IAEA Challenge on Materials for Fusion: Winning Submission

The IAEA Challenge on Materials for Fusion presented participants with a set of data files from a simulation of neutron damage in two materials of interest for use in a controlled nuclear fusion reactor: iron and tungsten. The data files consisted of the locations of the atoms after the "collisional cascade" of displacements caused by the "primary knock-on atom": the first atom to interact with the neutron.

Participants were invited to come up with innovative ways to visualize, analyze and explore the provided data. The winning submission was submitted by the group of Udo von Toussaint, J. Dominguez, M. Rampp and M. Compostella at the Max Planck Institut für Plasmaphysik, Garching Germany.

Winning Submission

Descriptor vectors

The approach taken in the winners' submission was to derive a 51-dimensional *descriptor vector* for each atom in the simulated crystal structure by expanding its local environment in a series of spherical harmonic and averaging this expansion over all rotations.

The descriptor vector acts as a "fingerprint" for the atomic environment which is not dependent on orientation and is robust to minor distortions of the crystal lattice due to the normal thermal motions that it undergoes.

In the below figure, clear differences in the descriptor vectors can be seen between the environment of a tungsten atom in an undamaged ("bcc") crystal (black) and that of an "interstitial" atom (red): one that has ended up being forced between existing lattice atom positions.

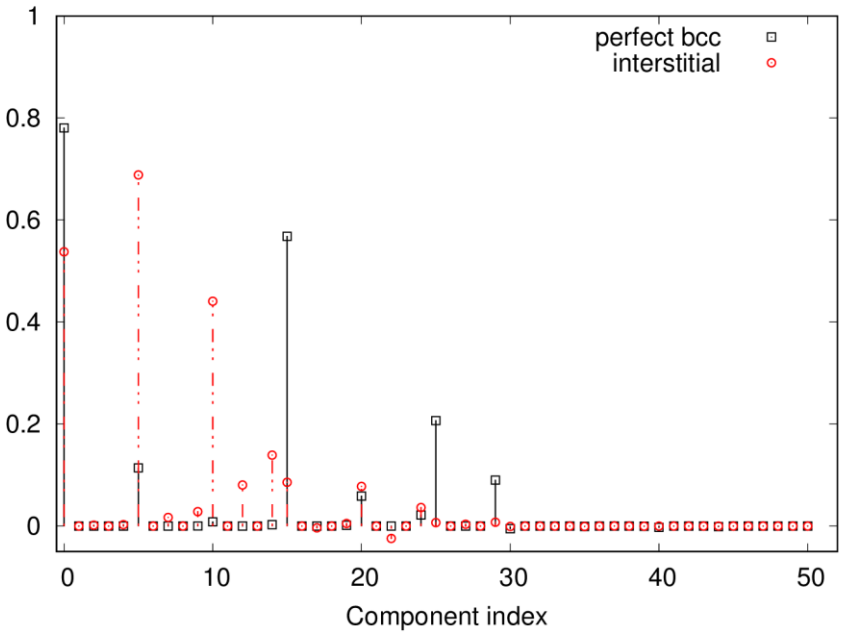

Figure 1. The components of the descriptor vectors for an atom in a perfect bcc crystal (black) and an interstitial atom in a bcc lattice (red).

Once a descriptor vector has been calculated for each atom, its "distance" from the reference vectors describing known environments can be calculated. The vast majority will be close to the

environment of an atom in a perfect crystal, but a few will be identified as being in a dissimilar environment: their "distance" from the perfect lattice environment is greater than some threshold value (0.15 in Figure 2 below; note the logarithmic y-axis). These can be assigned as "defects" and subject to further analysis.

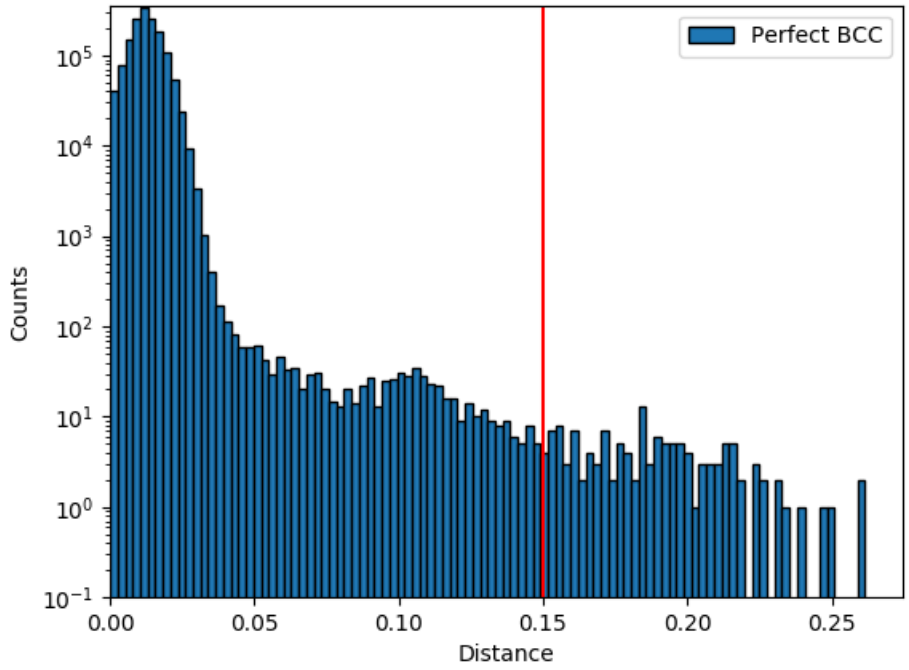

Figure 2. A histogram of distance of all atoms' descriptor vectors from the descriptor vector for an atom in a perfect bcc crystal (note the logarithmic scale). Atoms with a distance greater than 0.15 are identified as belonging to defects.

Identification of new defect structures

The process described above can be repeated for the distance between the descriptor vectors of candidate atoms and those of atoms in known defect environments (such as the interstitial vector "fingerprint" illustrated above). When known environments have been identified, the atoms remaining are further analyzed by applying the technique of Principal Component Analysis (PCA) to reduce the dimensionality of their descriptor vectors from 51 to two. The result can be readily visualized and, in the case of the data provided for the Challenge, showed the obvious clustering shown below (Figure 3)

In real space, these atoms were found to be present in a previously unidentified defect environment consisting of a triangular arrangement of three vacancies.

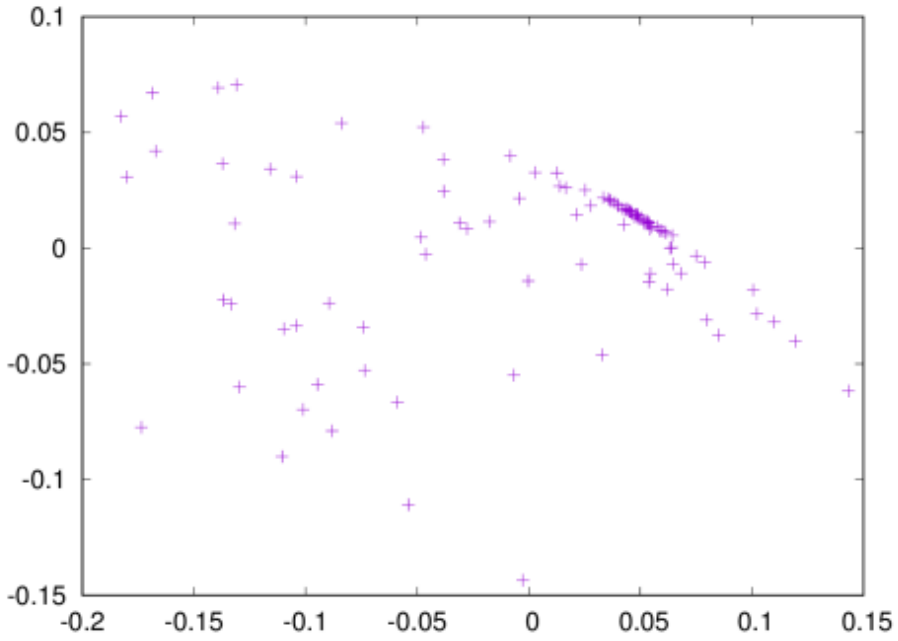

Figure 3. After dimensionality reduction by principal component analysis, a cluster of vectors corresponding to a previously-unknown defect type was identified.

Further visualization

A further processing algorithm was used to identify empty volumes in the simulated crystal, corresponding to missing atoms; this requires an efficient nearest-neighbor distance calculation. In visualizing the damaged crystal, voids such as these are given a blurry, cloud-like effect, other defects are coloured according to their type as identified by the algorithm described above. Atoms that are not in distorted environments are not shown at all. This makes the damage to the crystalline structure easy to visualize and quantify (Figure 4).

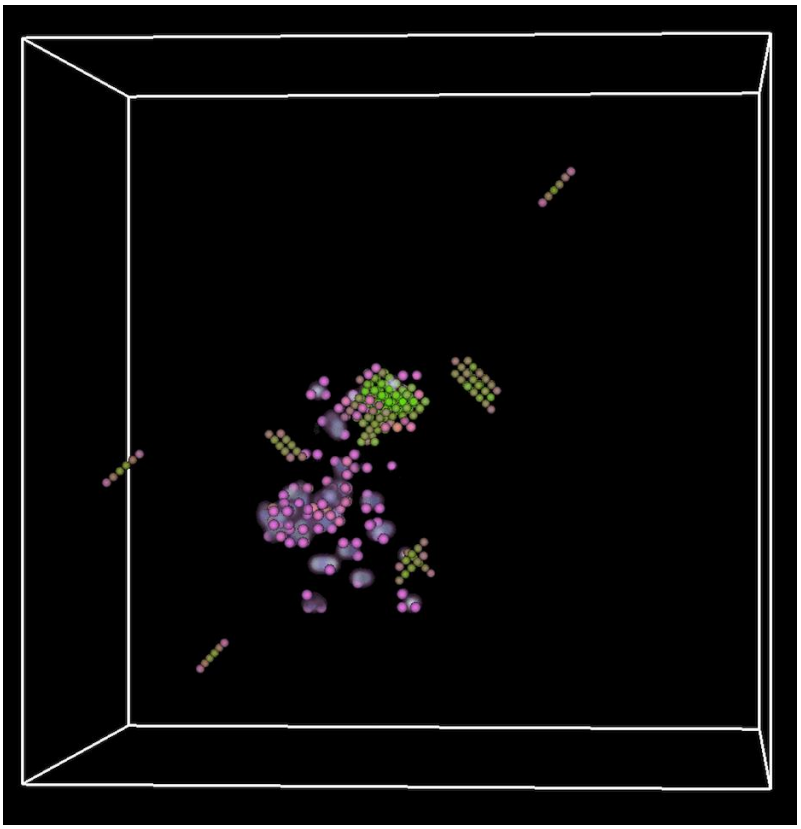

Figure 4. Visualization of the crystal defects is facilitated by presenting only atoms in distorted environments, coloured according to their defect type, and rendering vacancies (atoms missing from their lattice sites) as blurred regions.

Summary

The presented approach formulates a way of describing the local environment of each atom in the crystal as a 51-dimensional "descriptor vector" and then of measuring the similarity of these vectors to those describing an atom in (a) an undamaged crystal and (b) the environment of a known defect-type. At atom with a descriptor vector which is not similar to either of these types is possibly in a new type of defect. For such atoms, a process called principal component analysis (PCA) is performed. This reduces the multidimensionality down from 51 to 2, which means that the vectors can be plotted on a regular two-dimensional graph. In the data provided for this Challenge, the graph showed a clear clustering, corresponding to a previously unidentified defect type; in the real space of the crystal structure this corresponds to a triangular arrangement of vacancies (atoms missing from their lattice positions in the undamaged crystal).

Software used and references

1. QUIPPY and GAP (A.P. Bartok, G. Csanyi, arXiv: 1502.01366 and ref. therein)
2. Voro++ (Wigner-Seitz cell analysis of atomic volume)
3. KDTREE2 for nearest-neighbour calculations (https://github.com/jmhodges/kdtree2/ by M. B. Kennel and customised by the Challenge winners)
4. VisIt (https://visit.llnl.gov/) and custom scripts for visualisation

5. OVITO Open Visualization Tool (https://ovito.org): A. Stukowski, *Modelling Simul. Mater. Sci. Eng.* 18 (2010), 015012