision cascades occurring near a surface [208].

Recently, time-dependent DFT methods have started to become efficient enough to be able to calculate electronic stopping explicitly [371–376]. Although these methods are not fully reliable by themselves, they are being rapidly developed for other applications in physics and chemistry. Hence it is likely the approaches will soon give key insights into the exact nature of low-energy stopping power.

Finally, one has to naturally keep in mind that the fundamental key limitation of MD is the time scale it can handle. Since the basic time step is of the order of femtoseconds and cannot be increased from this, and MD iterates over the time step, any conventional MD approach is limited in the time it can handle. Several ways to speed up MD exist [377–384], but these generally apply only to equilibrium processes. On the other hand, for the topic of this Review, primary damage production, the MD time scale is not a serious limitation. Any primary damage production process is over in a few hundred ps at most, and MD generally can handle such time scales.

6.3. Limitations of any dpa model

The analytical dpa models described in Sections 4 and 5 are useful for quantifying the direct (ballistic) response of a material to incident energetic particles. Whereas the original NRT-dpa model is useful as a measure for quantifying and comparing radiation dose (exposure) in different environments, the refined models, arc-dpa and rpa, extend the usefulness of the original model by providing accurate estimates of the number of defects, or the level of radiation-induced mixing, respectively, and are also comparable between different irradiation environments.

However, there are several aspects of the primary radiation damage formation processes that these models do not capture. These include, on the one hand the heat spike behavior in energetic
cascades described in Section 2.2 which results in clustered damage formation, which is both qualitatively and quantitatively distinct from the sites of individual recoiling atoms in the ballistic, or NRT, picture. The other effect is that of damage accumulation, where the effects of irradiating a material saturated with defects will be markedly different from that of irradiating a pristine material. In this Section, we discuss the first of these effects; the latter will be reviewed in the following Section.

The heat spike phase of cascade evolution has two major consequences. One - the athermal recombination of a large fraction of the defects - is partly addressed in the arc-dpa efficiency factor. However, a second effect, the in-cascade formation of large defect clusters, cannot be captured by a model describing only the count of point defects. The morphology, i.e. the sizes, shapes and spatial distributions of defects, is not accounted for, yet it is precisely these properties that affect the subsequent thermal evolution of the irradiated material, and hence the long-term material response.

On the other hand, in the low-energy limit, with PKA energies approaching the TDE, the balance between damage formation and recombination is also more complex than what is captured in current BCA-based models. The TDE is non-isotropic, and defects are created with a certain non-monotonically increasing probability above the TDE [59]. The NRT-dpa, arc-dpa and rpa models are all for simplicity constrained to be exactly equal to 1 in the range $E_d < T_d < 2E_d/0.8$, although this is not strictly true for most materials (cf. Section 2.1). In metals especially, recombination results in a lower average value than 1 even at $T_d = 2E_d/0.8$. In Au, for example, the defect production around threshold values is over-estimated by a factor of more than 2 in the dpa models (see Fig. 19). The alternative form proposed in Eq. (13) may be able to account for this difference for the damage production.

The number of replaced atoms, on the other hand, can be significantly higher than unity for threshold energies, for example when the point defects are created as the result of a replacement collision sequence. In certain cases, the number of displaced atoms can be nonzero even when no stable Frenkel pair is created. Hence the rpa model underestimates this quantity around threshold energies, in some cases by a factor of about 5 compared to MD predictions.

The formation of large clusters also gives rise to a significant effect on the total number of point defects that survive a cascade, which for some materials results in a deviation in defect numbers from that predicted by both NRT-dpa and arc-dpa models. As mentioned in Section 2.2, the efficiency $\xi$ has been shown to saturate to a value of 0.2 − 0.5 for many metals. However, this observation does not hold for heavy metals such as Au, Pt and W as the PKA energy is increased beyond the energies covered in most available studies. In these materials, the large self-interstitial atom (SIA) clusters, that form athermally during the cascade evolution, prevent recombination, resulting in a super-linear increase in defect numbers as the PKA energy increases. These materials have a relatively high subcascade splitting threshold, at around 150 − 200 keV [105] (compared to 20 − 30 keV for Fe [385]). It is expected that the behavior above the subcascade splitting threshold will transition into a linear regime, as has been shown to happen e.g. for Fe (see Fig. 6). However, we are not aware of any direct demonstration of this for these heavy metals in the existing literature. The increase in the defect production efficiency (wrt. the NRT prediction) for the metals Au, Pt and W is illustrated in Fig. 19.

As noted in Sections 2.2 and 4.3, different interatomic potentials and simulation methods show significant disagreement concerning the formation of large defect clusters. This leads to a disagreement in the predictions of total numbers of defects in these materials for high PKA energies, as the defect cluster formation and recombination efficiency are directly linked. This can be seen in Fig. 19 for the different potentials for Au and W. The more general forms given in Eqs. (12) and (13) can accommodate for a super-linear increase in an analytical form. However, it is clearly important to reduce the uncertainty of the interatomic potential predictions in this high energy regime, or obtain new experimental data sets, before a reliable analytical description of the superlinear high-energy regime can be done.

A further aspect of the primary damage that cannot be captured in dpa models is the fact that the defects are not distributed homogeneously in the material. Rather, defects are created densely in the cascade region [386]. This affects the rates of thermal recombination and clustering [387], giving rise to a logical misalignment between MD-predicted primary damage and, e.g., rate theory models utilizing dpa measures as input. The spatial distribution of defects has been investigated to some extent, but few quantitative models have been proposed. Direct comparison with TEM experiments in W show some disagreement [388], with one possible reason being related to defect cluster sizes. MD simulations suggest that larger clusters may tend to form initially closer together than smaller clusters and point defects. However, limited statistics both from MD and from experiments make it impossible to probe this effect further based on currently available data.

Finally, we note that all the dpa models describe the primary damage state only, i.e. the state of the material some tens of picoseconds after a collision cascade has been initiated. On longer time scales, thermally activated defect migration almost invariably contributes to altering the microstructure (last frame in Fig. 5). These processes are active on much longer time scales compared to the primary damage formation, yet in many cases may significantly alter the experimentally measured values of the radiation damage. Recombination can significantly reduce the defect numbers from the arc-dpa value, while thermal diffusion may also enhance atomic mixing compared to the rpa value.

6.4. High-fluence effects and dpa concepts

Most of the discussion above has been concentrated on the damage production mechanisms in individual cascades. When the radiation fluence/dose increases, the damage produced by cascades will start to overlap, and several nontrivial effects may be the consequence of this, and in this respect metals and semiconductors show major differences. As noted in Section 2.3, elemental and
most compound semiconductors amorphize completely or partially when irradiated to high fluences [7,389]. By contrast, elemental pure bulk metals have never been made amorphous by any method, and irradiation is no exception. When pure metals, like Cu, are irradiated to high fluences, the damage production efficiency decreases and the damage level eventually saturates. In Cu this saturation level has been experimentally found to be at about 0.004 displaced atoms [390,391]. This, of course, means that at higher fluences a “dpa” value becomes completely disconnected from the actual number of defects, i.e. at a nominal dpa value of, say, 10 or 100, the defect concentration still remains at (roughly) 0.001–0.01.

The origin of the saturation can be fairly well understood from MD simulations. Simulations of collision cascades made on simulation cells with pre-existing point defects showed that when the concentrations of pre-existing defects exceeding about 1%, the final defect concentration decreased due to the irradiation [190]. A similar study on cascades overlapping the damage created by other methods, like (Discrete) Dislocation Dynamics and Monte Carlo simulations, can be and have also been used to close the gap in time scale and size scale, with the input and parameters from MD (e.g. Refs. [395–399]).

In spite of the large difference in dose rates, the simulations of massively overlapping cascades may give some qualitative insights on high damage levels. The simulation are very much higher than experimental ones under typical ion and neutron irradiation conditions, even though the local dpa levels are similar to experimental doses. The results obtained from these kind of simulations are thus directly comparable only with experiments, where the thermal migration of defects are low. The slow migration can be due to cryogenic temperatures or due to the defects becoming immobile because of their type or arrangement. Other methods, like (Discrete) Dislocation Dynamics and Monte Carlo simulations, can be and have also been used to close the gap in time scale and size scale, with the input and parameters from MD (e.g. Refs. [395–399]).

The primary damage state is of course not the end result of damage production in materials by irradiation. As mentioned, the MD simulations are very restricted in both time scale and size scale, which means that to directly compare the results, we are in need of some other method(s). Therefore, a lot of effort has been placed on multiscale modeling, that utilizes the data from MD as input, and extends it either in time scale or size scale, or ultimately in both.

Even though the focus of the current review lies on the primary damage, we provide here a brief overview of the physics and simulation tools that can be used to address materials modification after the primary event. Other review articles will address these issues in much greater detail; for a recent review see e.g. Refs. [243,400–405 and 406].

After the primary damage state is produced, in almost all cases the defects produced become mobile. Mobile defects can annihilate, build up larger defect clusters or add to the size of amorphous zones, make amorphous zones shrink in size, vanish at surfaces or grain boundaries, make dislocations, add to dislocation growth, contribute to dislocation climb, etc. The dislocations in turn can be mobile, interact with each other, multiply, etc. Just some aspects of this huge complexity is illustrated in Fig. 20. These effects are partly well understood (such as point defect migration in metals), partly quite poorly (like dislocation reactions).

The analytical models (such as the arc-dpa and rpa) might be soon extended to be able to describe cluster size distributions, but we do not see a way to extend them to long time scales and high damage doses, since this development has a complex temperature and damage overlap dependence. For a limited temperature range it may be doable to develop analytical models once a good understanding of all physics involved and large experimental or simulation data base is available.

The way forward to tackle the full complexity of damage buildup must be multiscale modeling, i.e. using several different layers of simulation tools in conjunction with each other. Some of these tools and a very rough description of which length and time scales they relate to, are presented in Fig. 21b and some physical phenomena related to the corresponding length and timescales in Fig. 21a. Very recently an approach was introduced which may make it easier to bridge the gap between atomistic and macroscopic modeling. In Ref. [407] a formalism was developed by which the atom–level mechanical property change can be directly transferred to finite element modeling. It is beyond the scope of this Review to discuss all of these methods in detail. Here we just note that although there are major challenges in developing these methods further, there is to the best of our knowledge no physics phenomenon that they
could not in any possible way handle. Some of the processes almost certainly are such that they cannot be described by any analytical equation. However, even in such case big data and artificial intelligence approaches may be used to describe them efficiently computationally. Thus with further advances in computer capacity, new algorithms and artificial intelligence, multiscale modeling should one day be able to provide predictive modeling of radiation effects starting from first principles over all time and length scales.

Acknowledgements

We thank the OECD/NEA for setting up the initial working group that initiated this work. This work was sponsored in part by the U.S. Department of Energy, Office of Fusion Energy Sciences, SJZ, RES, Office of Basic Energy Sciences, (grant # DEFG02-05ER46217), RSA, WJW. AS acknowledges support from the Academy of Finland through project No. 311472. SLD acknowledges support from the RCUK Energy Programme [grant number EP/P012450/1]. Grants of computer time from CSC - the Finnish IT Center for Science as well as the Finnish Grid and Cloud Infrastructure (persistent identifier urn:nbn:fi:research-infras-2016072533) are gratefully acknowledged. This work has been also partly carried out within the framework of the EUROfusion Consortium and has also received partial funding from the Euratom research and training programme 2014-2018 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jnucmat.2018.10.027.

References


A.V. Krasheninnikov, F. Banhart, Engineering of nanostructured carbon materials with electron or ion beams, Nat. Mater. 6 (10) (2007) 723.


M. Baldwin, R.P. Doerner, Helium backscattering and alpha particle loss on tungsten under fusion relevant plasma conditions, Nucl. Fusion 48 (2008), 035001.


B. D. Cai, C.M. Snell, K.M. Beardmore, N. Grønbech-Jensen, Simulation of defect production in high-energy collision cascades, a Declassified version of a report by W. Eckstein, in Ref. [408], eq. (4.4.3) on p. 55.


