**Advanced Material Characterization by Atom Probe Tomography and
ElectronMicroscopy**
Prof. Surendra Kumar Makineni
Materials Science
IISc Bangalore
Week-10
Lecture-29

Okay, so welcome to this class. In the last class, we briefly went through the composition profiles: 1D composition profiles, 2D composition profiles, and also we described proxigrams, which are proximity histograms. If you have an interface that is curved and not flat, you can accurately determine the composition profile across that interface by using proximity histograms, correct?

In the last two classes, we also described sample preparation by focused ion beam. I assume you have gone through those videos where you saw the step-by-step process of making needle specimens for the atom probe. So in this class, we will discuss more data analysis and the type of information we can obtain from the reconstructions. Now, what we call these are the Frequency distribution analysis.

Okay. So what is this frequency distribution analysis? Usually, these are helpful in cluster analysis. I will come to these terms later. Then decomposition effects.

In the microstructure, these may be related to phase separation and the formation of precipitates—decomposition effects. Also, these two can provide atomic-scale quantification, fine. Okay, so the first part described in the frequency distribution analysis is what we call binomial analysis. This binomial analysis, so this is nothing but a grid based analysis. frequency distribution okay so this is a grid based frequency distribution technique okay so here what we do is each voxel

I think the term voxel has been explained or described during I think the last few in the last few classes so each voxel usually here contains a constant total, a constant total number of atoms which is given by NB. Correct? So, what is the, so this is the each voxel. So, how the analysis is done?

So, for the analysis, it means that the frequency with which frequency with which atoms of some particular element, for example, A, atoms of some particular element A occur per block. Okay? And this is compared, this is compared to the theoretical distribution of A atoms that are present or occur by random. Okay?

in the system okay throughout the data set randomly throughout the data set okay so it is just to determine the frequency with which the particular element A which is present per block or per voxel what we can per block as compared to the A atoms which are randomly distributed throughout the data set. So, this is the comparison, okay? And this random distribution, this random distribution is nothing but the binomial probability distribution, okay?

Which is meant by Pb, okay? So, this random distribution of A atoms throughout the dataset is nothing but the binomial probability distribution, which is PV. So, if there is a material or if there is an alloy with the proportion of A, which is XA, then the corresponding binomial distribution means that The corresponding binomial distribution can be given by f_bn, which is equal to n_pbn, which is equal to n_nb factorial. These are mathematical terms.

There is no need to remember. But it is just to understand what this binomial distribution is. Factorial x_a n 1 minus x_a n_b minus n_a. Okay, so this is the term for the binomial distribution of A, which is present as the proportion or the fraction which is XA in the sample or in the dataset. Then the binomial distribution of this particular species A can be given by this formula.

Okay. So, if the construction, if given, if there is a reconstruction, it is divided into n blocks. I will give you the example for this, which is divided into the n blocks. Okay. Now, the FBM.

is the FBN is the is expected that expected the number of FBN is the expected number of blocks each containing n atoms of element A. Okay. So, this particular function FBN can be plotted in comparison to the, in comparison to the experimental frequency distribution, which is E n. Okay? So, this F b n is nothing but the expected frequency number of blocks of each containing n atoms of element A. Okay?

And this can be plotted in the comparison of experimental frequency distribution. Okay? So, this experimental frequency distribution, whatever you are getting from the needles, from the reconstruction, usually have a shift away from from the binomial distribution. If there is a shift as compared to the binomial distribution of this particular element A, it indicates that the atoms are arranged.

The atoms are arranged in a non-random manner. So, if these atoms are arranged in a non-random manner, it means that the A atoms, the A elements or what we call it is, it might be possible that there must be a multiphase which are presented or it might be the atomic scale clustering. or nanoscale castering, atomic scale or nanoscale castering existence, okay? So, this binomial distribution is for the random arrangement of A element within the data set, okay?

This is given by this. And if you plot the experimental frequency distribution, En, with respect to the binomial distribution, then if there is a shift in the there is a shift in the peak or there is a shift in the frequency distribution, it shows that the atoms are arranged in a non-random fashion. This hints to us that there might be the presence of multiphase alloys or there might be nanoscale clustering in the alloy. So, I will show you an example of

atomic-scale clustering in an aluminum-copper alloy. Okay? So, here you can see that these are the orange-colored copper atoms in an aluminum alloy. And this particular A is as-quenched meaning it is a solid solution, mostly a solid solution, then these are heat-treated at a certain temperature and time.

So, maybe you can label them as T1, T1, T1, T2. Okay? Now, here in this, each dataset is divided into equal blocks of 200 atoms, and these contain the same number of blocks. So, each dataset is divided into equal blocks of 200 atoms.

It means each block contains 200 atoms. And each data set has the same number of blocks in the data set. Okay. So in the as-quenched state, you can see there is a plot between the counts and the copper atoms per 1000-atom block. Okay.

So this x-axis gives the, usually it gives the concentration or the number of copper atoms per block. Number of copper atoms per block. Okay. And these are the counts. So if you see this black color curve, which is as-quenched.

The black color curve as compared to the binomial curve. The binomial curve, as I told you, the binomial distribution is nothing but the random distribution of copper atoms in the data set. So, if you see that this dashed line is a binomial distribution, and then there is data set 1, data set 2, data set 3, and for all three data sets, these frequency distributions have been plotted. So the black corresponds to as-quenched, meaning if you heat-treat the sample, it will be in the solid solution state and it has been quenched. So you can see that in the as-quenched sample, if you see the atom map, you can see that there is no spatial correlation.

But if you plot the frequency distribution, this particular frequency distribution, the black color curve, you can see that the experimental distribution, the black curve, gets widened and is shifted to the left. Okay. So, you can see that this particular curve, the black-colored curve, which is as-quenched in comparison to the binomial distribution, you can see that it has shifted to the left and also widened.

Okay? And now you can see that after thermal treatment, like at certain temperatures T1 and T2, okay? You can see that the deviation increases at T1. The deviation increases, and the peak shifts left more and more as the time increases. It means that it goes toward the lower copper concentration.

As I told you, on the x-axis, these are copper atoms per thousand blocks. You can see that as time progresses, the peak shifts toward the left. It means that this peak appears at a lower copper concentration. What does it indicate? It indicates more blocks...

contain fewer copper atoms. You can see that these are the counts, meaning the number of blocks. So, after thermal treatment, you can see that more blocks contain fewer copper atoms, and secondly, you can see that there is a long tail as compared to the as-quenched sample or the binomial distribution. So, there appears a long tail. It means that, long tail, it means that smaller number of blocks contains high copper concentration.

Correct? You can see that in these regions, There is a large tail that shows that the copper concentration is higher at these. So the very smaller number of blocks, around less than 500 number of blocks contains higher copper concentration. So this particular analysis shows that it reveals the copper segregation effects.

Copper segregation effects in the data set. You can see that these are the copper rich regions. Okay? So, this we can quantify by the frequency size distribution, frequency distribution. Fine?

Okay, now with introduction of these frequency size distribution or frequency distribution plots, what we can do is actually we can get the atomic distribution in the data sets. So, atom probe tomography or in the atom probe reconstruction, this is a very powerful tool where actually you can get the atomic interaction distances in three dimensions. It means that you can actually in three dimensions, you can get these atomic interaction distances. So, atomic interaction distance—these are nothing but atom-to-atom distance distribution. Atom-to-atom distribution.

So, with this, atomic interaction distances or atom-to-atom distance distribution in three dimensions—what we call it is first nearest neighbor distribution, which we refer to as 1NN nearest neighbor distribution (1NND). So, this is the most fundamental study where you can get the atomic distance with the amplitude reconstruction. So, this is the nearest.

So, 1 corresponds to the first nearest neighbor distribution. So, this technique is nothing but the interatomic spacing technique. Or interatomic separation in the immediate vicinity around each atom. So, it measures the interatomic spacing of the first nearest neighbor of each solute atom. Correct?

So, we can analyze the nearest neighbor distribution by including all the atoms or you can limit the distances between particular atoms. Okay? So, it might be the distance between AA or the distance between BB-type atoms. So, you can actually take either the analysis between all the atoms or

interatomic distances, or you can limit the distance between particular atoms. If there is an A element, you can actually get the nearest neighbor distribution of A to A or B to B.

And these distributions are generated to analyze the distance separating each atom and its closest nearest neighbor. Okay, so it generates the distribution to analyze the distances separating each atom from the closest nearest neighbor of that particular atom. So here, I am showing you an example.

So you can see that in this particular diagram, this is the distribution of the first nearest neighbor. as compared to the first nearest neighbor with a random distribution. Okay, so this is computed by the binomial distribution function. Okay, so this is the experimental one, which is actually present in the dataset. First nearest neighbor means these distributions consider that this particular red

So there are two kinds of atoms. One is red. One is red and one is blue. So blue is the blue number of atoms in the matrix or the blue atoms. Okay.

And the red color atoms assume that red are the tin and blue are copper. Fine. And you can see that this red color atom, if you draw a nearest neighbor, the first nearest neighbor, then you can see the distance between this particular, this atom and this first nearest neighbor atom. Okay.

And similarly, for each atom, it is calculated. For example, if you take this as a center, the first nearest neighbor is this particular atom. If you take this as a center, the first nearest neighbor will come with this atom. And these distances are actually measured and they're plotted as a frequency versus the nearest neighbor distance. So you can see that the pink color one is the random, but the red color one is the experimental one.

And you can see that these red color atoms, the tin atoms, it evolves as a highest frequency. The frequency increases at this particular nearest neighbor distance, which might be around 0.3. nanometer so at so most of the most of the atoms most of the atoms are at a distance of nearest to the 0.3 nanometers okay and the frequency is very high okay so and if you see that as the nearest neighbor distances increases you can see the distribution also becomes less okay and you will have a hump here it means that these number of atoms which is of 200 frequency these number of atoms has a distribution distance between 1 and 3

the nearest neighbor distance between 1 and 3 okay so this is how we can calculate the first nearest neighbor distribution and this distribution distance is the is nothing but the these are compared with the random distribution which is done by the binomial distribution function. It means that the binomial distribution function indicates the atoms are randomly distributed in a solid solution. Okay?

So, this particular example discusses the copper atom. Here, this is the tin atom. Here, you can see the copper atoms. So, it is the copper atoms with random distribution and copper atoms with experimental first nearest neighbor. So, for each atom, the nearest neighbor is calculated.

So, for this particular case, the nearest neighbor will be this; for this particular case, the nearest neighbor will be this. So, this is, and you can see that this is the random part, and this is your experimental part. You can see that it is shifted toward the left. It means that there is a tendency of segregation, a tendency of clustering of these copper atoms. Similarly, for the tin, it is also similar.

So, it means that the tin atoms in this solid are not randomly arranged. They are clustering together; they are coming together at certain locations, or we can call it as they are aggregated in the solid solution. Okay. So that's why there is a shift in the peak. There is a shift in the peak either for the copper or for the tin.

Okay. And you can see that, as I told you, this tin, this particular hump, it means that a very small number of tin atoms are far apart, approximately of sizes between 2 to 3 nanometers, and these numbers are very small. However, the numbers are very high, with a frequency of more than 600, where the distance between the tin atoms is 0.3 nanometers. And this distribution not only

for the first nearest neighbor, it can also be measured, it can also be plotted for the 5th nearest neighbor, 10th nearest neighbor, or 100th nearest neighbor. So depending on what type of nearest neighbor atoms you want, you can actually plot for the 1st, 5th, 10th, and 100th nearest neighbors. So we call it a kth nearest neighbor distribution. It means the examination of interatomic distances between each atom to its second nearest neighbor,

the third nearest neighbor, or the kth nearest neighbor atom. So, here I will show you the distribution of copper for the fifth nearest neighbor atom.

distribution as compared to the fifth nearest neighbor binomial random distribution. Fifth nearest neighbor means if you see this particular copper atom blue colour you can see that the closest is one, second closest is two, third closest is three, this is the fourth, and this is the fifth. So the distance between the origin atom and the fifth atom That particular distance is measured and is being plotted in the frequency versus KNN nearest neighbor distance. A similar analysis will be done for each atom.

So this is the first, second, then you will have a third, it might be a fourth, and this might be the fifth. So this particular distance... is measured and plotted as a frequency between the k and n nearest neighbor distances. Similarly, you can actually plot the 10th nearest neighbor distribution. Okay.

Okay, so with this, now the next is the cluster identification. Now we will talk about—we have talked about the frequency distribution—and you can see how these are compared with the random binomial distribution, and it shows that these some of the atoms are aggregated together or segregated together. So based on this, from the atom probe reconstruction, what we can do is cluster analysis, and this cluster analysis is done by several algorithms. which are based on certain assumptions.

certain assumptions. So what is the concept of cluster analysis? As the name suggests, these are the clusters. It means the identification of clusters or the agglomerated atoms in the dataset, and it can be done by using certain assumptions. So if there are two or more atoms

are associated or nearby one another, are associated with one another, then you can use several algorithms which can differentiate between The clusters or which can differentiate between the precipitates or phases. Okay. So, based on the algorithm or certain assumptions. So, the most common algorithm which is used is the maximum separation distance.

This is the most common assumption or most common algorithm used. Which is used to identify the clusters in the data set. What is the maximum separation algorithm? It is based on the assumption that the distance between the two solute atoms, the distance between the two solute atoms, between two solute atoms that are clustered together is smaller than the distance between the two solute atoms which are found anywhere in the specimen.

Smaller than the distance between the two solute atoms which are found anywhere in the data set. So, there is a term which is called D-maximum. Okay? So, if the distance between the two solutes which is present in the data set is less than the d max, then these are termed as a cluster, a cluster of these atoms. Okay?

So, it means that the algorithm, what it will do is, it will group, it will group together the atoms. that are separated by less than the distance between the atom, between the Dmax, less than the Dmax. That is, if there is a solute which is P and Q are separated by a distance dP and dPQ, then the cluster, the cluster algorithm will find out the clusters which are having a dPQ Less than the d max. These are identified as a cluster, one cluster.

So it means that here the p and q are considered in the same cluster. So what is the parameter which controls in this assumption or in this algorithm? The parameter which is controlled is d max. So with this parameter what we can do is we can filter the clustered from non-clustered atoms. On the basis of may be first nearest neighbor.

Okay. So here I am showing you an example. Okay. Here you can see that these are different kinds of atoms which are gray and blue. And you can see that if you assume that d max is chosen for the maximum separation distance between the

between the two solutes or between the two atoms, then you can see that if you point out this particular atom and if you draw a distance of d max and a circular or a sphere around this particular atom, whatever the atoms which is having a distance less than the d maximum, then these atoms are together called a one cluster. Similarly, this particular region—similarly, this particular region, okay? So, you can actually draw several clusters together. With this assumption that there is a Dmax.

Okay? So, a sphere of Dmax is drawn around each atom of the solute. If another solute atom occurs within this Dmax—if it is present within this Dmax—then we call it one cluster. Okay? So, this is one assumption.

And this process is repeated for each solute atom. This process is repeated for each solute atom. Finally, the clusters are assembled by connecting the atoms together via a network of common associations. So, you can see in the scene. For each atom, if you calculate, then you will have a cluster of these common associations.

Okay? So, these are termed as one cluster. Okay? But you know that in atom probe, we have a local change in density, atomic density, or also the APT resolution. These are dependent upon the trajectory aberrations and local magnification effects that might directly provide errors in the results or identification of the clusters.

So this particular method is based on the Dmax, the distance. So this analysis is based on the Dmax. Which is the distance, correct? Now, we can also identify the clusters based on the composition or the concentration. Okay?

So here, in this particular case, we can also do the cluster identification based on the composition, meaning composition-based. Okay? And here, instead of distance, Instead of distance, solute concentration or composition is used as a parameter to define the cluster. So a radius—usually what has been done is similar to these atoms which are present.

A radius of Dmax is prepared in the earlier method. But in this case, the sphere of radius Rmax is constructed around each solute atom, and the concentration Cs of solute atoms inside the volume of that particular R. R max is measured. If the concentration which is measured, C, if the concentration measured is greater than CS minimum, then the atom is considered as a clustered atom. So, this is another method where you can actually use the solute composition as a parameter to define the cluster.

So, the solutes and this rmax, the distance, the rmax for each solutum is plotted such as there will be a sphere and inside this sphere, you will have some solute concentration. If that particular solute concentration is greater than this Cs minimum, then we can tell it as

these solute atoms belongs to that, belongs to the cluster atom. So, this is one thing. Another thing is the radial distribution function.

The third thing is radial distribution function. We call it as a RDF analysis. RDF analysis is used to identify the structure in a material by examining the average local neighbourhood as a function of distance extended radially from each atom in the data set. So, RDF usually builds the description of local atomic neighborhood and provides average, it provides, so usually it provides the average radial

concentration profiles starting from each and every detected atom of chosen element of solute or the chosen atom. It actually provides the average radial concentration profiles. So, for example, if there is a probability of the probability density of finding an atom J at a radius of x, y and z coordinate. So, when an atom I is at the origin then the RDF is calculated as CER divided by C0 which is equals to NER divided by NE and R divided by C0.

So what are these terms? These terms are nothing but CER is the atomic composition of element E at a distance r. Atomic composition of element E at a distance r and C0 is the average composition of E. in the analyzed volume. Similarly, nR is the total number of atoms at a distance R, R from the selected or the chosen atom. And nER is the total number of E atoms of that particular element, atoms at a distance R.

So this is the third way of identifying the cluster energy which is called a radial distribution function. And this term radial distribution function and the cluster identification by the solute concentration and also the cluster identification based on the maximum suppression distance. These are the three basic methods to identify the cluster of any tomographic data. So after few classes what we do I will also cover some standard data sets and I will show you how these clusters is identified through these algorithms in the standard software. So with this I will end this class now.

And we will continue with this topic with certain other distribution maps which are very important to understand the crystallography or the interatomic spacings. And also we will cover the effect of detector efficiency on these cluster analysis. We will basically cover these topics. Thank you.