1 Introduction

The reactivity of materials always depends on the distribution of important groups of atoms. Here we will refer the important groups of atoms as clusters. Given the bulk material property, such as the element ratio, we want to compute the distributions of the clusters in order to quantitatively understand the reactivity. The ultimate goal of this cluster counting module is to statistically compute the cluster distributions for different conditions, such as different crystal structures, different rules for atom locations, and different counting and titration priorities. This ultimate goal can be further divided into four specific goals. First, given a crystal structure, we want to understand the lattice and be able to identify distinct clusters. Next, we want to be able to count different type of clusters given a structure configuration (which specifies the atom type at each site). Thirdly, we would like to generate random structure configurations for any given element ratio. Last, we want to titrate clusters one by one to avoid double counting clusters that share sites with each other. This report will cover the methods to achieve each goal and the cluster distribution results for Chabazite and Ferrierite for different Si/Al ratios under different rules.

2 Methods

2.1 Identify Distinct Clusters (understand the lattice/structure)

Clusters differ from each other based on atom distances and symmetries. To distinguish between different clusters, we took advantage of the Alloy Theoretic Automated Toolkit (ATAT) developed by Axel van de Walle. In ATAT, the corrdump program takes lattice parameters and site positions as input, determines the space group of the lattice, and find all symmetrically distinct clusters based on the space group. When analyzing distinct clusters, corrdump only count for sites which are possible for at least two types of elements. The sites which can accommodate only one type of element will help with symmetry analysis.

In addition, we also used ATAT corrdump program to generate a cluster list for a given structure, which can be further used for counting and titrating clusters. The cluster list contains the site coordinates for all the clusters in each cluster type for the given structure.

2.2 Count Clusters

There are usually two requirements for the clusters that we want to count: 1) having certain element at each site, and 2) being certain cluster type, such as Al-Al pairs in 6-membered-rings. Other rules, such as excluding certain cluster types, might also exist. For example, when there are 3 Al atoms not adjacent with each other in a 6-membered-ring, we may not want to count any Al-Al pair in this 6-ring. The codes to complete this task contain 2 parts.

1. If there is any excluding cluster type, the first step is to create a set of all the clusters that we want to exclude in one super cell. We can go through the cluster list for each excluding type, add every cluster as well as its sub-clusters into the excluding cluster set. For example, 3Al in 6-membered-rings is a particular 3-body cluster type. We will add each 3-body cluster in this type, as well as each 2-body cluster and 1-body cluster in this 3-body cluster, to the excluding set.

2. Then we will go through the cluster list for each of the cluster type that we are interested in. For every cluster in this type, if it has exactly the required element at each site(given by the structure configuration), and it doesn't belong to excluding cluster set, we will add the count for its cluster type by 1.

2.3 Randomly generate structure vectors with/without rules

When researchers design functional materials, the element ratio is usually the first property to control. Different element ratios will result in different cluster distributions. Moreover, for a given element ratio, the cluster distribution may vary with the distribution of elements within the structure. Thus the third goal here is to randomly generate numbers of structure configurations for a given element ratio to statistically analyze the cluster distributions. When structure configuration is generated, certain rules may apply. For example, Löwenstein's rule requires no first nearest neighbors for Al. Another rule is the probability of certain atoms at different sites may not be the same. For Ferrierite, there are four types of T-sites. The probability of Al at different T-sites may vary with the synthesis process. To generate random configurations with/without rules, we need 3 steps.

1. Randomly initialize a structure configuration(structure vector) which has the required atom ratios and the required site probability. The structure vector is a list of numbers (0 or 1). Each number represents the type of atom on the corresponding site. For example, we have a structure vector of [0,1,1,0] and there are two types of elements for each sites, Si(represented as 0) or Al(represented as 1). This structure vector represents a structure with 4 sites. The first site has the 0th element, which is Si, the second site has the 1st element, which is Al, the third site has Al and the fourth site has Si.

2. If there are any penalized cluster types, create a penalty dictionary. Assign a penalty factor for each penalized cluster type. The overall penalty for a structure configuration is the summation of the number of penalized clusters times the penalty factor for all penalized types.

3. Randomly swap atoms at two different sites. If different site types has different probabilities, only swap atoms on two sites with the same type. If the swap results in a not greater overall penalty, swap the atoms; otherwise, the swap probability follows the Boltzmann function of the penalty difference. Keep swapping until the overall penalty becomes stable.

2.4 Titrate Clusters

The ultimate goal of the cluster distribution analysis is to understand the reaction activity with the material structure. When two clusters share sites with each other, we count both clusters if they both meet the requirements. While in reactions, if one cluster reacts, the atoms in the cluster will participate in the reaction. The clusters that share site with the reacted cluster cannot further react. So we need a titration function to titrate the clusters one by one. After each titration, we mark the sites in the titrated cluster as used and won't titrate clusters sharing sites with the titrated cluster in the future. Another thing that we need to consider is that, not all the clusters types have the same reaction priority. Some types may react first and only all the clusters in these types are all used up, the other types of clusters can start to react. To complete this task, we need 3 steps.

1. If there is any excluding cluster type, create an excluding cluster list as in the counting function. Go through the cluster list for each excluding type, add every cluster as well as its sub-clusters into the excluding cluster set.

2. Define the titration group including the cluster types with the highest priority. Create an available cluster list with all the clusters that have the required atoms, being the type in the titration group, and not in the excluding cluster set. Then we start to titrate the clusters

one by one. After each titration, add 1 for the cluster type that the titrated cluster belongs to, change the sites in the titrated cluster to used sites in the structure configuration, and remove all the clusters that share sites with it from the available cluster list. Keep titrating until no cluster is in the available list.

3. Define another titration group for the cluster types with the second highest priority. Create an available cluster list as in step 2 based on the structure configuration that already cross out all the used sites in step 2. Then titrate as in step 2.

4. Keep repeat step 3 until all the titration groups have been titrated.

3 Results

3.1 Cluster types in Chabazite

In this report, the cluster types are named by two numbers by a hyphen. The first number represents the number of sites within the cluster, and the second number represents the order of the cluster type based on the maximum distance between 2 atoms within the cluster. The order for 1-body cluster is given arbitrarily by corrdump in ATAT. In chabazite, there is only one type of 1-body cluster (also referred to T-site). There are 23 distinct types of 2-body clusters with distances between two atoms not greater than 8.11A, including all the 2-body clusters shorter than 4NN in 8-membered-rings. The detailed description of the 23 clusters are listed in the Table 1.

Table 1: Two-body Clusters in Chabazite

cluster type	d(Å)	multiplicity/T-site	description
2-1	3.0919	0.5	1NN in 4/6 membered ring
2-2	3.1040	0.5	1NN in 4/6 membered ring
2-3	3.1042	0.5	1NN in 4/8 membered ring
2-4	3.1121	0.5	1NN in 4/8 membered ring
2-5	4.3856	1	2NN in 4 membered ring
2-6	4.3869	0.5	2NN in 4 membered ring
2-7	5.3660	1	2NN in 6 membered ring
2-8	5.5341	1	2NN in 8 membered ring
2-9	5.7360	1	2NN in 8 membered ring
2-10	6.1961	0.5	Across multiple rings
2-11	6.1961	0.5	3NN in 6 membered ring
2-12	6.2022	0.5	Across multiple rings
2-13	6.3423	1	Across multiple rings
2-14	6.9302	0.5	Across multiple rings
2-15	7.2008	0.5	3NN in 8 membered ring
2-16	7.2152	1	Across multiple rings
2-17	7.3315	1	Across multiple rings
2-18	7.3377	1	3NN in 8 membered ring
2-19	7.4789	0.5	3NN in 8 membered ring
2-20	7.8414	0.5	4NN in 8 membered ring
2-21	8.0974	0.5	4NN in 8 membered ring
2-22	8.0985	1	Across multiple rings
2-23	8.1089	0.5	Across multiple rings

As we can see in the table, there is only one cluster type(2-7) of 2NN and only one cluster type(2-11) of 3NN in 6-membered-rings for chabazite.

3.2 Cluster types in Ferrierite

Ferrierite has a more complex structure. It has four types of 1-body clusters(T-sites). There are 30 distinct types of 2-body clusters with distances between two atoms not greater than 6.58A, including all the 2-body clusters shorter than 3NN in 6-membered-rings.The detailed description of the 30 clusters are listed in the Table 2.

Table 2: Two-body Clusters in Ferrierite

cluster type	d(Å)	multiplicity/T-site	description
2-1	3.1097	0.4444	1NN in 5/10 membered ring
2-2	3.1129	0.2222	1NN in 5/8 membered ring
2-3	3.1144	0.1111	1NN in 5 membered ring
2-4	3.1422	0.4444	1NN in 5/8 membered ring
2-5	3.1655	0.2222	1NN in 5 membered ring
2-6	3.1685	0.2222	1NN in 5/6 membered ring
2-7	3.2064	0.1111	1NN in 6 membered ring
2-8	3.2159	0.2222	1NN in 5 membered ring
2-9	4.4266	0.1111	2NN in 5 membered ring
2-10	4.4281	0.2222	2NN in 5 membered ring
2-11	4.5930	0.4444	2NN in 5 membered ring
2-12	4.6599	0.4444	2NN in 5 membered ring
2-13	4.8211	0.4444	2NN in 5 membered ring
2-14	4.8243	0.4444	2NN in 5 membered ring
2-15	4.8425	0.4444	2NN in 5 membered ring
2-16	4.8437	0.4444	2NN in 5 membered ring
2-17	4.9118	0.4444	2NN in 5 membered ring
2-18	5.1158	0.4444	2NN in 5 membered ring
2-19	5.2933	0.4444	2NN in 5/10 membered ring
2-20	5.3833	0.2222	2NN in 6 membered ring (T2-T4)
2-21	5.5937	0.4444	2NN in 8/10 membered ring
2-22	5.6045	0.4444	across multiple rings
2-23	5.7326	0.1111	2NN in 6 membered ring (T2-T2)
2-24	5.7727	0.2222	2NN in 10 membered ring
2-25	5.7748	0.4444	2NN in 8 membered ring
2-26	5.7885	0.2222	across multiple rings
2-27	5.9070	0.0556	3NN in 6 membered ring (T4-T4)
2-28	6.4278	0.4444	across multiple rings
2-29	6.5585	0.2222	across multiple rings
2-30	6.5684	0.1111	3NN in 6 membered ring (T2-T2)

There are two cluster types (2-20 and 2-23) of 2NN, and two cluster types (2-27 and 2-30) of 3NN in 6-membered-rings.

3.3 Al-Al pairs in 6-ring in Chabazite

The titration results of Al-Al pairs for different Si/Al ratios under different rules in Chabazite are shown in Figure 1. For each set of rules, 10,000 structure configurations were generated with the Monte Carlo swap algorithm. Each configuration has 972 T-sites. Here, the Al-Al pairs include 2NN(cluster type 2-7) and 3NN(cluster type 2-11) in 6-rings. These two cluster types are titrated with equal priorities. The three lines are results based on different rules. All of them follow the configuration rule of no 1NN(Löwenstein Rule). For the concern about 3 non-adjacent Al in 6-rings, we performed three sets of rules, 1)doing nothing about them, 2) allowing 3 non-adjacent Al in 6-rings in structure configurations while excluding them in the titration process (green line), and 3) avoiding 3 non-adjacent Al in 6-rings in the generation of structure configurations (red line). We can see that, the total number of Al-Al pairs after we excluded the 3Al in 6-rings(green line) decreases compared with that of no titration rules (blue line). Each 3 non-adjacent Al in 6-rings will result in one 2NN (cluster type 2-7) cluster in the titration. The excluding of 3 Al in 6-rings should bring down the total number of Al-Al pairs. The total number of Al-Al pairs after we penalizing 3 non-adjacent Al in 6-rings (red line) increases compared that with no penalization for 3 Al in 6-rings (blue line). When the 3 non-adjacent Al in 6-rings is penalized, the 2NN in 6-rings should decrease. The increase of the total number of Al-Al indicates an increase in the 3NN in 6-rings.

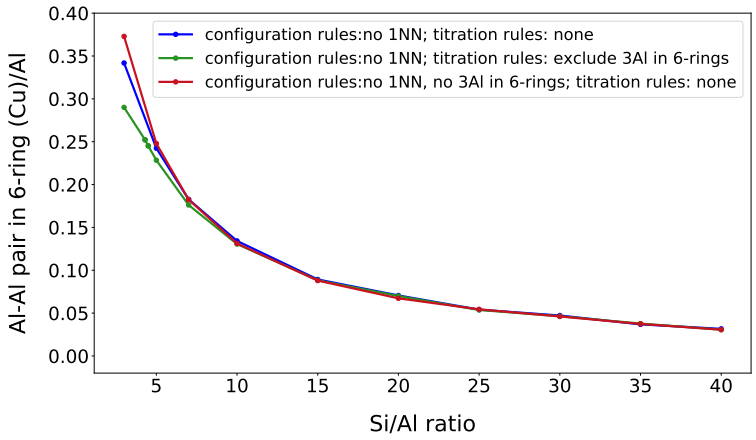

Figure 1: Titration results of Al-Al Pairs in 6-rings for Different Si/Al Ratio in Chabazite

3.4 Al-Al pairs in 6-ring in Ferrierite

The titration and counting results of Al-Al pairs for different Si/Al ratios in Ferrierite under Löwenstein's rule are shown in Figure 2. There are 2 types of 2NN and 2 types of 3NN in 6-rings for Ferrierite. Based on free energy calculation, only T2-T4 2NN(cluster type 2-20) and T4-T4 3NN(cluster type 2-27) are titrated/counted. In titration, they have the equal priorities. We generated 10,000 random structure configurations with the Monte Carlo swap algorithm. Each configuration contains 972 T-sites. For each configuration, we count once and titrate 1,000 times. The mean value for the 10,000 configurations are used as the final results.

When two clusters share sites with each other, if one is titrated, the other one will become invalid. Thus the titrated Al-Al pairs/Al are lower than the counted ones. Both T2-T4 and T4-T4 decrease with the increase of Si/Al ratio in counting and titration.

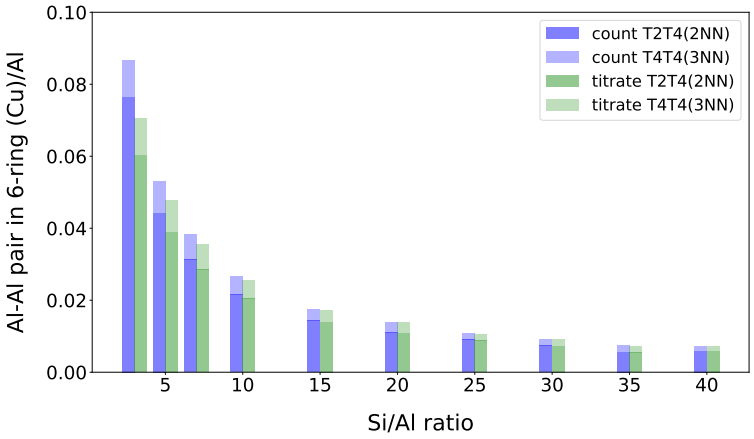

Figure 2: Count & Titration results of Al-Al Pairs in 6-ring for Different Si/Al Ratio in Ferrierite (configuration rules: no 1NN; titration rules: T2-T4 and T4-T4 with equal priorities)

Figure 3 shows the titration results of Al-Al pairs in 6-rings for Ferrierite when Si/Al ratio is 15. Based on the observations in experiments(should double check details with Sichi), the distribution of Al on different types of T-sites are not completely random. When Si/Al ratio is 15, around 85% of Al will be on T4 sites. In addition to Löwenstein's' rule, we add a second rule (shown as A in blue) which controls the probability of Al on T4 sites from 0.75 to 0.95 during the generations of random configurations, in order to evaluate the effects of different T4 probabilities on the Al-Al distributions. As shown with the blue bars, T2-T4(2NN) in 6-rings (darker blue) decreases while T4-T4(3NN) in 6-rings (lighter blue) increases with the increase of Al on T4 sites. The total number of Al-Al pairs increase with the increase of Al on T4 sites. Another observation in experiments is that there is no T4-T4 pair in 6-rings. So we add another rule which penalizes T4-T4 pairs in 6-rings. As shown in green bars, the total number of Al-Al pairs greatly decreases compared with that without the no T4-T4 rule. When the T4 probability is not greater than 0.85, there is no T4-T4 (shown in darker green) cluster. However, when T4 probability is greater than 0.85, T4-T4 pairs start to show up. It's not possible for Ferrierite structure configurations to have no 1NN and no T4-T4 with Si/Al ratio of 15 when T4 probability is larger 0.85. This is one possible reason that 85% of Al are on T4-sites in experiments. T2-T4 (shown in lighter green) continues to decrease with the increase of T4 probability. The total number of Al-Al pairs in 6-rings is 0.023/Al when the configuration rules are no 1NN, T4 probability of 0.85 and no T4-T4 in 6-rings for Si/Al ratio of 15. This is close to the experimental observation of Cu/Al in Ferrierite with Si/Al ratio of 15.

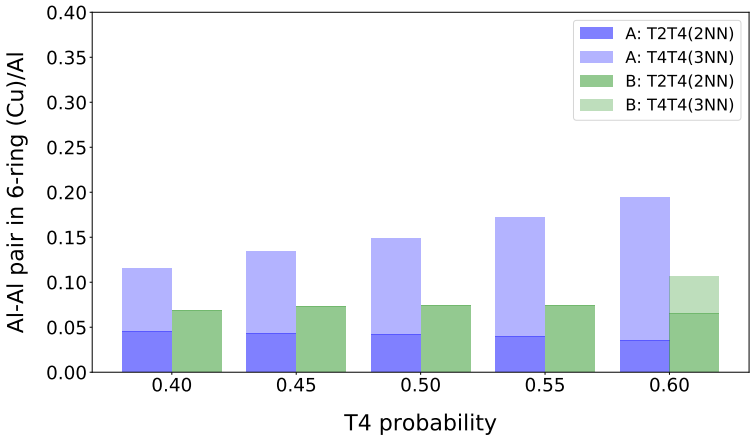

Figure 3: Titration Results of Al-Al Pairs in 6-ring for Ferrierite (Si/Al Ratio = 15). (A) configuration rules: no 1NN, and certain T4 probabilities; titration rules: T2-T4 and T4-T4 with equal priorities. (B) configuration rules: no 1NN, certain T4 probability, and no T4T4(3NN) in 6ring; titration rules: T2-T4 and T4-T4 with equal priorities.

Figure 4 shows the titration results of Al-Al pairs in 6-rings for Ferrierite when Si/Al ratio is 9. When Si/Al ratio is 9, around 50% of Al are observed on T4 sites. Shown as A in blue, we add the second rule (shown as A in blue) which controls the probability of Al on T4 sites from 0.4 to 0.6 in the random configurations in addition to Löwenstein's rule. As shown with the blue bars, we make the same observations with the Si/Al ratio of 15. T2-T4(2NN) in 6-rings (darker blue) decreases while T4-T4(3NN) in 6-rings (lighter blue) increases with the increase of Al on T4 sites. The total number of Al-Al pairs increase with the increase of Al on T4 sites. Again, we add another rule which penalizes T4-T4 pairs in 6-rings. As shown in green bars, the total number of Al-Al pairs greatly decreases compared with that without the no T4-T4 rule. When the T4 probability is not greater than 0.55, there is no T4-T4 (shown in darker green) cluster. However, when T4 probability is greater than 0.55, T4-T4 pairs start to show up. It's not possible for Ferrierite structure configurations to have no 1NN and no T4-T4 with Si/Al ratio of 9 and T4 probability of 0.55. T2-T4 (shown in lighter green) doesn't change much with the increase of T4 probability. The total number of Al-Al pairs in 6-rings is 0.074/Al when the configuration rules are no 1NN, T4 probability of 0.85 and no T4-T4 in 6-rings for Si/Al ratio of 9. This is close to the experimental observation of Cu/Al in Ferrierite with Si/Al ratio of 9.

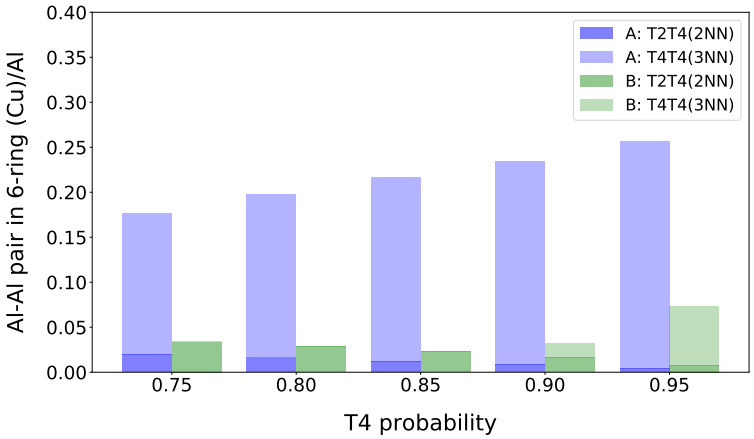

Figure 4: Titration Results of Al-Al Pairs in 6-ring for Ferrierite (Si/Al Ratio = 9). (A) configuration rules: no 1NN, and certain T4 probabilities; titration rules: T2-T4 and T4-T4 with equal priorities. (B) configuration rules: no 1NN, certain T4 probability, and no T4T4(3NN) in 6rings; titration rules: T2-T4 and T4-T4 with equal priorities.

4 Conclusions

Different configuration rules have been tested. Löwenstein's rule(no 1NN) and no 3 non-adjacent Al in 6-rings were used for Chabazite. Löwenstein's rule(no 1NN) , constant Al probability on different T-sites and no T4-T4 in 6-rings were used for Ferrierite.