On the use of a cluster identification method and a statistical approach for analyzing atom probe tomography data for GP zones in Al-Zn-Mg(-Cu) alloys

Sohail Shah, Elisabeth Thronsen, Frédéric De Geuser, Constantinos Hatzoglou, Calin Marioara, Randi Holmestad, Bjørn Holmedal

On the use of a cluster identification method and a statistical approach for analyzing atom probe tomography data for GP zones in Al-Zn-Mg(-Cu) alloys

Sohail Shah[a*], Elisabeth Thronsen [b], Frederic De Geuser[c], Constantinos Hatzoglou[a], Calin D. Marioara[d], Randi Holmestad[b], Bjørn Holmedal[a]

[a]*Department of Materials Science and Engineering, Norwegian University of Science and Technology (NTNU), N-7491 Trondheim, Norway*

[b]*Department of Physics, Norwegian University of Science and Technology (NTNU), N-7491 Trondheim, Norway*

[c]*University Grenoble Alpes, CNRS, Grenoble INP, SIMaP, Grenoble F-38000, France*

[d]*SINTEF Industry, N-7465, Trondheim, Norway*

Abstract

Early-stage clustering in two Al-Mg-Zn(-Cu) alloys has been investigated using atom probe tomography and transmission electron microscopy. Cluster identification by the iso-position method and a statistical approach based on the pair correlation function have both been applied to estimate the cluster size, composition and volume fraction from atom probe datasets. To assess the accuracy of the quantification of synthesized distributions of clusters of different mean size, a synthesized virtual dataset was used accounting for a simulated degraded spatial resolution. The quality of the predictions made by the two complementary methods are discussed, considering the measured and virtual datasets.

Keywords

Aluminum alloys, Atom probe tomography (APT), Transmission electron microscopy (TEM), clusters, Pair correlation function (PCF), Natural ageing (NA)

1 Introduction

Atom probe tomography (APT) has a unique combination of three dimensional (3D) spatial resolution down to the nanometer scale along with high chemical sensitivity [1, 2]. This enables the study of small precipitates or clusters in metals. Quantification of cluster size and composition at different stages of precipitation is essential and remains at the forefront of research in understanding and designing the next generation of alloys.

Several approaches for cluster identification have been developed to define and characterize clustering and precipitation over the years [1, 3-13]. In most cases, they consist of three data mining steps; (i) the clustering algorithm identifies solute clusters in the data by considering solute segregation[14], (ii) an enveloping algorithm searches for atoms of other elements that are within each solute cluster and (iii) an eroding algorithm erodes the matrix-particle interface that may have formed as a result of the enveloping step.

These methodologies generally demand the user to manually determine multiple parameters, leading to inconsistent results and challenges in reproducing them. Also, to correctly identify certain clustering phenomena, a visual inspection of the results is often used to assess the

suitability of the method. This too is strongly user dependent and can give inconsistent results. Usually, a well-defined parameter estimation procedure is used as seen for the case of the maximum separation method (MSM)[15] to improve reproducibility and consistency in the results. Studying clustering and precipitation through atom probe is rendered more complicated due to the local magnification effects due to differences in evaporation fields between different phases (matrix and particles) [16]. This effect is most severe for high evaporation field particles [5]. The advantage of cluster identification is that it enables the extraction of local information by providing information on each identified cluster (composition, size).

Statistical approaches such as radial distribution functions [17, 18], on the other hand, aim at describing the distribution of solute atoms as a whole and do not consider each cluster separately. This gives element specific averaged information on the cluster state of the matrix, both in terms of amplitude and correlation length. This enables a direct comparison of the degree of clustering between different datasets without the need of user-defined parameter.

Recently, Zhao *et al.* [19] coupled this parameter free statistical approach to an interpretation model enabling a fit of the RDFs in order to extract the composition of second phase particles and estimate volume fraction, number density and particle size. The methodology has been previously applied to Al-Cu-(Li-Mg) [20] alloys, Al-Mg alloy system [21] and the Al-Zn-Mg [19] system. This enables a statistical significance of the results and reproducibility, while giving the average values from the dataset analyzed.

Here, we combine two different approaches: a local cluster identification method, i.e. the iso-position method (IPM) and a statistical methodology based on RDF or rather on pair correlation functions (PCF) [12, 19]. On the basis of these two methodologies applied to atom probe datasets of two Al-Zn-Mg-(Cu) alloys, as well as with the help of simulated datasets, we aim at better assessing the robustness of the experimental parameters obtained by APT on very small clusters.

2 Experimental Section

6 mm thick plates from two commercial alloys, received from Benteler automotives are used in this study. The major difference between the two alloys is that one contains copper and is slightly richer in solute content, while the other is a copper-free alloy. The accurate composition measurements of the two alloys are given in Table 4.

Table 1. Alloy composition of the investigated alloys [22].

		Cu	Fe	Mg	Si	Ti	Zn	Zr	Al
7003	wt.%	0.01	0.22	0.73	0.09	0.02	5.68	0.15	Bal.
	at.%	0.00	0.11	0.84	0.09	0.01	2.43	0.05	
7046	wt.%	0.28	0.2	1.31	0.08	0.03	6.47	0.15	Bal.
	at.%	0.12	0.1	1.52	0.08	0.00	2.79	0.05	

Small blocks of samples sized 3 cm × 3 cm were cut and solution heat treated (SHT) at 480ºC for half an hour in a salt bath followed by a quick water quench. The samples were grounded and polished for micro-hardness measurements using a 1000 gf Vickers indenter. The hardness values reported are an average of 5 individual measurements for each condition. The alloys were allowed to naturally age (NA) for 3 months at room temperature, after which they were used for APT and TEM characterizations.

All samples for APT were prepared by following a standard two-step electropolishing process [12]. A minor contamination of copper (Cu) was seen as a capping layer at the tip of the APT needle similar as reported by Famelton *et al.* [23]. This could be due to the Cu present within the sample itself or the Cu tube holding the APT needle which contaminates the electrolytic solution. The reconstructed volume containing this Cu enriched region was removed from the subsequent data analysis.

APT experiments were carried out on the Local Electrode Atom Probe (LEAP) 5000XS by Cameca instruments. It is a straight flight path instrument with a detection efficiency of 80%. For all analysis, the base temperature was set to 30K and the laser energy is adjusted to obtain an equivalent pulse fraction of 20% of the DC voltage to avoid preferential evaporation (Ref) (i.e. around 100-130 pJ). A detection rate of 0.5% and a pulse frequency of 250 kHz were used for all samples analyzed. Datasets containing 25-35 million ions were collected for each alloy. Reconstructions were made using the structural information according to Gault et al. [24] by using the Integrated Visualization and Analysis Software (IVAS) by Cameca Instruments Inc. The Norwegian Atom Probe App software [25] developed by C. Hatzoglou has been additionally used for post processing of the experimental as well as for the simulated data.

The cluster identification method used, i.e., the IPM method, has been developed at the Groupe de Physique des Materiaux (University of Rouen, France) and is based on the chemical concentration (Mg + Zn) and atomic distance criteria. A detailed explanation of the methodology and its criteria estimation can be found elsewhere [5, 12].

In the 7xxx Al alloy system, the clusters exhibit higher atomic density as compared to that of the matrix, and with morphological distortion. This is mainly due to the local magnification effects typically observed due to the difference in evaporation fields between cluster/matrix, previously reported [26, 27] and also seen in the same alloys in a previous study [22, 28]. However, the cluster dimension along the evaporation direction (Z direction) is less biased by local magnification effects [16, 29], since it is the least biased direction in the APT. A more precise estimate of the size of clusters, as identified by iso-position method, can be based on a directional estimate in the Z-direction, the so-called Extent_z parameter [13], which for one cluster is defined as

$$E_z = |Z_{max} - Z_{min}| \quad (1)$$

Here Z_{max} and Z_{min} are the outermost atoms in the z direction of the identified cluster. Hence, E_z is the maximum distance between outermost cluster atoms in this direction. The E_z is very similar to the way size measurements from TEM results often are obtained.

Another standard way of estimating size for spherical particles in APT is using the radius of gyration (or diameter), as obtained from a cluster identification methodology. The radius of

gyration, is the root mean square of the distance between the center of mass and the coordinates of the constituent atoms classified as a cluster[12, 30]. Other methodologies like the best fit ellipsoid [31] can also be used to quantify size measurements for non-spherical particles, but are not required in this study. Sizes reported in this study are all based on diameters extracted from the Guinier radius and from E_z. The Guinier diameter is simply twice the Guinier radius calculated from the cluster identification. The size and compositional errors are calculated based on the standard deviation. The error in the number density is calculated by dividing the number density by the square root of the total number of clusters.

The PCF method is based on the RDF calculated for each solute (Zn, Mg and Cu) as the central atom. The RDF $n_{a\text{-}b}(r)$ is defined as the number of atoms of species 'b' at a distance 'r' of an atom 'a'. The RDF enables the computation of the average composition $C_{a\text{-}b}$ (r) of element 'b' at a distance of 'r' of the element given as:

$$C_{a-b}(r) = \frac{n_{a-b}(r)}{\sum_b n_{a-b}(r)} \tag{2}$$

If $\overline{C_a}$ is the bulk composition of element 'a', then in an isotropic system, $\overline{C_a}.C_{a-b}(r)$ is the correlation of the spatial composition of element 'a' and element 'b' given as:

$$\overline{C_a} \cdot C_{a-b}(r) = < C_a(\overrightarrow{r_1}) \cdot C_b(\overrightarrow{r_2}) > \tag{3}$$

Where $r = |\overrightarrow{r_2} - \overrightarrow{r_1}|$. When 'r' approaches large values (larger distances), the convolution approaches $\overline{C_a}.\overline{C_b}$. The pair correlation function can then be introduced by subtracting this final value:

$$\gamma_{a-b}(r) = \overline{C_a} \cdot C_{a-b}(r) - \overline{C_a} \cdot \overline{C_b} \tag{4}$$

The $\gamma_{a-b}(r) = 0$ occurs when there is no correlation, and this is true for large 'r'. The PCF can be rewritten as

$$\gamma_{a-b}(r) = < \Delta C_a(\overrightarrow{r_1}) \cdot \Delta C_b(\overrightarrow{r_2}) > \tag{5}$$

This highlights the fluctuation in composition ($\Delta C_a = C_a(\overrightarrow{r_1}) - \overline{C_a}$) which is the difference from the average, is relevant to the PCF. A normalized correlation function can further be introduced:

$$\gamma_{a-b}(r) = \Delta C_a \cdot \Delta C_b \cdot \gamma^0_{a-b}(r) \tag{6}$$

Where $\gamma^0_{a-b}(0) = 1$ and $\gamma^0_{a-b} = 0$ for large 'r'. If we consider only a single type of ion, we obtain the autocorrelation:

$$\gamma_{a-a}(r) = \Delta C_a^2 \cdot \gamma^0_{a-a}(r) \tag{7}$$

As mentioned by Zhao *et al.* [19], the mean squared composition fluctuation is immune to the local magnification effects due to difference in the electric field required to induce the evaporation phenomenon of the different phases (but not to spatial resolution effects).

If we further assume a two-phase system with a homogenous composition in element 'a', with precipitate composition C_p^a in a matrix of composition C_m^a and that both share the same atomic volume, the volume fraction of the precipitates can be written as:

$$f_v = \frac{\overline{C_a} - C_m^a}{C_p^a - C_m^a} \tag{8}$$

The average of the product of compositional fluctuations can be then calculated as:

$$\overline{\Delta C_a \cdot \Delta C_b} = f_v(1 - f_v)(C_p^a - C_m^a)(C_p^b - C_m^b) \tag{9}$$

For a single type of ions, this can be simplified as:

$$\overline{\Delta C_a^2} = f_v(1 - f_v)\left(C_p^a - C_m^a\right)^2 = (C_p^a - C_a)(C_a - C_m^a) \tag{10}$$

The value of the PCF at origin ('r'=0) in combination with the matrix composition [32] help extract the precipitate composition and volume fraction without being affected by local magnification effects. The normalized correlation function for a sphere of radius 'R' can also be calculated assuming a homogenous density with a sharp interface with the matrix as [33, 34]:

$$\gamma_0^{sphere}(r, R) = \begin{cases} 1 - \frac{3r}{4R} + \frac{r^3}{16R^3} \ (r \leq 2R) \\ 0 \ (r > 2R) \end{cases} \tag{11}$$

Assuming a log-normal size distribution of spheres and integrating Equation **Erreur ! Source du renvoi introuvable.**, the experimental data can be fit to extract the size of the clusters. A 20% dispersity in the log-normal distribution is assumed and a contribution of an excluded volume is also added.

The PCF is plotted as $\gamma(r) \cdot r$ allowing for an easier comparison of the differences in size between different conditions, as the position of the maximum of $\gamma(r) \cdot r$ is related to the length of correlation (i.e., to the size of the clusters).

The TEM specimens were prepared by grinding bulk specimens to ~ 100 µm thickness before they were punched out to 3 mm discs and subsequently electropolished using an electrolyte mixture of 1/3 HNO_3 and 2/3 C_3CH_3OH. The temperature was kept at 25±5°C at an applied voltage of 20 V. High angle annular dark-field scanning TEM (STEM) images were collected using a double corrected JEOL ARM200CF operated at 200 kV with convergence semi-angle and inner collector angle 27 and 48 mrad, respectively.

3 Results

3.1 Hardness

The hardness of the supersaturated solid solution (SSSS) for the 7003 alloy starts at around 46 HV while it is higher for the 7046 alloy, close to 61 HV. The hardness progressively increases for both alloys during NA, as shown in Figure 5.

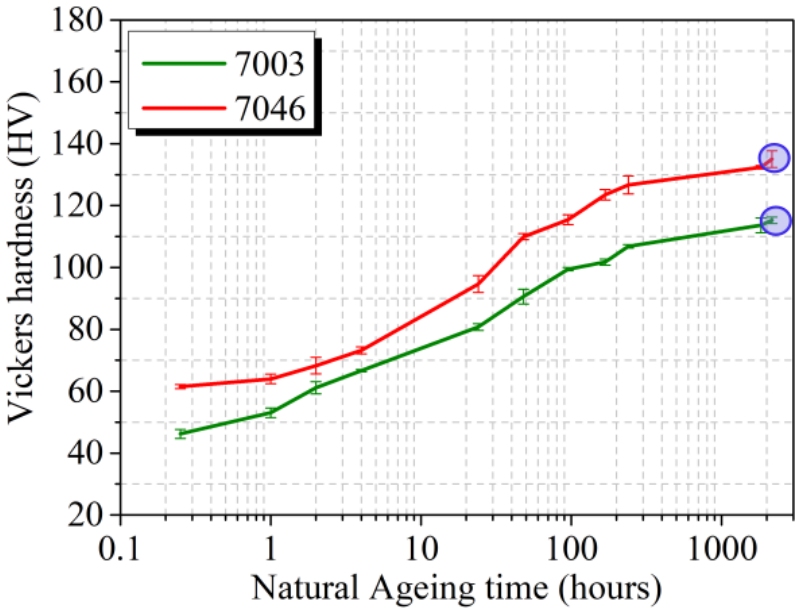

Figure 1. Evolution of hardness as a function of ageing time for both 7003 and 7046 alloys. The blue circles indicate the condition which was chosen for APT analysis.

The hardness measurements taken after 3 months of NA, revealed a considerable hardness increase to around 115 HV for the 7003 alloy and 135 HV for the 7046 alloy. Further hardness measurements revealed that the hardness for both alloys later saturated at around 125 HV and 155 HV for the 7003 and 7046 alloy, respectively, after one year NA. Nevertheless, 3 months of NA was chosen as the conditions to characterize the microstructure of these alloys.

3.2 APT results processed by the IPM methodology

The APT datasets were used to extract smaller region of interest (ROI) that were used for visualizing the clustering, as shown through Figure 2 (a and b) for 7003 alloy and Figure 2(d and e) for 7046 alloy. The IPM was used to extract size distributions, as seen in Figure 2 (c and f) for 7003 and 7046 alloys, respectively. Size measurements in this study refer to the diameter of the precipitates measured using Equation (1) similar to our previous studies [22, 35]. Additionally, the radius of gyration has been used to elucidate the mean diameter. The radius of gyration is the radius of a body having its mass concentrated at a single distance from its center of mass [1, 2]. The size distributions shown is an average of the size distributions obtained for the two datasets for each alloy.

For both alloys, the size distribution fits well with a log-normal distribution. Table 4 shows the quantitative measurements extracted from the IPM. It is quite evident that the two datasets for each alloy show similar size and composition of the clusters. The average size of the clusters for the 7003 alloy is around 1.8 nm, while it is slightly lower, around 1.5 nm, for

the 7046 alloy, as measured by the E_z. The average Zn/Mg ratio for the 7003 alloy is around 2 while it is around 1.5 for the 7046 alloy, as indicated in Table 4. The measured Extent$_X$ (E_x) is consistently larger than E_z, as seen in Table 2 for all four datasets. This is mainly due to the spatial resolution aspect of APT dominating the local magnification effects and will be discussed in detail in later sections. Since Extent in x and y are very similar, only E_x is reported in Table 2 and **Erreur ! Source du renvoi introuvable.**. The measured E_z is in close agreement with the size measurements by TEM (shown in the later section) and hence is regarded as more relevant in terms of describing the size using the IPM in this case.

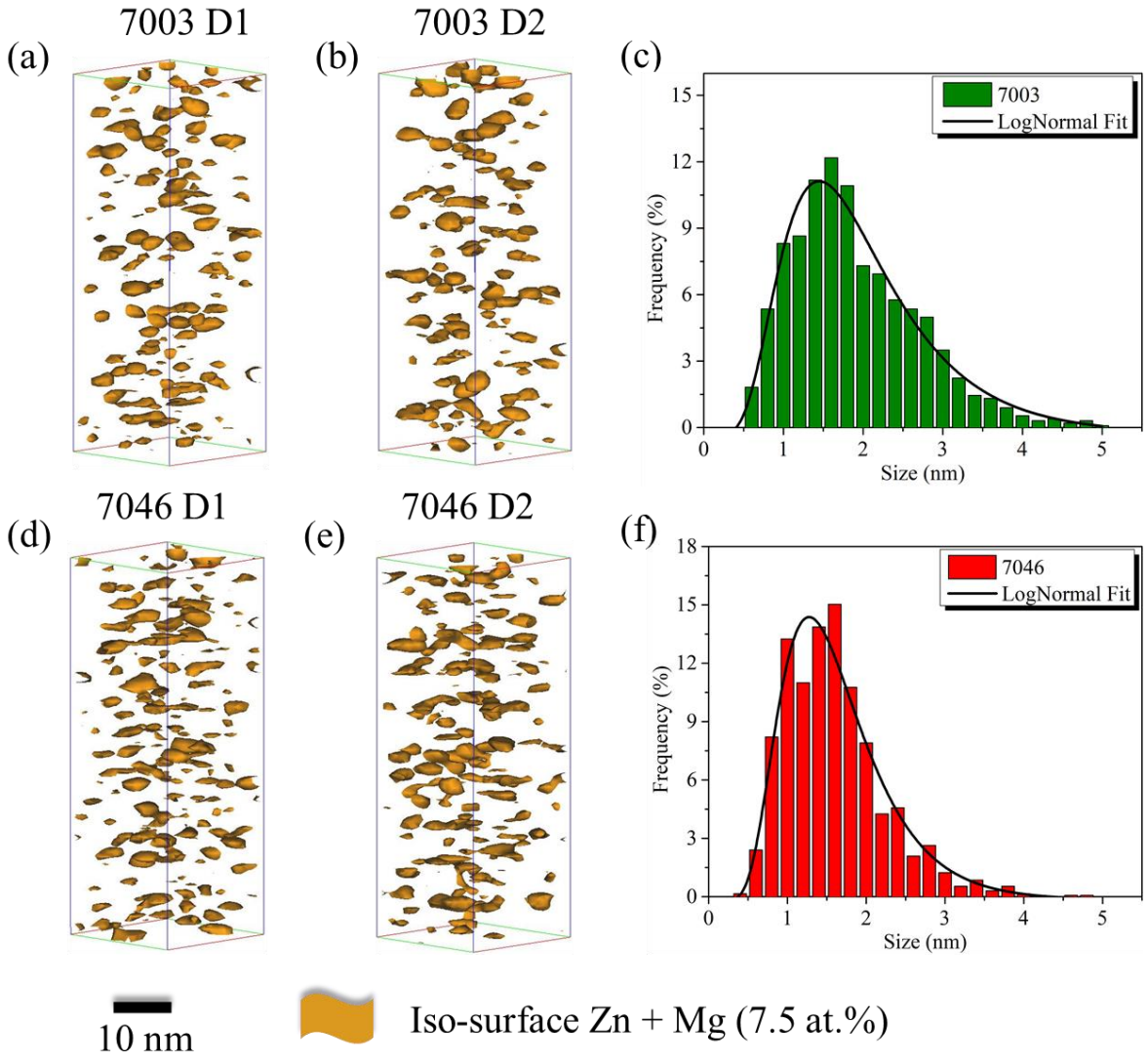

Figure 2. APT sub-volumes of dimensions 25 × 25 × 70 nm^3 extracted from the APT datasets showing clustering in the 7003 alloy for (a) dataset 1 and (b) dataset 2. Similar cluster volumes for the 7046 alloy are shown for (d) dataset 1 and (e) dataset 2. The average size distribution based on E_z of the two datasets of the two alloys is shown in (c) and (f).

Table 2. Average (mean) results of cluster analysis by IPM for APT datasets of 7003 and 7046 alloys. Compositions and Zn/Mg ratios are calculated by class size of 0.2 nm the

average is computed. The uncertainties are calculated through $2\sqrt{\frac{C(1-C)}{N}}$, where, C is the concentration of a particular element and N the total number of atoms.

	Mean Diameter (nm)	E_x (nm)	E_z (nm)	Cluster Composition (at.%)			Number Density ($10^{24}/m^3$)	Zn/Mg
				Zn	Mg	Cu		
7003 D1	2.45 ± 0.60	2.59 ± 1.53	1.75 ± 0.91	11.77 ± 0.32	5.82 ± 0.22	0.09 ± 0.02	3.19 ± 0.08	2.03 ± 0.14
7003 D2	2.53 ± 0.59	2.69 ± 1.45	1.88 ± 0.84	10.77 ± 0.29	5.35 ± 0.21	0.04 ± 0.01	3.25 ± 0.08	2.02 ± 0.10
7046 D1	2.15 ± 0.45	2.29 ± 1.10	1.50 ± 0.64	11.59 ± 0.41	7.61 ± 0.15	0.19 ± 0.10	2.34 ± 0.07	1.52 ± 0.08
7046 D2	1.96 ± 0.39	1.63 ± 1.02	1.21 ± 0.57	12.39 ± 0.45	8.09 ± 0.67	0.32 ± 0.19	1.51 ± 0.10	1.54 ± 0.17

3.3 APT results processed by the PCF methodology

The calculations of zinc (Zn-Zn), magnesium (Mg-Mg), copper (Cu-Cu) correlations along with the cross correlations for the three elements are shown in Figure 5. The first maximum of the curves, indicative of the size of the objects, is similar for all the correlations. This indicates a single distribution of clusters containing both Zn and Mg (and Cu for 7046). This size is consistently smaller for 7046 than for 7003, in agreement with the IPM results. The amplitude of the Zn-Zn PCF is larger for 7003 than for 7046, while the Mg-Mg PCF amplitude is smaller for 7003 than for 7046. These observations confirm a higher Zn/Mg ratio for the 7003 than for the 7046. The Cu correlations in Figure 3 containing are shown only for the 7046 alloy, since 7003 is a Cu-free alloy. Figure 3(f) shows higher scatter for the Cu-Cu plot due to lower signal from the Cu atoms.

More quantitative results can be obtained from the PCF by fitting the data to a 2-phase modelled dataset (matrix+clusters) containing spherical particles of homogeneous composition and with sharp interfaces (we shall further see that while this might be a reasonable assumption for larger precipitates, it is a coarse approximation for small clusters). The spherical shape is a good approximation being confirmed by the TEM analyses presented in the next section. We further assume a log-normal distribution of the diameter. This model implies that all correlations should be proportional, i.e., they should all have the same shape $\gamma_0(r)$, but with a different amplitude $\overline{\Delta C^2}$. For more efficient fitting, we thus fitted the shape to the Zn-Zn correlation and assumed it was fixed for the other correlations, with only the amplitude changing. Despite the limitations of the interpretation model (spheres,

homogeneous composition, sharp interface), the quality of the fit is excellent, enabling the parameters of the model (size, compositions) to be extracted.

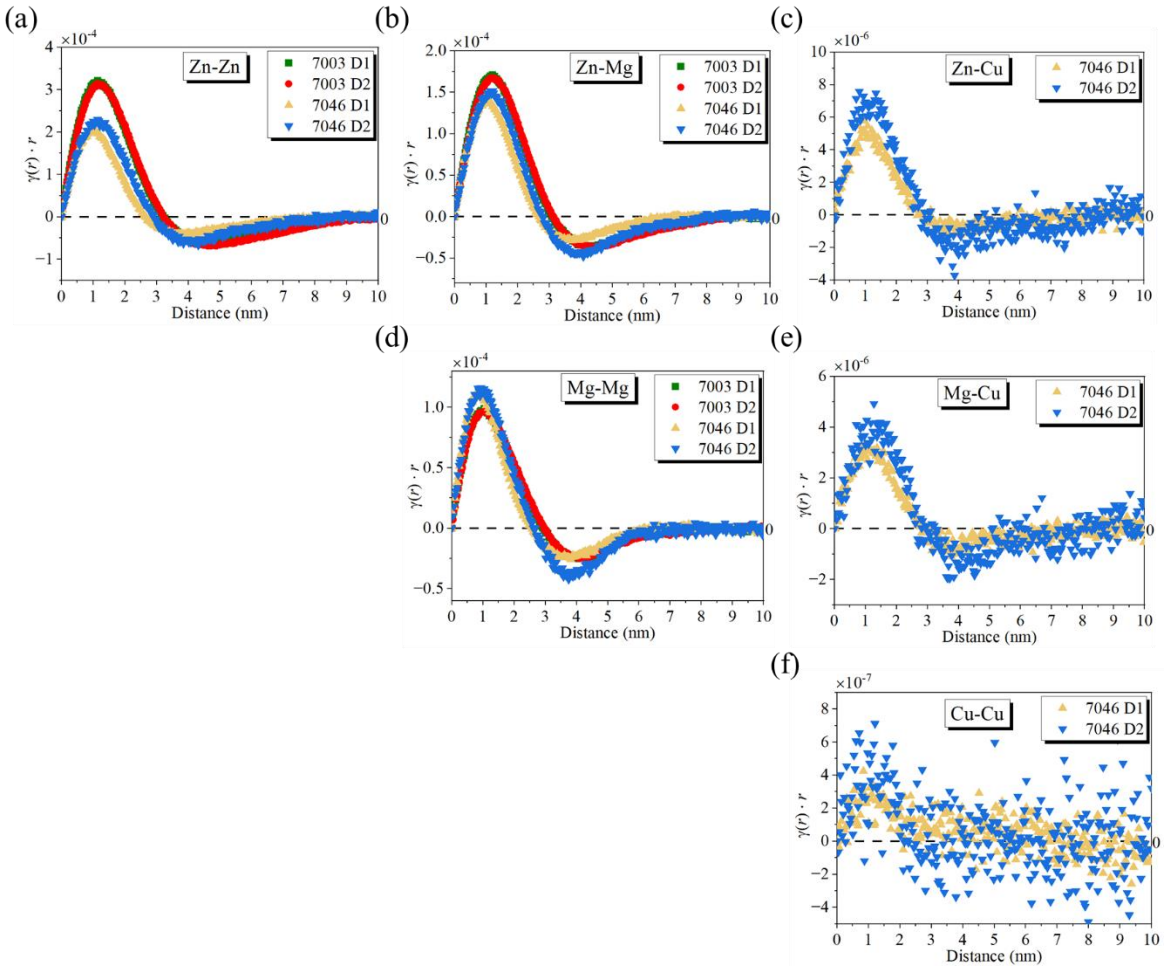

Figure 3. Pair correlation functions for (a) Zn-Zn, (b) Zn-Mg, (c) Zn-Cu, (d) Mg-Mg, (e) Mg-Cu and (f) Cu-Cu pairs from the APT datasets shown in Figure 2.

Table 3. Mean cluster size (diameter), composition, number density and Zn/Mg ratio obtained by fitting the PCF methodology data in Figure 5 assuming a log normal distribution of spherical clusters.

	Mean diameter (nm)	Cluster Composition (at.%)			Number Density ($10^{24}/m^3$)	Zn/Mg
		Zn	Mg	Cu		
7003 D1	2.98	9.76	5.03	0.14	5.9	1.97
7003 D2	3.15	8.69	4.52	0.11	6.3	1.95
7046 D1	2.61	7.76	5.34	0.3	13.2	1.47
7046 D2	3.1	8.02	5.32	0.36	7.7	1.53

The parameters extracted from the fit are shown in Table 3. These parameters show similar trends than the IPM data, in that the 7003 sizes are consistently larger than the 7046, the cluster compositions are in reasonable agreement and the Zn/Mg ratio is close to 2 for 7003 and 1.5 for 7046. There are also systematic discrepancies between the IPM and PCF results,

which can be explained by the way these parameters are computed and will be discussed further in the paper.

3.4 Cluster observed by HAADF-STEM

Atomically resolved HAADF-STEM was employed to investigate the clusters in the alloys. Figure 4. HAADF-STEM images from (a) 7003 and (b) 7046 showing a dense population of GPI zones. Corresponding size distribution estimation has been done for (c) 7003 and (b) 7046 based on 50-60 clusters. The finer clusters without structure were not included in the statistics due to uncertainty of overlap.

a and b show that both 7003 and 7046 exhibit a dense population of GPI zones. Visually, the GPI zones have a larger diameter in the 7003 alloy (Figure 4. HAADF-STEM images from (a) 7003 and (b) 7046 showing a dense population of GPI zones. Corresponding size distribution estimation has been done for (c) 7003 and (b) 7046 based on 50-60 clusters. The finer clusters without structure were not included in the statistics due to uncertainty of overlap.

a) than the ones in the 7046 alloy (Figure 4. HAADF-STEM images from (a) 7003 and (b) 7046 showing a dense population of GPI zones. Corresponding size distribution estimation has been done for (c) 7003 and (b) 7046 based on 50-60 clusters. The finer clusters without structure were not included in the statistics due to uncertainty of overlap.

b). This is in accordance with our previous results on the same alloy [22]. To investigate this in a quantitative matter, the diameters of at least 300 GPI zones were measured per condition. The GPI zones in 7003 had an average diameter of (1.87±0.17) nm, while in 7046 the average diameter was (1.67±0.11) nm.

Size distributions for the 7003 and 7046 alloys are shown in Figure 4(c and d). Here, only the structured GPI zones that could clearly be identified were used to estimate the distribution (around 50-60 GP zones). The unstructured clusters, usually the clusters smaller than 1 nm in diameter, were challenging to include in the statistics due to the uncertainty of overlapping with lower-lying clusters. While this should not impact the trend indicating slightly larger clusters in the 7003 alloy than in the 7046 alloy, it should be kept in mind that it means that the mean diameters estimated from HAADF-STEM are likely overestimated.

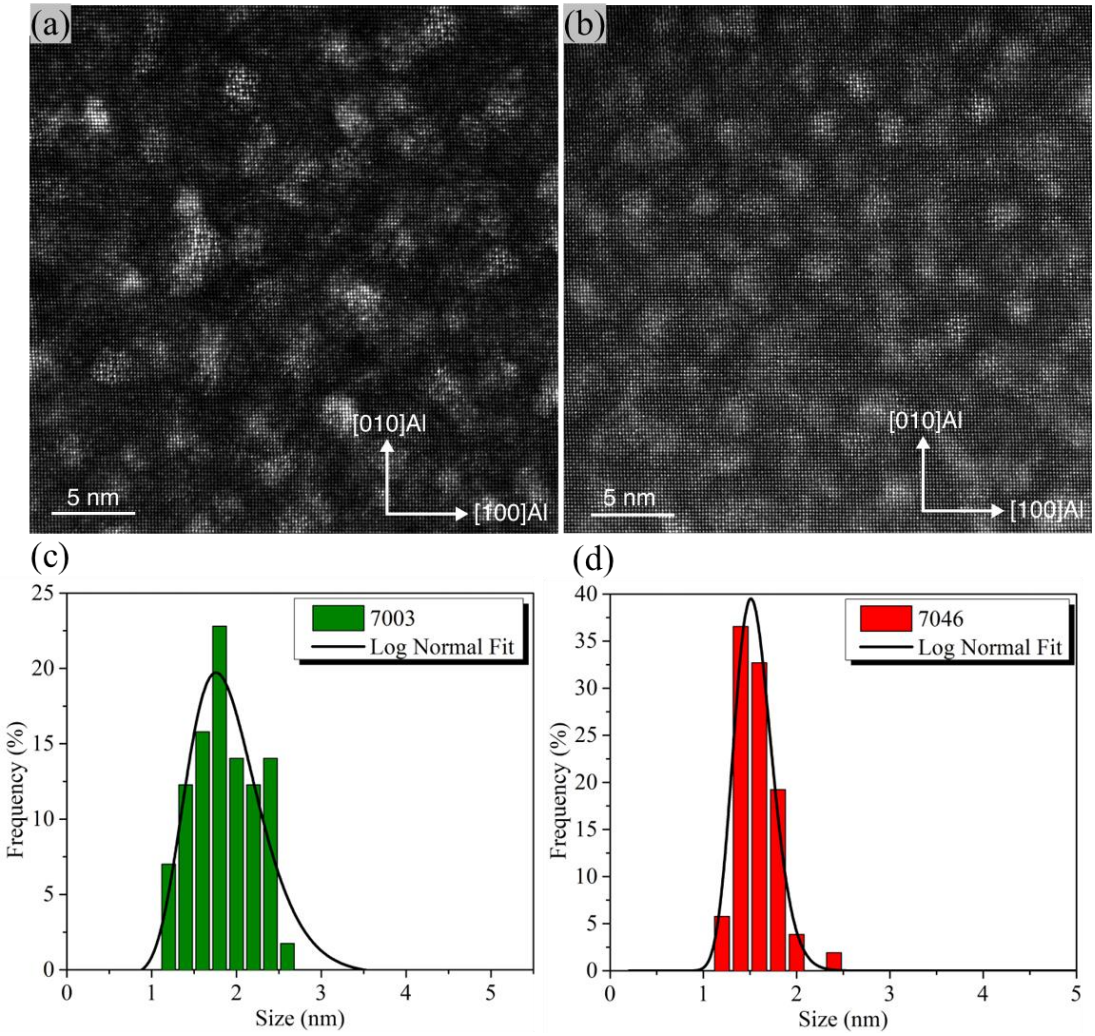

Figure 4. HAADF-STEM images from (a) 7003 and (b) 7046 showing a dense population of GPI zones. Corresponding size distribution estimation has been done for (c) 7003 and (b) 7046 based on 50-60 clusters. The finer clusters without structure were not included in the statistics due to uncertainty of overlap.

3.5 The IPM and the PCF methodology applied on simulated datasets

Experimental artefacts associated with an APT dataset include:

(i) A reduced detection efficiency of 80% (similar to the LEAP 5000XS used to analyze the experimental datasets in Figure 2).
(ii) Positioning uncertainty of atoms with different resolutions.

The reduced detection efficiency of 80% was accounted for in the simulated datasets. APT volumes were simulated with box dimensions $20 \times 20 \times 100$ nm^3. A dataset with a mean particle diameter of 1.8 nm was synthesized, picked up from the learnings of the 7003 alloy datasets. This is labeled as V1.8. A standard deviation of around 20% of the mean diameter and a log-normal distribution were applied. The particles were Zn and Mg enriched but without Cu, like the experimental datasets of the 7003 alloy. Only Al, Zn and Mg elements were used to generate the volumes. The particles introduced into the volumes had a sharp interface. Once all particles have been introduced, a random spatial distortion of each atom

position was made to obtain a more realistic dataset as compared to an experimental measurement. Hence, the initially abrupt particle interface was blurred.

First, an 'ideal resolution' of 0.2 nm in depth and 0.5 nm in lateral resolution was applied to the entire dataset. This implies that the position of atoms was randomized by 0.2 and 0.5 nm in the evaporation and lateral directions respectively. The concentration of solute in the particles is set higher than what is extracted from the experimental data. Next, the resolution is degraded until a match in the PCFs is observed when compared with the experimental PCFs. We call this dataset the 'degraded resolution' and labels it V1.8DR.

The number densities of both datasets were set to $1E^{25}$ m^{-3}. The size distribution for V1.8 is shown in Figure 5(b) while the corresponding PCFs are seen through Figure 5(c-e). The amplitude of the PCF of V1.8 is much greater than other datasets (experimental and simulation) while the amplitude of PCF of V1.8DR is very similar to that of the experimental datasets (7003 D1 and D2).

Table 4. Average (mean) results of particle analysis done by the IPM on simulated volumes.

Condition	Guinier Diameter (nm)	E_x (nm)	E_z (nm)	Cluster Composition (at.%)		Number Density ($10^{24}/m^3$)	Zn/Mg
				Zn	Mg		
V1.8	2.32 ± 0.24	2.36 ± 1.11	2.24 ± 1.01	16.56 ± 0.52	8.17 ± 0.32	5.65 ± 0.02	2.03 ± 0.08
V1.8DR	1.93 ± 0.20	2.40 ± 1.012	1.89 ± 0.86	12.92 ± 0.37	6.37 ± 0.21	3.50 ± 0.01	2.03 ± 0.07

Table 5. Mean particle diameter, composition, number density and Zn/Mg ratio obtained by the PCF methodology by fitting the data in Figure 5, assuming a log normal distribution of spherical particles.

Condition	Diameter (nm)	Particle Composition (at.%)		Number Density ($10^{24}/m^3$)	Zn/Mg
		Zn	Mg		
V1.8	2.24	18.76	9.43	6.82	1.99
V1.8DR	2.90	10.50	5.12	5.14	2.05

The quantitative data extracted by the IPM for V1.8 and V1.8DR are shown in Table 4. Both these volumes have a higher concentration of solute input in the particles; 20 at.% Zn and 10 at.% Mg. Since both the ideal resolution and the degraded resolution volumes resulted in concentrations much lower than the input concentrations, it is clear from the IPM results that the apparent composition of the clusters is greatly affected by the resolution. Number densities of detected clusters are also highly affected, for both volumes, as seen in Table 4.

When the resolution is degraded, the size measurements for V1.8DR by the IPM are underestimated, as seen in Table 4, using the E_z parameter, being smaller than expected. This is a non-trivial effect, which indicates that a substantial proportion of the interfacial cluster atoms has been attributed to the matrix during IPM. Since the applied spatial resolution is better in Z than in XY, the E_x values are consistently larger than E_z.

Because of the broader extension of the clusters due to the spatial resolution, the PCF methodology computes a larger average size than prescribed by input the dataset. The degraded resolution also strongly affects both the amount of detected solutes in the clusters and their number densities c.f. **Erreur ! Source du renvoi introuvable.** and Table 4.

The PCF methodology confirms the effect of the degraded resolution with similar results as compared to IPM for V1.8, as seen in Table 4 and Table 5. Most importantly, it should be noticed that V1.8DR seems to mimic the behavior of the 7003: it gives similar results on the detected clusters with both IPM and PCF, the overall shape and amplitude of the PCF curves are similar, and the volume was obtained with initial cluster sizes compatible with HAADF-STEM.

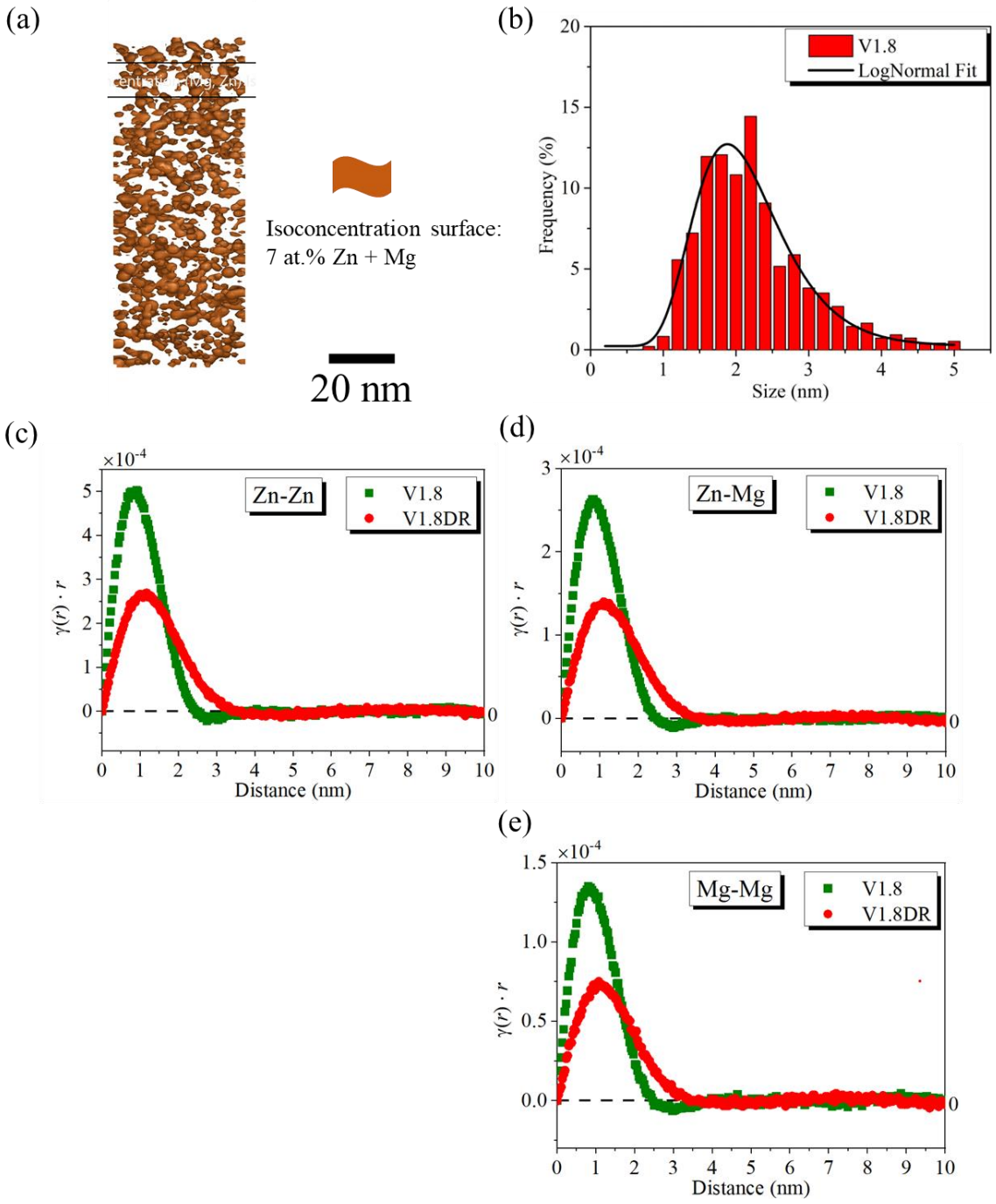

Figure 5. Simulated volume (a) V1.8 showing iso-concentration surface of 7 at.% Zn+Mg to visualize the particles inserted. Size distribution obtained from IPM for (b) V1.8. PCF for (c) Zn-Zn, (f) Zn-Mg and (g) Mg-Mg for both volumes.

4 Discussion

4.1 The apparent discrepancies in the clusters parameters as seen by the methodologies

We have looked at the clustering state in the 7003 and 7046 alloys by APT (IPM and PCF) and by HAADF-STEM. We found similar trends among the methods: the clusters seem consistently smaller in the 7046 than in the 7003. The Zn/Mg ratio was found to be about 2 for 7003 and 1.5 for 7046 with both APT methods. There are however systematic discrepancies which can be at least partially explained both by the difference in approach between IPM and PCF based methods, and by the nature of APT data. These discrepancies are:

i) The IPM method finds consistently higher solute content in the clusters than the PCF methodology.
ii) The IPM method finds consistently lower number density of clusters than the PCF methodology.
iii) Both APT methods find larger clusters than HAADF-STEM.

These discrepancies should be discussed in the context of very small clusters which do not present a sharp interface with the matrix, partly or entirely because of a finite spatial resolution limiting the positioning accuracy.

It should be kept in mind that the IPM methodology and the PCF methodology are two very different approaches. The IPM is a cluster identification method. Its objective is to label each atom in the dataset as belonging either to the clusters or to the matrix. The result depends on the choice of parameters, which are chosen through an optimization process aiming at minimizing the amount of “false positives” whereby atoms from the matrix are wrongly labeled as clusters. It requires as input an estimate of a minimal size for the clusters. An erosion step guarantees that the clusters do not include a shell of matrix. This ensures that compositions computed on the detected clusters are measured in a condition, neither including too small clusters nor interfacial area of the clusters that includes matrix atoms from a diffuse interface.

This is in contrast with the PCF method, which considers the dataset as a whole and assumes that the clusters have a homogeneous composition. Hence, when the cluster/matrix interface is actually diffuse and the cluster size is small, the average composition measured by the PCF fitting method is likely to be smaller, explaining discrepancy i).

Discrepancy ii) is indirectly related to i) and can also be explained by the selective approach (IPM) vs global approach (PCF): the smaller clusters from the dataset may be part of the

cluster distribution fitted by the PCF approach, while being labeled as too small to be a cluster by IPM, thus lowering the detected number density. Whether this is a good or a bad thing depends on the situation. The IPM parameters could be set to detect more clusters, but very likely at the cost of an apparent lower cluster composition.

Discrepancy iii) is more informative on the metrology of such small clusters by APT than on the methods themselves. The size histograms obtained by HAADF-STEM are likely to be very good estimates of the size of the clusters. If anything, they might be slight overestimations since the smaller diffuse clusters were left out of the calculation. This means that the clusters apparent size in APT is too large. Since the datasets reconstructions were properly calibrated and since the local magnification effect (higher density in these objects) should have a reverse effect, this can be only attributed to a spatial resolution blurring of the interfaces [36].

The application of the two methodologies to simulated datasets enabled to better estimate the instrumental effect on the cluster parameters by attempting to create realistic simulated datasets that are compatible with our experimental results.

4.2 Studies of Simulated Volumes

4.2.1 Number density measurements

From the ideal resolution dataset 64% of the particles are detected with the IPM, while when applied to the degraded resolution dataset (V1.8DR) only 39% of particles are detected. Hyde *et al.*[5] found it challenging to detect clusters smaller than 1 nm in diameter using the same methodology, while the detection of clusters was even worse using the MSM, despite having a higher solute concentration contrast between the matrix and the clusters. This is in line with the current observations as demonstrated in Figure 5(b) with the size distribution showing no clusters detected below 1 nm in diameter.

The results for the 'degraded resolution' suggest that the number densities measured are even lower when a lower resolution is accounted for. The spatial distortions in lateral and depth resolution make it further challenging to detect the clusters smaller or close to 1 nm in diameter. To distinguish the smallest particles (with diameter smaller than about 1 nm) from random solute fluctuations becomes problematic, and thus a lot of fine particles are not detected by the IPM. With the degraded resolution, the smaller clusters are simply not there anymore due to the blurring caused by the spatial resolution effects.

The statistical method using the PCF on the other hand, estimates the number density from the volume fraction of particles based on the average volume of particles as $f_v/\langle V \rangle$. There is only a slightly higher estimated volume fraction for the cases with degraded resolution. However, the average particle diameter estimated for V1.8DR is higher than for V1.8. Hence, similar number densities are predicted by the PCF methodology for the ideal resolution and the degraded spatial resolution. As already discussed, this difference between IPM and PCF is due to different approaches (selective for IPM and global for PCF).

4.2.2 Size measurements

When considering the ideal resolution dataset (V1.8), we see that both methods (IPM and PCF) overestimate the size of particles by around 28% c.f. Table 4 and Table 5. However, when considering the degraded resolution dataset (V1.8DR), a smaller average cluster diameter is estimated by the IPM (~1.9 nm c.f. Table 4), while the PCF methodology

overestimates the size by almost 60% (c.f. Table 5) The IPM mitigates the effect of spatial resolution and blurring of atoms by utilizing the erosion step which removes the interface atoms from the clusters. This in turn reduces the size of clusters and can explain the trend of lower size measurements for V1.8DR. Note that this dataset is very similar to the experimental one in terms of PCF amplitudes and concentration estimates.

The PCF methodology, on the other hand, slightly overestimates the cluster diameter with degraded spatial resolution. This can be explained by the scatter to the atomic positions, which for the PCF methodology leads to an increase in the net volume within each cluster. The volume fraction measured by the PCF methodology also increases for V1.8DR as compared to their ideal resolution counterparts, which explains the increase in the mean diameter estimated by the PCF methodology.

4.2.3 Composition estimates

The simulated volumes V1.8 and V1.8DR had a higher input concentration of solutes in the particles (Zn = 20 at. % and Mg = 10 at. %) than the measured ones. This was mainly done to achieve a similar amplitude of the PCFs after spatial degradation as compared to the experimental data. The amplitudes of V1.8DR and of the experimental dataset (7003 D1 and D2) are very similar, as seen in Figure 6.

For both the IPM and the PCF methodology, the Zn and Mg content in the clusters are close to the input composition of particles for V1.8. However, when the spatial resolution is degraded, the Mg and Zn contents in the clusters become underestimated by both methodologies for V1.8DR, as seen in Table 4 and Table 5.

Let us first consider the PCF approach. The amplitude of PCFs for the degraded datasets is smaller than its ideal resolution counterpart, as seen in Figure 6. Note that Equation (10) for the particle compositions involves the bulk composition as well as the matrix composition. C_p^a is very sensitive, because of the factor of $1/(\bar{C} - C_m)$ in Equation (10). This increases the uncertainty and makes the composition of solute in particles challenging to accurately measure by this method. The computation of the particle composition thus strongly depends on the matrix and bulk compositions used. The values of ΔC_{Mg}^2 , ΔC_{Zn}^2 and ΔC_{Zn-Mg}^2 at r=0 can be used in Equation (10) to extract these as the two unknowns (C_p^{Mg} and C_p^{Zn}), although matrix compositions can be influenced strongly by the lateral scatter.

Although IPM estimated the composition of particles in V1.8 close to the input concentrations, care must be taken, since the particles smaller than about 1 nm in diameter, remain undetected due to the spatial degradation and possibly the particles no longer existing (become a part of the random fluctuation). These small particles then become poorly represented in terms of the statistics. Hence, the composition of these particles has a high uncertainty, and the estimated composition is dominated by the larger particles, which have a defined interface and composition. However, in the synthesized samples considered here, all particles had the same stoichiometry, hence this issue did not affect the estimates in our cases.

The IPM underestimates the composition for the degraded resolution volumes. This is expected, since the lateral scatter can result in the artificial introduction of a significant level of matrix (Al) into the defined particle. A similar trend was seen by Hyde *et al.* [37] when using MSM to determine the composition for simulation of trajectory aberrations. This results in a small solute cluster being significantly diluted due to the scatter caused by the degraded resolution parameters.

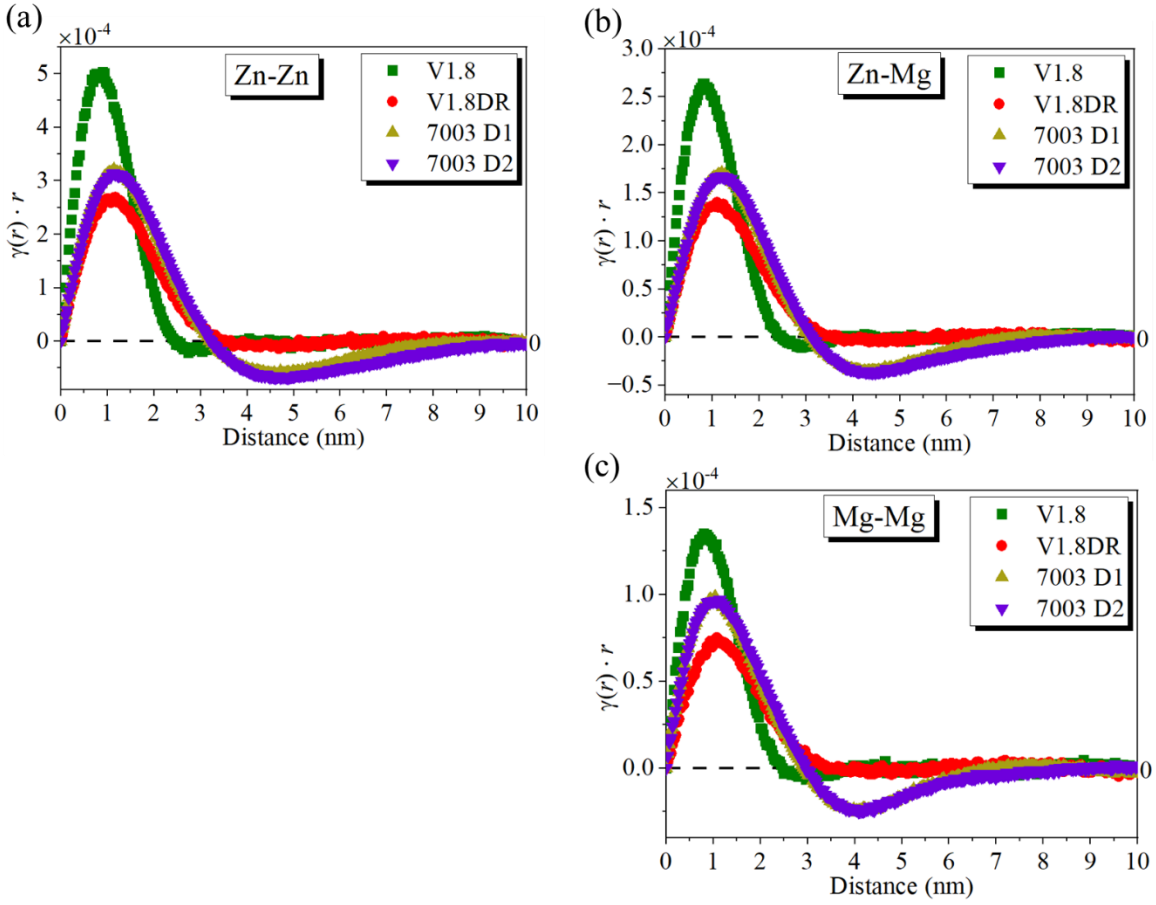

Figure 6. PCF plots of (a) Zn-Zn, (b) Zn-Mg and (c) Mg-Mg for the simulated and experimental (7003) datasets. The degraded resolution dataset of the simulated volume is a close match in terms of PCF's compared to the 7003 datasets.

Interestingly, the concentration of solute in particles measured by both IPM and PCF for V1.8DR is very similar to the experimental datasets, as seen in Table 3. Although the solute concentration is underestimated when compared to the input values in the dataset, we see that the degraded resolution (V1.8DR) is indeed a good representation of the actual experimental dataset. This in turn indicates the underestimation of solute concentration in the experimental datasets. It is well known that the local magnification effects cause trajectory aberrations, ultimately leading to the underestimation of solute inside clusters [38]. We see a similar trend here confirmed by V1.8DR, wherein a higher concentration of solute in particles is heavily underestimated, causing erroneous composition measurements. The Zn/Mg ratio estimated for both simulated volumes is close to two, as seen from Table 4 and Table 5, indicating that

the quantification of the amount of Zn and Mg inside particles is affected proportionally and hence, the true ratio is expected to be preserved.

4.3 Experimental data interpretation

The analysis conditions were optimized for the experimental data to avoid preferential evaporation. A high pulse fraction and low temperature were used to analyze all datasets [39]. No sign of molecular dissociation [40] was seen. Based on the understanding gained from the simulated volumes, the average compositions of GP zones for 7003 and 7046 alloys estimated by both PCF methodology and IPM are underestimated. The apparent particle composition was lower than the input value (20 at. % Zn and 12 at. % Mg) for the simulated volume, along with a reduced measured number density, mainly due to particles below 1 nm not being detected.

In principle, the shapes of the PCF's for Zn, Mg and Zn-Mg should all be the same, since they origin from the same particles (spheres). This is very well reproduced in the simulated volume, as seen in Figure 5(b and c). Both simulated volumes (V1.8and V1.8DR) have the same shape as seen from Figure 5(b and c). They overlap each other and follow the 'pseudo-binary' assumption.

The experimental dataset, on the other hand, shows slightly different shapes of the PCF's. Since we have only a single population of clusters, (GP zones predominantly) we should ideally expect all PCF's to be of the same shape. In particular, the mean size of the GP zones would be slightly smaller if it was based on the fit from Mg rather than that of Zn. Several works on Al-Zn-Mg alloys [41, 42] have reported this effect, where the extension of the Zn enrichment in precipitates is higher than that of Mg. This can be interpreted as an APT artefact, wherein subtle changes in the evaporation field of the matrix atoms within particles changes the electric field distribution and hence causes these chromatic aberrations [43].

However, the use of the Zn PCF in the calculations cannot solely explain the overestimation of the mean diameter by the PCF methodology. A decrease of the estimated diameter of around 7% is seen when using the PCF of Mg instead of Zn. The spatial resolution is element specific, as we see in the case of PCF's plotted for different elements for the same dataset in Figure 7. This is more of an instrument effect that is not seen in the simulated volumes. Additionally, this will also affect the results obtained by the IPM. The spatial resolution plays an important role in determining the size, as also pointed out by other authors [36, 43].

It can be seen by comparing Figure 5(a) and Figure 5(b and c) that there is a part of the PCF which is negative. This "dip" is more prominent in the PCF curves from the measured experimental datasets than in the simulated volumes. The interparticle interaction qualitatively gives this effect and the dip can be well reproduced by accounting for an 'excluded volume' [44] in the PCF methodology. This is because two particles cannot overlap. Hence the position of each particle is not independent. However, the dip can also be due to depletion in solute around the particles as part of the solute diffusion, which is also not accounted for. Anyhow, the correction for the dip decreases the estimate of the mean diameter only slightly, by a few percent. Hence, it will neither affect much the estimates of particle compositions.

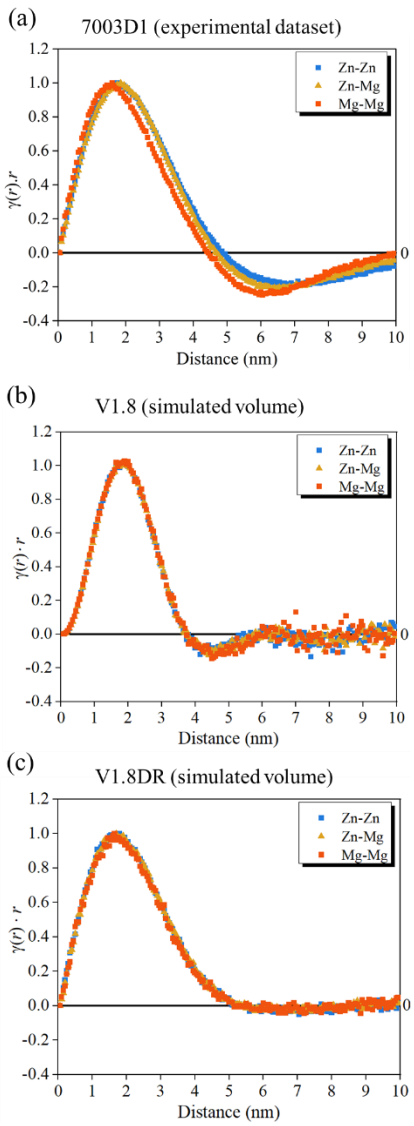

Figure 7. Normalized PCF for Zn-Zn, Zn-Mg and Mg-Mg obtained from (a) 7003 experimental dataset and simulated volumes (b) V2DR and (c) V3DR.

The IPM gives representative size measurements with Guinier diameter, E_x as well as E_z for the simulated volumes as seen in **Erreur ! Source du renvoi introuvable.** and Table 4. This is not the case for experimental datasets though. TEM measurements agree with the IPM size measurements (E_z) for the experimental datasets. The E_x is overestimated for all four datasets and on average is around 1.7 times the E_z. The considered clusters are low field precipitates giving rise to higher atomic density, hence the opposite was expected, i.e., a contraction in the XY plane and potentially to appear smaller. However, this is not the case. This opposite trend is attributed to the local magnification effects and trajectory overlaps due to the difference in evaporation field between the matrix and particles which are seen in the experimental datasets [29]. In other words, the spatial resolution is locally worse.

Number density measurements, as highlighted in the above section, is heavily underdetermined by IPM for V1.8. From the TEM results, many disordered clusters smaller

than 1 nm are seen, which are not taken into account while estimating the size distribution as seen in Figure 4(c and d) due to the challenge of de-convoluting the overlapping clusters.

Both methodologies have shown reasonable agreements on size and composition measurements. The important understanding from the simulated volumes in an attempt to mimic the experimental datasets showed that both methodologies underestimate the concentration of solute within particles. Features less than 1 nm usually go undetected, and consequently the number densities are underestimated. Care must be taken, since number densities measured by IPM are sensitive to the size of the particles, while the estimate by the PCF is based on the average volume of the particles.

5 Conclusions

By using a combination of APT data interpretation methodologies, supported by simulated datasets and HAADF-STEM experiments, we have shown that reliable APT measurements of the size and compositions from small clusters issued from the early stages of decomposition in Al-Zn-Mg alloy are challenging. Because of their small size, the particles/clusters are subject to significant distortions due to trajectory aberrations, leading to an increase of their apparent size as well as a decrease of their apparent solute content. While these effects are likely to be system dependent, we have estimated in our case that the clusters solute content dropped 40%. The effect on the apparent size is more complex to evaluate because it is strongly anisotropic, due to the anisotropy of the local magnification effect. It is clear, however, that these effects are most severe for the smaller clusters, so that particular care should be taken when interpreting apparent size and compositions of clusters from APT measurements of naturally age hardened or under aged conditions of AlZnMg aluminum alloys.

References

1. Gault, B., et al., *Atom probe microscopy*. Vol. 160. 2012: Springer Science & Business Media.
2. Kelly, T.F. and M.K. Miller, *Atom probe tomography.* Review of scientific instruments, 2007. **78**(3): p. 031101.
3. Vaumousse, D., A. Cerezo, and P. Warren, *A procedure for quantification of precipitate microstructures from three-dimensional atom probe data.* Ultramicroscopy, 2003. **95**: p. 215-221.
4. Stephenson, L.T., et al., *New techniques for the analysis of fine-scaled clustering phenomena within atom probe tomography (APT) data.* Microscopy and Microanalysis, 2007. **13**(6): p. 448-463.
5. Hyde, J.M., et al., *Analysis of radiation damage in light water reactors: comparison of cluster analysis methods for the analysis of atom probe data.* Microscopy and Microanalysis, 2017. **23**(2): p. 366-375.
6. Lefebvre, W., et al., *3DAP measurements of Al content in different types of precipitates in aluminium alloys.* Surface and Interface Analysis: An International Journal devoted to the development and application of techniques for the analysis of surfaces, interfaces and thin films, 2007. **39**(2-3): p. 206-212.
7. Felfer, P., et al., *Detecting and extracting clusters in atom probe data: A simple, automated method using Voronoi cells.* Ultramicroscopy, 2015. **150**: p. 30-36.

8. Johnson, C.A. and J.H. Klotz, *The atom probe and Markov chain statistics of clustering.* Technometrics, 1974. **16**(4): p. 483-493.
9. Lefebvre, W., T. Philippe, and F. Vurpillot, *Application of Delaunay tessellation for the characterization of solute-rich clusters in atom probe tomography.* Ultramicroscopy, 2011. **111**(3): p. 200-206.
10. Samudrala, S., et al., *A graph-theoretic approach for characterization of precipitates from atom probe tomography data.* Computational Materials Science, 2013. **77**: p. 335-342.
11. Zelenty, J., et al., *Detecting clusters in atom probe data with Gaussian mixture models.* Microscopy and Microanalysis, 2017. **23**(2): p. 269-278.
12. Lefebvre, W., F. Vurpillot, and X. Sauvage, *Atom probe tomography: put theory into practice*. 2016: Academic Press.
13. Miller, M. and R. Forbes, *Atom-probe tomography: the local electrode atom probe. Springer.* 2014.
14. Lervik, A., et al., *Atomic structure of solute clusters in Al–Zn–Mg alloys.* Acta Materialia, 2021. **205**: p. 116574.
15. Williams, C.A., et al., *Defining clusters in APT reconstructions of ODS steels.* Ultramicroscopy, 2013. **132**: p. 271-278.
16. Vurpillot, F. and C. Oberdorfer, *Modeling atom probe tomography: A review.* Ultramicroscopy, 2015. **159**: p. 202-216.
17. De Geuser, F., W. Lefebvre, and D. Blavette, *3D atom probe study of solute atoms clustering during natural ageing and pre-ageing of an Al-Mg-Si alloy.* Philosophical Magazine Letters, 2006. **86**(04): p. 227-234.
18. Sudbrack, C.K., R.D. Noebe, and D.N. Seidman, *Direct observations of nucleation in a nondilute multicomponent alloy.* Physical Review B, 2006. **73**(21): p. 212101.
19. Zhao, H., et al., *Parameter free quantitative analysis of atom probe data by correlation functions: Application to the precipitation in Al-Zn-Mg-Cu.* Scripta Materialia, 2018. **154**: p. 106-110.
20. Ivanov, R., A. Deschamps, and F. De Geuser, *A combined characterization of clusters in naturally aged Al–Cu–(Li, Mg) alloys using small-angle neutron and X-ray scattering and atom probe tomography.* Journal of Applied Crystallography, 2017. **50**(6): p. 1725-1734.
21. Medrano, S., et al., *Cluster hardening in Al-3Mg triggered by small Cu additions.* Acta Materialia, 2018. **161**: p. 12-20.
22. Shah, S., et al., *Effect of cyclic ageing on the early-stage clustering in Al–Zn–Mg (-Cu) alloys.* Materials Science and Engineering: A, 2022: p. 143280.
23. Famelton, J., et al., *Xenon plasma focussed ion beam preparation of an Al-6XXX alloy sample for atom probe tomography including analysis of an α-Al (Fe, Mn) Si dispersoid.* Materials Characterization, 2021. **178**: p. 111194.
24. Gault, B., et al., *Advances in the calibration of atom probe tomographic reconstruction.* Journal of Applied Physics, 2009. **105**(3): p. 034913.
25. *Atom Probe Tomography (APT) - Department of Materials Science and Engineering - NTNU*. 2022; Available from: https://www.ntnu.edu/ima/research/apt.
26. Hatzoglou, C., et al., *A chemical composition correction model for nanoclusters observed by APT-Application to ODS steel nanoparticles.* Journal of Nuclear Materials, 2018. **505**: p. 240-248.
27. Lawitzki, R., P. Stender, and G. Schmitz, *Compensating local magnifications in atom probe tomography for accurate analysis of nano-sized precipitates.* Microscopy and Microanalysis, 2021. **27**(3): p. 499-510.
28. Thronsen, E., et al., *The evolution of precipitates in an Al–Zn–Mg alloy.* Journal of Materials Research and Technology, 2023. **23**: p. 5666-5680.
29. Vurpillot, F., A. Bostel, and D. Blavette, *Trajectory overlaps and local magnification in three-dimensional atom probe.* Applied Physics Letters, 2000. **76**(21): p. 3127-3129.

30. Miller, M. and E. Kenik, *Atom probe tomography: A technique for nanoscale characterization.* Microscopy and Microanalysis, 2004. **10**(3): p. 336-341.
31. Karnesky, R.A., C.K. Sudbrack, and D.N. Seidman, *Best-Fit Ellipsoids of Atom-Probe Tomographic Data to Study Coalescence of Gamma Prime (L1_2) Precipitates in Ni-Al-Cr.* arXiv preprint cond-mat/0703452, 2007.
32. Geuser, F.D. and W. Lefebvre, *Determination of matrix composition based on solute-solute nearest-neighbor distances in atom probe tomography.* Microscopy research and technique, 2011. **74**(3): p. 257-263.
33. Guinier, A., G. Fournet, and K.L. Yudowitch, *Small-angle scattering of X-rays.* 1955.
34. Philippe, T., et al., *Clustering and nearest neighbour distances in atom-probe tomography.* Ultramicroscopy, 2009. **109**(10): p. 1304-1309.
35. Shah, S., et al., *Precipitation, mechanical properties and early slant ductile fracture in cyclic and naturally aged Al-Zn-Mg (-Cu) alloys.* Materials & Design, 2022. **222**: p. 111026.
36. De Geuser, F. and B. Gault, *Metrology of small particles and solute clusters by atom probe tomography.* Acta Materialia, 2020. **188**: p. 406-415.
37. Hyde, J., et al., *A sensitivity analysis of the maximum separation method for the characterisation of solute clusters.* Ultramicroscopy, 2011. **111**(6): p. 440-447.
38. Hatzoglou, C., B. Radiguet, and P. Pareige, *Experimental artefacts occurring during atom probe tomography analysis of oxide nanoparticles in metallic matrix: Quantification and correction.* Journal of Nuclear Materials, 2017. **492**: p. 279-291.
39. Saxey, D., *Correlated ion analysis and the interpretation of atom probe mass spectra.* Ultramicroscopy, 2011. **111**(6): p. 473-479.
40. Jin, S., et al., *Effects of atom probe analysis parameters on composition measurement of precipitates in an Al-Mg-Si-Cu alloy.* Ultramicroscopy, 2022. **235**: p. 113495.
41. Engdahl, T., et al., *Investigation of fine scale precipitates in Al–Zn–Mg alloys after various heat treatments.* Materials Science and Engineering: A, 2002. **327**(1): p. 59-64.
42. Stiller, K., et al., *Investigation of precipitation in an Al–Zn–Mg alloy after two-step ageing treatment at 100 and 150 C.* Materials Science and Engineering: A, 1999. **270**(1): p. 55-63.
43. Marquis, E.A. and F. Vurpillot, *Chromatic aberrations in the field evaporation behavior of small precipitates.* Microscopy and microanalysis, 2008. **14**(6): p. 561-570.
44. Kruglov, T., *Correlation function of the excluded volume.* Journal of applied crystallography, 2005. **38**(5): p. 716-720.