Title
Potential Energy Surface Exploration of Metal Catalytic Clusters

Author
Zhai, Huanchen

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Potential Energy Surface Exploration of Metal Catalytic Clusters

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Chemistry

by

Huanchen Zhai

2019

ABSTRACT OF THE DISSERTATION

Potential Energy Surface Exploration of Metal Catalytic Clusters

by

Huanchen Zhai

Doctor of Philosophy in Chemistry

University of California, Los Angeles 2019

Professor Anastassia N. Alexandrova, Chair

Metal sub-nano clusters are important materials for catalysis of chemical reactions such as dehydrogenation of hydrocarbons. However, their potential energy surface (PES), which is responsible for explaining relative stability of different cluster geometries, can be complicated. The complexity can be greatly reduced by describing the PES by low-energy isomers, which can be found by the standard mathematical method called global optimization. We have proposed global optimization acceleration schemes using Force Field (FF) fitting, or deep neural network (DNN) fitting, for gas phase metal clusters. Both these models can be trained to give an approximation to

the PES at the density functional theory level. For surface supported clusters, we found that Basin Hopping (BH) global optimization can usually give satisfactory results. To address the difficulty in performing structural search for surface supported clusters at high hydrogen coverage, we proposed a revised BH approach with core-shell separation scheme. The application of this new approach shows that the cluster shape can be very different from that of their adsorbate-free counterparts.

After the structure of isomers is found, it is also necessary to define the geometry similarity between two isomers with the same chemical formula. Structure similarity measurement is an important part of any PES exploration techniques. We have investigated existing algorithms and introduced two new algorithms, namely, atomic matching based on depth-first search and bipartite model atomic matching, for gas phase and surface supported clusters, respectively.

Additionally, isomers may interconvert across barriers, i.e., exhibit fluxionality, during catalysis. To study the big picture of the fluxional behavior for surface supported clusters, we model such process as isomerization graph using bipartite matching algorithm, harmonic transition state theory (HTST), and paralleled nudged elastic band (NEB) method. Detailed inspection shows that, at temperatures typical for catalysis, the cluster geometry changes frequently within several regions in the graph, while transition across regions is less likely. This local fluxionality picture provides a new perspective on understanding finite-temperature catalytic processes.

The dissertation of Huanchen Zhai is approved.

Benjamin Joel Schwartz

Philippe Sautet

William M. Gelbart

Anastassia N. Alexandrova, Committee Chair

University of California, Los Angeles

2019

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEGMENTS

First I would like to thank my advisor Professor Anastassia N. Alexandrova, for providing me with the opportunity to study at UCLA and all kinds of support that I need during my PhD career. As a successful young scientist, she is knowledgeable and professional on leading and performing computational chemistry researches. Her guidance always gives me a clear direction and encourages me to think creatively. She is also attentive to my future career, offering me many chances to stand out in the community and always being respectful of my own decision.

I would also like to thank my committee members, Professor Philippe Sautet, Professor Benjamin J. Schwartz, and Professor William M. Gelbart. Their insightful comments and advices on my PhD work and career plan are important for me to accomplish the degree. In particular, the weekly joint group meeting of Prof. Sautet group and Prof. Alexandrova group becomes the most memorable event during my PhD careers, from which I have learned many valuable scientific ideas.

I greatly appreciate the help that I got from all members in Prof. Alexandrova group, especially graduate students Mai-Anh Ha, Borna Zandkarimi, postdoc Dr. Han Guo, and visiting student Jiyuan Liu. The conversation with them offered me many useful insights and ideas.

My research would be impossible without the funding from Air Force Office of Scientific Research, the computational resources from Department of Defense High Performance Computing Modernization Program (HPCMP), Extreme Science and Engineering Discovery Environment (XSEDE), and UCLA IDRE.

Sections 1.2.1, 1.2.2, 1.3, 1.6, 3.1.1 and part of Conclusions are adapted with permission from the following journal article. Copyright 2019 American Chemical Society.

Zhai, H.; Ha, M.-A.; Alexandrova, A. N. AFFCK: Adaptive Force-Field-Assisted Ab Initio Coalescence Kick Method for Global Minimum Search. *J. Chem. Theory Comput.* **2015**, *11*, 2385–2393.

Sections 1.1, 1.2.5, 1.4, 1.5, 1.7, 1.8, 3.1.2 and part of Conclusions are adapted with permission from the following journal articles. Copyright 2019 American Chemical Society.

Zhai, H.; Alexandrova, A. N. Ensemble-Average Representation of Pt Clusters in Conditions of Catalysis Accessed through GPU Accelerated Deep Neural Network Fitting Global Optimization. *J. Chem. Theory Comput.* **2016**, *12*, 6213–6226.

Zhai, H.; Alexandrova, A. N. Fluxionality of Catalytic Clusters: When It Matters and How to Address It. *ACS Catal.* **2017**, *7*, 1905–1911.

Sections 2.3, 3.2, Chapter 4, and part of Conclusions are adapted with permission from the following journal article. Copyright 2019 American Chemical Society.

Zhai, H.; Alexandrova, A. N. Local Fluxionality of Surface-Deposited Cluster Catalysts: The Case of Pt_7 on Al_2O_3. *J. Phys. Chem. Lett.* **2018**, *9,* 1696–1702.

Sections 2.1, 2.2, 2.4 and part of Conclusions are adapted from the following journal article. Copyright 2019 Wiley-VCH.

Zhai, H.; Sautet, P.; Alexandrova, A. N. Global Optimization of Adsorbate Covered Supported Cluster Catalysts: the Case of $Pt_7H_{10}CH_3$ on α-Al_2O_3. *ChemCatChem* **2019**, accepted.

VITA

Education

Sept 2017	Advanced to PhD candidate, UCLA
Mar 2017	Master of Science in chemistry, UCLA
2011 — 2015	Bachelor of Science in physics, Shandong University, China

Publications

Huanchen Zhai, Philippe Sautet and Anastassia N. Alexandrova. *ChemCatChem* (2019): accepted. https://doi.org/10.1002/cctc.201901830

Huanchen Zhai, Francisco Muñoz and Anastassia N. Alexandrova. *Journal of Materials Chemistry C* 7 (2019): 10700-10707.

Huanchen Zhai, and Anastassia N. Alexandrova. *The Journal of Physical Chemistry Letters* 9 (2018): 1696-1702.

Huanchen Zhai, and Anastassia N. Alexandrova. *ACS Catalysis* 7 (2017): 1905-1911.

Huanchen Zhai, and Anastassia N. Alexandrova. *Journal of Chemical Theory and Computation* 12 (2016): 6213-6226.

Huanchen Zhai, Mai-Anh Ha, and Anastassia N. Alexandrova. *Journal of Chemical Theory and Computation* 11 (2015): 2385-2393.

Huanchen Zhai, and Shi Ying Lin *Chemical Physics* 455 (2015): 57-64.

Avik Halder, Mai-Anh Ha, Huanchen Zhai, Bing Yang, Michael J. Pellin, Sönke Seifert, Anastassia N. Alexandrova, and Stefan Vajda. *ChemCatChem* (2019): accepted. https://doi.org/10.1002/cctc.201901795

Elisa Jimenez-Izal, Huanchen Zhai, Ji-Yuan Liu, and Anastassia N. Alexandrova. *ACS Catalysis* 8 (2018): 8346-8356.

Eric T. Baxter, Mai-Anh Ha, Ashley C. Cass, Huanchen Zhai, Anastassia N. Alexandrova, and Scott L. Anderson. *The Journal of Physical Chemistry C* 122, (2018): 1631-1644.

Rui Shan Tan, Huanchen Zhai, Wei Yan, Feng Gao, and Shi Ying Lin. *The Journal of Chemical Physics* 146 (2017): 164305.

Rui Shan Tan, Huanchen Zhai, Feng Gao, Dianmin Tong, and Shi Ying Lin. *Physical Chemistry Chemical Physics* 18 (2016): 15673-15685.

Conference Talks

Jun 27 2018	International Conference on Theoretical Aspect of Catalysis 2018
Mar 6 2018	American Physical Society Match Meeting 2018
Nov 18 2016	Workshop at Institute for Pure & Applied Mathematics

Chapter 1

Global optimization of gas phase clusters

Global optimization is a basic computational technique for exploring the Potential Energy Surface (PES) of clusters, or any multivariable function in the most general sense. Typically, the configuration space of catalytic clusters (such as Pt_7) can be very large, which means that a thorough exploration is usually not feasible. However, based on the nature of the system under consideration, different approaches can be developed to make the exploration focus on the most relevant region of configuration space, thus saving a tremendous amount of computational time. In this chapter, we will first talk about the global optimization of the gas phase cluster.

1.1 General procedure of global optimization

Global optimization in a computational chemistry context means finding the lowest energy structure of a certain chemical system, such as a metal cluster. The global minimum structure is regarded as the most important one, but we also care about other low-energy isomers most of the time. The basic way to do this can be described as three steps:

1. Generate random initial structures.

2. Perform geometry optimization to find some local minima. In some methods like Basin Hopping and Generic Algorithm, more new structures can be generated based on previous found local minima. Structures can be discarded if the single point energy or geometry optimization does not converge.
3. Compare the energy of local minima to identify the global minimum and low energy isomers.

Step 1 is often trivial in most global optimization packages. In the cases when the global optimization method does not have a strong dependence on the initial guess, one can use some very simple or even brute-force methods to generate the initial guess. However, there are still some motivations for finding better initial guesses: (i) a reasonable initial guess will greatly save the time-cost of local optimization; (ii) a bad initial guess will possibly cause the non-convergence in electronic structure calculations; (iii) in the case when fitting technique is used in the global optimization, the initial guess will have an influence on completeness of sampling, which means that bad initial guesses can reduce the quality of the fitting and correspondingly the quality of the global optimization. In Section 1.2 we will introduce several approaches for finding good initial structures.

Most global optimization methods differ in step 2, which is the most time-consuming step. Over decades many efficient global optimization techniques have been developed, including Genetic Algorithm (GA) based methods,[1–5] Particle Swarm Optimization (PSO),[6,7] Simulated Annealing (SA),[8] and Basin Hopping (BH).[9,10] Many of these methods incorporate the Density Functional Theory (DFT) level local geom-

etry optimization procedure as an internal step of global optimization, for applications on metal clusters.[11]. However, for medium- to large-sized transition metal clusters, the DFT local optimization step itself is very time-consuming, especially when hybrid functionals are used. In addition, in a high-level parallel programming environment, some of the aforementioned global optimization methods are less favorable due to their evolutionary feature. In Sections 1.3 and 1.4 we will introduce two fitting based global optimization approaches, which can greatly reduce the time required for DFT optimization. An extension of BH method for surface supported system will be discussed in Chapter 2.

Step 3 also seems to be trivial, since only a comparison based on energy should be performed and we immediately find the low-energy isomers. However, it is possible that we will have multiple copies of the same local minima. If the total number of these local minima is not big, we can remove duplicates by visualizing them. But for larger systems, the configuration space can be very big, which means that an automatic and reliable structure similarity measurement scheme is usually required. This topic will be covered in Chapter 3.

Global optimization for large systems (such as the systems with a surface support) can in general be very time-consuming, usually requiring the use of high-performance computing facilities. It is thus important to have an integrated code package including all necessary algorithms, which can automatically search for low-energy isomers for a given system. The package should be able to support high-level parallel computation and the queueing system in any typical supercomputer. An introduction to the

code that we have developed and used for the work in this dissertation is given in Appendix.

1.2 Initial structure generation

One of the major differences between the general global optimization and the global optimization in a computational chemistry context is that for chemical problems some special techniques can be applied to generate starting points very close to the minimum. Then the global optimization in a computational chemistry context will benefit from this in two ways:

1. Since the initial structures are more reasonable, the local and global minima will be found within shorter time.
2. The convergence problem of Self-Consistent Field (SCF) iteration will be alleviated for more reasonable structures.

So the first step of nearly any global optimization method, is to generate some initial structures (atomic coordinates), which will be the starting point of the global optimization procedure. Researchers have developed many different constraints for constructing reasonable initial structures (namely, structures with a chemically accessible geometry). Here we will revisit three kinds of very common strategies, and then propose a new random structure generation method. The three existing methods are CK (Coalescence Kick), RV (Random Vector) and BC (Bound Checking). Based on these methods, we proposed two new algorithms, namely, CK with point group symmetry and Bond Length Distribution Algorithm (BLDA).

Before introducing these methods, we need to define an important quantity, $l_{1,i}$, which is the shortest distance between one selected atom i and all other atoms in the cluster. Note that $l_{1,i}$ can be different for each atom i in the cluster. Formally,

$$l_{1,i} = \min_{j \neq i} \text{distance}\left(\text{atom}_i, \text{atom}_j\right)$$

In addition, if we consider each $l_{1,i}$ as an instance (sample) of a variable l_1, then for each cluster we will have n numbers of l_1, where n is the number of atoms in the cluster. When m structures of a certain cluster are considered, we will be able to have mn instances of l_1. From these mn samples we can estimate the distribution of l_1.

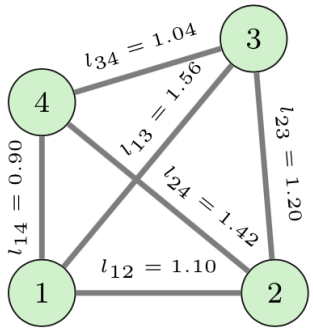


Figure 1.1 All interatomic distances in a four-atom system.

Here we use a four atom system to show how we can find the l_1s. Figure 1.1 shows all six interatomic distances in a four-atom system. Now we can write a symmetric interatomic distance matrix (shown in Figure 1.2). If we sort independently all rows of the matrix, then the values in the second column are l_1s. The values in the third column are l_2s. The quantities l_1 and l_2 will play an important role in understanding the difference in structure generation methods.

$$\begin{bmatrix} l_{11} & l_{12} & l_{13} & l_{14} \\ l_{21} & l_{22} & l_{23} & l_{24} \\ l_{31} & l_{32} & l_{33} & l_{34} \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix} = \begin{bmatrix} 0.00 & 1.10 & 1.56 & 0.90 \\ 1.10 & 0.00 & 1.20 & 1.42 \\ 1.56 & 1.20 & 0.00 & 1.04 \\ 0.90 & 1.42 & 1.04 & 0.00 \end{bmatrix} \Rightarrow \text{sort each row independently} \Rightarrow \begin{bmatrix} 0.00 & 0.90 & 1.10 & 1.56 \\ 0.00 & 1.10 & 1.20 & 1.42 \\ 0.00 & 1.04 & 1.20 & 1.56 \\ 0.00 & 0.90 & 1.04 & 1.42 \end{bmatrix}$$

$$l_1 = \{0.90, 1.10, 1.04, 0.90\}$$

Figure 1.2 Obtaining l_1 from symmetric inter-atomic distance matrix for a four-atom system.

1.2.1 Coalescence kick method

In CK method (shown in Figure 1.3),[12] the positions of all atoms are first randomly generated in a large box, and then checked for connectivity. If the distance between any two atoms is within a predefined value $d_{\max}$, which is normally equal to or a little greater than the sum of covalent radii of the two atoms, the two atoms are considered to be within one fragment. The initial random structures are also checked to make sure that no two atoms are as close as the $d_{\min}$ distance. After that, the coalescence procedure is repeated until all atoms are in one fragment. During the coalescence procedure all fragments are pushed to the center of mass simultaneously by a small step length.

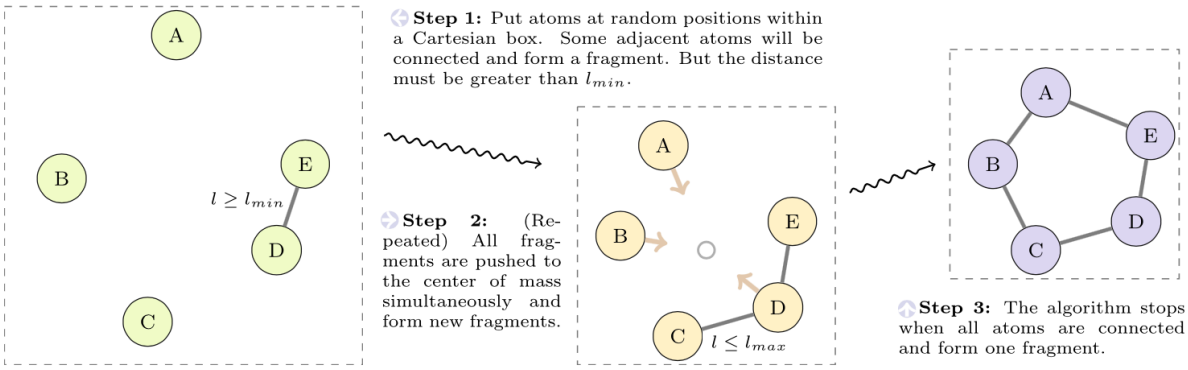


Figure 1.3 The CK method illustrated using a cluster made up of 5 atoms, as an example. A, B, C, D, E denote the five atoms in the cluster. Gray solid line connected atoms are considered in one fragment.

Now in terms of the previously defined l_1, the whole point of CK method is to put the following constraints on initial structures: $d_{\min} \leq l_{1,i} \leq d_{\max}$, for each atom i. This only gives the upper and lower bounds of variable l_1. To obtain more detailed information of l_1, we can generate a number of random clusters using CK method, then plot the distribution of all l_1 numbers.

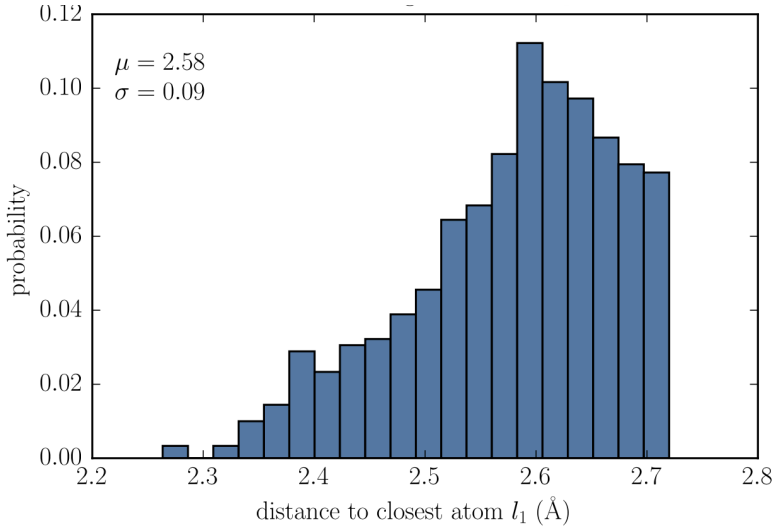


Figure 1.4 l_1 distribution of 300 gas phase Pt_6 clusters generated using CK method.

For example, here we consider a set of configurations of Pt_6 generated using the AFFCK package. In AFFCK, d_{min} is defined to be kd_{max}, where k is a fixed factor. The default value for k is 0.67. And d_{max} is set to be double the covalent radius of Pt. Here we have used the value 1.36 Å for the covalent radius of Pt, which is slightly greater than the normal value 1.28 Å, because d_{max} is used as the upper bound of the Pt-Pt distance. So in our case, d_{max} = 2.72 Å and d_{min} = 1.82 Å. From the distribution graph (Shown in Figure 1.4) we can see that in fact no distance touches the lower bound. In contrast, a large number of distances are close to the upper bound, which gives a discontinuous boundary. This discontinuous feature is one of the most important motivation for us to propose the new structure generation algorithm BLDA (introduced in Section 1.2.5). However, if we look at the mean value and standard deviation, we found that the mean is very close to the expected value 2 × 1.28 = 2.56 Å and the deviation is relatively small. These are good features of CK method.

1.2.2 Coalescence kick method with point group symmetries

The original CK method has a disadvantage that the randomly generated structures seldom possess any symmetry, because the probability that atoms would be put at the position of some symmetry elements is very low. In contrast, some of the local minima may possess symmetry. Hence, we can speed up the geometry relaxation for some of these clusters by preparing symmetric starting structures.

To compensate for this disadvantage, without losing the unbiasedness, one can generate one portion of the structures without symmetry and the remaining structures with all possible symmetry types allowed for a given composition. For each symmetry type, the probability for the atoms lying on each special area – such as symmetry elements, their cross positions, and areas divided by the symmetry elements – are set to be equal. In the process of generating symmetric CK structures, we put only some of the atoms at random position in the Cartesian box, whereas the position of other atoms is calculated by performing symmetry operations. The symmetry is preserved during the coalescence, which then looks like a simple scaling.

Consider the C_{3h} point group symmetry as an example to demonstrate how we can generate random structures with equal probability for atoms lying on each special area. As shown in the Figure 1.5, areas I-VI are equivalent areas, which means that if we have one atom in area I, we should create five replicas of it in areas II-VI, by applying the symmetry operations. So in these areas atoms should be repeated six times. “A” denotes the C_3 rotation axis, where atoms should be repeated two times. “B” denotes the center point, where atoms do not need to be repeated. “C” denotes the

mirror plane, where atoms should be repeated three times. Thus we altogether have four types of special areas.

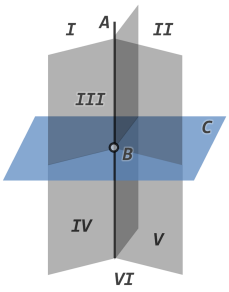


Figure 1.5 Special areas (I to VI), axis (A), point (B) and plane (C) of a system that has C_{3h} point group symmetry. Atoms located at these positions should be repeated certain times.

Given the total number of atoms of one specific atomic type, we should first determine how many atoms will be put in each of these special areas, which forms a combination of integers. If the total number of atoms is not very large, there will be only a few possible combinations. Then we can select one of such combinations randomly each time.

If the system contains different types of atoms, each atomic type should be treated independently when generating initial symmetry. Currently in AFFCK code developed in our group, all but cubic symmetry point groups (T_d, T, T_h, O_h, I_h) are implemented. The user, however, is responsible for determining which points groups are possible for the given cluster composition, and manually list all or some of them in the input file. The user also has to decide the percentage of structures that should have symmetry (one of the listed ones, chosen at random with equal probabilities); the rest of the population will be generated without symmetry constraints. This

method, of course, still leaves a possibility that geometries generated without symmetry will eventually optimize to symmetric ones, and that symmetric structures would optimize to asymmetric ones. A typical setting that we used for the structures definitely having symmetry was 10-20%.

1.2.3 Bond checking method

Another simpler way to generate random structures is by explicitly checking the bound of l_1.[13] For example, this method has been used in the Generic Algorithm implementation in Atomic Simulation Environment (ASE) package.[14] Specifically, the structure is generated by adding one atom at a time to the current structure. At each adding step, the position of one atom is first randomly generated, then checked to make sure that l_1 is within the required range $d_{\min} \leq l_{1,i} \leq d_{\max}$, for all atom i. In the ASE package, the values $d_{\min} = 0.8\ d_{\text{cov}}$ and $d_{\min} = 1.6\ d_{\text{cov}}$ are used, where d_{cov} is the sum of covalent radii.

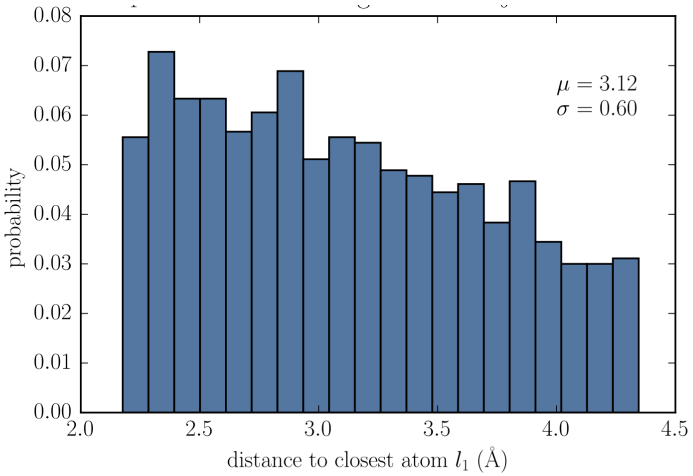


Figure 1.6 l_1 distribution of 300 gas phase Pt_6 clusters generated using BC method.

For example, here we consider a set of configurations of Pt_6 generated using ASE package. The covalent radius of Pt used in ASE package is 1.36 Å.

From Figure 1.6 we can see that this method gives a very diverse distribution with both edges cut sharply. As a result, the average value is away from the sum of covalent radii and the deviation is large. What can be expected is that if a single point energy calculation is performed in these structures, some SCF non-convergence will be observed and the energies of these structures will be higher than CK generated structures.

1.2.4 Random vector method

Another choice is to fix l_1 to a certain number. A convenient choice of this number is the sum of covalent radii of the two involved atoms. In other words, we will have a delta function as the l_1 distribution. This is implemented in the CLUSTER code, developed by Kanters *et. al.*[3]

In practice, we need to consider how to fix l_1 as desired. In the CLUSTER code, the restriction is applied using an algorithm called Random Vector. Specifically, the cluster is created by each time adding one atom to the current structure. The position of that atom is determined by first selecting a random direction, to which the new atom will be placed. After that, we only need to figure out a length. With the length and direction, we can thus construct a random vector, along which the new atom can be placed. The length of the vector is the minimal length such that the all of distances between this new atom and any other existing atoms are greater than the sum of

covalent radii. Since this selected length is the minimal length, we know that l_1 of this new atom must be equal to the selected length. The procedure is shown in Figure 1.7.

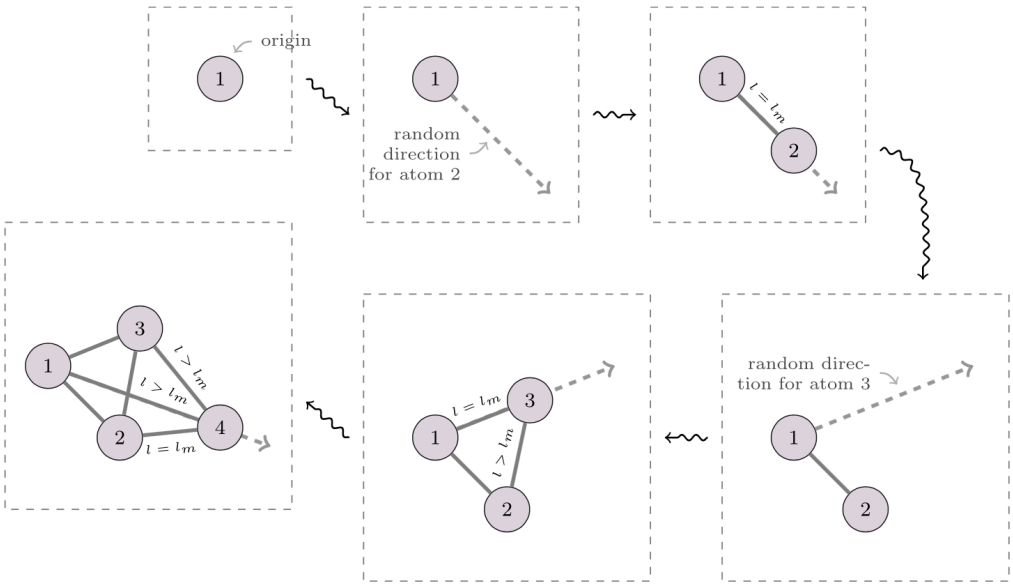


Figure 1.7 The RV method illustrated using a cluster made up of 4 atoms.

1.2.5 Bond length distribution method

We may improve the aforementioned methods by requiring that the initial structures have a similar distribution to that of the expected local minima. Figure 1.8 shows the l_1 and l_2 distribution of low-energy local minima of Pt_6 clusters found using CK method, where l_2 denotes the distance between each atom and its next nearest neighbor. (In Figure 1.8, the DFT optimization is done with Turbomole 6.6, PBE0 hybrid functional, and def2-TZVP basis. The dashed lines are corresponding normal distribution curves fitted based on the mean and standard deviation value of l_1 and l_2 data.)

It is obvious that for Pt_6 clusters both l_1 and l_2 distances approximately obey the normal distribution. In addition, the mean values of l_1 and l_2 (2.50 Å and 2.54 Å) are very close to the sum of covalent radii of Pt atoms (1.28 × 2 = 2.56 Å). In light of this observation, we may require that the l_1 and/or l_2 parameters of initial structures be selected from a normal distribution. We call this new generation algorithm based on this statistical restriction the Bond Length Distribution Algorithm (BLDA).

When only l_1 is restricted, the algorithm is referred to as first-type BLDA (F-BLDA). When both l_1 and l_2 are restricted, the algorithm is referred to as second-type BLDA (S-BLDA). It is clear that when l_2 is restricted to obey a normal distribution, each atom in the generated cluster will have a coordination number of at least two. While this is true for most metal clusters, for other clusters (like Boron clusters [15]) this assertion is very likely to fail. Therefore, we note that F-BLDA can be used for most gas phase atomic cluster system, and S-BLDA is more focused on metal clusters, which tend to have compact geometries in which their delocalized bonding is optimized. In the work presented in this dissertation, S-BLDA has been used as initial structure generation method for Pt_9 and Pt_{13}. An simple extension of it has been used for surface supported Pt_7 clusters.

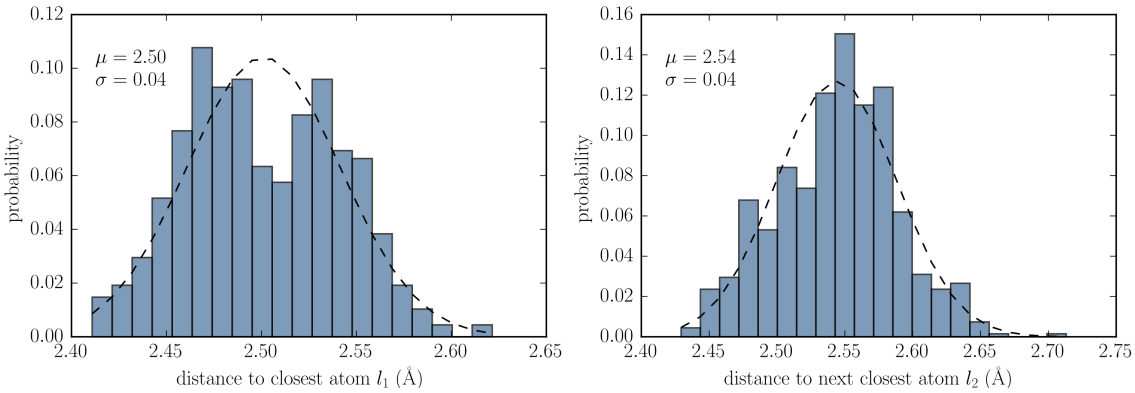


Figure 1.8 l_1 and l_2 distribution of 113 Pt_6 low-energy local minima (energy lower than 1.0 eV with respect to global minimum energy).

Now the most important thing is that we need to find a generation procedure so that the generated structures satisfy the first or second type BLDA requirement. For first type situation, the algorithm is a direct extension of the RV method. In particular, there are three major differences between RV and the first type BLDA.

Random direction generation. In the RV method, the random direction is generated in a Cartesian box. In this way, some particular directions, such as those close to the vertex of Cartesian box, will be sampled more. Ideally we want to sample all points on the unit sphere evenly. One early study [16] used a filtering technique by calculating the distance to the center, to exclude those points outside the sphere but in the cube. We note that this filtering technique introduces extra steps and is thus inefficient.

A more efficient way [17] is first choosing u and v to be random within (0, 1). Then

$$\varphi = 2\pi u, \qquad \theta = \cos^{-1}(2v - 1)$$

gives the spherical coordinates for points uniformly distributed over the unit sphere. The differential element of solid angle is defined as

$$\mathrm{d}\Omega = \sin\theta\,\mathrm{d}\varphi\mathrm{d}\theta \;=\; -\mathrm{d}\varphi\mathrm{d}\cos\theta$$

Therefore, when $\cos\theta$ is picked uniformly in $(-1, 1)$, and φ is picked uniformly in $(0, 2\pi)$, then Ω will be uniformly distributed over $(-2\pi, 2\pi)$.

Starting point of random direction. In the RV method, the starting point of random direction is always fixed at the origin, namely, the position where we put the first atom. This is not ideal, because in this way every time when we add a new atom we are not adding it in all directions with equal possibility. The positions that are far away from the origin will be less sampled. Therefore, in the BLDA algorithm, we choose the starting point of the random vector to be the geometry center of current structure.

Bond Length Determination. In the RV method, the minimal bond length between the new atom and all old atoms is fixed. Before adding a new atom, we need to check the distance between the new atom and each old atom. In the first type BLDA algorithm, for each old atom, we now pick up a value from the required l_1 normal distribution. Instead of a fixed value, this picked value will be the bond length between the new atom and the old atom. For another old atom, we will pick another value form the normal distribution. After all old atoms are checked, the largest calculated length of random vector will be used, because other shorter lengths will make some atoms too close to each other. The core step of adding one atom to the existing structure is illustrated in Figure 1.9. In this figure, before placing atom 4, a random direction is

first selected. For each old atom i, a distance λ_i is picked from the required distribution norm(μ_1, σ_1). Then a trial atom 4 is placed along the selected direction with a distance λ_i from atom i. Finally, among all trial atoms, the one with maximal distance from geometry center is used.

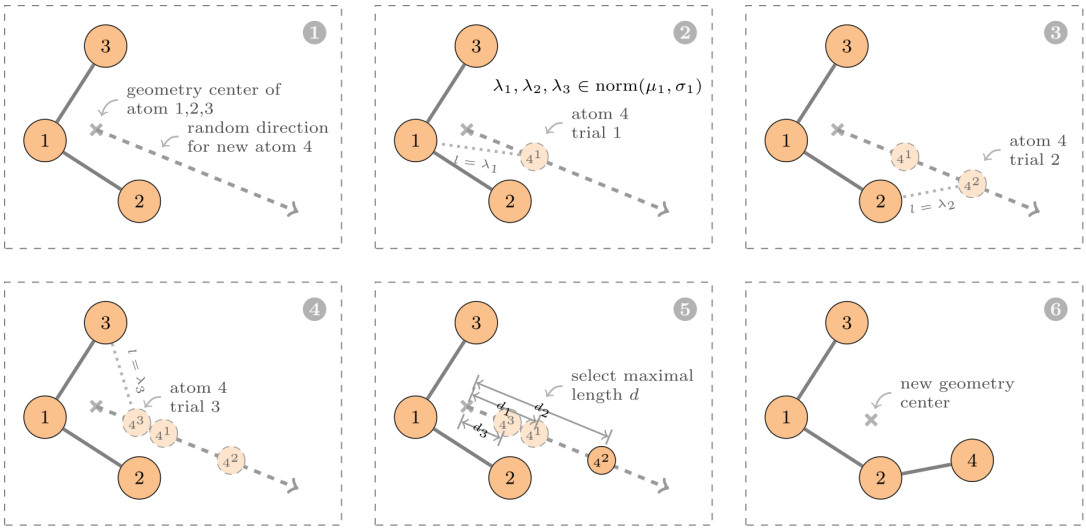


Figure 1.9 One intermediate step of F-BLDA: adding an atom to the existing structure.

It should be noted that since some new bonds (with bond length not picked from any normal distribution) can be generated with a small probability in this generation procedure, strictly speaking, the bond lengths resulted from this procedure will not be exactly normal-distributed. Nevertheless, in practice, we can see that it can give a very good approximation to the desired normal distribution.

Comparing to F-BLDA, the implementation of S-BLDA is more complicated. Nevertheless, it is still based on the procedure of adding one atom at a time. When placing the first two atoms, we follow the same procedure as that in F-BLDA. Starting from the third atom, the step of adding one new atom is divided into four sub-steps:

Selection. Select any two atoms A, B from existing atoms. At the same time, pick two lengths, λ_1 and λ_2 from norm(μ_1, σ_1) and norm(μ_2, σ_2), respectively. Here norm(μ, σ) represents a normal distribution with mean μ and standard deviation σ. These two normal distributions are the required l_1 and l_2 distribution, respectively.

Circling. The new atom is required to have a distance of length λ_1 to atom A, and a distance of length λ_2 to atom B. This restricts the new atom to pick a position on a circle in three dimensional space. The normal vector and radius of the circle plane can then be determined.

Projection. Select a random direction vector in three dimensional space. Then project this random vector to the circle plane so that we can determine the exact position on the circle for the new atom to pick.

Checking. For each of old atoms other than A and B, denoted as C_i, pick a length λ_i from norm(μ_1, σ_1). The distance between the new atom and atom C_i is required to be greater than λ_i. If for any atom C_i this is not satisfied, go back to sub-step (3) and select another random direction. If it still fails, try at most 10 times from sub-step (3) to (4). If it fails more than 10 times, go back to sub-step (1) and try another two atoms A and B.

1.2.6 Comparison between different approaches

Figure 1.10 shows the performance of different initial structure generation methods, in terms of the electronic energy distribution of generated structures. We can see that BC generates very bad structures, since its l_1 distribution is very diverse and away

from the true distribution. CK has a relatively low deviation but worse than RV. RV with zero l_1 deviation is worse than the first type BLDA. And the second type BLDA, with the distribution of l_2 considered, is much better than the aforementioned methods. In Figure 1.10, energies are relative to the energy of singlet global minimum (−716.077972 a.u.). The DFT is done with Turbomole 6.6, PBE0 hybrid functional.

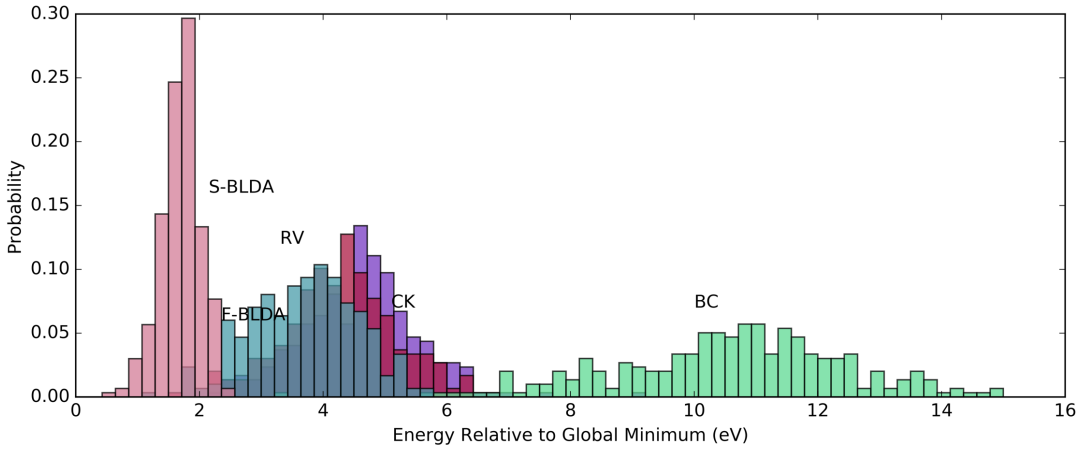


Figure 1.10 Single point energy (singlet) distribution of CK, RV, BC, first and second type BLDA generated structures.

1.3 Force field fitting

In this section, we introduce a global optimization approach, with an intermediate cruder optimization step using a classical force field and the CK method for initial structure generation. This force field is being developed on-the-fly for each specific system. Hence, the name of the method is Adaptive Force Field-Assisted *ab initio* Coalescence Kick (AFFCK). In principle, the initial structure can also be generated by any other method, but at the time when AFFCK is developed in our group, better generation methods such as BLDA has not been proposed.

The main steps of the AFFCK method can be divided into two phases: the force field construction phase and the optimization phase. In the force field construction phase, we first generate random bound structures using CK method, but without doing geometry optimization. The generated structures are checked for geometric similarity, and the structures similar to previous ones in shapes are discarded. Then we use some *ab initio* or DFT method to calculate the single point energy of these structures. The structures and their corresponding energies are subsequently used to fit the Force Field (FF) formula. Typically, ca. 3% of the structures with highest energies are discarded before fitting. We use a simple energy function that is linear with respect to all the parameters. Hence all the parameters can be determined via solving a set of over-determined linear equations. The typical number of structures used in a fitting (for clusters containing 8-10 atoms, as in our test systems) ranges between 1,000 and 5,000, although many fewer are actually needed to find the global minimum.

For the optimization phase, we also start from generating random structures via CK method without optimization. The number of structures is now much larger than that of the first phase (typically 10,000). We apply the pre-relaxation procedure to these newly generated structures. In the pre-relaxation procedure, structures can be optimized to their "local minima" according to the classical FF function that we have fitted in the first phase. The geometry optimization in the pre-relaxation procedure is facilitated by the nonlinear Conjugate Gradient (CG) method. Many of the initial structures are likely to descend to the same "local minima", and therefore, we perform the similarity check and discard redundant structures. At the final step of the second

phase, the unique, lowest-energy FF "local minima" are optimized by an *ab initio* or DFT method.

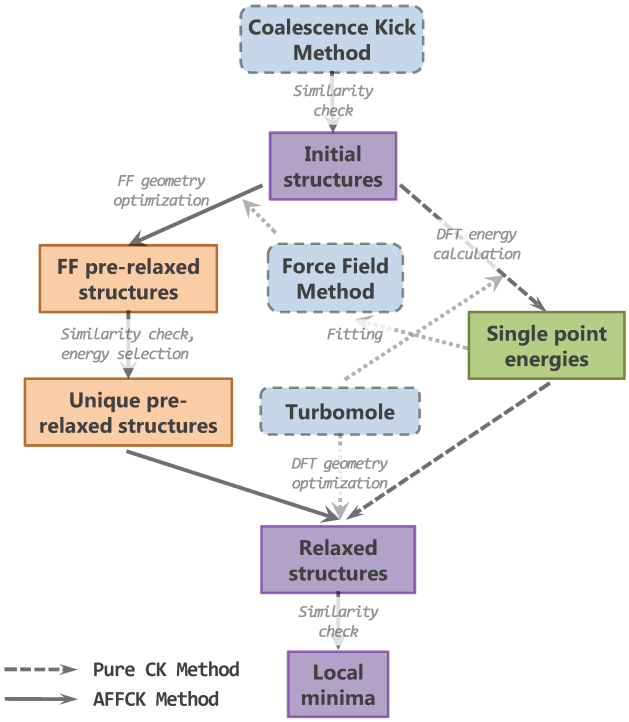


Figure 1.11 Basic steps of the AFFCK method and the pure CK method, and their relationships.

Since the pre-relaxed FF structures are usually close to equilibrium structures (as will be shown in Section 1.6), the time cost for the final quantum chemical optimization can be greatly reduced. The accuracy of the fitted FF function can be verified by comparing with *ab initio* or DFT results (geometries and energy rankings). Since the FF parameters are system-specific, which is different from the general FF method, the accuracy of the FF step is quite high, as will be demonstrated. Figure 1.11 shows the basic steps of the CK and AFFCK methods, and their relationships.

The FF function in the AFFCK method includes three terms (including one constant term):

$$E_{\mathrm{FF}} = \sum_{\text{pairs}} \sum_{\text{types } i} \left(\frac{A_{0i}}{l^9} - \frac{B_{0i}}{l^6} \right) + \sum_{\text{bonds}} (A_1\, l^2 + B_1\, l + C_1) + E_{\mathrm{const}}$$

where l is the distance between two atoms. The first term denotes the van der Waals (vdW) interaction. We take the 9-6 Lennard-Jones potential for van der Waals interaction, though more terms for different types of weak interactions could be added. Our first approach was to use the vdW terms for all pairs of atoms not connected by bonds. However, we found that a better fitting result was obtained when vdW interactions were computed for all pairs of atoms, regardless of them being connected by a bond, but depending on how many bonds are formed by the two atoms in a pair collectively, vdW parameters are different (fitted independently). For example, in Table 1.1, we list vdW parameters for pairs of atoms that collectively form fewer than 4 bonds, 4, 5, 6, and more than 6 bonds to other atoms, for Pt_8. The second term denotes bond stretching energies. For the first two terms, if the relevant atoms have a different combination of element types, we will use a different set of parameters. Though this is a rather simple FF form, it serves its purpose of a crude preliminary relaxation, as will be shown shortly, and it is very fast to build on-the-fly for every system under consideration. Fine-tuning of FF is not the goal here.

To any structure to be evaluated by the function above, a bond analysis is applied first. We estimate the typical bond length between two atoms by summing their covalent radii, and we accept the maximum bond length by a tolerance value of 0.45 Å, and the minimum bond length is set to be 0.4 Å (data from *jmol* [18] settings).

The parameters in the function can be obtained by solving a set of linear equations. Since it is over-determined, it can be solved by the least squares method. To determine the most suitable number of parameters of each term, we can run several tests for different number of parameters. The standard deviations of energies can be calculated to indicate which configuration of parameters is better.

The pre-relaxation is done with the nonlinear CG method. For the line search part of the CG method, we apply the Newton-Raphson algorithm, which requires the first and second derivatives of the energy function with respect to all coordinates. Since our function is relatively simple, those derivatives are not hard to obtain in analytical forms.

1.4 Deep neural network fitting

In previous section we have discussed the possibility that FF fitting can be used to assist global searching. However, the restricted form of the FF function may prevent its extension to larger systems. Neural Network (NN) has been widely applied to the fitting of molecular PES for many years.[19] Nevertheless, most studies of this fitting approach focus on small-sized clusters and molecules, such as Si_5,[20] BeH_3,[21] and FH_2O.[22] Recently, a high-dimensional NN fitting method (atomistic NN) has been proposed, which is based on expressing the total energy as the sum of atomic energies.[23,24] The new atomistic NN approach has been successfully applied to the PES fitting of a variety of systems, including Zn_NO_N (N=1-40).[25] Very recently, it has been shown that a combination of BH and atomistic NN approach (NN-BH) can be used for global optimization for large-sized metal clusters, such as Au_{58} [26] and Na_{20}-

[27]. Despite the success of atomistic NN method, many-body expansion method has also been combined with NN for PES fitting, and some applications to small molecules, such as H_2O_2 [28] and C_2H_3Br [29], are reported.

To the best of our knowledge, all aforementioned NN PES fitting approaches are based on the traditional fully connected NN with shallow structure (with typically three to four layers). However, it is shown [30] that NN with insufficient depth may require large number of training samples to tune the parameters. On the other hand, large scale fully connected NN is difficult to train and can potentially cause over-fitting issues. During recent years, several novel approaches have been proposed to solve the Deep Neural Network (DNN) training problem. For example, large Convolutional Neural Network (CNN), as an example of locally connected DNN, is shown to have an impressive performance on image classification.[31] Alternatively, Deep Belief Network (DBN) [32] is an example of DNN that can be pre-trained using unsupervised learning techniques. In addition, benefited from recent progress on GPU computing, [33] the training of DNN can be largely accelerated.

In this section, we discuss a general approach to find global and low energy local minima of metal clusters at the DFT level, using locally connected DNN fitting combined with the BLDA method for structure generation. Unlike some other NN based global optimization methods, the training data are not obtained by performing full local optimization. Instead, the limited-step local optimization is performed so that the PES can be moderately explored, and meanwhile the time cost of local optimization with DFT can be greatly reduced. The NN is constructed in a truncated many-body-expansion way, but several mixing layers are added so that the NN potential

can represent a more general function than the original truncated many-body one. The fitted NN is then used to optimize some newly generated random structures. As a final step, automatically selected NN-relaxed structures are fully optimized using DFT to yield the final results.

1.4.1 Truncated many-body approximation

The NN architecture used in this work is designed based on the truncated many-body expansion approach,[34] with only the k-body terms considered. In this work, $k = 4$ is used. Based on our experience, NN constructed using higher order terms can make the training very time-consuming and may cause over-fitting issues. On the other hand, if only lower order terms are included, poor fitting accuracy may be expected.

For simplicity, consider a homogeneous atomic cluster, say, Pt_N ($N \geq k$). The Cartesian coordinates of the atoms are denoted as $\boldsymbol{r}_i$ ($i = 1, 2, \ldots, N$). Given any set X, $\mathcal{S}_k(X)$ will be the set of all subsets of X that are composed of k elements. Obviously, if the number of elements in X, denoted as $|X|$, is equal to N, then $|\mathcal{S}_k(X)| = C(N, k)$, which is the number of combinations of k elements from X.

The original PES can be expressed as a function of $\{\boldsymbol{r}_i\}$, namely, $F(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N)$. Using the many-body expansion, if only the k-body terms are retained, then the original PES is approximated by $F'(\{\boldsymbol{r}_i\}) = \text{sum}(F_k(\{\boldsymbol{r}_j\}))$, where k is a fixed parameter, $\{\boldsymbol{r}_j\} \in \mathcal{S}_k(\{\boldsymbol{r}_i\})$, and the sum is over all $\{\boldsymbol{r}_j\}$ in $\mathcal{S}_k(\{\boldsymbol{r}_i\})$. Now the undetermined function F_k can be fitted using the traditional fully connected NN.

In practice, when N is large and k is kept to be a moderate number (as in our case, four), this procedure may cause large fitting error since many high order terms are truncated. To alleviate this problem, we replace the summation by an undetermined function G. This will introduce some non-linear mixing among k-body terms, by which the fitting accuracy can be improved. The original PES is then more accurately approximated by $F''(\{\boldsymbol{r}_i\}) = G(\{F_k(\{\boldsymbol{r}_j\})\})$, where G is a function of $C(N, k)$ variables. Since F_k is merely an intermediate variable, it no longer has to be a scalar function. Therefore, in our approach it is generalized to be a vector function $\boldsymbol{F}_k$, and this will introduce more flexibility in the constructed neural network. Therefore, our final neural network framework to approximate the real PES is $F'''(\{\boldsymbol{r}_i\}) = G(\{\boldsymbol{F}_k(\{\boldsymbol{r}_j\})\})$, where $\boldsymbol{F}_k$ and G can be fitted simultaneously, by locally connected NN. We call this approach the Many-Body Expansion Neural Network plus Mixing (MBE-NN+M) approach.

It is known that Cartesian coordinates are not suitable to be directly used as input of NN.[23] Therefore, two transformations are performed on the input sample data in Cartesian coordinates $\{\boldsymbol{r}_j\}$. The original coordinates $\{\boldsymbol{r}_j\}$ (after k-body selection procedure) are first transformed to interatomic distances $\{a_m\}$, where m = 1, 2, …, C(k, 2) are the indices of all interatomic distances in the k-body fragment. After that, an exponential function is applied to help better describing the PES:[35] $b_m = \exp(-a_m/L)$, where $\{b_m\}$ is the input of NN and L is a fixed parameter. The exponential function basically amplifies the short-distance areas in the PES. In this work, we choose L = 4.0 Å. The output of NN is a single value representing the energy. A linear transformation is used so that the energies are mapped to interval [0, 1] for NN training, and

the corresponding inverse transformation is performed when interpreting NN predictions.

1.4.2 Neural network architecture

The DNN architecture used in this work consists of three parts: MBE, mixing, and pooling. Local connectivity, parameter sharing, and average-pooling [36] are three main features used in this NN, which make it different from the traditional fully connected NN.

The MBE part of NN is used to fit $\boldsymbol{F}_k$. Based on the MBE-NN+M approach, the input of NN is a second-order tensor (or two-dimensional array) of the size (C(N, k), C(k, 2)), where the second dimension represents C(k, 2) variables of the function $\boldsymbol{F}_k$, and the first dimension indicates that there are C(N, k) k-body terms, each denoted as $\boldsymbol{F}_k(\{\boldsymbol{r}_j\})$. To fit $\boldsymbol{F}_k$, the NN connection is localized in the second dimension and the layer size of the first dimension is kept fixed. The weights and biases of NN connection are shared among different indices of the first dimension, so that the fitted function form of $\boldsymbol{F}_k$ is kept consistent among different k-body terms. Since $\boldsymbol{F}_k$ is a vector function, the size of the output value (as a vector) of each $\boldsymbol{F}_k$ term, denoted as f, must be determined when constructing NN. In this work, we choose f = 2. Therefore, the MBE part is composed of four layers with the sizes: (C(N, k), C(k, 2)) - (C(N, k), 40) - (C(N, k), 70) - (C(N, k), 60) - (C(N, k), 2).

The mixing part is used to fit G. Within this part the NN connection is localized in the first dimension and the size of the second dimension is kept fixed. The parameters

of NN connection in this part are shared among different indices of the second dimension. In this work, the mixing part is composed of two layers with the sizes: (C(N, k), 2) - (40, 2) - (10, 2).

The last part is used to transform the output of mixing part to a single value, representing the energy. The average-pooling approach is used, which means that we take the average value of all elements in the matrix of the previous layer as the final output. In this work, the pooling part is composed of one layer of the size: (10, 2) - (1).

To describe the nonlinearity of the unknown function, suitable activation functions must be used in NN architecture. In previous NN PES fitting work by other authors, common choices are the hyperbolic tangent and the sigmoid function.[24] However, it has been found that these traditional activation functions are less efficient in DNN.[37] It has also been observed that in some DNNs, the earlier layers tend to learn less efficiently comparing with later layers, when trained by backpropagation learning algorithm.[38] Therefore, in this work, the first three hidden layers are activated by hyperbolic tangent while the remaining layers except the output layer are activated by the softplus function. We found that this hybrid approach is able to overcome the intrinsic disadvantage of both activation functions and increase the training efficiency. The overall architecture of NN used in this work is shown in Figure 1.12. Note that the actual numbers of NN nodes (layer sizes) are larger than or equal to the number of nodes shown in the figure. In the figure, the layer sizes for the Pt_{13} case (N = 13, k = 4) are indicated in the bottom. There is no node connection across the stacked planes.

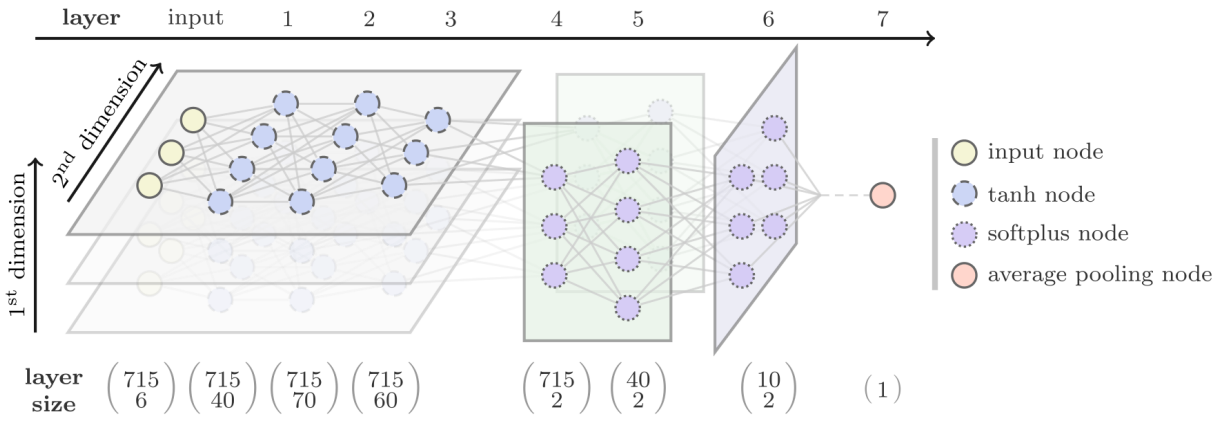


Figure 1.12 Architecture of the deep neural network used for fitting PES.

1.4.3 Neural network implementation

The NN can be constructed using *theano* framework,[39] which is a *python* library that allows to accelerate the evaluation and differentiation of mathematical expressions both numerically and symbolically. Based on this framework, efficient C/openMP codes and/or CUDA codes for GPU can be automatically generated at runtime. We note that this feature is significant for saving time during training our large-scale deep NN. It is illustrated in Figure 1.13 that for a typical DNN PES fitting task in this work, acceleration of computation with GPU can be more than 50 times compared to that done with only CPU. Based on this observation, in this work, the GPU Tesla K40s is used for all NN training tasks. In Figure 1.13, the test case is a 7-layer NN PES fitting for Pt_{13} energy data. During each epoch, 50 000/5 000 random structures and their corresponding energies are used for training/testing, respectively. For the first architecture (OpenMp/24), the CPU model is Intel Xeon E5-2697v2 (2.7 GHz) and the 24 cores OpenMP parallelism scheme is used. For the other two architectures, a single CPU core and one indicated GPU accelerator are used.

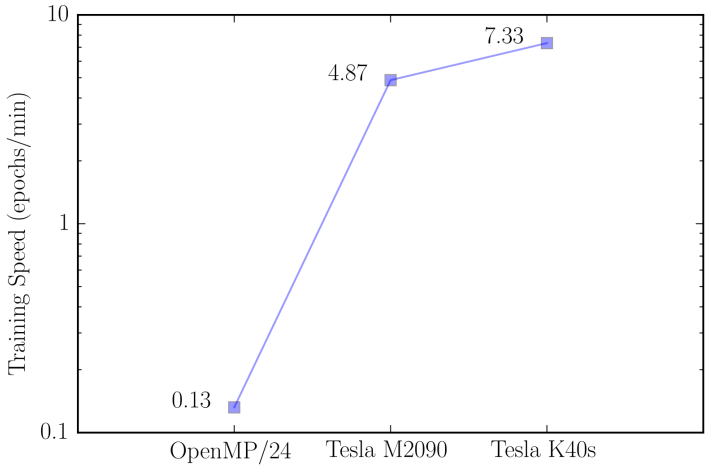


Figure 1.13 Training speed measured in number of epochs per minute under different machine architectures.

The Mini-Batch Stochastic Gradient Descent with Momentum (MB-SGD-M) method [40] is used for NN training. This is a widely tested training method for large-scale NNs. The step decay approach is used to improve convergence. The step length for parameter updating at epoch i is determined by $s_0\, r/(r + i)$, where s_0 is the initial step length and r is the step decay factor. In our work, we choose $s_0 = 0.1$ and $r = 60$. When NN training finished, Limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method [41] is used for NN based geometry optimization.

1.4.4 Neural network assisted global optimization

Based on the aforementioned neural network fitting methods, we are able to propose a new NN-fitting-based and highly parallel global optimization scheme shown in Figure 1.14. In this figure, the paralleled algorithms are indicated by multiple arrows

and the parallelism models are shown in parentheses. The DFT optimization is paralleled in two levels. Different atomic configurations are independently paralleled at job level. For every atomic configuration, the OpenMP or MPI parallelism is utilized as implemented in the DFT package. The overall scheme (NN-PGOPT) consists of three major steps (The AM-DFS method for structure filtering will be introduced in Chapter 3):

1) **Initial structure preparation.**

    a. Random structure generation (S-BLDA)

    b. Filtering out duplicates (AM-DFS)

2) **Local optimization.**

    a. Limited-step geometry optimization (DFT)

    b. PES fitting (MBE-NN+M and MB-SGD-M)

    c. PES optimization (L-BFGS)

    d. Filtering out duplicates (AM-DFS)

    e. Full-step geometry optimization (DFT)

3) **Global minimum identification.**

    a. Filtering out duplicates (AM-DFS)

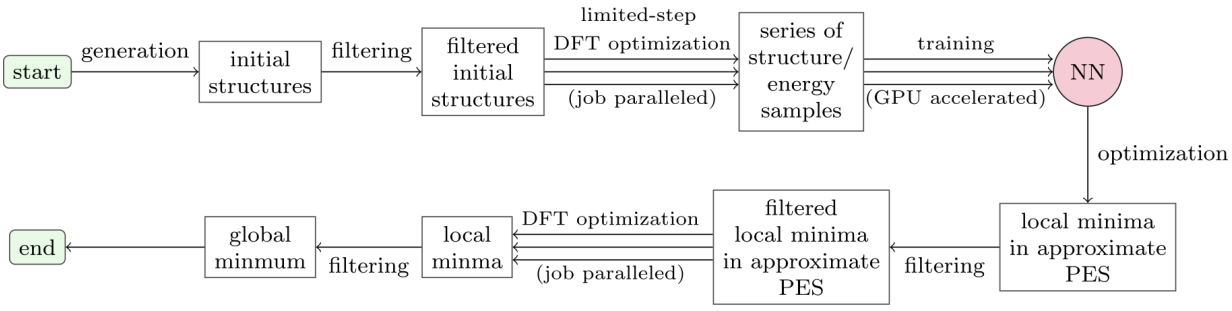


Figure 1.14 Overall flow-chart of NN-PGOPT scheme.

1.5 Superposition approximation and ensemble average

In the most general sense, global optimization is the technique to find the most stable configuration (called global minimum) at absolute zero temperature for a system. At realistic temperature, it is expected that the global minimum will become less important; the energy separations between local minima, however, will have a significant effect on properties. A quantitative analysis of the contributions of the isomers, and the changes of properties in realistic temperature is thus necessary. This can be achieved by an ensemble-average representation of isomers. In what follows, we assume that the interconversion of isomers is not strongly kinetically hindered, and that at high temperatures, eventually, the thermodynamic equilibrium would be reached, with every isomer being populated according to its free energy. Kinetic trapping effect can be understood in the cluster isomerization process, which will be subject of Chapter 4.

At finite temperature, the ensemble average of any property A can be approximately computed by taking the weighted average of that property of each geometrically unique isomer A_k:

$$\langle A \rangle \approx \sum_k P_k A_k$$

where P_k are the occurrence probabilities of isomer k. The probabilities P_k are obtained from the superposition approximation,[42] where the partition function of the system is expressed as the sum over individual partition function of isomers

$$Z = \sum_k Z_k, \quad P_k = \frac{Z_k}{Z}$$

In order to compute Z_k, the electronic, vibrational and rotational degrees of freedom of the system will be considered.[43] The translational contribution is almost the same for each isomer, which is ignored here. Therefore, the partition function of isomer k is written as

$$Z_k = Z_{\text{elec},k} Z_{\text{vib},k} Z_{\text{rot},k}$$

where the electronic partition function is

$$Z_{\text{elec},k} = g_{\text{spin},k} e^{-\beta E_k}$$

where $g_{\text{spin},k}$ is the spin degeneracy, which is equal to the multiplicity of the isomer. E_k is the ground state electronic energy of the isomer. It is assumed that for small enough clusters the quantum confinement effect is large enough to make the ground and excited electronic states well-separated in energy, and contributions to the electronic entropy due to the population of excited states being insignificant. This fact is generally supported by the lack of multi-reference character of the cluster wave functions in many selected Pt clusters that we tried (CASSCF test).

The vibrational partition function is (using quantum harmonic oscillator approximation)

$$Z_{\mathrm{vib},k} = \prod_{i}^{3N-6} \frac{e^{-(\beta\hbar\omega_{ki})/2}}{1 - e^{-\beta\hbar\omega_{ki}}}$$

where ω_{ki} are $3N$ - 6 vibrational frequencies of the isomer k. Note that the Zero-Point Energy (ZPE) and vibrational entropy effect are included in this term. The rotational partition function is (in high-temperature approximate form)

$$Z_{\mathrm{rot},k} = \frac{\sqrt{\pi}}{\sigma_k}\left(\frac{2}{\beta\hbar^2}\right)^{\frac{3}{2}}\sqrt{I_{k1}I_{k2}I_{k3}}$$

where σ_k is the order of rotational subgroup of the point group symmetry of isomer k and I_{k1}, I_{k2}, I_{k3} are three eigenvalues of inertia tensor.

To study the structure transition at certain temperature, the heat capacity can be computed as the following

$$C_V = k_B\beta^2\left[-\frac{1}{Z^2}\left(\sum_k W_k Z_k\right)^2 + \frac{1}{Z}\sum_k (W_k^2 Z_k + V_k Z_k)\right]$$

where

$$W_k = \frac{1}{Z_k}\frac{\partial Z_k}{\partial \beta} = -\frac{3}{2\beta} - \sum_i \frac{\hbar\omega_{ki}}{e^{\beta\hbar\omega_{ki}} - 1} - \frac{1}{2}\sum_i \hbar\omega_{ki} - E_k$$

$$V_k = \frac{\partial W_k}{\partial \beta} = \frac{3}{2\beta^2} + \sum_i \frac{(\hbar\omega_{ki})^2 e^{\beta\hbar\omega_{ki}}}{(e^{\beta\hbar\omega_{ki}} - 1)^2}$$

1.6 Application on Pt_8 cluster

We illustrate the performance of AFFCK on the Pt_8 cluster, using DFT for single point energy evaluation. AFFCK may be extended to any DFT package. We have tested it on plane wave DFT software such as the Vienna Ab-initio Simulation Package (VASP) [44–47] and the Turbomole [48] program to perform DFT calculations. All calculations presented in this section used the PBE0 [49–51] hybrid functional (with spin polarization), in view of its known adequate performance for both clusters of B [52] and Pt [53,54] and computational affordability. The def2-TZVP [55] basis set was employed. The level of theory in use can be adjusted to what is more suitable for the studied systems, and this is one of the levers for increasing the accuracy of the search, if desired. For the purpose of testing AFFCK, UPBE0/def2-TZVP was found to be optimal, providing both reasonable computational time and sampling of CK structures (>96% converged).

1.6.1 General force field fitting approach

The AFFCK method was applied to Pt_8 cluster to find the global minimum and local minima structures. To illustrate the accuracy and efficiency of AFFCK method, we also perform the pure CK search on the cluster from the same initial guessed structures.

We have done several tests to determine how many parameters we should include in our set for a good FF fitting. The quality of fitting can be evaluated by calculating the standard deviation σ_m of the energies:

$$\sigma_m = \sqrt{\frac{\sum_k^m (E_{0k} - E_{\mathrm{FF}k})^2}{m}}$$

where m is the total number of initial structures, E_{0k} and $E_{\mathrm{FF}k}$ are energies obtained by *Turbomole* and the fitted FF formula, respectively. The smaller σ_m is, the closer our fitted energies are to the DFT energies. Since the parameters are determined by solving a set of over-determined linear equations, σ_m will always decrease when we increase the number of parameters. The significance of parameters can thus be measured by looking at how much σ_m decreases while we introduce new parameters. The parameter test results are showed in Table 1.1. In the table, all the parameters are calculated using the energy unit Hartree and length unit Å. "N/A" denotes that these parameters are not included. The van der Waals interaction terms are divided according to the total number of bonds that the interacting atoms have, and "as (n)" denotes that these parameters are taken the same as those of n bonds case. $\Delta\sigma_m$ is the difference of σ_m comparing to the previous line. The three numbers in "stretch energy terms" column are A_1, B_1, C_1, respectively, and the two numbers in "van der Waals interaction terms" columns are A_{0i}, B_{0i}, respectively (see Section 1.3 for the definition of parameters). From the listed $\Delta\sigma_m$ we can see that we cannot improve the quality of the formula much when splitting van der Waals terms corresponding to total bonds number less than 6. Nevertheless, we choose to keep these parameters since their fitted values are all physical.

Table 1.1 The fitted values of parameters and corresponding σ_m when choosing different set of parameters in the FF formula, for singlet Pt_8 cluster (PBE0 functional).

# of parameters	constant term	stretch energy terms	van der Waals interaction terms					σ_m/eV	$\Delta\sigma_m$/eV
			<4	4	5	6	>6		
3	-954.4	N/A	as (5)	as (5)	151.0 14.01	as (5)	as (5)	0.744	
6	-954.4	0.068 -0.37 0.50	as (5)	as (5)	126.0 12.46	as (5)	as (5)	0.696	0.048
8	-954.3	0.067 -0.36 0.48	as (5)	as (5)	122.5 13.37	as (>6)	123.7 11.59	0.649	0.047
10	-954.3	0.067 -0.36 0.48	as (4)	123.7 13.95	122.7 13.21	as (>6)	123.8 11.56	0.645	0.004
12	-954.3	0.076 -0.41 0.55	68.66 10.96	129.3 14.99	124.7 14.06	as (>6)	127.4 12.6	0.642	0.003
14	-954.2	0.067 -0.36 0.46	132.4 15.9	143.9 16.13	132.2 14.47	134.4 13.59	120.0 10.66	0.593	0.049

To find the accurate global minimum for the Pt_8 cluster, we consider different electron spin multiplicities (singlet, triplet, and quintet) independently. This approach was used in both pure CK, and in AFFCK calculations, which means that, for the AFFCK method, we have different sets of FF parameters fitted for different electronic states.

Table 1.2 lists all these fitted parameters as well as the number of initial structures used for fitting. We can also find that when total number of bonds gets larger, the coefficients of their corresponding van der Waals interaction term get smaller. The reason is that atoms with larger total number of bonds lie in the center of the cluster. 1,000 initial guess structures for each spin multiplicity were generated, and the DFT single point energy calculations converged for 956, 963, and 927 structures, for singlets, triplets, and quintets, respectively. We chose 97% of them to fit the FF formula, while structures with top 3% highest energies were excluded.

Table 1.2 The fitted values of parameters, number of initial structures used in the FF formula, and corresponding σ_m for Pt_8 cluster energies (PBE0 functional) with different spin multiplicities. Notations and units are the same as Table 1.1.

spin multiplicity	# of initial structures used	constant term	stretch energy terms	van der Waals interaction terms					σ_m/eV
				<4	4	5	6	>6	
singlet	927	-954.2	0.067 -0.36 0.46	132.4 15.90	143.9 16.13	132.2 14.47	134.4 13.59	120.0 10.66	0.593
triplet	934	-954.3	0.060 -0.32 0.40	95.85 11.86	132.8 14.09	130.3 13.20	127.9 12.16	117.0 9.66	0.435
quintet	899	-954.3	0.061 -0.33 0.42	89.22 11.47	132.9 14.07	130.1 13.20	127.5 12.16	117.7 9.77	0.432

The fitted FF formula can then be used to find the pre-relaxed structures of the Pt_8 cluster, minima within the FF formalism, as stated in previous sections. For each spin multiplicity, 10,000 additional guessed structures are generated using the CK coalescence approach, without DFT calculations, and pre-relaxed using nonlinear CG method and FF. Then for each spin multiplicity, 300 unique pre-relaxed structures with relatively low FF energies are selected, using the similarity measurement method introduced in Section 3.1.1. The selection parameters are d_{rel} = 0.03 Å and d_{max} = 0.7 Å. These structures are then re-evaluated and re-optimized using DFT.

Now by comparing the approximate FF and DFT energies of these pre-relaxed structures, we can have an idea of the accuracy of AFFCK energies with respect to DFT ones. Figure 1.15 shows the energies of initial and pre-relaxed Pt_8 clusters, calculated with FF and PBE0, for different spin multiplicities. In this figure, x axes show the serial number of the structures. Green and yellow dots are FF (fitted) and DFT energies of initial structures, respectively; red and blue dots are FF (predicted) and DFT energies of pre-relaxed structures, respectively. The structures are sorted by their FF energies. Some initial structures with very high energies are excluded in order to show more details of the main part of the plot. Energies relative to -25,982.924 eV, which is the energy of the global minimum that we find, were used. We can find that the fitting error for singlet state are relatively larger, comparing to other two states, which agrees with the σ_m values that we have listed in Table 1.2. Despite this, our simple formula was shown to make a good approximation of DFT energies for Pt_8 clusters, and the fitted parameters obtained from initial structures with high energies did work for structures that have lower energies as well.

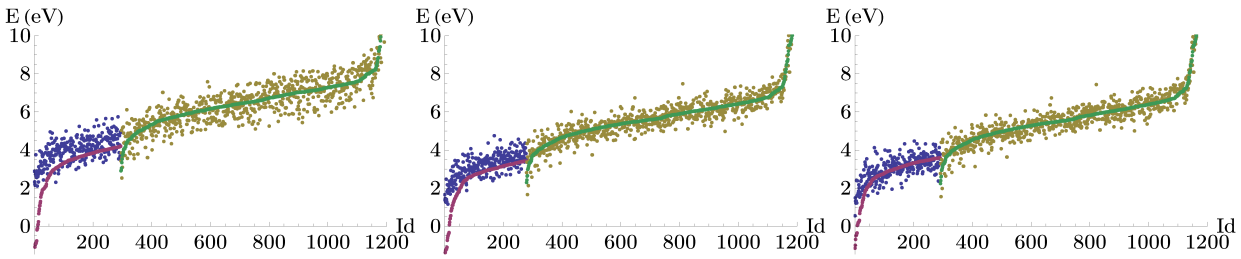


Figure 1.15 FF and DFT (PBE0 functional) energies of singlet (left), triplet (middle) and quintet (right) Pt_8 clusters.

The pre-relaxed structures then undergo further optimization by DFT. We can compare the number of steps taken for optimizing pre-relaxed structures (by AFFCK method) and initial structures (by pure CK method), to verify the efficiency of the FF procedure in the AFFCK method. For the pure CK method, the optimizations were converged for 956, 963, and 927 (out of 1,000) structures, for singlets, triplets, and quintets, respectively, and for AFFCK method, 295, 279, 291 (out of 300) structures converged, for singlets, triplets, and quintets, respectively. Figure 1.16 shows the number of steps taken for optimizing these structures with different spin multiplicities. In this figure, blue and green bars stand for the AFFCK method (optimizing from pre-relaxed structures) and the pure CK method (optimizing from initial structures), respectively. The mean numbers of optimization steps were 96, 97, 92 (pure CK) and 78, 74, 76 (AFFCK), for singlets, triplets, and quintets structures respectively. Averagely 20% of the DFT optimization time was saved per structure, and after considering the number of structures that were used, 76% of the total optimization time was saved. From these we can conclude that the AFFCK method indeed accelerated the DFT optimization significantly by the pre-relaxing procedure (which itself is very fast).

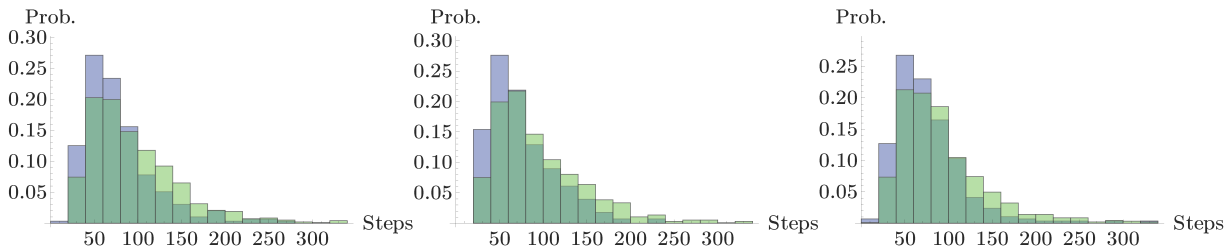


Figure 1.16 Frequency histograms for the distribution of the number of optimization (DFT with PBE0 functional) steps of converged singlet (left), triplet (middle) and quintet (right) Pt_8 clusters.

The effectiveness of the pre-relaxing procedure can be further examined by looking at the "energy shift" from initial to final relaxed structures. When energies of the structures to be optimized are closer to that of minima structures, the optimization will be faster. In this case, the "energy shift" will be shorter. Figure 1.17 shows the energy distribution of Pt_8 structures at different calculation stages. In this figure, energy shift by the AFFCK and pure CK methods are showed in upper and lower three plots, respectively. All energies are calculated using DFT with PBE0 functional. Red, yellow, blue, and green parts represent initial, pre-relaxed, DFT relaxed (from pre-relaxed ones), and DFT relaxed (from initial ones) structures, respectively. Energies relative to -25,982.924 eV, which is the energy of the global minimum that we find, were used. It is obvious that the AFFCK method shortens the "energy shift" by improve the quality of initial guessed structures by the CK method.

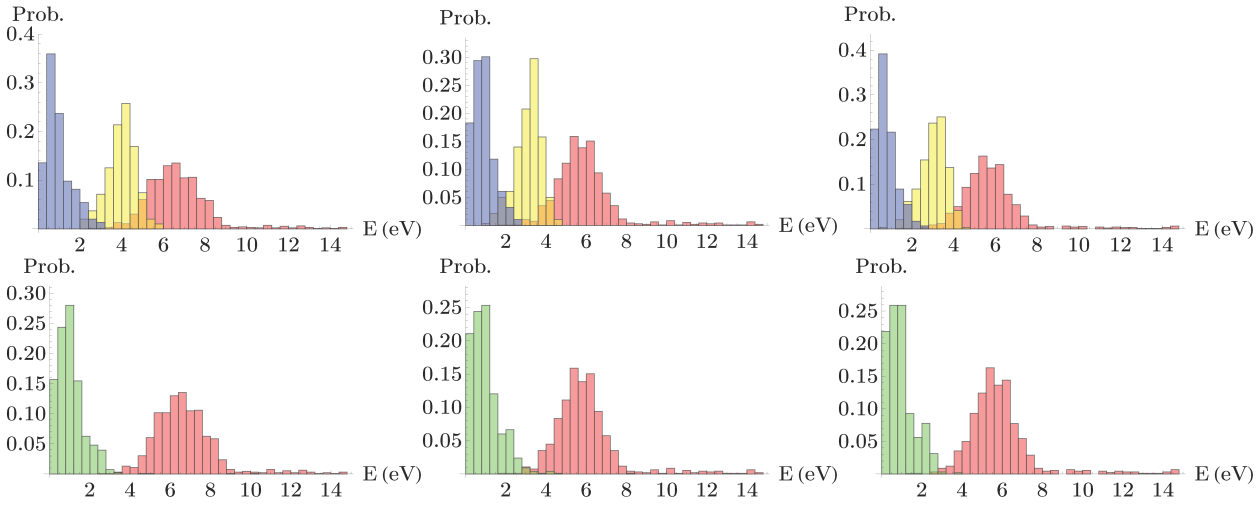
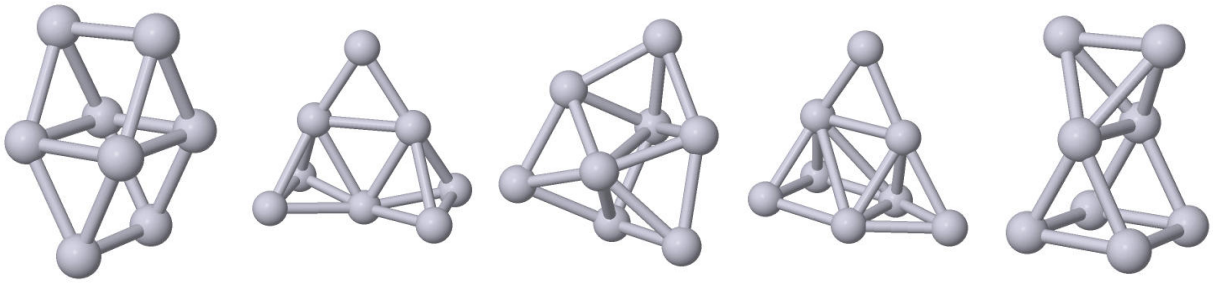
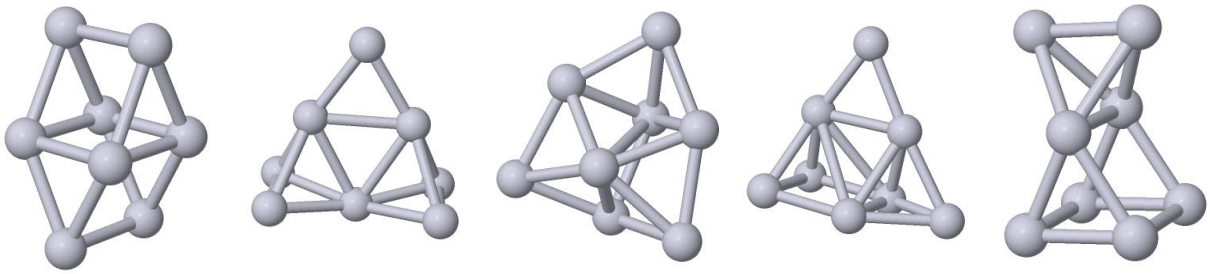


Figure 1.17 Frequency histograms for energy distributions of singlet (left), triplet (middle) and quintet (right) Pt_8 clusters at different calculation stages.

In addition to the effectiveness, AFFCK is also proved to be accurate for the Pt_8 calculation. The finally relaxed structures using the AFFCK and pure CK methods were filtered to find all the unique local minima. The structure selection parameters that used are d_{rel} = 0.015 Å and d_{max} = 0.2 Å. Some similar structures were further filtered manually. The first five minima found by each method are shown in Figure 1.18. Although we optimized much fewer (300) structures in the AFFCK method, the local minimum structures found by the two methods agree with each other. Furthermore, the energy for global minima **I** found by AFFCK is lower than that of **I'** found by pure CK, i.e. with our rather modest CK settings pure CK failed to find the quintet structure **I**. There is no doubt that increasing the number of trial structures in pure CK would fix the problem. However, in our test roughly three times more structures were used in CK versus AFFCK, demonstrating the effectiveness and utility of the FF pre-relaxation. Other local minima that we have found are shown in Figure 1.19.

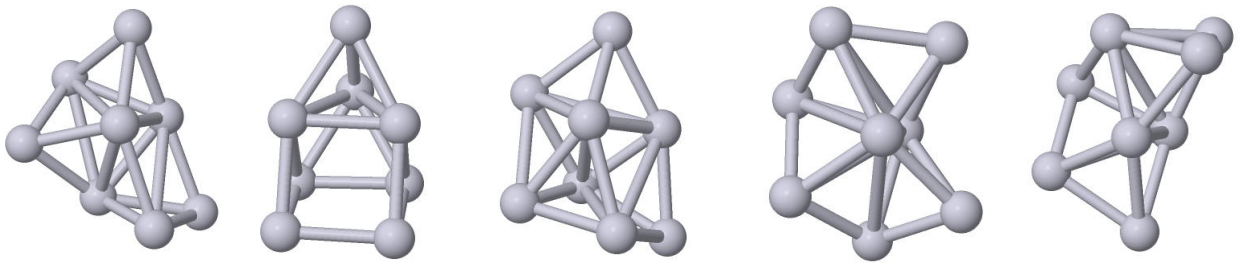


I, quintet, D_{2d} ΔE = 0.00 eV | **II**, quintet, C_{2v} ΔE = 0.12 eV | **III**, quintet, C_s ΔE = 0.13 eV | **IV**, quintet, C_s ΔE = 0.17 eV | **V**, quintet, C_{2v} ΔE = 0.17 eV



I', triplet, D_{2d} ΔE = 0.06 eV | **II'**, quintet, C_{2v} ΔE = 0.12 eV | **III'**, quintet, C_s ΔE = 0.13 eV | **IV'**, quintet, C_s ΔE = 0.17 eV | **V'**, quintet, C_{2v} ΔE = 0.17 eV

Figure 1.18 First five local minima of Pt_8 clusters found by AFFCK (upper) and pure CK (lower) method. ΔE are energies relative to I, calculated using DFT with PBE0 functional.



VI, quintet, C_1 ΔE = 0.23 eV | **VII**, quintet, C_s ΔE = 0.26 eV | **VIII**, quintet, C_1 ΔE = 0.30 eV | **IX**, quintet, C_2 ΔE = 0.32 eV | **X**, triplet, C_1 ΔE = 0.32 eV

Figure 1.19 Other higher energy isomer of the global minimum Pt_8 cluster I. ΔE are energies relative to I, calculated using DFT with PBE0 functional.

1.6.2 Symmetry specific calculation

In previous section we perform all calculations without any symmetry specific treatments. Here we further investigate the utility of the symmetry generation procedure, which has been discussed in Section 1.2.2, with the Pt_8 cluster serving as an example. We used the same FF formula as that discussed in Section 1.3. Instead of generating 10,000 structures all belonging to the point group C_1 for pre-relaxing, we generated 2,100 of those structures with specific point group symmetries, namely, 100 structures for each of C_s, C_i, C_2, C_3, C_4, C_{2v}, C_{3v}, C_{4v}, C_{2h}, C_{3h}, C_{4h}, D_2, D_3, D_4, D_{2h}, D_{3h}, D_{4h}, D_{2d}, D_{3d}, D_{4d}, and S_4 point group symmetry. These 2,100 structures were pre-relaxed for different spin multiplicities independently. Then we selected 100 structures with lower energies for DFT optimization for each spin multiplicity (i.e. total of 300, as in the previous situation). Figure 1.20 shows the new energy shifts for this symmetry specific AFFCK (SS-AFFCK) treatment, with a comparison with the AFFCK method without symmetry treatment. In this figure, energy shift by normal AFFCK and symmetry specific AFFCK method are showed in solid and dashed edges, respectively. All energies are calculated using DFT with PBE0 functional. Red, yellow, blue, brown, and purple parts represent initial, pre-relaxed (normal AFFCK), DFT relaxed (normal AFFCK), pre-relaxed (SS-AFFCK), and DFT relaxed (SS-AFFCK) structures, respectively. Energies relative to -25,982.924 eV, which is the energy of the global minimum found by normal AFFCK, were used. We can conclude that the symmetry treatment did not refine the final energy distribution, however, it pushed the pre-relaxed structures closer to the low energy region. Figure 1.21 shows comparison of the number of steps taken in DFT optimization between the SS-AFFCK and normal AFFCK methods. In this figure, Purple (solid edge) and blue (dashed edge) bars stand for the SS-AFFCK method and the normal AFFCK method, respectively. We found that the

structures treated by SS-AFFACK converged rapidly, with a great increase of number of structures converging in less than 20 steps. Figure 1.22 shows some of the relaxed structures found by the SS-AFFCK method. We note that the first three local minima (**I**", **II**", and **III**") have also been found with this treatment. **IX*** is the mirror symmetric configuration of **IX**. Other isomers found here generally have higher energies than those found by the normal AFFCK method, but many of them are in higher order symmetry (especially **XI**-T_d, **XV**-O_h), and some of them might be saddle points. We also note that **XV** was reported in previous literatures.[56,57]

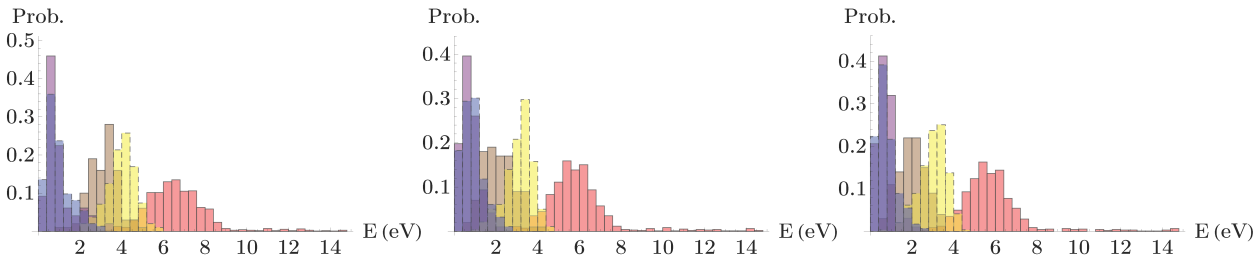


Figure 1.20 Frequency histograms for energy distributions of singlet (left), triplet (middle) and quintet (right) Pt_8 clusters at different calculation stages.

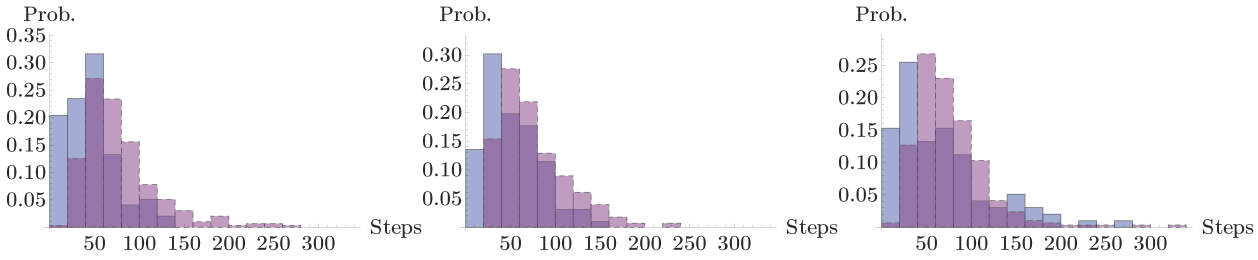
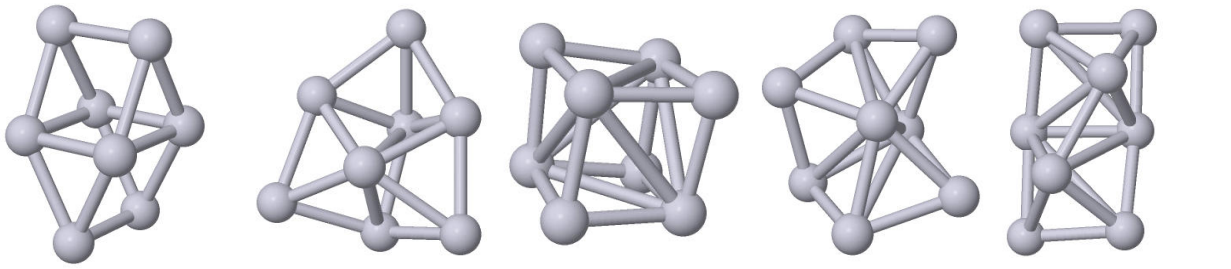
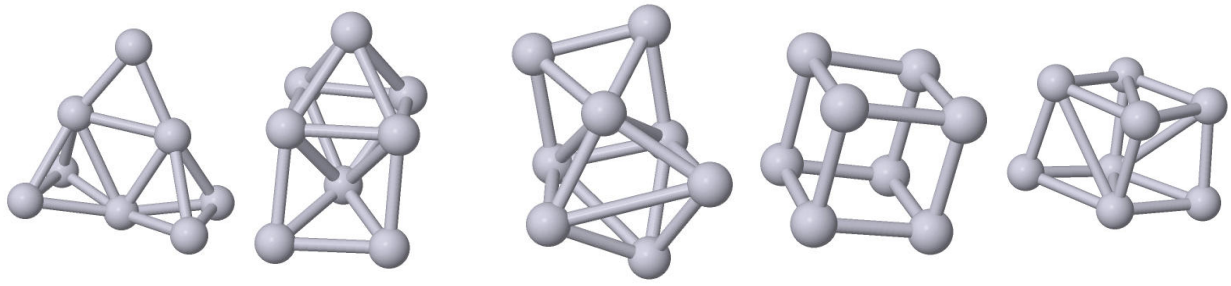


Figure 1.21 Frequency histograms for the distribution of the number of optimization (DFT with PBE0 functional) steps of converged singlet (left), triplet (middle) and quintet (right) Pt_8 clusters.



I", quintet, D_{2d} ΔE = 0.00 eV | **III"**, quintet, C_s ΔE = 0.13 eV | **XI**, quintet, T_d ΔE = 0.25 eV | **IX***, quintet, C_2 ΔE = 0.32 eV | **XII**, quintet, C_2 ΔE = 0.34 eV



II', triplet, C_{2v} ΔE = 0.38 eV | **XIII**, singlet, C_s ΔE = 0.45 eV | **XIV**, quintet, C_s ΔE = 0.46 eV | **XV**, quintet, O_h ΔE = 0.49 eV | **XVI**, quintet, C_{2v} ΔE = 0.52 eV

Figure 1.22 Some representative high energy isomer of the global minimum Pt_8 cluster I, found by SS-AFFCK method. ΔE are energies relative to I, calculated using DFT with PBE0 functional.

1.7 Application on Pt_9 cluster

1.7.1 DFT parameters and fitting errors

The global optimization is performed for Pt_9 as an example of an application of the NN-PGOPT scheme. A separate direct searching is also performed, where the steps 2a - 2d mentioned in Section 1.4.4 are skipped, in order to test the performance without NN fitting.

The DFT optimization is done using Turbomole 6.6 [58,59] with Tao, Perdew, Staroverov, and Scuseria hybrid (TPSSh) [60] functional, since the nonempirical

TPSSh hybrid functional has been found to have a good performance for late transition metals clusters.[61] To save DFT optimization time, the randomly generated structures are first optimized with def2-SV(P) and then with the def2-TZVP basis set,[62] in the direct searching case. For NN-PGOPT scheme, the NN is fitted for def2-SV(P) energies, then the full-step DFT optimization is performed with def2-TZVP.

500 unique initial structures are generated using S-BLDA approach with parameters μ_1 = 2.50 Å, μ_2 = 2.55 Å, and $\sigma_1 = \sigma_2$ = 0.04 Å for singlet, triplet, quintet, septet, and nonet multiplicities, respectively. The threshold value d = 0.25 Å is used to filter the S-BLDA generated structures. 10 DFT optimization steps are performed on these initial structures to yield 4240, 4770, 4670, 4210, and 4080 structure/energy samples for each multiplicity, respectively. For each multiplicity, the samples span an energy range of around 5.0 eV. Five independent NNs are fitted for different multiplicities, and an additional reference NN is first fitted for singlet to generate an initial guess of NN parameters. 82%, 9%, and 9% of samples are used for training, testing and validation, respectively. Since the cluster structure will be the same when the coordinates of any two atoms of the same element are swapped, these atoms can be randomly shuffled to generate more samples, so that the NN can learn the permutation invariance of atoms of the same element. In this way, 200,000, 20,000, and 20,000 samples are generated for training, testing and validation, respectively. The mini-batch size and the momentum are set to be 50 and 0.7, respectively, as the parameter of MB-SGD-M method.

Since $N = 9$ and $k = 4$, the layer sizes of the NN used for Pt_9 PES fitting are (126, 6) - (126, 40) - (126, 70) - (126, 60) - (126, 2) - (40, 2) - (10, 2) - (1). Therefore, the total number of parameters (including weights and biases) used in this NN is 40 (6 + 1) + 70 (40 + 1) + 60 (70 + 1) + 2 (60 + 1) + 40 (126 + 1) + 10 (40 + 1) = 13022. The number of epochs, fitting errors and time are summarized in Table 1.3. In this table, the parameter guess column indicates that the final parameters of which NN are used as the initial guess of parameters. The errors are measured in Root Mean Square Error (RMSE). We note that within each epoch, all training samples are used once for training the NN with random order and then all testing samples are used once to give the testing error. After all epochs, the validation samples are used once to give the validation error. The parameters with the lowest testing error are selected as the final parameters.

After the PES fitting, 5000 unique initial structures are generated with the same parameters as those used in the first step, for each multiplicity, and then optimized to their local minima using the corresponding NN. Some NN-relaxed local minima with interatomic distances that are too long or too short, are considered as extrapolated structures and are excluded. Then the remaining NN-relaxed local minima are filtered with threshold value $d = 0.25$ Å. Finally, 300 filtered NN-relaxed local minima with relatively lower energies are selected to perform the full-step DFT optimization with def2-TZVP basis set.

Table 1.3 The number of epochs, average fitting errors per structure and total time used in training NN for Pt_9 PES.

NN	parameter guess	epochs	training error (meV)	testing error (meV)	validation error (meV)	total time (hour:min)
reference	random	1400	236	262	252	3:57
singlet	nonet	500	230	243	246	1:25
triplet	reference	500	151	158	153	1:25
quintet	triplet	500	137	147	146	1:25
septet	quintet	500	136	146	142	1:25
nonet	septet	500	129	131	132	1:25

1.7.2 Time efficiency

A separate direct global optimization without NN fitting is also performed for Pt_9 for comparison. During the direct approach, 500 unique initial structures are generated for each multiplicity using the same method and parameters as those in NN-PGOPT scheme. 30 additional converged local minima from singlet, triplet, quintet and septet are added to triplet, quintet, septet and nonet initial structures, respectively, since the converged minima of one multiplicity can be a good initial guess to be relaxed for a different multiplicity. Full-step DFT optimization with def2-SV(P) basis set is then performed on these initial structures. The threshold value d = 0.10 Å is used to filter the found local minima for each multiplicity individually. After the filtering, 264, 300, 300, 294, and 212 unique lower energy local minima are selected for DFT optimization with def2-TZVP basis set, respectively. The final results are filtered with the threshold value d = 0.25 Å.

The CPU hours used during the DFT optimization for the direct scheme and NN-assisted scheme are summarized in Table 1.4. We note that the total time for NN optimization and filtering for all structures and multiplicities is about 10 hours using a single CPU core. The total time for GPU accelerated NN fitting is shown in Table 1.3. Therefore, the time cost for NN fitting, optimization and filtering is negligible comparing to the DFT optimization time. In total, 23.5% CPU hours can be saved using NN-PGOPT scheme for Pt_9 global optimization, comparing to the direct approach. We note that at the first stage with def2-SV(P) basis set, since only the first 10 steps of DFT optimization are performed using NN-PGOPT scheme, 62% CPU hours are saved. However, because the quality of NN-relaxed structures is worse than that of the fully DFT-relaxed structures, at the second stage with def2-TZVP basis set, 67% more CPU hours are required for NN-PGOPT scheme. Nevertheless, in terms of the total time, NN-PGOPT is faster, given that the same number of initial structures is used.

Table 1.4 The CPU core hours used during the DFT optimization for the direct scheme and NN-assisted scheme.

approach	**basis set**	**singlet**	**triplet**	**quintet**	**septet**	**nonet**	**total**
	def2-SV(P)	11096	7794	8769	8750	13939	50348
direct	def2-TZVP	6066	4129	4039	4339	2819	21392
						total	**71740**
	def2-SV(P)	4344	2895	3271	3493	5182	19185
NN-PGOPT	def2-TZVP	9515	6787	6324	6735	6369	35730
						total	**54915**

1.7.3 Local optima searching performance

The final low-energy isomers found using the direct approach and NN-PGOPT are listed in Figure 1.23 and Figure 1.24, respectively. In the two figures, only structures with energy less than 0.40 eV relative to the putative global minimum (the leftmost structure) are listed. Green, purple, blue, orange, and brown lines represent the energies of singlet, triplet, quintet, septet and nonet multiplicities, respectively. The point group symmetry (if not C_1) is indicated for each structure. We note that the number of isomers in the selected energy range (0.0-0.4 eV), which corresponds to a catalyst condition related region, is the same for the two approaches. And the first four isomers with lower energies are also consistent. The inconsistency in higher energy region is due to the complexity of the PES and that 500 initial structures are not enough to generate a complete list of local minima. However, the requirement of completeness is both unnecessary and computationally too expensive.

Figure 1.25 shows some NN-relaxed structures with relatively low energy re-evaluated at DFT level. The single point energies (relative to the leftmost structure, which is 0.38 eV higher in energy than the putative global minimum in Figure 1.24) are re-evaluated at the TPSSh/def2-TZVP level of theory. Since these structures are not real local optima at the DFT level, their energies are generally higher than the DFT-relaxed optima shown in Figure 1.23 and Figure 1.24. However, the geometry of these NN-relaxed structures is already very close to their DFT-relaxed counterparts. For example, the NN-relaxed structures 1, 5, 2, 7, 8, and 3 in Figure 1.25 are very similar to the local optima 1, 2, 3, 4, 5, and 6 in Figure 1.24. We note that since the distances

between these NN-relaxed structures and their corresponding DFT optima vary, the energy order is not expected to be consistent.

Figure 1.26 shows the energy and step distribution during the first stage (with def2-SV(P) basis set) of the direct approach. The average energy of randomly created initial structures is 3.7 eV higher than the energy of the global minimum, while the average energy of the additional initial structures borrowed from local minima of other multiplicities is only 0.3 eV higher than the energy of the global minimum. This indicates that the effect of multiplicity change on energy is much smaller than that of geometry change. The average number of steps is 50 for the full optimization with def2-SV(P) basis set. Note that in the NN-PGOPT approach only the first 10 of these steps are performed, but we can only save 62% of the time. This is because the beginning steps generally take longer time.

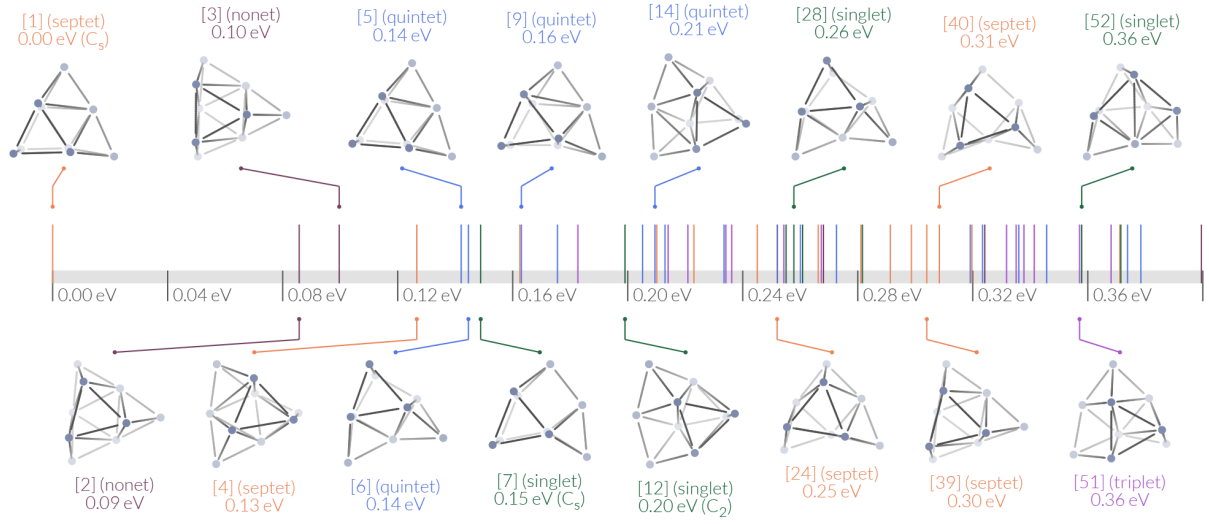


Figure 1.23 The energy and some of the structures of 58 low energy optima found at TPSSh/def2-TZVP level of theory using direct approach.

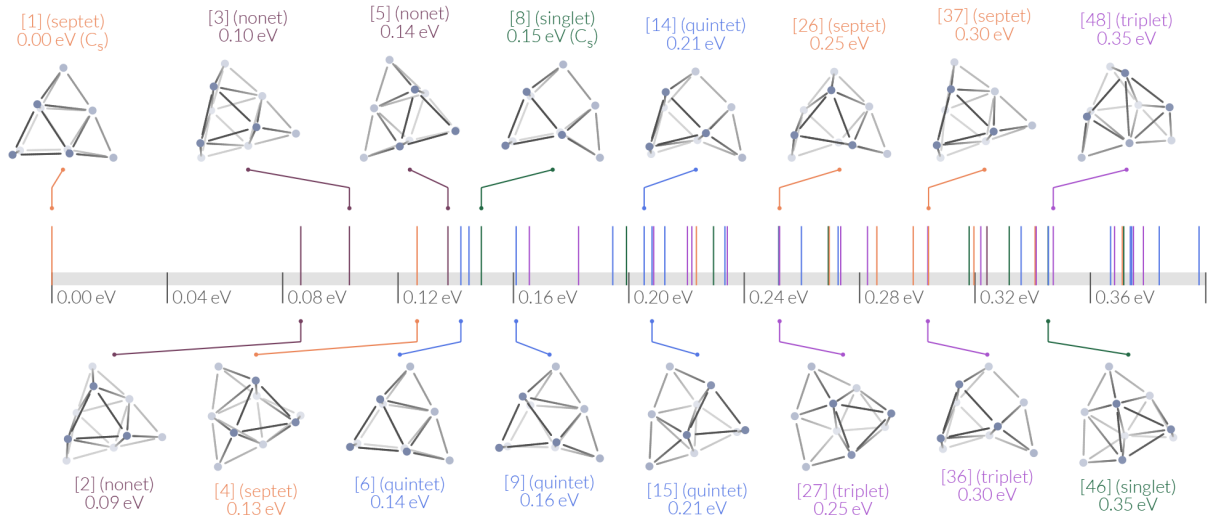


Figure 1.24 The energy and some of the structures of 58 low energy optima found at TPSSh/def2-TZVP level of theory using NN-PGOPT approach (after final full-step DFT optimization).

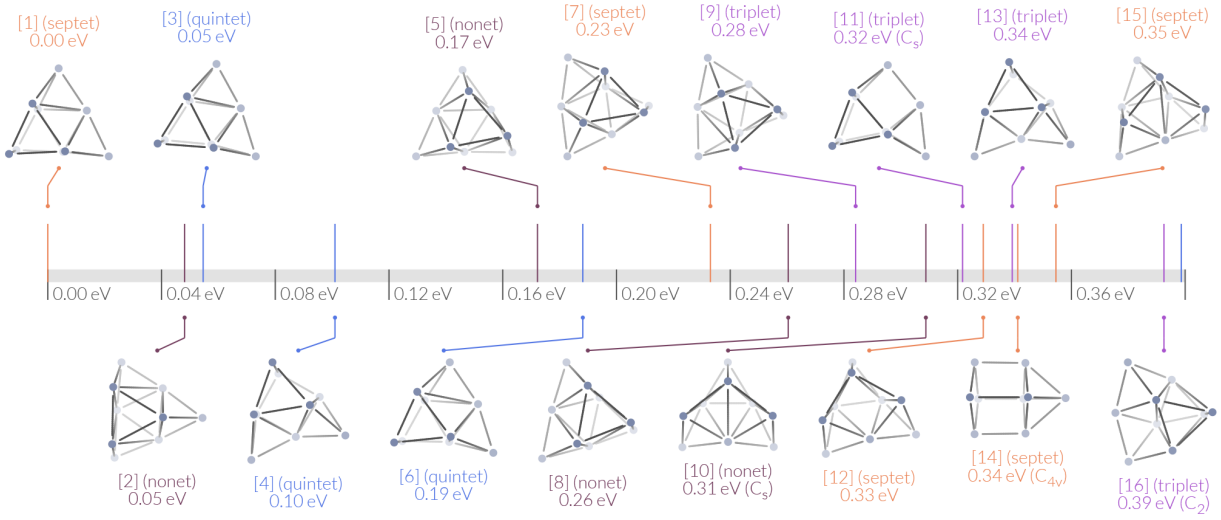


Figure 1.25 NN-relaxed low energy structures using NN-PGOPT approach (before final full-step DFT optimization).

Figure 1.27 and Figure 1.28 show the energy and step distribution during the second stage (with def2-TZVP basis set) of the direct and NN approach, respectively. The

average relative energy of NN-relaxed and def2-SV(P) fully relaxed structures are 1.5 and 1.2 eV, respectively, which means that the quality of NN-relaxed structures is slightly worse. However, the average relative energy of final re-relaxed structures for NN and direct approach are 0.8 and 1.0 eV, respectively. This indicates that in general the NN assisted approach is able to generate final structures that are more concentrated in low energy region.

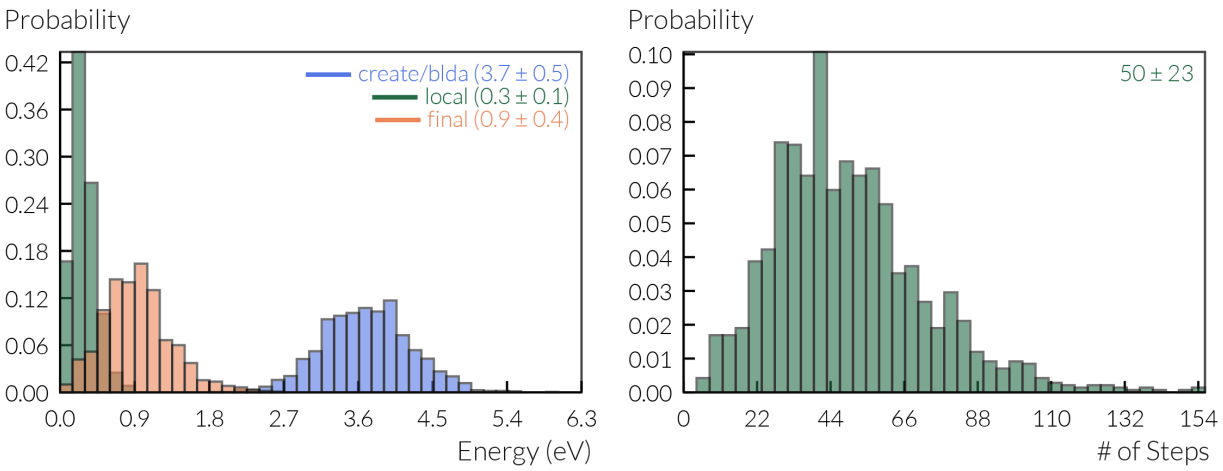


Figure 1.26 Left: the energy distribution of S-BLDA generated initial structures (blue), additional initial structures borrowed from local minima of another multiplicity (green) and relaxed local minima at TPSSh/def2-SV(P) level of theory (orange). Right: the TPSSh/def2-SV(P) full-step local optimization step distribution. The mean and standard deviation of the distribution are also indicated in the figure.

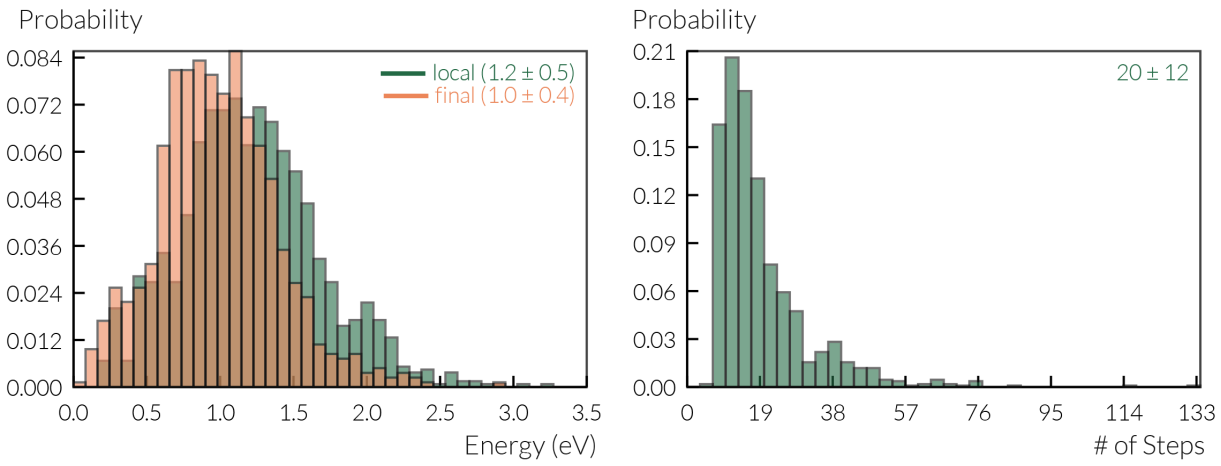


Figure 1.27 Left: the energy distribution (at def2-TZVP level) of def2-SV(P)-relaxed local minima structures (green), and re-relaxed local minima at TPSSh/def2-TZVP level of theory (orange). Right: the TPSSh/def2-TZVP local optimization step distribution.

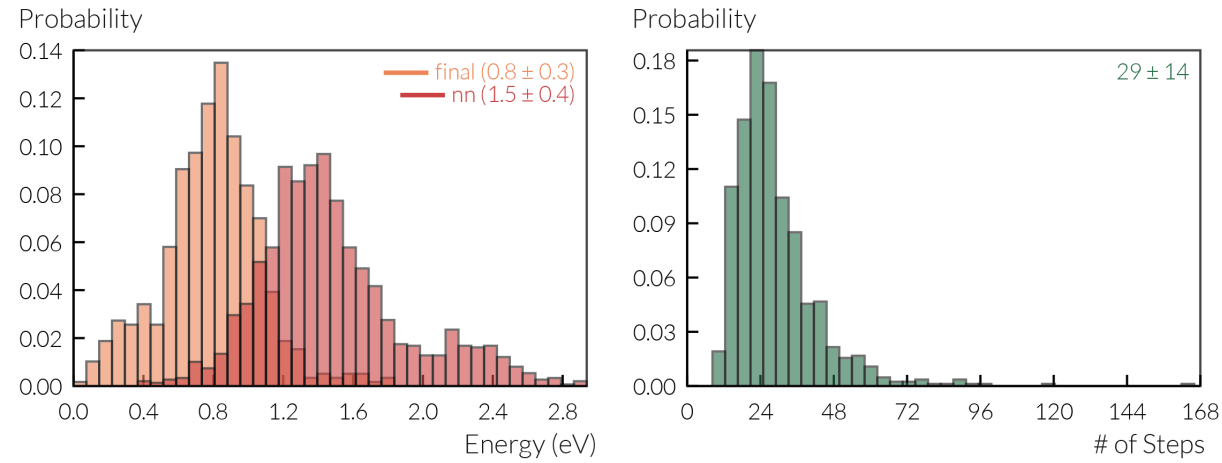


Figure 1.28 Left: the energy distribution (at def2-TZVP level) of NN-relaxed local minima structures (red), and re-relaxed local minima at TPSSh/def2-TZVP level of theory (orange). Right: the TPSSh/def2-TZVP local optimization step distribution.

1.7.4 Putative global minimum

In this work the septet tricapped octahedron (*A*), shown in Figure 1.29, is found to be the putative global minimum for Pt_9. The TPSSh and PBE energies relative to the energies of respective lowest energy structure are also listed in Figure 1.29. The TPSSh energy is evaluated using Turbomole 6.6 with def2-TZVP basis set. The PBE energy is evaluated using VASP 5.4.1 with 20 Å cubic cell and energy cutoff 500 eV. Singh and Sarkar [63] have also reported this structure (with nonet multiplicity) as the lowest energy isomer. However, Kumar and Kawazoe [64] and Chaves *et. al.* [65] have reported the planar structure with four squares (*B*) as the lowest. Winczewski *et. al.* [66] have reported the defected tetrahedron (*C*). We re-optimized the three structures using both TPSSh and PBE functional. Based on the calculated energies, the difference in global minimum structures is due to the different choice of DFT functional.

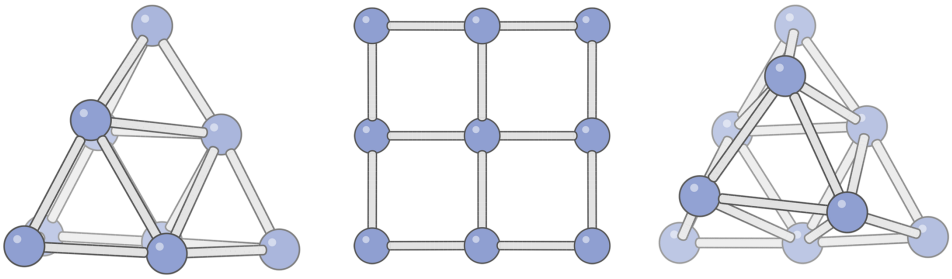


Figure 1.29 The putative global minima proposed in this work (*A*) and other literatures (*A*, *B*, *C*). *C* is also found in this work.

1.8 Application on Pt_{13} cluster

1.8.1 DFT parameters and fitting errors

The Pt_{13} global optimization is performed as another example of the application of NN-PGOPT scheme. The DFT optimization is done using spin-polarized DFT program implemented in Vienna Ab-initio Simulation Package (VASP) 5.4.1 [47] with Projector Augmented-Wave (PAW) method [67] and Perdew-Burke-Ernzerhof (PBE) [50] functional. The cubic cell of size 15×15×15 $Å^3$ is used. The multiplicity is automatically relaxed during the geometry optimization. Only the gamma point is sampled. To save time, the DFT optimization at both the first and second stage of NN-PGOPT scheme is performed with energy cutoff 280 eV. The final local optima are re-optimized with energy cutoff 400 eV but no significant geometry or energy change is observed.

1000 unique initial structures are generated using the S-BLDA approach with the same parameters as the Pt_9 case. 20 DFT optimization steps are performed to yield 19,860 structure/energy samples. After discarding some high energy structures, 19,741 samples spanning an energy range of 5.44 eV are used. 82%, 9% and 9% of samples are atomically shuffled to generate 200,000, 15,000, and 15,000 samples for training, testing and validation, respectively. The mini-batch size and the momentum are set to be 50 and 0.7, respectively, as the parameter of the MB-SGD-M method.

Since $N = 13$ and $k = 4$, the layer sizes of the NN used for Pt_{13} PES fitting are (715, 6) - (715, 40) - (715, 70) - (715, 60) - (715, 2) - (40, 2) - (10, 2) - (1). Therefore, the total number of parameters (including weights and biases) used in this NN is 40 (6 + 1) + 70 (40 + 1) + 60 (70 + 1) + 2 (60 + 1) + 40 (715 + 1) + 10 (40 + 1) = 36582. The NN is trained for 700 epochs but no error improvement has been observed since the 410th

epoch. The final training, testing, and validation errors are 268, 298, and 303 meV, respectively.

After the PES fitting, 5000 unique initial structures are generated and then optimized to their local minima using the corresponding NN. Extrapolated structures and duplicates are excluded. Finally, 500 NN-relaxed local minima with relatively lower energies are selected to perform the full-step DFT optimization.

1.8.2 Local optima searching performance

Some of the final low energy isomers found using the NN-PGOPT approach are listed in Figure 1.30. Although the PES of Pt_{13} is expected to be much more complicated than that of Pt_9, fewer low energy isomers are found within 0.4 eV region. This indicates that the energy of the global minimum is significantly lower than most of other local isomers. This is different from the situation of Pt_9 and may have important practical consequences, which is investigated in detail in the next subsection. Figure 1.31 shows the energy distribution of initial, partially DFT-relaxed, NN-relaxed, and fully DFT-relaxed structures. In this figure, all energies are evaluated at the DFT/PBE level of theory. From the average values shown in the figure, we can infer that the constructed NN has the ability to predict new lower-energy structures than its input, because the average energy of NN-relaxed structures is 0.9 eV lower than that of the partially relaxed. Therefore, DFT local optimization time is saved by the combination of partial DFT optimization and NN fitting.

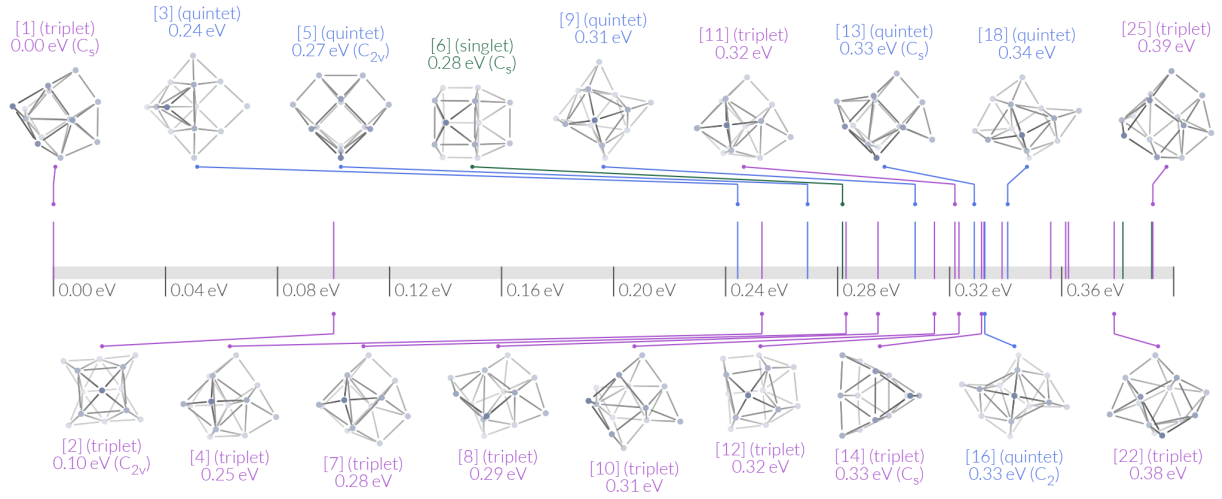


Figure 1.30 The energy and some of the structures of 25 low energy optima found at the DFT/PBE level of theory using the NN-PGOPT approach.

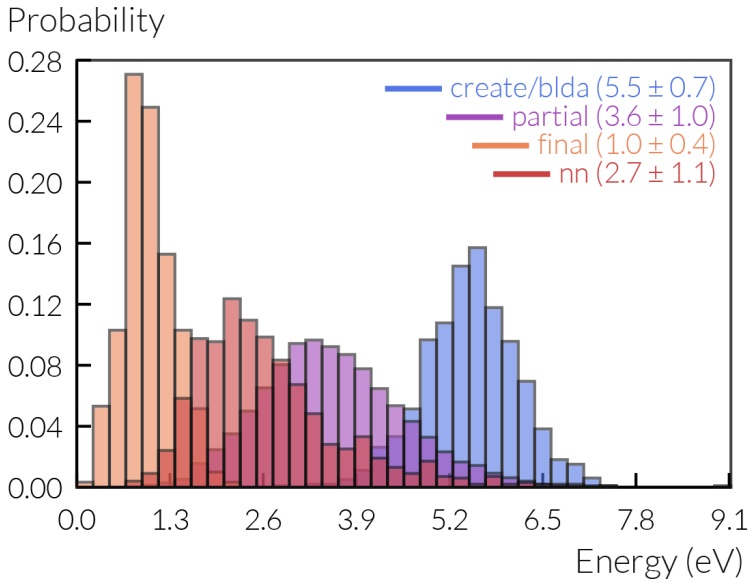


Figure 1.31 The energy distribution of S-BLDA generated initial structures (blue), all intermediate structures during the 20-step DFT optimization (purple), NN-relaxed local minima structures (red), and DFT re-relaxed local minima (orange).

1.8.3 Putative global minimum

In this work the triplet tricapped pentagonal prism (*A*) is found to be the putative global minimum for Pt_{13}, shown in Figure 1.32. In this figure, the shown PBE relative energies are evaluated using VASP 5.4.1 with 15 Å cubic cell and energy cutoff 400 eV. Single point energies at the TPSSh/aug-cc-PVTZ level of theory are calculated for each of these structures using MOLPRO 2015.1. [68] Many other publications have also reported this structure as the putative global minimum. [65,69,70] However, Sun *et. al.* and Zhang *et. al.* have reported the square pyramid (*B*) as the global minimum. [71,72] As shown in Figure 1.30 and Figure 1.32, the square pyramid structure is 0.10 eV higher in energy than the lowest one. Since the same DFT parameters are used, it may suggest that our search is more complete.

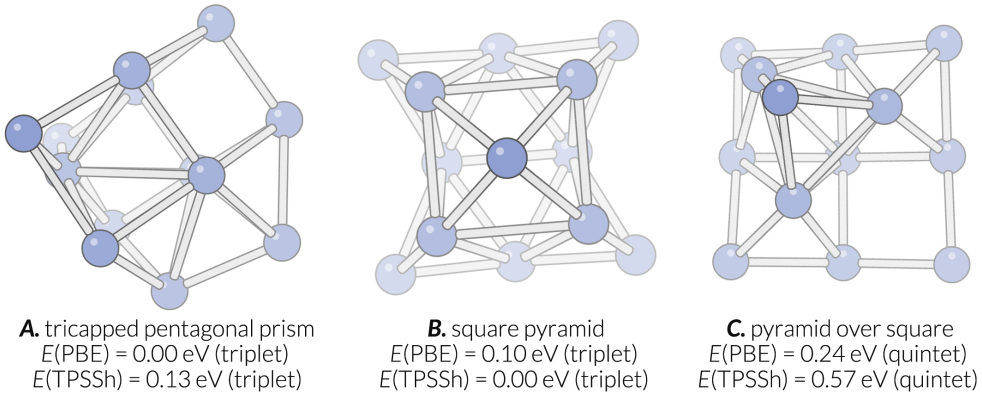


Figure 1.32 The putative global minima proposed in this work (*A*) and other literatures (*A*, *B*) and a third-lowest isomer found in this work (*C*).

1.8.4 Ensemble-average representation of Pt clusters

Both Figure 1.29 and Figure 1.32 show that different DFT functionals could result different relative order of isomer energies, and thus change the geometry of putative global minima. This may be partially due to the complexity of PES of both Pt_9 and

Pt_{13}. The energy distribution of the two clusters investigated show different patterns, especially if we only focus on PBE energies. Specifically, there are many more Pt_9 isomers distributed in the low energy range (0 to 0.25 eV). To understand the impact of these observations, we apply the ensemble-average representation introduced in Section 1.5 to the two clusters.

Figure 1.33 and Figure 1.34 show the calculated probabilities of first five low energy isomers of Pt_9 and Pt_{13} clusters at finite temperatures, respectively. For Pt_9, the PBE global minimum at zero temperature (planar square, quintet) is no longer the most populated one at $T > 200$ K, mainly due to the fact that the second lowest isomer (tricapped octahedron, septet) has a higher multiplicity and lower symmetry, which are favored at high temperature. For Pt_{13}, the structure favors its global minimal shape (tricapped pentagonal prism, triplet) till $T = 1000$ K. After that, the third lowest isomer (pyramid over square, quintet) becomes more populated. We note that the second lowest isomer is not favored because of its slightly higher symmetry.

The probability change between different temperature indicates that there can be a structure transition for Pt_9 (with PBE functional) at low temperature ($T < 200$ K), and for Pt_{13} at relatively high temperature ($T \sim 1000$ K). The structural stability over a wide temperature range, predicted here for Pt_{13} may play an important role in the explanation of catalyst selectivity.

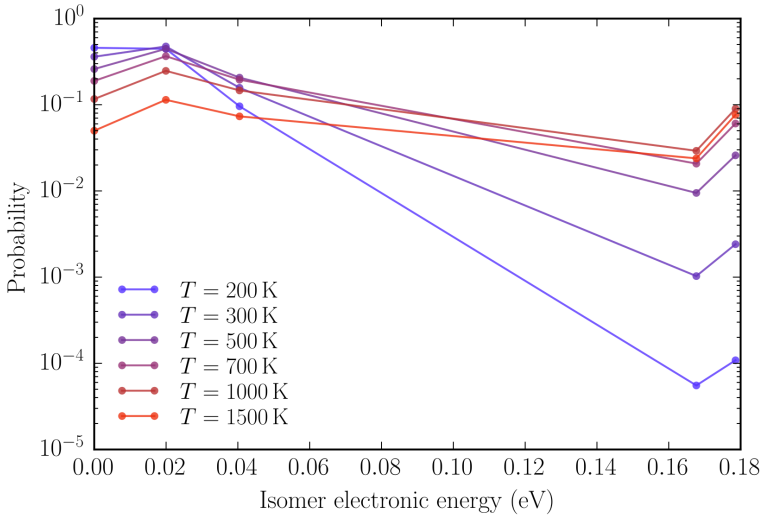


Figure 1.33 The occurrence probabilities of first 5 low energy isomers of Pt_9 clusters at different temperatures, with energy and frequencies evaluated using PBE functional.

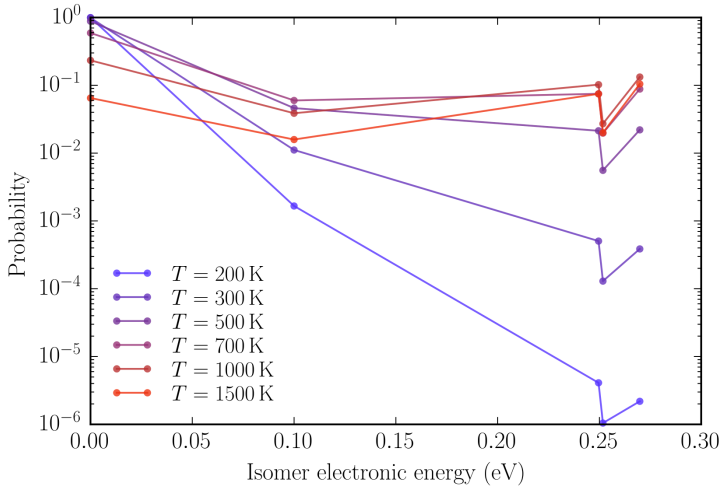


Figure 1.34 The occurrence probabilities of first 5 low energy isomers of Pt_{13} clusters at different temperatures, with energy and frequencies evaluated using PBE functional.

Figure 1.35 and Figure 1.36 show the heat capacity of the two cluster isomer ensembles as a function of temperature, for electronic and all degrees of freedom, respectively. From the heat capacity contributed from the electronic degree of freedom, we can clearly identify a peak at T = 100 K for Pt_9 and a peak at T = 950 K for Pt_{13}. This agrees with our previous observation on probabilities. We note that the peak for Pt_9 is too small to be observed in the total heat capacity, and the peak for Pt_{13} in total heat capacity is shifted to T = 720 K. This can be ascribed to the vibrational and rotational entropy effects. The dashed lines in Figure 1.35 and Figure 1.36 show that a significant underestimation of heat capacity (and also some other properties) would be expected if we only consider several low energy isomers. Nevertheless, the first ten low energy isomers will be a good approximation if one focuses on phenomena at room temperature, for these systems. Figure 1.36 shows that the high temperature limits of total heat capacity for Pt_9 and Pt_{13} are close to 22.5 and 34.5, respectively. This is related to the total number of degree of freedom that we have considered, namely, $3N$ - 6 from the vibrational and 1.5 from the rotational ones, where N is the total number of atoms. Electronic degrees of freedom, however, do not contribute at high T.

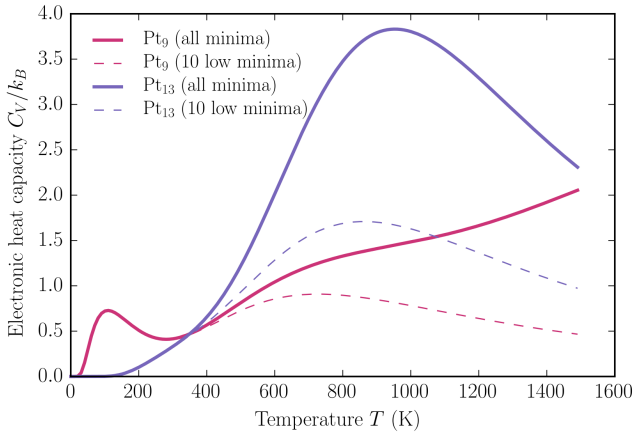


Figure 1.35 The heat capacity contributed from the electronic degree of freedom for isomer ensemble of Pt_9 (red) and Pt_{13} (blue), when all found local minima (solid line) and only the first ten low energy isomers (dashed line) are considered.

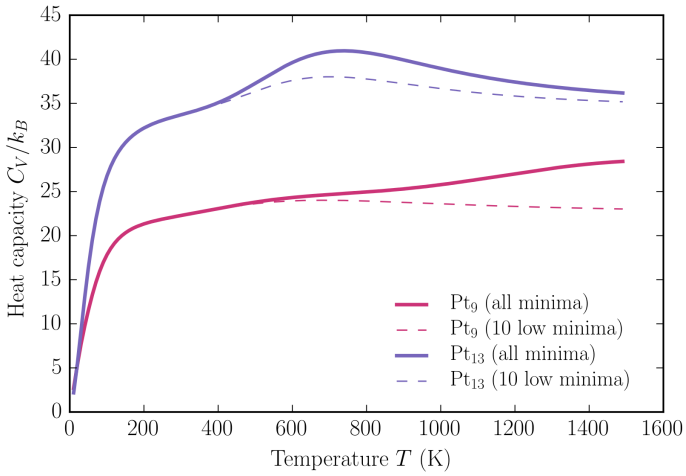


Figure 1.36 The heat capacity contributed from the electronic, vibrational, and rotational degree of freedom for isomer ensemble of Pt_9 (red) and Pt_{13} (blue), when all found local minima (solid line) and only the first ten low energy isomers (dashed line) are considered.

Finally, in Figure 1.37 we note that the electronic property of the clusters at finite temperatures can be different from the ones evaluated on just the global minimum.

For example, The Vertical Ionization Potentials (VIPs, $E(Pt_n^+)$- $E(Pt_n)$) of Pt_9 and Pt_{13} show different trends when temperature increases. The VIP of Pt_{13} changes relatively slowly, due to fact that the energies of Pt_{13} isomers are sparsely distributed.

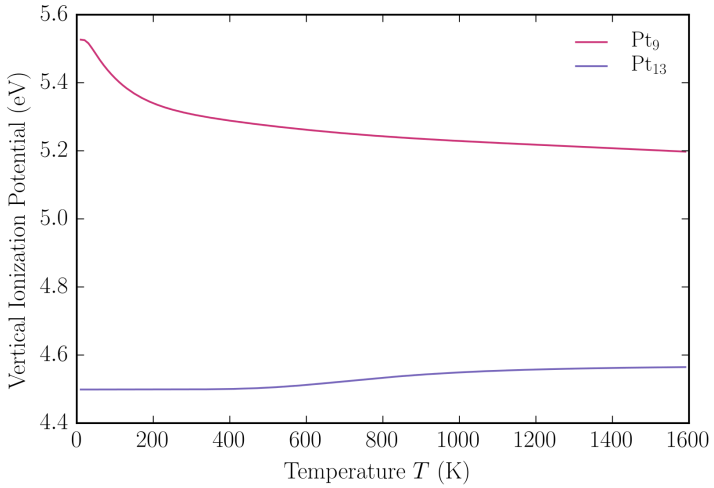


Figure 1.37 The ensemble-averaged vertical ionization potential of Pt_9 (red) and Pt_{13} (blue), evaluated at different temperatures.

Chapter 2

Global optimization of surface supported clusters

Despite the simplicity of the gas phase cluster model, in most cases, it is not the most realistic model for industrial catalysis. Surface-supported transition metal cluster, as a type of heterogeneous catalysis system, is more attractive in modern industry. For example, oxide surface supported catalytic sub-nanometer platinum clusters have found wide application in industrial reactions involving (de-)hydrogenation of hydrocarbons, such as naphtha reforming and alkane non-oxidative dehydrogenation [73–79] Industrial catalysts correspond to low loading of Pt supported on alumina, with a corresponding particle size of 0.6 to 1.1 nm. [80]

In general, the global optimization of the surface supported cluster system is more difficult than that of gas phase clusters. One of the reason is that the typical model system contains hundreds of atoms, including the surface slab. However, the surface atoms are often far less mobile than the cluster atoms during the relaxation process. They also have different symmetry properties (space group symmetries) from the cluster atoms. This heterogeneity feature makes it difficult to extend the fitting technique to surface supported clusters, for global optimization purpose. In contrast, the BH approach can be easily extend to heterogeneous systems. Therefore, in this chapter, we will focus on the BH based global optimization approaches.

2.1 Basin hopping approach

The traditional BH global optimization is based on the canonical Monte Carlo (MC) simulation, [9] where at each step the coordinates of the current structure are perturbed by a random distance (called perturbation), and the resultant structure is then geometrically optimized to a local minimum. The optimized structure can be either discarded or accepted, and the decision is based on the probability dependent on energy difference from the previously accepted structure and the temperature. The core step in BH approach is the perturbation. For heterogeneous systems, we can easily perturb only the cluster part, but keep surface atom fixed.

2.2 Global optimization with hydrogen coverage

In a typical computational study of catalytic clusters, after the shape of catalytic clusters is identified by global optimization, the structures of global minimum and possibly some other low-energy isomers can be selected as reference states, for studying active sites for various reactants and corresponding reaction pathways. [79,81–83] However, despite the relatively low computational cost, this scheme may not be reliable if the low-energy bare-cluster isomers do not appear to be the most active ones, or if a big shape change occurs upon the adsorption of reactants. In particular, this will be the case when the catalytic cluster itself is highly fluxional, or if adsorbates affect the cluster shape. [84,85] In this situation, the actual geometry change can only be tracked through a global optimization for the whole system, including the cluster, reactants and hydrogen coverage. This is a highly challenging task as the configuration space is significantly larger than that for a routine global optimization. It also requires adapted algorithms to preserve the identity of the reactant adsorbate during global optimization.

In 2011, an early computational study by Sautet and coworkers presents a surface stability diagram for a Pt_{13} cluster supported on γ-Al_2O_3 indicating the H coverage as a function of H_2 pressure and temperature. It shows that a high H coverage can be reached and that the cluster structure changes from biplanar to cuboctahedral geometry when high hydrogen coverage is above 18, [84] where molecular dynamics (MD) simulations have been used to obtain the low-energy isomer of hydrogen adsorbed structure. It was followed in 2013 by a study of the C_1 and C_2 intermediates resulting from the dissociative adsorption of ethane on the supported Pt cluster, showing a strong influence of the nature of the intermediate on the equilibrium H coverage and on the cluster shape. [76] Very recently, using the similar MD approach Chizallet et al studied PtSn clusters supported on γ-Al_2O_3, at different hydrogen coverages. The study shows that both Pt/Sn ratio and number of hydrogens play an important role in cluster shape and properties. [85] In 2015, a computational study performed by Pacchioni et al indicates that, at high hydrogen coverage, hydrogen atoms can transfer from Ru_{10} cluster to surface, where the support surfaces are anatase TiO_2 and tetragonal ZrO_2 (101) surfaces, but the complete global optimization was not used in the study. [86] Recently, Sun and Sautet proposed a modified genetic algorithm for global optimization of gas phase Pt_{13} cluster with high hydrogen coverage, which shows the importance of considering low-energy metastable cluster isomers in reactions. [87]

Here we present a method for global optimization of surface-supported Pt clusters under high hydrogen coverage, based on a modified BH scheme. In our approach for the global optimization of adsorbate covered supported cluster, we first improved the

initial structure generation algorithm based on bond length distribution by introducing constraints on the bonding element types. Then the core-shell separation is used in the perturbation step in BH method, to help keeping the integrity and reasonableness of the perturbed structure.

2.2.1 Initial structure generation

The method of initial structure generation for starting BH global optimization is based on BLDA introduced in Section 1.2.5, which can successfully applied to adsorbate-free Pt_7 supported on α-alumina. In BLDA, the structure is constructed by each time adding one atom or fragment to the existing structure. The position of new atom is determined by selecting the distance between the new atom and one old atom from a normal distribution. In this work, in order to create the cluster structure with hydrogen coverage, we first add Pt atoms (core part), then add CH_3 and H atoms (shell part) (Figure 2.1), because it is known that these adsorbates bind on the surface of Pt clusters and do not migrate inside the small Pt cluster. To further improve the initial guess, in this case, the ways Pt atoms and adsorbates are added are slightly different. The new Pt atom is required to be connected to at least two atoms already in the system (including existing Pt atoms and surface atoms), while the new CH_3 fragment or H atom is required to be connected to at least one Pt atom in the cluster. Thus the H atoms can take both atop and bridge sites, or somewhere in-between.

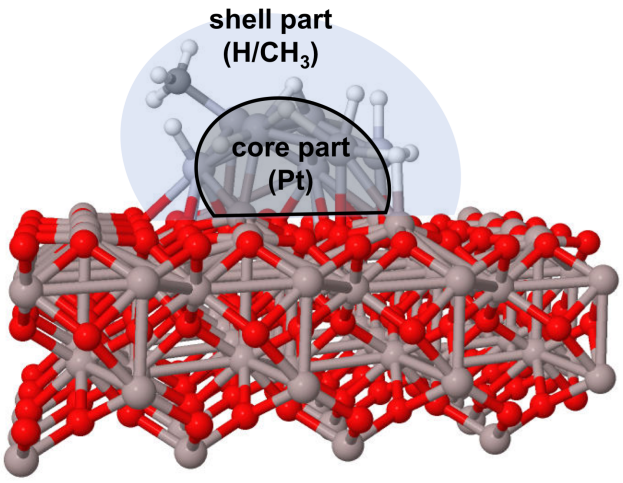


Figure 2.1 Core-shell separation of $Pt_7H_{10}CH_3$ structure supported on the α-Al_2O_3 (0001) surface slab model.

2.2.2 Revised basin hopping approach

Several issues arise when the traditional BH global optimization is applied to a surface supported cluster under high coverage. Taking our testing system ($Pt_7H_{10}CH_3$ on α-Al_2O_3) as an example, first of all, the perturbed $Pt_7H_{10}CH_3$ structure can easily become disconnected, if the default perturbation distance for Pt is of the same order of magnitude as the Pt-H bond length. Secondly, the CH_3 fragment may lose integrity during the random perturbation. We also note that the α-Al_2O_3 surface atoms are not very mobile during the optimization. It is thus reasonable to not move any oxide surface atoms during the perturbation. However, the perturbed Pt atoms may then be at a too short distance from the surface atoms, which would possibly introduce some difficulties in the SCF procedure in DFT calculation.

We have introduced a partial FF optimization to solve the short-distance problem, which is applied after the perturbation. The FF used in this work is the Lennard-Jones (LJ) potential, given by the following formula:

$$E = 4\epsilon \sum_{i<j} \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right]$$

Typically, the parameters ϵ and σ should be determined for each pair of atoms. However, in our perturbation procedure with core-shell separation, each time only atoms with one certain element type are allowed to move (first Pt, then C, and finally H), so the parameter ϵ only scales the energy, which is then scaled again by the optimization step length. Therefore, if we further assume that only one-type of pair-interaction (Pt-Pt, Pt-C, and Pt-H) dominates the energy in every perturbation phase, we can ignore the parameter ϵ (namely, setting $\epsilon = 1$). Note that here the propose of FF is to generate reasonable perturbed structures for further accurate quantum mechanics treatment, not for accurate energy evaluation, especially when the FF optimization is only partially converged. Therefore, we think that this rough approximation should be appropriate. The equilibrium separation σ is selected to be the sum of covalent radii of the pair of elements.

The FF optimization is partial in the sense that it will terminate after no pair of atoms has a distance smaller than 0.7 times the sum of covalent radii, which means that this partial optimization will not lead the structure to a local minimum on the FF PES, but will only make it good enough for starting a DFT optimization. At very

short pair-distance, the force of the system is dominated by the pairwise nuclear repulsion, which is classical and can be reasonably approximated by the LJ potential.

Similar to some other BH implementations, we introduced the site swap operation (with a 25% occurrence rate) for CH_3 fragment, before the actual perturbation. During this operation, the CH_3 fragment is randomly swapped with a hydrogen atom, in order to moderately sample all possible CH_3 binding sites. In addition, in order to keep the core-shell structure of the system, we perform the perturbation of atoms in core part and shell part separately. As a first step, each H/CH_3 is labeled by the index of cluster atom to which it is bound. Then all cluster atoms are randomly perturbed, with a magnitude of distance equal to 0.4 times the covalent radius of the involved element (Pt, in our case). Then the cluster and surface parts are partially optimized using FF to make sure that no two cluster atoms are too close to each other (with the surface part kept fixed). After this, H/CH_3 are moved by exactly the same distance and direction as the cluster atom they bind to. In the Pt cluster case, this process guarantees that the Pt-H and Pt-C bonds are unchanged after the perturbation of Pt atoms. Then we apply the perturbation on H/CH_3 coordinates, with a magnitude of distance equal to 0.6 times the covalent radius of the elements (H or C). Here, the factor for the perturbation distance is larger for light elements, since light element atoms are in general more mobile than heavy cluster atoms. The partial FF optimization is then applied again, this time for the whole system (with the bare-cluster and surface parts kept fixed). Finally, we need to detect whether the structure becomes disconnected after the perturbation. If any atom or fragment is separated from the

main structure, it will be added back using the initial structure generation algorithm discussed above.

Figure 2.2 provides a detailed explanation of this procedure, using a local minima structure as the starting structure, as an example. In this figure, isomer B1 (see Section 2.4) (a) is selected as a starting structure. First, we will make perturbation on Pt atoms, so all CH_3/H atoms are removed, but their relative position to nearest Pt atom is recorded. The bare Pt cluster is shown in (b). After making perturbation on Pt atoms (like the procedure in any traditional Monte Carlo algorithm), the structure is shown in (c). Now it is likely that some Pt atoms become too close to each other. So a partial FF optimization is applied to (c), giving structure (d). Note that during the FF optimization the surface atoms are all kept fixed. Then we add back CH_3 (represented as a single C atom in the figure, since C-H bond length will not be perturbed), according to previously recorded relative position, giving structure (e). Making movement on CH_3 gives structure (f). Now we add back all H atoms according to their relative position to nearest Pt, and then make perturbation on H atoms (not including H atoms in CH_3), the resulting structure is shown in (g). Note that the perturbation step length for H is slightly larger than that for Pt, so we can see that some H atoms go far away from the cluster and some H atoms become too close to Pt. Now we first detect all H atoms that are not connected to the cluster, and remove them. For the remaining H atoms, a partial FF optimization is applied (with all surface atoms and Pt atoms kept fixed), which gives structure (h). Finally, the removed unconnected H atoms are added back using our structure generation (sub-) algorithm. The structure after perturbation is shown in (i).

Note that the structure generation algorithm that we have used in this work creates a new structure by every time adding one atom (or molecule). So this sub-algorithm for adding one atom to a given structure perfectly fits the need here. Because of this last step, when a H atom goes away from its original position and then added back to a new position, its binding site is automatically swapped. In more general case, the site swap operation can also happen for CH_3, with a small occurrence rate (25%). This procedure is implemented between (a) and (b) in Figure 2.2. The site swap is implemented by first remove one or two CH_3/H fragments, then add them back to random new sites, according to the procedure defined in the structure generation algorithm (see Figure 1.9), by randomly selecting a Pt atom, a Pt-H or Pt-C distance from normal distribution, and a direction from the sphere around that Pt. The algorithm does not distinguish atop and bridge sites internally, but mathematically the sphere around all Pt atoms will cover all possible atop, bridge and hollow sites, and somewhere between them. The local minima structures found using this algorithm also show that we have a good sampling on different binding site types. In our algorithm, multiple adsorbates are allowed to bind to the same Pt atom, and the coordination number of a certain H will not have influence on whether this Pt will be selected for binding new adsorbates. However, the collision detection procedure will be applied every time an atom is added to the existing structure, so when the coordination number of a certain Pt is very high, it is unlikely that the new structure will pass the collision detection procedure, and another Pt (possibly with lower coordination number) will be selected for binding adsorbates.

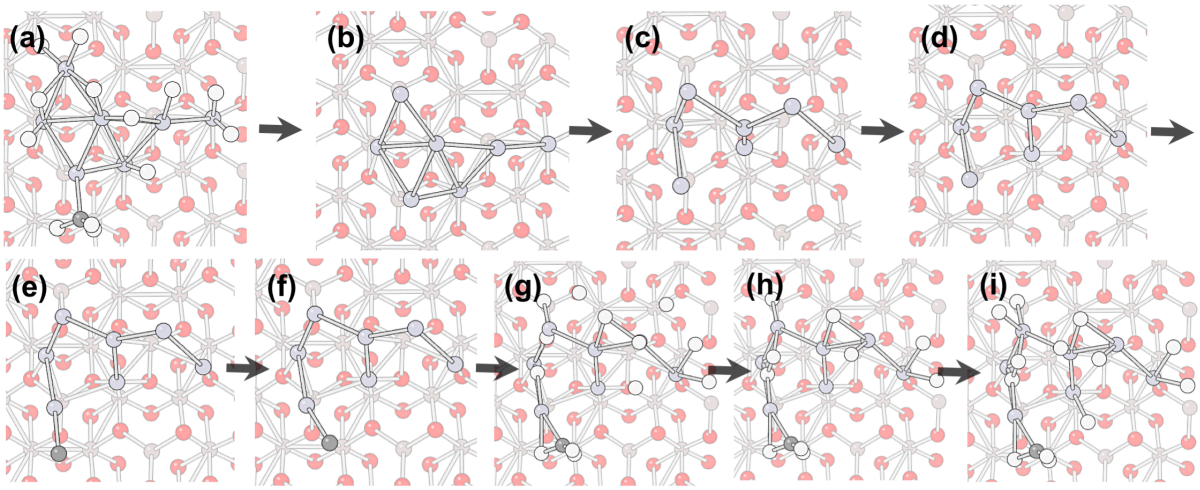


Figure 2.2 The perturbation procedure with core-shell separation, starting from a relaxed structure.

2.2.3 Binding site sampling

Due to the large number of atoms in the system under consideration ($Pt_7H_{10}CH_3$) and high mobility of hydrogen atoms compared with the cluster core, or pure Pt cluster systems, we note that in practice the above BH process may actually miss the true global minimum or some important low-energy local minima, if only thousands of structures are optimized and sampled. Instead, it is much easier for the BH to find the low-energy core part (whose composition is Pt_7 in our test system). Upon the optimized core part, the H/CH_3 adsorbates can occupy different binding sites, producing a large number of configurations. Therefore, it would be possible to get better global minimum candidates by keeping some Pt core parts found by BH and sampling more binding sites for H/CH_3.

In our approach, a random binding site sampling is performed for several low-energy isomers obtained from BH optimization. For each isomer, we randomly change the

binding site of the CH_3 fragment, and/or several H atoms, to generate new structures. These new structures are then optimized to nearest local minima to see if they can give new lower-energy isomers.

2.3 Application on *α*-alumina supported Pt_7 cluster

A parallel version of BH method has been implemented and used to find low-energy local minima for *α*-alumina supported Pt_7 cluster. The sampled statistical ensemble is kept at the temperature of 1,500 K, to ensure a big chance to cross the barrier. The BH algorithm has been implemented for both detailed balanced version and non-detailed-balanced version, where the latter is used in this work. In the non-detailed-balanced version, the detailed balance principle is violated, but this will enable the algorithm to reach the global minimum more rapidly. The initial structures, which are generated using second-type BLDA, can be very close to the equilibrium structure, thus improving the ergodicity at low-energy region, which is especially important when the total number of moves is reduced to achieve better parallel scalability. In this work, 10 parallel Monte Carlo walkers are used, with each performing 200 moves. At one move only the cluster part of the system will be perturbed. After each move VASP is invoked as an external program to perform local optimization using the non-linear CG numeric algorithm, which is a step bringing the structure to its closest local minima. The alumina surface support is modeled as a 3×3 unit cell slab model, with lattice constants $a = 4.807$ Å and $c = 13.117$ Å.

Figure 2.3 shows low-energy isomers found using BH approach. In this figure, first, we included 20 isomers with energy lower than 0.6 eV relative to global minimum

(GM) found using the BH approach. Second, some intermediates found along the optimized paths, which are necessary to build the isomerization graph, are included (see Chapter 4 for details).

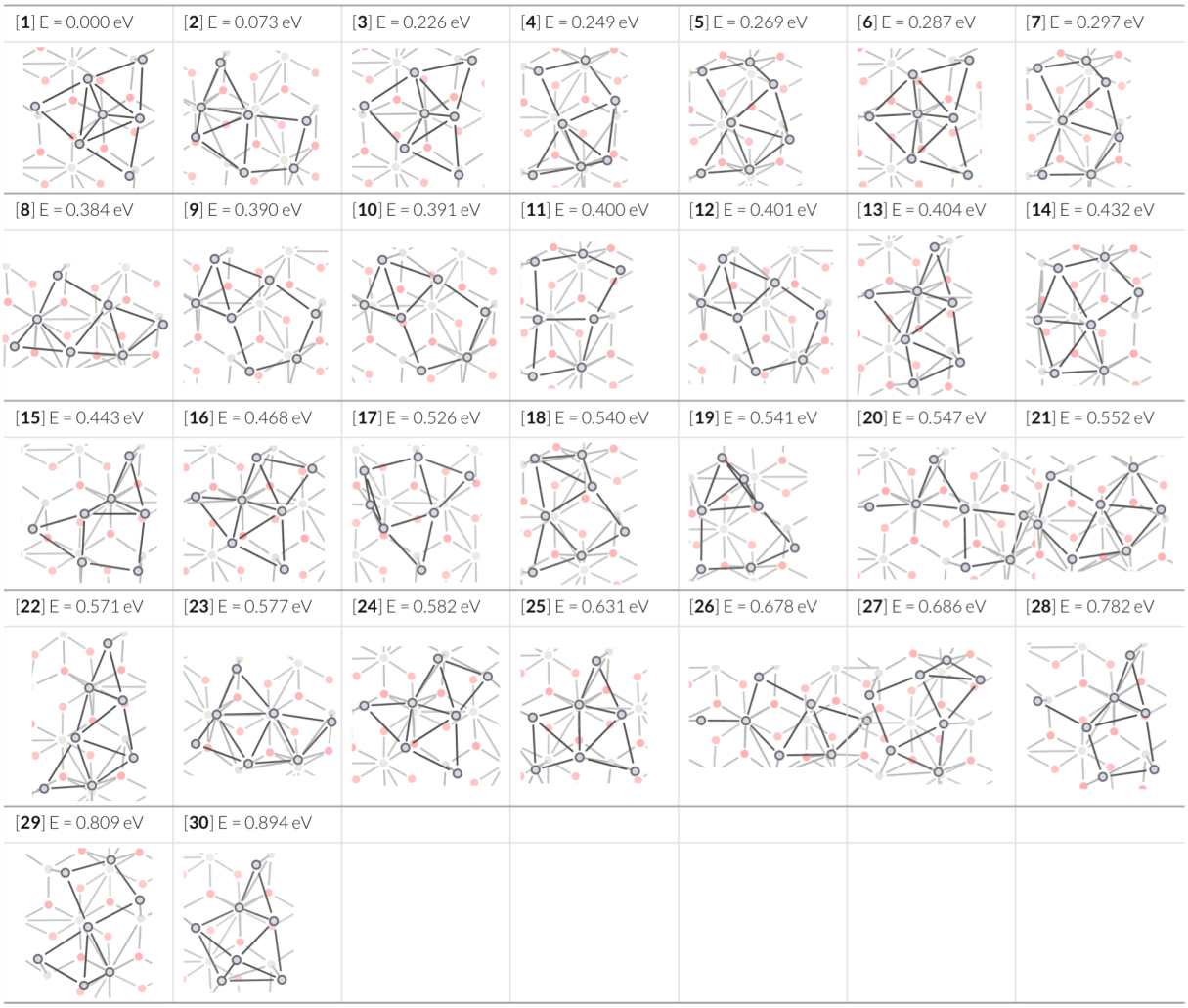


Figure 2.3 The low-energy isomers for Pt_7 cluster supported on Al_2O_3.

2.4 Application on Pt_7 cluster with hydrogen coverage

The $Pt_7H_{10}CH_3$ on α-Al_2O_3 serves as a testing system for the revised BH approach for hydrogen coverage. The Pt_7 size is both realistic for a sub-nanometer size cluster and

its global optimization with 11 adsorbates is computationally feasible. An initial coverage of 9 H is selected, since this corresponds to 1.3 H per Pt, a realistic coverage at typical temperature (800 K) and H_2 pressure (1-10 bar) used in catalytic reforming reactions.[76] Methane is used as a model alkane and a first intermediate for its dehydrogenation ($H+CH_3$) is added to the hydrogenated cluster, hence providing the $Pt_7H_{10}CH_3$ formula.

The DFT calculation is performed using Vienna Ab initio Simulation Package (VASP) 5.4.4, [47] with the Projector Augmented-Wave (PAW) method [67] and the Perdew-Burke-Ernzerhof (PBE) functional. [50] The 4×4 slab model for α-Al_2O_3 (0001) with lattice constants a = 4.807 Å and c = 13.117 Å is used. After adding cluster part, the vacuum gap along direction normal to the surface is at least 15 Å. To reduce the computational cost, the thickness of slab model is $c/2$. The lower half of the surface model is fixed during geometry optimization, and only Γ point is sampled in k-space. dDSC vdw correction [88,89] is used in the geometry relaxation step of the global optimization. The kinetic energy cutoff is set to be 500 eV for final results.

The global optimization of $Pt_7H_{10}CH_3$ deposited on α-Al_2O_3 surface was performed by 10 parallel BH walkers, with each walker moving 300 steps. After the BH optimization, more than 1,800 additional geometry optimization were performed for binding site sampling. In total, 1,124 geometrically unique local minima were found for the system, spanning an energy range of 3.0 eV. The energy distribution of all found local minima is shown in Figure 2.4(a). There are 58 local minima in the low-energy range (up to 0.4 eV above the energy of putative global minimum (GM)). In our previous study, we found that, for the bare Pt_7 cluster deposited on α-Al_2O_3, there are only 11

unique local minima within the 0.4 eV energy range. The reason for the increase in number of low-energy isomers appears to be the geometric variety of the adsorbed H/CH_3. In fact, many of the $Pt_7H_{10}CH_3$ structures share the same Pt_7 core shape. The 58 local minima can be classified into 6 categories (labeled by A, B, C, D, E and F), according to their Pt_7 core geometries, as shown in Figure 2.4(b) and Figure 2.5. In the two figures, different colors and letters (A to F) correspond to different Pt_7 core structures. Table 2.1 lists the number of structures in each core shape category. We can see that none of these core shapes dominates the low-energy region.

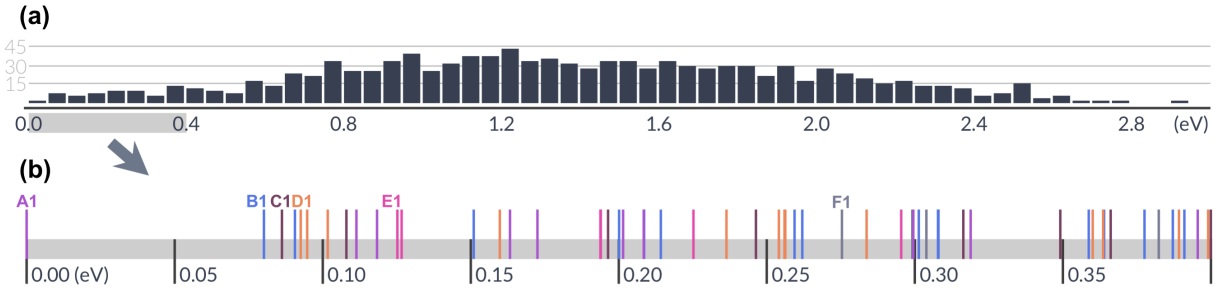


Figure 2.4 (a) Energy distribution of geometrically unique local minima of $Pt_7H_{10}CH_3$ clusters with respect to the putative global minimum, found by global optimization. The shaded gray area is expanded in part (b). (b) Energy spectrum of low-energy isomers within the energy range of 0 to 0.4 eV with respect to the energy of the putative global minimum (isomer A1).

Table 2.1 Number of different isomer structures for each Pt_7 core shape within the energy range of 0 to 0.4 eV with respect to the energy of the putative GM (isomer A1).

Pt_7 core shape	A	B	C	D	E	F
Number of structures	10	16	9	13	6	4

Figure 2.5 shows the lowest-energy structures in each Pt_7 core categories. In this figure, Energy (ΔE, in eV) is relative to the energy of putative GM (isomer A1). Charge (ΔQ) is the sum of Bader charge in the atoms of the cluster part ($Pt_7H_{10}CH_3$). It is worth noting that the Pt_7 core structures in $Pt_7H_{10}CH_3$ have very low similarity to those of bare Pt_7, which implies that the adsorption of H/CH_3 has a significant influence on the geometry of the Pt_7 cluster. In particular, structures B1, C1 and D1 have a singled out Pt atom, coordinating with only one other Pt atom. However, this kind of structure is not stable in the absence of adsorbates (does not appear in low-energy region). Obviously, the change of geometry has to be related to the number of H atoms adsorbed, since less compact Pt_7 core structure should allow better Pt-H interactions. The low similarity in bare and adsorbed Pt_7 clusters indicates that taking bare Pt_7 structures as reference states to help exploring the PES of adsorbed clusters is incorrect, as it misses some important local minima and core shapes.

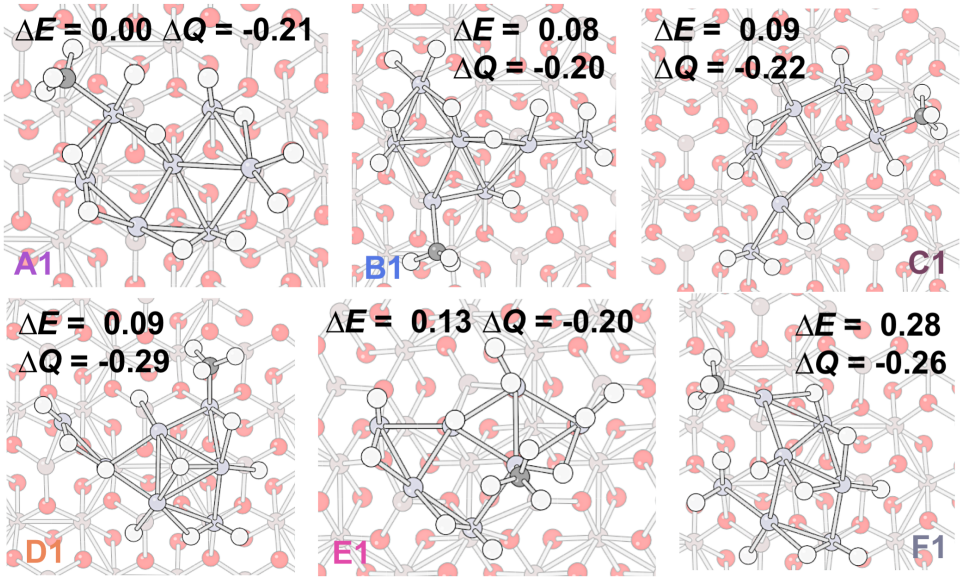


Figure 2.5 Isomer structures of each Pt_7 core shape (A to F) with lowest-energy for the alumina-supported $Pt_7H_{10}CH_3$ cluster found by global optimization.

Figure 2.6 shows the occurrence probability of all core shapes at different temperature, based on electronic energies (vibrational entropy contribution to the probability has been ignored, as an approximation), and Boltzmann statistics. Constructing such "hot" phase diagrams for catalytic systems, in order to elucidate the structures that emerge within the state of the catalyst at temperatures of interest and remove the focus on the GM alone, has been introduced by Zhang, Hermans, and Alexandrova. [90] It is necessary because often, higher energy structures in the ensemble are responsible for the catalysis. [78,79,87,91–94] Since many structures are close in energy in the studied case, many of them would be populated and constitute the state of the catalytic system at finite temperatures. Below 200 K, the ensemble primarily consists of the structures of type A, with negligible contributions from other structures. As the temperature increases the probability of other core types increases. In particular, in

the window of 700-900 K, typical for Pt-catalyzed dehydrogenation reactions, isomers of the types A-E contribute on the nearly equal footing. In each type of Pt core structure, several isomers are present due to different positions of H atoms. Any or all of these isomers could be relevant for the catalytic mechanism.

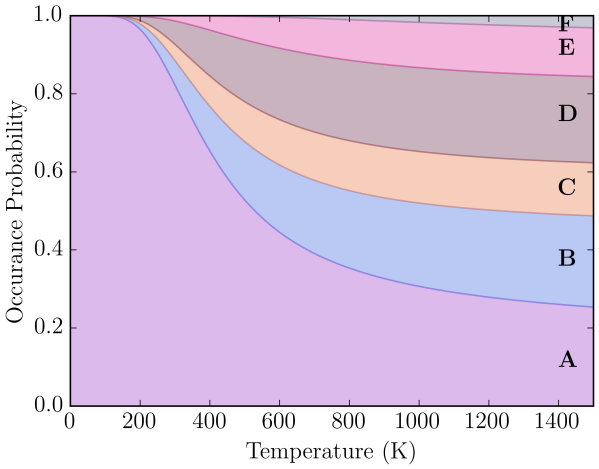


Figure 2.6 Occurrence probability of Pt_7 core shapes (A to F) for $Pt_7H_{10}CH_3$ on alumina at different temperature, based on Boltzmann distribution of electronic energies.

Since the Pt_7 core shapes A and B are among the dominant, and also have slightly larger number of accessible local minima, we focus on these two categories for the analysis of H binding sites. Figure 2.7 and Figure 2.8 show the geometries and charges of some of the low-energy isomers belonging to the Pt_7 shapes A and B. In the two figures, energy (ΔE, in eV) is relative to the energy of putative GM (isomer A1). Charge (ΔQ) is the sum of Bader charge in atoms of cluster part ($Pt_7H_{10}CH_3$). In every structure, H atoms occupy bridge and atop sites, while CH_3 is always on an atop site (with one exception being B3). In some structures, such as the putative GM A1, and B4, there are several H atoms binding to the CH_3 binding site, and this can be an

initial structure for the recombination and desorption of CH_4. On the other hand, in some other local minima, such as A5 and B1, there are no adjacent H atoms near CH_3. How these local minima can participate in CH_4 formation reaction, or inversely, be produced in the course of CH_4 binding and dissociation, is then related to the H atom mobility in this system, which will be covered in future part of this work.

Calculated Bader charges on atoms in the cluster are shown in Table 2.2 and Table 2.3. In the two tables, charge *Q*(cluster) is the sum of the charge in all atoms of cluster part ($Pt_7H_{10}CH_3$). In general, the charge transfer between the oxide surface and the cluster is small, with the cluster receiving ca. 0.2 electrons. In addition, the charge transfer between H/CH_3 adsorbates and Pt_7 is not significant, in most cases, indicative of covalent Pt-H and Pt-C bonding in the system, which is expected. One exception is the local minimum D1, where the summed charge on all ten H atoms is significant (-0.48). This is explained by the H atom in the unique hollow site on the face of a 3-Pt triangle (as shown in Figure 2.5). The H atom in the hollow site has a charge of -0.19, which is very different from the charges of the H atoms occupying the atop or bridge sites.

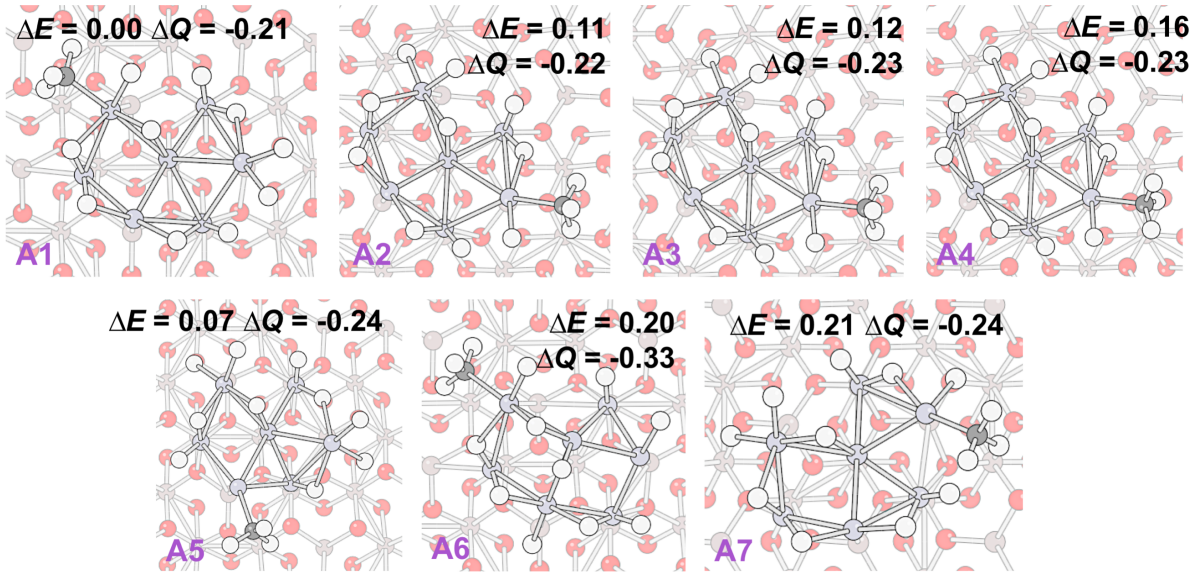


Figure 2.7 Low-energy structures of the Pt_7 core type A, of $Pt_7H_{10}CH_3$ clusters found by global optimization.

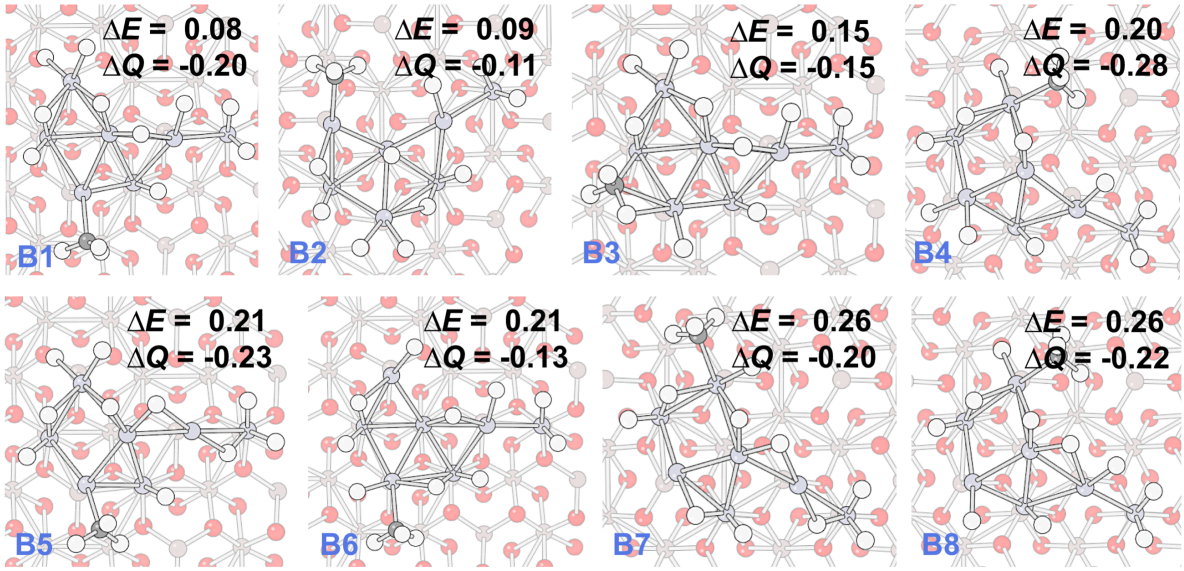


Figure 2.8 Low-energy structures of the Pt_7 core shape type B, of $Pt_7H_{10}CH_3$ clusters found by global optimization.

Table 2.2 Bader charges on different parts of cluster for isomers of shape A-F with lowest energy.

isomer	$Q(CH_3)$	$Q(H_{10})$	$Q(Pt_7)$	Q(cluster)	isomer	$Q(CH_3)$	$Q(H_{10})$	$Q(Pt_7)$	Q(cluster)
A1	-0.08	-0.05	-0.09	-0.21	D1	-0.02	-0.48	+0.21	-0.29
B1	-0.06	-0.10	-0.05	-0.20	E1	-0.02	-0.25	+0.07	-0.20
C1	-0.07	-0.05	-0.10	-0.22	F1	-0.09	-0.17	0.00	-0.26

Table 2.3 Bader charges on different parts of cluster for some typical low-energy isomers of shape A and B.

isomer	$Q(CH_3)$	$Q(H_{10})$	$Q(Pt_7)$	Q(cluster)	isomer	$Q(CH_3)$	$Q(H_{10})$	$Q(Pt_7)$	Q(cluster)
A1	-0.08	-0.05	-0.09	-0.21	B1	-0.06	-0.10	-0.05	-0.20
A2	-0.03	-0.11	-0.08	-0.22	B2	-0.10	+0.20	-0.21	-0.11
A3	-0.03	-0.13	-0.07	-0.23	B3	-0.05	-0.08	-0.02	-0.15
A4	-0.04	-0.10	-0.09	-0.23	B4	-0.07	-0.12	-0.09	-0.28
A5	-0.07	+0.01	-0.18	-0.24	B5	-0.07	-0.11	-0.05	-0.23
A6	-0.09	-0.26	+0.02	-0.33	B6	-0.04	-0.01	-0.08	-0.13
A7	-0.10	0.00	-0.14	-0.24	B7	-0.07	-0.03	-0.10	-0.20
					B8	-0.07	-0.02	-0.13	-0.22

Born-Oppenheimer (DFT) MD simulations in the NVT ensemble at 700 K, have been performed on selected isomers to further study the fluxionality of the cluster structure. Note that, within the necessarily limited amount of sampling and lack of rare-event sampling techniques, MD is a method that would never recover the isomeric diversity clearly accessible to the system, as it would sample only a rather local configuration environment. MD is used in this work to qualitatively and visually assess the relative motilities of atoms in the system. Figure 2.9 and Figure 2.10 are based on the MD trajectory starting from isomer A1, and Figure 2.11 and Figure 2.12 are based on the MD trajectory starting from isomer B1. In Figure 2.9 and Figure 2.11, darker color means higher probability for the atom taking certain *XY* position. The darkness scales for (b) and (c) are different. Each simulation lasts 10 ps, with the 1 fs time step. The first 1 ps of equilibration is not used for analysis.

We found that H atoms can be highly mobile at this temperature, while the Pt atoms are rather fixed at their positions. The relocated H atoms are outlined in Figure 2.10 and Figure 2.12. In both trajectories, some H atoms (circled in Figure 2.10 and Figure 2.12) can change from bridge site to atop site, or the reverse. The *XY* position of H atoms along the trajectories (shown in Figure 2.9 and Figure 2.11) spans a much larger area than that of Pt atoms, meaning that the H atoms are able to explore adjacent sites during this relative short simulation time.

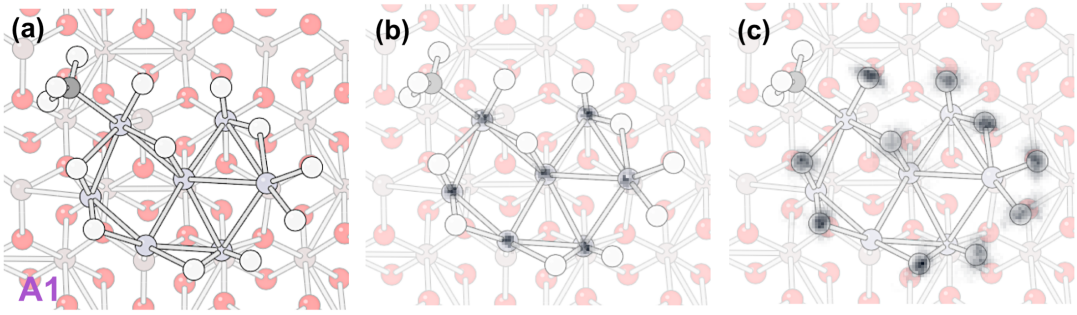


Figure 2.9 The probability of occurrence in *XY* plane of (b) Pt atoms and (c) adsorbed H atoms, during the MD trajectory starting from the isomer A1 (shown as a stationary structure in (a)).

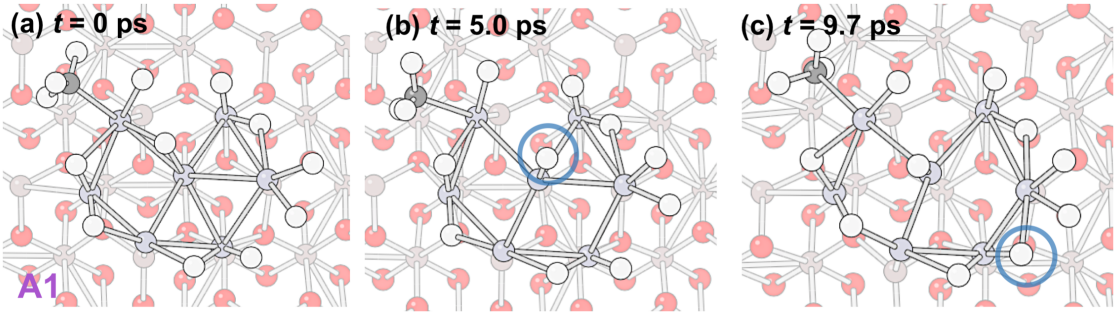


Figure 2.10 Cluster structure at simulation time *t* = (b) 5.0 ps and (c) 9.7 ps, during the MD trajectory starting from the isomer A1 (shown in (a)).

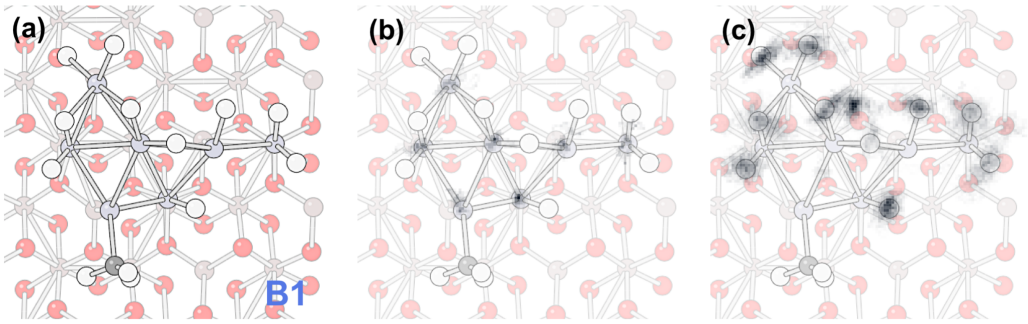


Figure 2.11 The probability of occurrence in *XY* plane of (b) Pt atoms and (c) adsorbed H atoms, during the MD trajectory starting from the isomer B1 (shown in (a)).

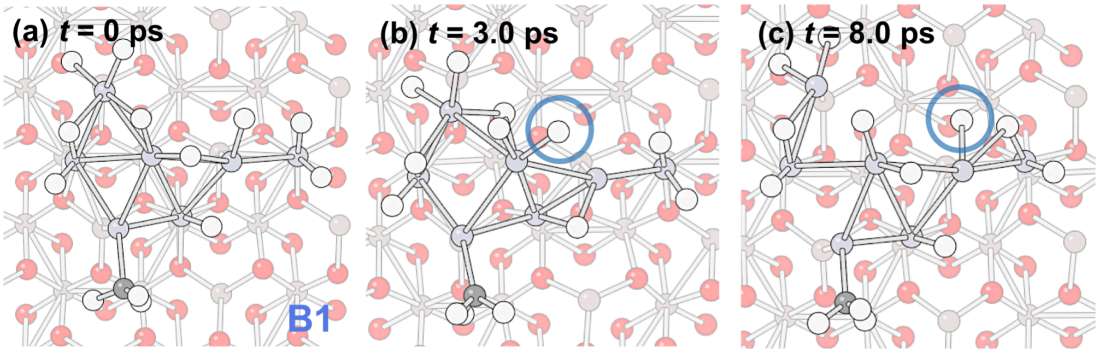


Figure 2.12 Cluster structure at simulation time t = (b) 3.0 ps and (c) 8.0 ps, during the MD trajectory starting from isomer B1 (shown in (a)).

To further examine the mobility of H atoms, we calculated Mean Square Displacement (MSD) averaged over the Pt and H atoms (Figure 2.13). The MSD of the H atoms shows very different patterns for the trajectory starting from A1, comparing with that from B1, which is mainly due to the different PES property near isomers A1 and B1. As can be seen from Figure 2.9, for the trajectory starting from A1, the XY projections of different H atoms do not overlap. Therefore, the H atoms are mostly transiting between the bridge and the nearest atop sites but do not interchange or shift any

further. For the trajectory starting from B1, some irreversible change of the H binding site takes place. As a result, the structures shown in Figure 2.12 are quite different from each other. Hence, the MSD along the trajectory becomes very large at later simulation times. Nevertheless, in both cases we can see that the MSD averaged over all H atoms is significantly larger than that averaged over all Pt atoms. Therefore, we conclude that H atoms rearrange much more easily than Pt atoms in this system.

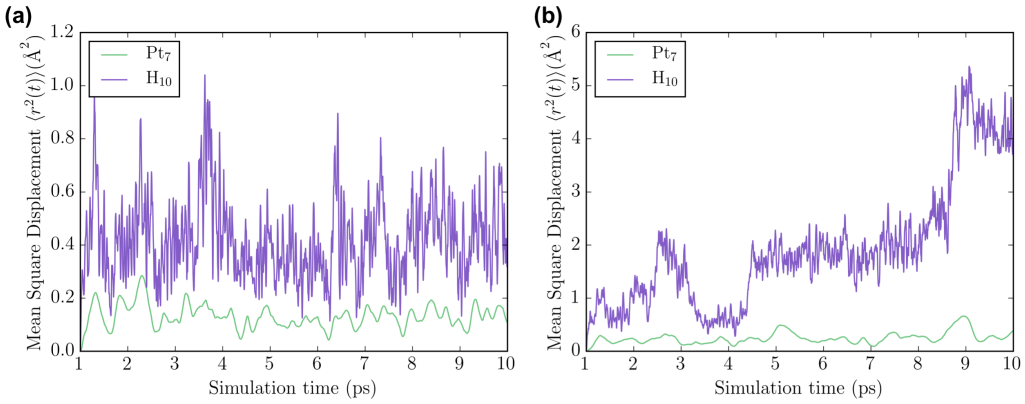


Figure 2.13 The MSD averaged over Pt (green) and H (purple) atoms, as a function of simulation time, calculated on MD trajectory at 700K starting from isomer (a) A1 and (b) B1.

Chapter 3

Structure similarity measurement

Since metal clusters can have very irregular shapes, automatic and accurate similarity measurement of cluster geometries is important for large scale global optimization. As explained in Section 1.4.4, structure similarity measurement can be utilized in multiple stages of the global optimization approach. In particular, it can be useful in filtering initial structures and detecting duplicates in relaxed isomers. The number of duplicates of low-energy isomers can also be used as an indicator of the convergence of global search.

3.1 Gas phase cluster case

The simplest way is comparing two structures based on their DFT energies.[3] Since energy is not available for initial structures, and for some large-size clusters different structures may have very close energies, this approach is not always applicable. Another approach is based on the difference of sorted atom-atom adjacency matrices,[95] or similarly, integrals over the distribution of atomic distances and angles.[71] This relatively simple method has been used in AFFCK. However, we will show that this approach is not exact for gas phase clusters. Other approaches use some fingerprint functions or descriptors, such as spherical harmonics, [96] spherically averaged scattering intensity [97] and experimental elemental or molecular properties. The effectiveness of fingerprint approaches will largely depend on the selected functions, and

may not be equally applicable to all cases. In particular, elemental and molecular properties are not generally applicable or available for metal clusters. Therefore, there is a need to develop a simple, reliable and non-energy-dependent structure similarity measurement method. In this section, we will propose a new method based on atomic matching, implemented using depth-first search.

3.1.1 Sorted interatomic distances

The structure similarity measurement method used in AFFCK is based on the sorted list D_i of all interatomic distances for the structure configuration. In this approach, two structures will be considered similar if

$$\frac{\sum_k \left|D_i(k) - D_j(k)\right|}{\sum_k D_i(k)} < d_{\text{rel}}, \qquad \text{and}$$

$$\max_k \left(\left|D_i(k) - D_j(k)\right|\right) < d_{\max}$$

are satisfied simultaneously, where i and j denote the two structures to be checked, k runs over the interatomic distances, d_{rel} and $d_{\max}$ are two thresholds for accumulated relative difference for all interatomic distances and maximum difference of individual distances, respectively. The selected values of parameter d_{rel} and $d_{\max}$ are different during different steps of AFFCK, see Section 1.6.1 for detailed numbers.

However, we note that two different structures can have exactly the same atomic distance distribution, or equivalently, the same sorted adjacency matrices, as illustrated in Figure 3.1, and this is also true for the angle distribution, especially when many distances or angles in the cluster are identical. In Figure 3.1, we can see that when the distances are matched, the corresponding atoms cannot be matched at the

same time. In this particular case, for planar quadrilateral the three “1.0” distances connect atoms 1,2,3 and 4. But for triangular pyramid the three “1.0” distances connect atoms 1,2 and 3. Therefore, the requirement of a one-to-one atomic matching can be used to distinguish the two structures.

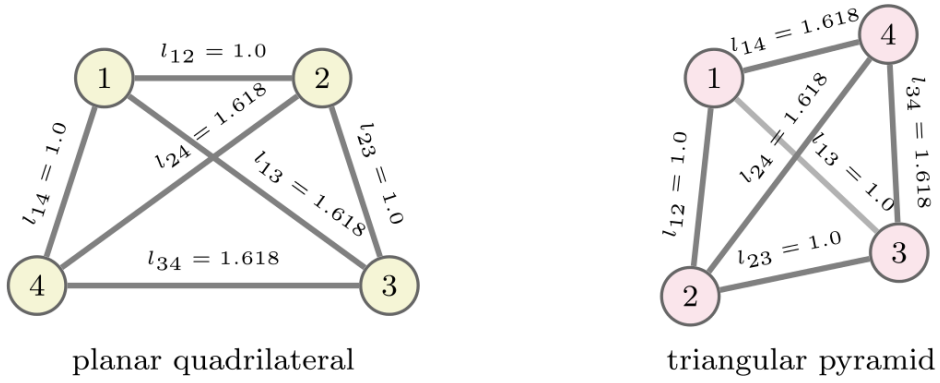


Figure 3.1 An example of two different structures with the same interatomic distance distribution.

3.1.2 Atomic matching based on depth-first search

The example shown in Figure 3.1 indicates that, in order to distinguish two isomers with exactly the same interatomic distance distributions, we need to also find a one-to-one atomic matching. Based on a best atomic matching, the corresponding interatomic distances can be compared between two structures. If any of the matched pairs of two interatomic distances have a difference larger than a threshold value, then we can assert that the two structures are different. We note that the interatomic distances based method cannot distinguish structures that are mirror images of each other. However, the mirror images also have the same energy at the DFT level, thus are unnecessary to be separated.

Now the problem is how to find the best atomic matching between two structures. A direct way is to enumerate all possible sequences of atom indices for one structure and match them to the sequence of the other. Since there are numerous permutations of atom indices for clusters composed of ten or more atoms, this is computationally too expensive. However, given that the threshold value d for confirming similarity is small, the comparison can be performed when an incomplete atomic matching is available. If the maximum difference of matched interatomic distances based on the given atomic matching fragment is larger than the threshold, then all atomic matching with that fragment will definitely fail. Using this rule, only a small portion of permutation space is needed to be checked. In practice, we use the Depth-First Search (DFS) [98] algorithm to enumerate the sequences of atoms for one structure, and match each sequence to the ordered atomic sequence of the other structure. Only atoms of the same element will be matched. Whenever a new testing atomic matching fragment is formed, the corresponding difference of matched interatomic distances is checked against the threshold value. If the difference is larger than the threshold value, all sequences involving the current fragment will be excluded from the searching space. Otherwise, we may find one atomic matching, based on which all differences of matched interatomic distances are smaller than the threshold value. In this case, we assert that the two structures are similar. If no such kind of atomic matching can be found, we claim that the current structure is unique. Based on our testing, using this approach on a single CPU core, thousands of structure comparisons for Pt_{13} can be accomplished within seconds. We denote this new, fast and deterministic structure comparison method the Atomic Matching Depth-First Search (AM-DFS) method.

3.2 Surface supported cluster case

For surface supported clusters, the cluster coordinates no long have rotational symmetry. As a result, method based on interatomic distances cannot be efficiently applied to the surface supported cluster case. Instead, space group symmetry of the surface atom coordinates can play an important role in cluster similarity measurement.

3.2.1 Mean distance between clusters

We define the difference between the geometry of two isomers i, j as

$$\Delta_{ij} = \frac{1}{n} \min_{\hat{S}\in \mathrm{G(surf)}} \min_{\hat{P}\in P_n} \sum_{k=1}^{n} \left\| \left(\hat{S}\hat{P}\boldsymbol{x}^i\right)_k - \boldsymbol{x}_k^j \right\| \equiv \min_{\hat{P}\in P_n} \Delta_{ij}\left(\hat{P}\right)$$

where n is the number of atoms in the cluster (for Pt_7, $n = 7$), G(surf) is the subgroup of the space group of the support material (for example, Al_2O_3), including only those symmetric transforms, namely, $\hat{S}$, that keep z coordinates invariant, P_n is the set of all permutations $\hat{P}$ of a sequence 1,2, …, n, $\boldsymbol{x}^j_k$ is the x, y, z coordinates (as a vector) of the kth atom in isomer j, and $\|\cdot\|$ denotes the 2-norm of a vector, which is also the Euclidean length of the vector. The geometric difference Δ_{ij} defined above is a measurement of linear Euclidean distance between two isomers based on optimal symmetry transformations, which is also known as mean distance (MD) between the two isomers.

3.2.2 Bipartite model for atomic matching

If the structural difference Δ_{ij} is directly evaluated using by its defining formula, the time complexity will be O(n!), which is non-polynomial. We propose to evaluate this using an efficient matching algorithm in graph theory, called Kuhn-Munkres (KM) algorithm (also known as Hungarian algorithm),[99,100] which has been widely applied and adapted to similarity measurement and alignment problem of gas phase clusters.[101–105] There are two major differences between the KM algorithm implementation adapted for this work and previous gas-phase-focused applications: (i) the MD Δ_{ij} is minimized over symmetry operations related the space group of the support material, as shown in our definition; (ii) in the application of finding Minimal Energy Pathway (MEP) (see Chapter 4), not only the optimal, but also sub-optimal atomic matches are used for optimization.

A weighted bipartite can be denoted as $B = (X, Y, E, W)$, where X and Y are two vertex sets satisfying that $|X| = |Y| = n$, $E \subseteq X \times Y$ is edge set (in this work we always have $E = X \times Y$), and $W: E \rightarrow \mathbb{R}$ defines weights of all edges. A matching M is a subset of edges no two of which share a common vertex. A matching M is best if $|M| = n$ and the sum of weights of edges in M is minimal. The bipartite model for solving cluster atomic matching is demonstrated in Figure 3.2.

Kuhn-Munkres algorithm finds the best solutions for bipartite matching with O(n^3) time complexity. This means that the value of $\min_{\hat{P} \in P_n} \sum_{k=1}^{n} \left\| \left(\hat{S}\hat{P}\boldsymbol{x}^i \right)_k - \boldsymbol{x}_k^j \right\|$ and the corresponding $\hat{P}$ can be evaluated in polynomial time.

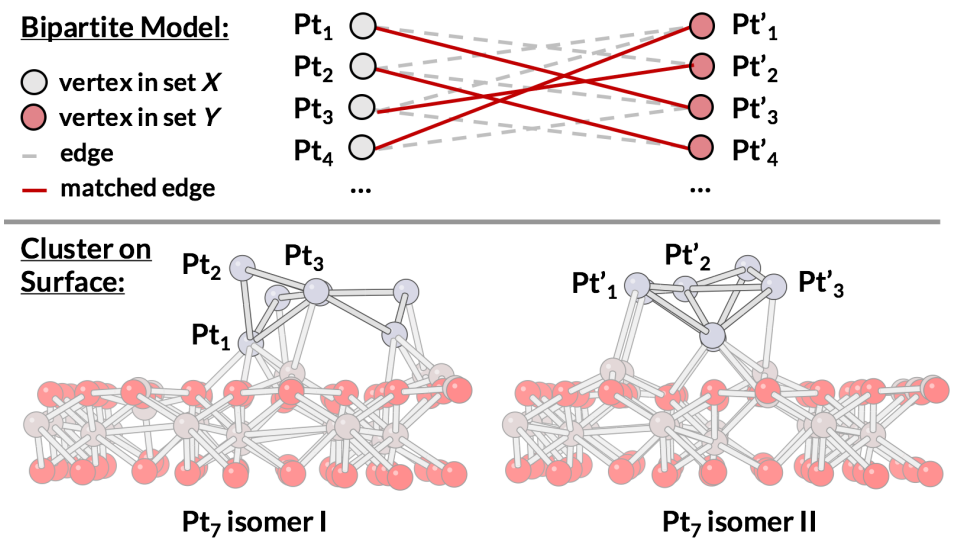


Figure 3.2 We model the atoms in two isomers as vertices in two sub-graphs (X and Y) in the bipartite graph, and the weights of edges are defined by the Euclidean distances between each of all possible pairs of the atoms.

Chapter 4

Isomerization graph construction

Using global optimization algorithms, discussed in previous chapters, combined with DFT, the PES of catalytic cluster systems can be explored, from which the geometric and electronic structures of some of the low-energy isomers can be identified and studied. However, many important finite-temperature effects, such as cluster isomerization process,[81,91] cannot be derived simply from the static (thermodynamic) isomer representation (see Section 1.5). The kinetics of cluster isomerization is another important component contributing to the accessible system dynamics. But finding efficient computational methods to study these dynamic effects remains a challenge.

In this chapter we use as example the Pt clusters deposited on α-alumina, of interest to catalytic dehydrogenation. A particularly interesting cluster size regime is around seven or eight Pt atoms, where special catalytic activity was observed.[79,92,106] In catalytic dehydrogenation of ethylene, for example, Pt_7 on α-Al_2O_3 showed a special activity.[79] Clusters of this size have a significant number of isomers, the number of which grows exponentially as the number of atoms increases. But only a few of these isomers are low enough in energy to be thermally accessible at temperatures of dehydrogenation (400-600 K).[79] And a comprehensive view of the isomerization process will help us understand cluster fluxional behavior, which is believed to play an important role in catalysis.[107,108]

Recently, some novel computational methods have been developed for investigating the isomerization process of molecules and clusters. Anharmonic downward distortion following (ADDF)[81,109] and reactive global optimization (RGO)[110] are low-frequency modes searching based methods, which have been successfully applied to gas phase systems and heterogeneous system, where for the latter case the surface is approximately represented by point charges.[110] However, it is known that the surface can form directional and strong bonds to the atoms of the cluster, and thus greatly impact the cluster geometry, electronic structure, and stability. Hence, both the surface and the cluster need to be represented accurately. In this case, the existing methods become inefficient, since they cannot be applied to large systems, particularly because of the computational expense of the vibrational frequency calculations (based on finite difference algorithm). An alternative is a two-step scheme, where the low-energy isomers are first obtained from global optimization, and then the transition paths between each pair of these isomers can be calculated, for example using the nudged elastic band (NEB) method.[111] Since NEB calculations are completely uncoupled, the scheme can be highly parallelized. This is the basic computational framework used in this chapter.

4.1 Atomic matching dependence of pathway optimization

However, there is one major problem for studying a cluster like Pt_7, or any other system that consists of many atoms of the same element. For a given pair of two different isomers, determination of an initial "guessed" path as the starting point of NEB will no longer be trivial. To determine an initial path, a matching (hereafter "atomic matching") between atoms of the two isomers (namely, a map from the atomic

indices of isomer *A* to the atomic indices of isomer *B*) has to be established. Then, the coordinates of intermediate structures along the path can be calculated by linear interpolation between coordinates of every pair of matched atoms. If the atomic matching is not optimal, the minimal energy path (MEP)[112] given by converged NEB generally will not be optimal either. For example, in Figure 1 we show two possible paths that connect the isomers #19 and #11 that resulted from a global optimization run. In this figure, the atoms that move significantly during isomerization are labeled as 1 and 2 in the isomer #19. Along the upper pathway (green dashed lines), atom 1 goes towards center. Along the lower pathway (red solid lines), atom 1 goes away from center. It turns out that the second pathway has a much lower barrier height. Black circles, red dots, and brown dots represent Pt, O, and Al atoms, respectively. For every pair of Pt_7 isomers, the number of possible atomic matches is 7! = 5040. Obviously, it is computationally unaffordable to traverse all these possible matches to find the optimal isomerization path. (Note that here we do not consider the matching of surface atoms. We assume that the displacement relative to their individual equilibrium positions is small for surface atoms, so that surface atoms will directly match to the closest surface atoms in the other isomer.)

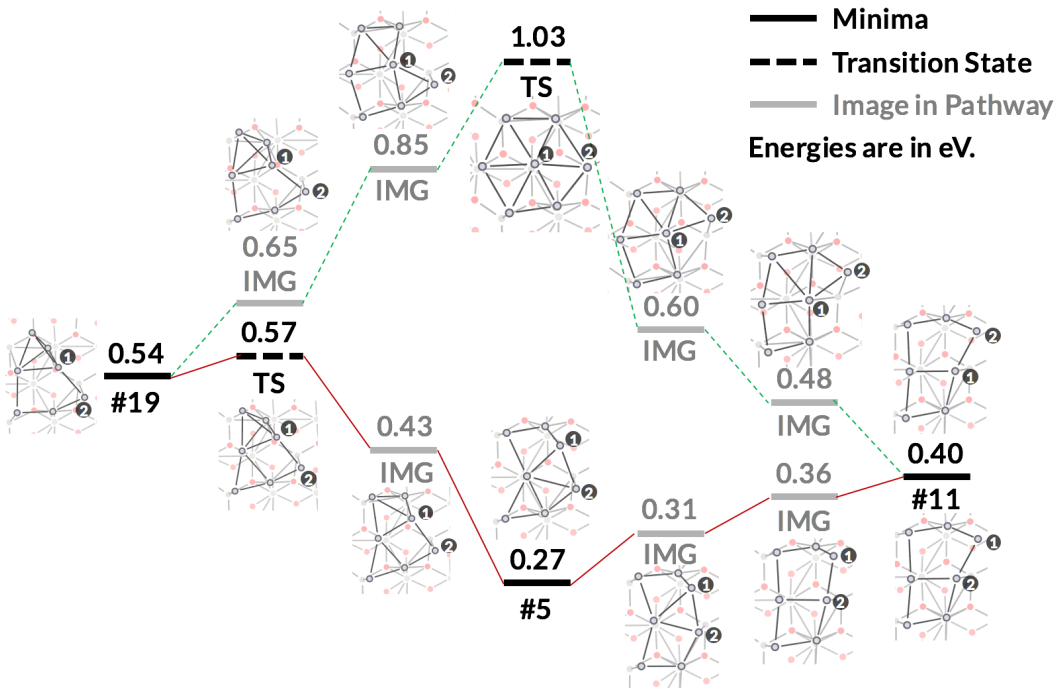


Figure 4.1 The two pathways of isomerization from the isomer #19 to the isomer #11 (numeration comes from the global optimization, starting from the global minimum as isomer #1), with different atom matching schemes.

4.2 Extended Kuhn-Munkres algorithm

One possible solution is performing NEB calculations only for "promising" atomic matching. The good atomic matching can be found (as a side product) using the mean distance Δ_{ij} between two isomers defined in Section 3.2.1. We further assume that if the MEP between two isomers is a direct path (without any intermediates), the geometry of the two isomers must be similar enough. This assumption is based on the following argument. The actual distance measured along MEP (MEP distance) is at least as long as the linear distance defined in Eq. (1). Therefore, the MEP distance is generally long for isomers with large Δ_{ij}. Under this case, either there will be more

likely some intermediates along the MEP, or the energy barrier of a direct transition will possibly be high. Then it is reasonable to consider only k permutations $\hat{P}$ of atomic indices, which gives k minimal $\Delta_{ij}(\hat{P})$, where k is a small constant. Kuhn-Munkres algorithm (extended for finding k best solutions) for bipartite matching has $O(kn^3)$ time complexity (k is set to 4 in this work).

4.3 Isomerization graph for α-alumina supported Pt_7 cluster

The 30 low-energy local minima (considered as isomers) shown in Figure 2.3 for Pt_7 cluster supported on Al_2O_3 have been used for isomerization study. In order to study the cluster isomerization at the temperature that is typical for catalysis (about 700 K), we consider isomers with energy less than 0.6 eV. Using similarity measurement defined in Section 3.2.1, we can now calculate pathway between pairs of isomers with Δ_{ij} lower than a threshold (in this work, 1.5 Å) by NEB optimization. For each pair, if there are multiple $\hat{P}$ satisfying $\Delta_{ij}(\hat{P}) < 1.5$ Å, at most four different $\hat{P}$ that give relatively lower $\Delta_{ij}(\hat{P})$ will be used. Some additional higher-energy isomers are found to be intermediates along the transition paths between these low-energy isomers, and they are also included in Figure 2.3.

A 15 Angstrom vacuum gap along surface z axis has been added to avoid image interactions. The energy cutoff of basis set expansion is set to 400 eV. The dipole correction is enabled for NEB and frequency (finite difference method) calculations. For path optimization, 5 images are used between a pair of isomers. The normal NEB is performed first. Then for those converged paths without any intermediates, climbing-

image NEB (CI-NEB) is used to find the transition state geometry and the corresponding transition state energy. If the path contains intermediates, the isomers corresponding to the intermediates are identified. Eventually the path between any two isomers can be decomposed to one or multiple direct paths. We found 48 such direct paths which are listed in Table 4.1. Among these paths, 29 direct paths are part of minimal energy paths, which are shown in Figure 4.2.

Based on NEB optimized MEP, an isomerization graph (Figure 4.3) can be constructed. Isomers are represented as vertices. If the optimized path contains no intermediates (a direct path), it corresponds to an edge in the graph; otherwise, it corresponds to a path in the graph. When only edges in MEP are retained in the graph, the graph becomes a minimum spanning tree (MST, a general tree is a connected n-vertex graph with $n - 1$ edges) shown in Figure 4.2. In Figure 4.2 and Figure 4.3, isomers, transition paths, isomer energies and transition barrier energies are denoted as circles, arrows (edges), numbers in bracket and numbers on arrows, respectively. Isomers are labeled from 1(lowest energy, GM) to 30 (highest energy). In MST, there is a unique path between any two isomers, which represents the MEP between the two isomers. The rough boundaries of the region I and region II (discussed in text) are labeled in red and green, respectively. The energy and geometry of transition states in direct paths are shown in Supporting Information Figure 4.4. In Figure 4.4, the two numbers in bracket indicate the index (in Figure 2.3) of the two isomers connected by the direct path. The energies are relative to the global minimum energy. An alternative way to visualize the connection between isomers is the disconnectivity graph [113] shown in Figure 4.5. In this figure, isomers that are considered within

region I and region II in 700 K are labeled as red and green, respectively. The energy (relative to global minimum) scale are shown as the vertical bar to the right.

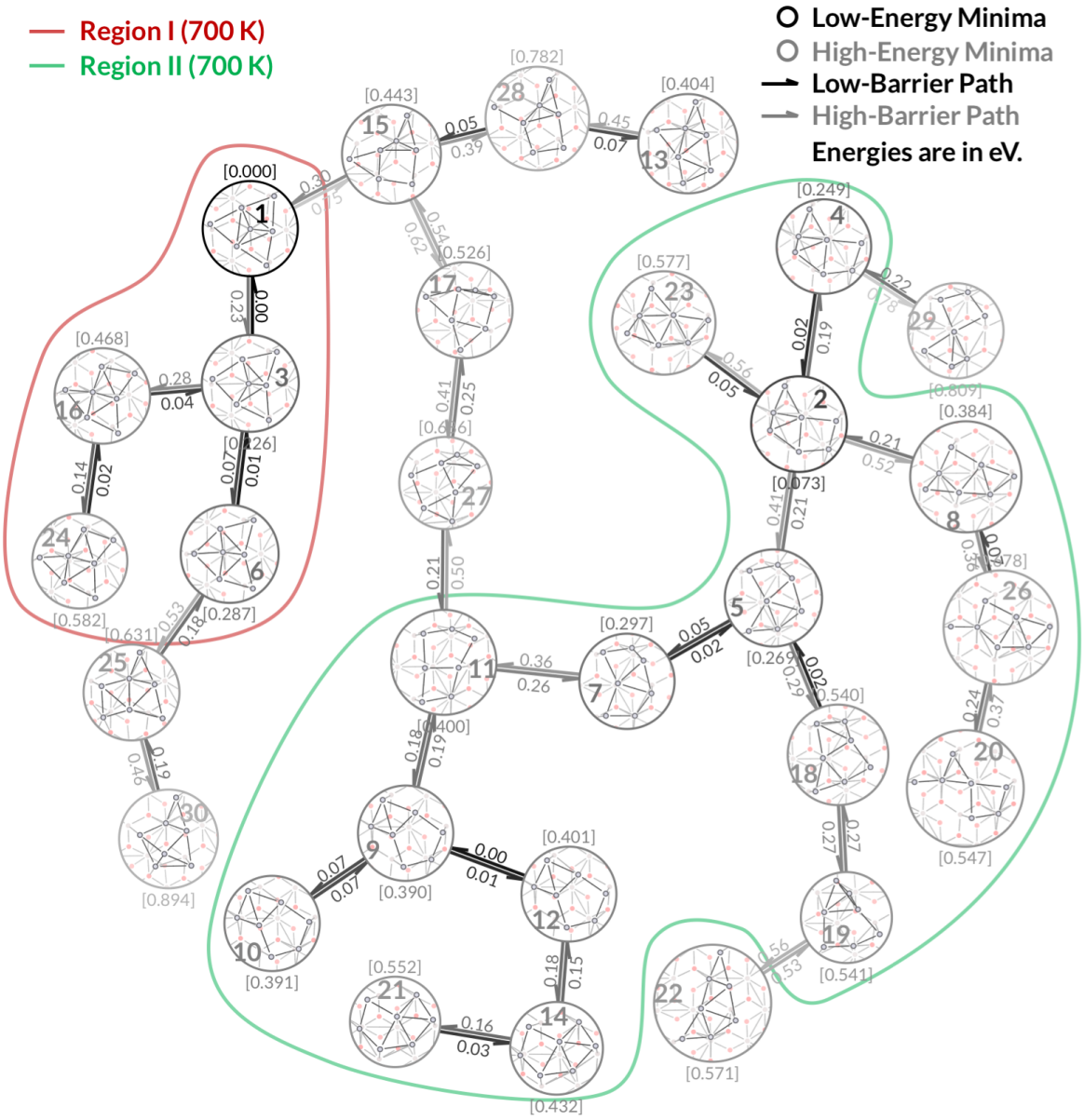


Figure 4.2 The minimal energy paths (MEPs) in the minimum spanning tree (MST) derived from isomerization graph (Figure 4.3).

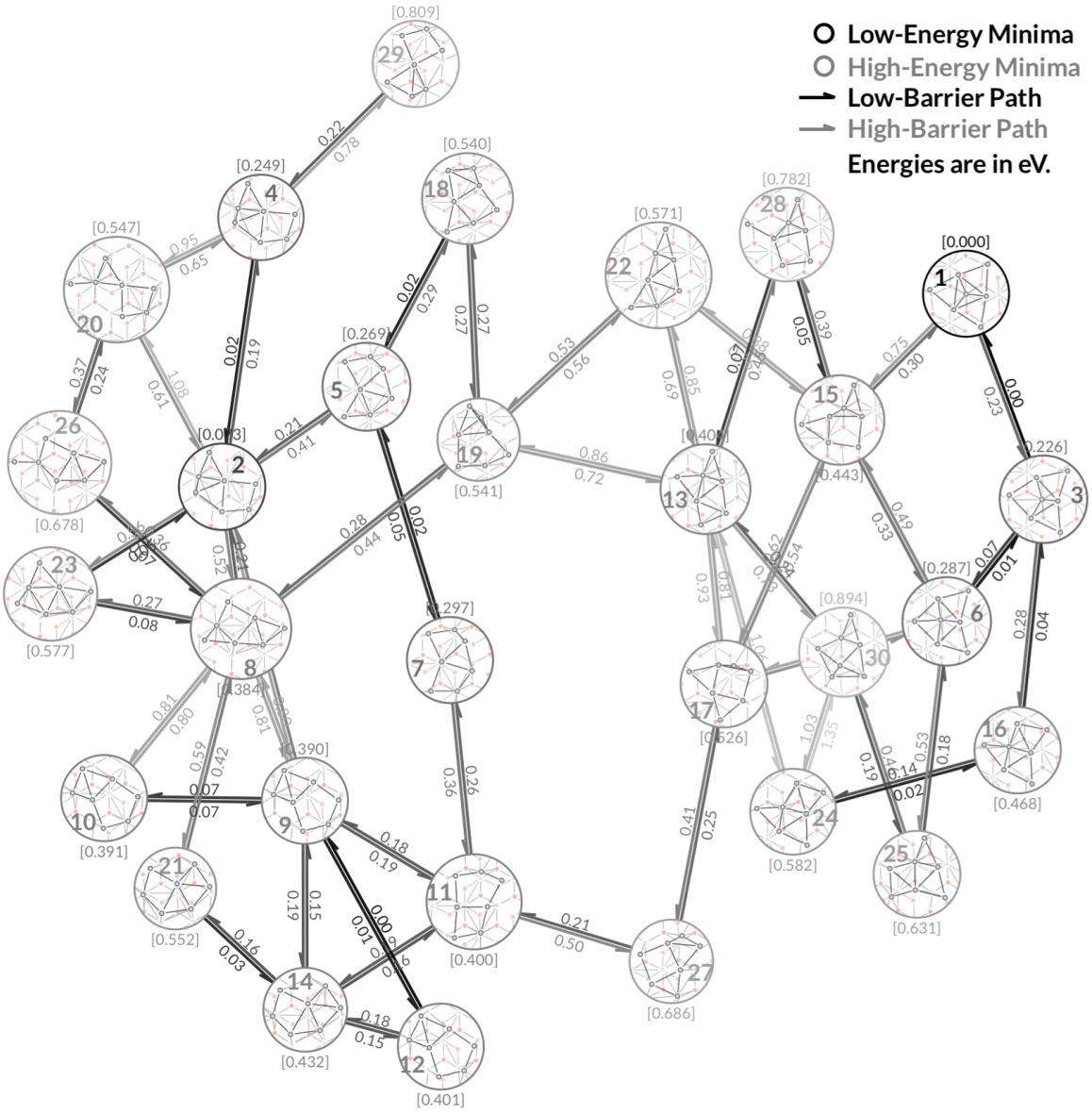


Figure 4.3 The Pt_7 on Al_2O_3 isomerization graph, including all direct NEB paths between low-energy isomers, shown as 48 edges in the graph.

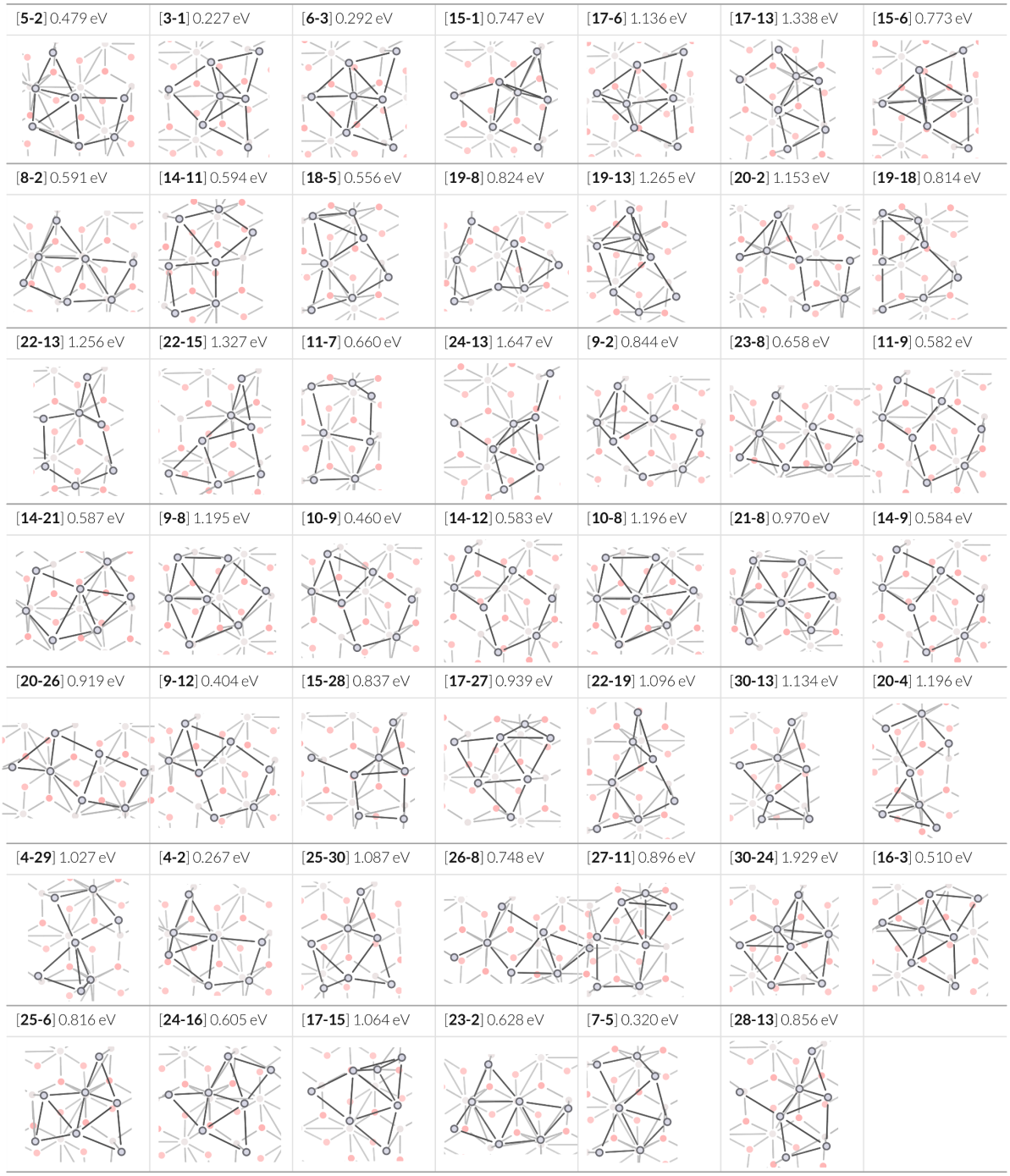


Figure 4.4 The energy and geometry of transition states in direct paths between Pt_7 cluster supported on Al_2O_3 isomers considered in this work.

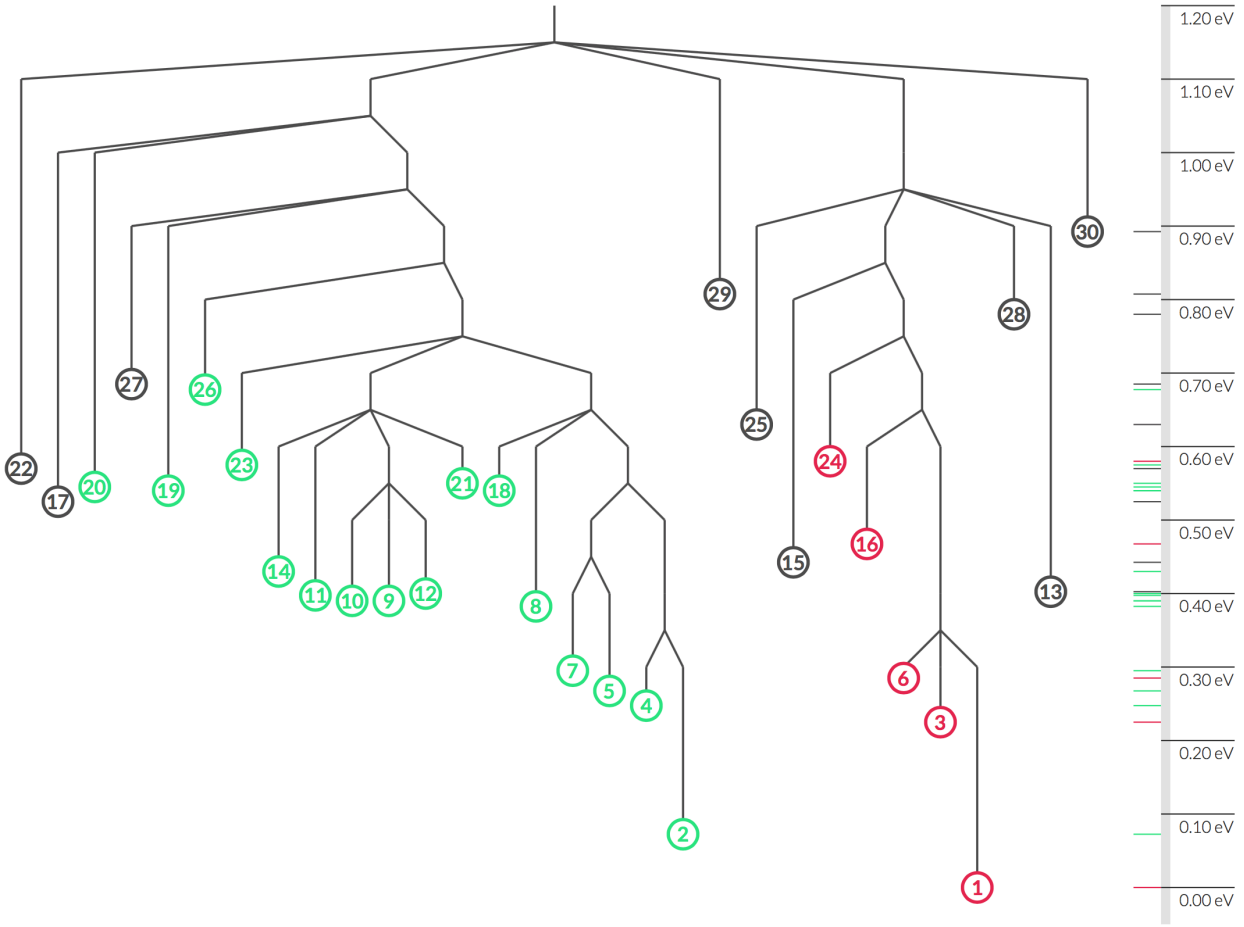


Figure 4.5 The disconnectivity graph of Al_2O_3 supported Pt_7 cluster isomers, showing connectivity between low-energy isomers.

For every direct path, according to harmonic transition state theory (HTST) the rate constant for isomer transition is estimated to be [111]

$$k_{\mathrm{HTST}} = \frac{\prod_i^{3N} \upsilon_i^{\mathrm{init}}}{\prod_i^{3N-1} \upsilon_i^{\mathrm{TS}}} \mathrm{e}^{-\frac{E^{\mathrm{TS}} - E^{\mathrm{init}}}{k_B T}}$$

where N is the total number of atoms, including cluster atoms and upper half of the surface atoms (which are not fixed during the relaxation, N = 7 (cluster) + 135 (surface) = 142 in this work). $\upsilon_i^{\mathrm{init}}$ and υ_i^{TS} are frequencies of the vibrational normal modes

at the initial state and transition state, respectively. E^{init} and E^{TS} are energies of the two states, respectively. The rate constants are evaluated at 450 K and 700 K and listed in Table 4.1. In this table, the columns are index of the initial isomer, index of the final isomer, energy of initial isomer, energy of final isomer, energy of transition state, barrier height, rate constant at 450 K, rate constant at 700 K, and notes respectively. The notes indicate whether it is a transition inside one region ("Region I" or "Region II"), or a transition at region boundary (RB) ("RB I" or "RB II"), or other types (empty). Note that this table includes not only the transition shown in MST (Figure 4.2), but also other direct paths in the original isomerization graph (Figure 4.3). The extra paths are not part of MEP, and their barrier energies and rate constants can be high.

We note that in HTST harmonic approximation of the PES has been made in the vicinity of isomers and transition states. This makes HTST suitable for probing finite temperature effects, which are not included in simple local minima representation of the PES.

Table 4.1 The barrier heights and HTST rate constants of all direct paths between isomers in Figure 2.3.

$\#_{\text{init}}$	$\#_{\text{final}}$	E_{init} (eV)	E_{final} (eV)	E_{TS} (eV)	barrier (eV)	$k_{450\text{K}}$ (1/s)	$k_{700\text{K}}$ (1/s)	Notes
5	2	0.2688	0.0727	0.4793	0.2105	5.65×10^{10}	3.93×10^{11}	Region II
2	5	0.0727	0.2688	0.4793	0.4066	1.84×10^{9}	7.80×10^{10}	Region II
3	1	0.2263	0.0000	0.2266	0.0004	1.06×10^{13}	1.07×10^{13}	Region I
1	3	0.0000	0.2263	0.2266	0.2266	3.37×10^{10}	2.72×10^{11}	Region I
6	3	0.2871	0.2263	0.2921	0.0050	4.23×10^{12}	4.43×10^{12}	Region I
3	6	0.2263	0.2871	0.2921	0.0658	1.17×10^{12}	2.15×10^{12}	Region I
15	1	0.4426	0.0000	0.7465	0.3040	5.10×10^{9}	8.39×10^{10}	RB I
1	15	0.0000	0.4426	0.7465	0.7465	9.96×10^{4}	9.65×10^{7}	RB I

17	6	0.5257	0.2871	1.1356	0.6099	1.02×10^{7}	2.79×10^{9}	RB I
6	17	0.2871	0.5257	1.1356	0.8486	2.39×10^{4}	5.92×10^{7}	RB I
17	13	0.5257	0.4038	1.3375	0.8118	1.84×10^{5}	3.25×10^{8}	
13	17	0.4038	0.5257	1.3375	0.9337	5.05×10^{3}	2.74×10^{7}	
15	6	0.4426	0.2871	0.7726	0.3300	1.70×10^{10}	3.55×10^{11}	RB I
6	15	0.2871	0.4426	0.7726	0.4855	3.95×10^{8}	3.45×10^{10}	RB I
8	2	0.3837	0.0727	0.5914	0.2077	2.18×10^{12}	1.47×10^{13}	Region II
2	8	0.0727	0.3837	0.5914	0.5186	1.37×10^{9}	1.62×10^{11}	Region II
14	11	0.4316	0.3999	0.5944	0.1628	1.09×10^{11}	4.90×10^{11}	Region II
11	14	0.3999	0.4316	0.5944	0.1944	1.70×10^{11}	1.02×10^{12}	Region II
18	5	0.5404	0.2688	0.5556	0.0152	6.68×10^{12}	7.68×10^{12}	Region II
5	18	0.2688	0.5404	0.5556	0.2868	1.06×10^{10}	1.48×10^{11}	Region II
19	8	0.5409	0.3837	0.8237	0.2828	3.95×10^{9}	5.34×10^{10}	Region II
8	19	0.3837	0.5409	0.8237	0.4401	2.97×10^{8}	1.71×10^{10}	Region II
19	13	0.5409	0.4038	1.2646	0.7236	1.81×10^{5}	1.42×10^{8}	RB II
13	19	0.4038	0.5409	1.2646	0.8608	1.27×10^{4}	3.51×10^{7}	RB II
20	2	0.5468	0.0727	1.1528	0.6060	5.55×10^{6}	1.47×10^{9}	Region II
2	20	0.0727	0.5468	1.1528	1.0801	7.30×10^{1}	1.53×10^{6}	Region II
19	18	0.5409	0.5404	0.8142	0.2732	4.63×10^{10}	5.74×10^{11}	Region II
18	19	0.5404	0.5409	0.8142	0.2737	4.23×10^{10}	5.26×10^{11}	Region II
22	13	0.5706	0.4038	1.2564	0.6859	1.55×10^{3}	8.61×10^{5}	
13	22	0.4038	0.5706	1.2564	0.8527	4.77×10^{3}	1.23×10^{7}	
22	15	0.5706	0.4426	1.3274	0.7569	4.37×10^{2}	4.66×10^{5}	
15	22	0.4426	0.5706	1.3274	0.8849	5.18×10^{3}	1.79×10^{7}	
11	7	0.3999	0.2974	0.6603	0.2604	2.92×10^{10}	3.21×10^{11}	Region II
7	11	0.2974	0.3999	0.6603	0.3630	1.81×10^{9}	5.12×10^{10}	Region II
24	13	0.5821	0.4038	1.6471	1.0650	1.24×10^{2}	2.26×10^{6}	RB I
13	24	0.4038	0.5821	1.6471	1.2433	2.34×10^{0}	2.20×10^{5}	RB I
9	2	0.3900	0.0727	0.8441	0.4541	5.23×10^{7}	3.43×10^{9}	Region II
2	9	0.0727	0.3900	0.8441	0.7714	4.37×10^{5}	5.32×10^{8}	Region II
23	8	0.5771	0.3837	0.6578	0.0807	8.40×10^{12}	1.77×10^{13}	Region II
8	23	0.3837	0.5771	0.6578	0.2742	1.30×10^{11}	1.63×10^{12}	Region II
11	9	0.3999	0.3900	0.5823	0.1824	1.26×10^{11}	6.76×10^{11}	Region II
9	11	0.3900	0.3999	0.5823	0.1923	2.11×10^{10}	1.24×10^{11}	Region II
14	21	0.4316	0.5524	0.5873	0.1557	7.55×10^{10}	3.17×10^{11}	Region II
21	14	0.5524	0.4316	0.5873	0.0348	1.85×10^{10}	2.55×10^{10}	Region II
9	8	0.3900	0.3837	1.1946	0.8045	2.60×10^{4}	4.30×10^{7}	Region II
8	9	0.3837	0.3900	1.1946	0.8109	3.47×10^{5}	6.07×10^{8}	Region II
10	9	0.3909	0.3900	0.4603	0.0694	5.78×10^{12}	1.09×10^{13}	Region II
9	10	0.3900	0.3909	0.4603	0.0702	1.85×10^{12}	3.52×10^{12}	Region II
14	12	0.4316	0.4010	0.5832	0.1517	8.67×10^{10}	3.51×10^{11}	Region II
12	14	0.4010	0.4316	0.5832	0.1823	5.32×10^{10}	2.85×10^{11}	Region II
10	8	0.3909	0.3837	1.1956	0.8047	8.64×10^{4}	1.43×10^{8}	Region II
8	10	0.3837	0.3909	1.1956	0.8119	3.67×10^{5}	6.49×10^{8}	Region II

21	8	0.5524	0.3837	0.9699	0.4175	1.82×10^{7}	8.53×10^{8}	Region II
8	21	0.3837	0.5524	0.9699	0.5863	2.57×10^{8}	5.69×10^{10}	Region II
14	9	0.4316	0.3900	0.5844	0.1528	1.06×10^{11}	4.32×10^{11}	Region II
9	14	0.3900	0.4316	0.5844	0.1943	2.75×10^{10}	1.64×10^{11}	Region II
20	26	0.5468	0.6783	0.9186	0.3718	3.14×10^{9}	9.64×10^{10}	Region II
26	20	0.6783	0.5468	0.9186	0.2404	3.61×10^{10}	3.30×10^{11}	Region II
9	12	0.3900	0.4010	0.4040	0.0140	1.12×10^{12}	1.28×10^{12}	Region II
12	9	0.4010	0.3900	0.4040	0.0031	2.65×10^{12}	2.73×10^{12}	Region II
15	28	0.4426	0.7824	0.8373	0.3947	4.45×10^{9}	1.69×10^{11}	
28	15	0.7824	0.4426	0.8373	0.0549	7.28×10^{12}	1.21×10^{13}	
17	27	0.5257	0.6863	0.9394	0.4136	1.79×10^{9}	8.08×10^{10}	
27	17	0.6863	0.5257	0.9394	0.2531	2.37×10^{10}	2.44×10^{11}	
22	19	0.5706	0.5409	1.0961	0.5255	1.18×10^{5}	1.49×10^{7}	RB II
19	22	0.5409	0.5706	1.0961	0.5551	5.17×10^{6}	8.59×10^{8}	RB II
30	13	0.8941	0.4038	1.1341	0.2400	8.51×10^{10}	7.76×10^{11}	
13	30	0.4038	0.8941	1.1341	0.7303	6.41×10^{5}	5.35×10^{8}	
20	4	0.5468	0.2495	1.1956	0.6488	2.77×10^{6}	1.09×10^{9}	Region II
4	20	0.2495	0.5468	1.1956	0.9461	1.91×10^{2}	1.16×10^{6}	Region II
4	29	0.2495	0.8094	1.0266	0.7771	1.57×10^{5}	2.01×10^{8}	RB II
29	4	0.8094	0.2495	1.0266	0.2172	9.18×10^{11}	6.79×10^{12}	RB II
4	2	0.2495	0.0727	0.2669	0.0175	1.47×10^{13}	1.72×10^{13}	Region II
2	4	0.0727	0.2495	0.2669	0.1942	3.56×10^{12}	2.13×10^{13}	Region II
25	30	0.6314	0.8941	1.0873	0.4559	3.10×10^{8}	2.07×10^{10}	
30	25	0.8941	0.6314	1.0873	0.1931	6.67×10^{10}	3.95×10^{11}	
26	8	0.6783	0.3837	0.7479	0.0696	1.17×10^{12}	2.22×10^{12}	Region II
8	26	0.3837	0.6783	0.7479	0.3643	2.13×10^{9}	6.10×10^{10}	Region II
27	11	0.6863	0.3999	0.8956	0.2093	3.98×10^{10}	2.74×10^{11}	RB II
11	27	0.3999	0.6863	0.8956	0.4957	3.97×10^{7}	3.82×10^{9}	RB II
30	24	0.8941	0.5821	1.9287	1.0346	5.56×10^{2}	7.65×10^{6}	RB I
24	30	0.5821	0.8941	1.9287	1.3466	2.62×10^{-1}	6.36×10^{4}	RB I
16	3	0.4679	0.2263	0.5098	0.0419	4.13×10^{11}	6.07×10^{11}	Region I
3	16	0.2263	0.4679	0.5098	0.2836	8.96×10^{9}	1.22×10^{11}	Region I
25	6	0.6314	0.2871	0.8157	0.1843	2.93×10^{11}	1.60×10^{12}	RB I
6	25	0.2871	0.6314	0.8157	0.5286	4.79×10^{7}	6.24×10^{9}	RB I
24	16	0.5821	0.4679	0.6053	0.0232	8.89×10^{12}	1.10×10^{13}	Region I
16	24	0.4679	0.5821	0.6053	0.1374	2.12×10^{11}	7.50×10^{11}	Region I
17	15	0.5257	0.4426	1.0643	0.5386	1.08×10^{8}	1.55×10^{10}	
15	17	0.4426	0.5257	1.0643	0.6217	1.15×10^{7}	3.52×10^{9}	
28	13	0.7824	0.4038	0.8558	0.0734	1.18×10^{12}	2.33×10^{12}	
13	28	0.4038	0.7824	0.8558	0.4521	1.88×10^{8}	1.21×10^{10}	
7	5	0.2974	0.2688	0.3197	0.0223	3.83×10^{12}	4.70×10^{12}	Region II
5	7	0.2688	0.2974	0.3197	0.0509	2.64×10^{12}	4.22×10^{12}	Region II
23	2	0.5771	0.0727	0.6283	0.0512	1.82×10^{12}	2.92×10^{12}	Region II
2	23	0.0727	0.5771	0.6283	0.5556	1.62×10^{7}	2.70×10^{9}	Region II

To further examine the local fluxionality picture of isomerization of Pt_7 clusters, we performed molecular dynamics (MD) simulations. The MD simulation can properly include the effect of anharmonicity, which is missing in our isomerization graph build upon HTST. The NVT ensemble simulation is performed at 450 K and 700 K, respectively, for the 5.0 ps duration with the time step being 0.5 fs. A Nose-Hoover thermostat is used. The system was considered to be equilibrated after 0.5 ps, and after that time, information was collected. The geometry of the cluster part of the system at any moment of the simulation can be very different from the 0 K isomers, due to the complexity of PES and thermal expansion effect. To analyze the trajectories, mean distance defined in Section 3.2.1 is used to measure the difference between geometry of the given cluster at every MD instant and geometries of several adjacent 0 K isomers. Figure 4.6 (a) and (b) shows the trajectories starting from the isomer #1 simulated at 450 K and 700 K, respectively. In this figure, the trajectories from 0.0 to 0.5 ps are used to establish the equilibrium and thus not included. For trajectory at each temperature, the snapshot geometry at every time step is compared against several isomer geometries inside the region I (isomers #1, #3, #6, #16 and #24) and outside the region I (#15). The geometrical difference between two structures is defined in Section 3.2.1. Note that during the simulation, the position of surface atoms also changes. These (usually small) changes are not included in the evaluation of the difference. As the system evolves, the cluster loses its geometric identity, and may approach other isomers instead. It is shown that the (locally) equilibrated system is roughly a mixture of the isomers #1 and #3 at 450 K. The similarity to other isomers is small throughout the trajectory. However, when we simulate the system at 700 K, we can see much greater isomerization and geometric mixing, and, for example, at 3.9 ps of

the trajectory the cluster is equidistant from the isomers #1, #3, #6, and #16. At this moment, the geometry is still far away from isomer #15 (with a distance of 1.0 Å). This implies that, at 700 K the locally fluxional region I (including isomers #1, #3, #6, and #16) may be a better representation of the system evolving from the isomer #1, than a fixed representation characterized by the geometry of isomer #1 alone. Similarly, Figure 4.7 (a) and (b) shows the trajectories starting from isomer #2 simulated at 450 K and 700 K, respectively. In this figure, for trajectory at each temperature, the snapshot geometry at every time step is compared against several isomer geometries inside the region II (isomers #2, #4, #5, #7, #11) and outside the region II (#27). Although the energy of isomers #1 and 2 only differs by 0.072 eV, they are connected by a long path (including isomers #27, #17 and #15) with high barrier in isomerization MST and belong to different regions (Figure 4.2). It is shown that the system is fluxional among isomers #2, #4 and #5 at 450 K, and among #2, #4, #5, and #7 (all from region II) at 700 K, which are clearly indicated in Figure 4.7 (a) at 1.5 ps and Figure 4.7 (b) at 1.2 ps. This agrees with our previous analysis based on isomerization graph, and the regions I and II are actually connected subgraphs of the isomerization MST. This outcome is remarkable, because, in some cases of catalysis, it provides possible mechanism for the system to retain specific, highly active metastable states for some period of time. In the case of Pt_7, the timescale for local fluxionality is about 1 ns (estimated from the cross-region transition rate in 700 K).

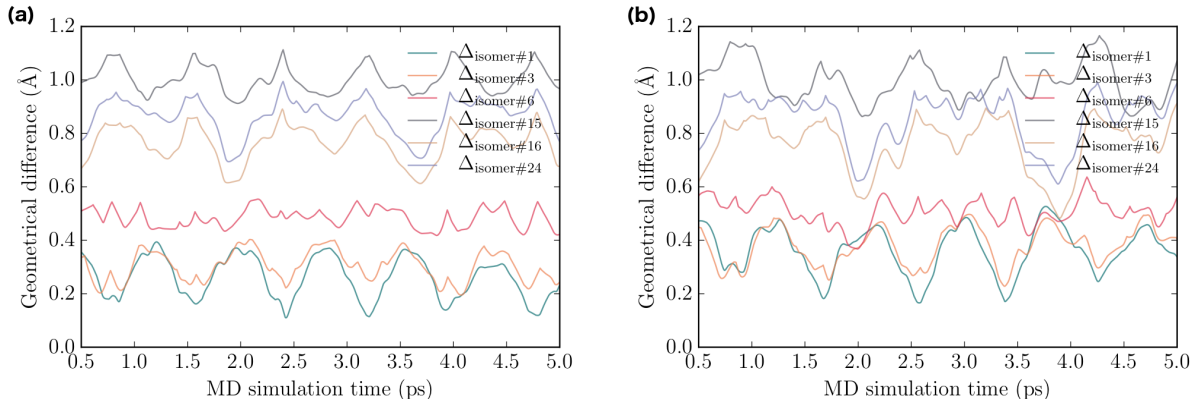


Figure 4.6 The MD trajectories starting from isomer #1 simulated at (a) 450 K and (b) 700 K.

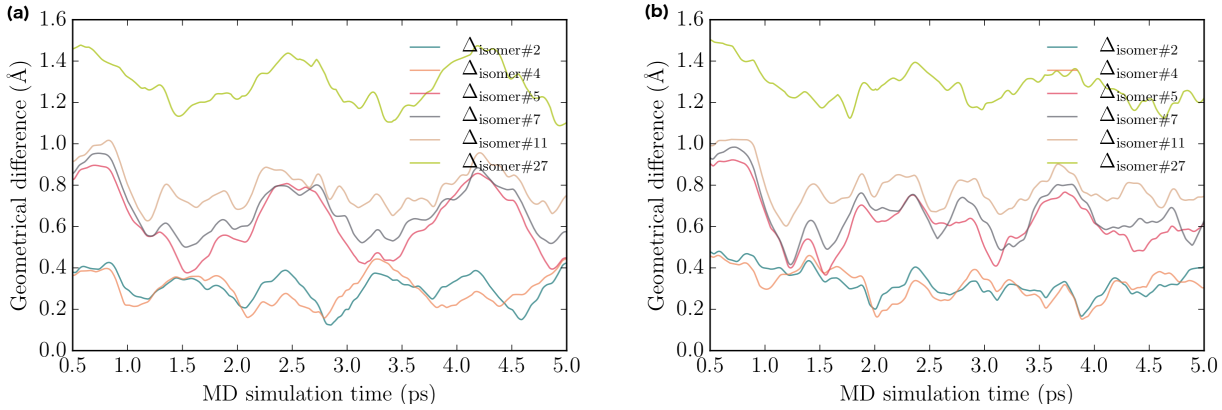


Figure 4.7 The MD trajectories starting from isomer #2 simulated at (a) 450 K and (b) 700 K.

Finally, detailed analysis on MD trajectories allows us to estimate anharmonicity effects in the system. We note that the geometries in MD are at least 0.11 Å far away from the isomer #1 or #2 at 450 K, and 0.17 Å far away from the isomer #1 or #2 at 700 K, respectively, which is a manifestation of the thermal expansion effect. In addition, there are some periodic patterns in Figure 4.6 and Figure 4.7, which indicates

that some detailed comparison can be made between HTST rate constants and information extracted from MD simulations. From MD trajectories in Figure 4.6 and Figure 4.7, rate constants of the fastest process (namely, transitions between isomers #1 and #3 in region I and between isomers #2 and #4 in region II) can be approximately estimated by plotting the difference between the two lowest lines in Figure 4.6 and Figure 4.7, and locating maxima and minima, shown in Figure 4.8 and Figure 4.9. In these two figures, The maxima (labeled by red vertical lines) correspond to instants when the MD geometry is closest to isomer #1 (or #2) and farthest to isomer #3 (or #4), and the minima (labeled by green vertical lines) correspond to instants when the MD geometry is closest to isomer #3 (or #4) and farthest to isomer #1 (or #2). The rate constants are estimated by the inverse of average time span between adjacent maxima and minima. The rate constants estimated from HTST and MD trajectories for transition between these fast processes are summarized in Table 4.2.

For transition between isomer #1 and #3, HTST predicts very different rates for the two directions, mainly due to the 0.226 eV energy difference between 0 K isomers #1 and #3. However, the geometrical difference between isomers #1 and #3 is only 0.26 Å. At finite temperature, the average electronic energies corresponding to these two isomer basins may become much closer as their geometries become more similar after thermal expansion. For transition between isomers #2 and #4, the MD predicted rates are close for the two directions as well, which is again explained by their very similar geometries. Finally, considering the nature and short duration of the MD simulation, we note that there are some important factors that may cause MD predicted rates listed in Table 4.2 to be inaccurate, as explained below. The vibrational frequencies

of the cluster part of the system are ca. 50 cm^{-1} or 1.5 to 3.0 ps^{-1}, which can couple with the fast isomerization process. In addition, when several fast isomerization processes happen simultaneously, we cannot distinguish them by the simple analysis here. More accurate analysis should be based on free energy sampling methods, such as free energy perturbation,[114] umbrella sampling [115] and metadynamics,[116] which is beyond the scope of the work. To sum up, anharmonicity should be important for these fast processes because of the small geometrical difference between isomers and coupling between vibrational motion and isomerization process.

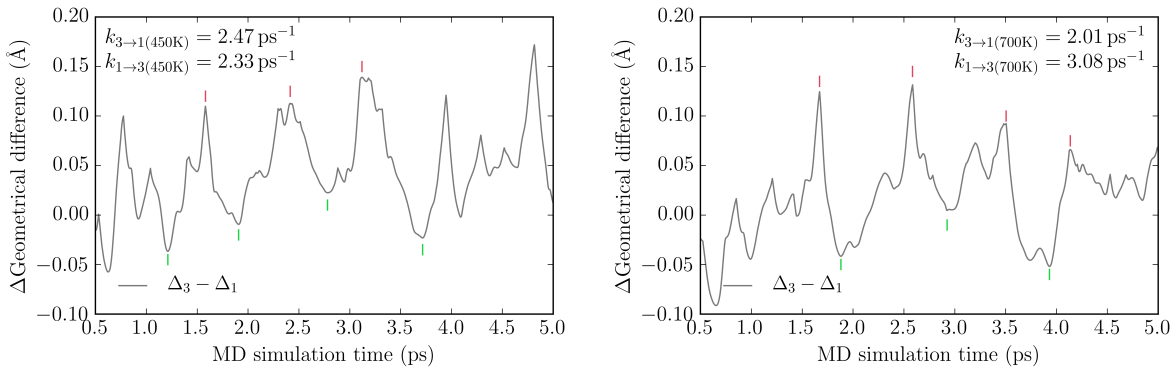


Figure 4.8 The difference between geometrical differences measured from isomers #3 and #1 based on MD trajectories starting from isomer #1 simulated at (a, left) 450 K and (b, right) 700 K.

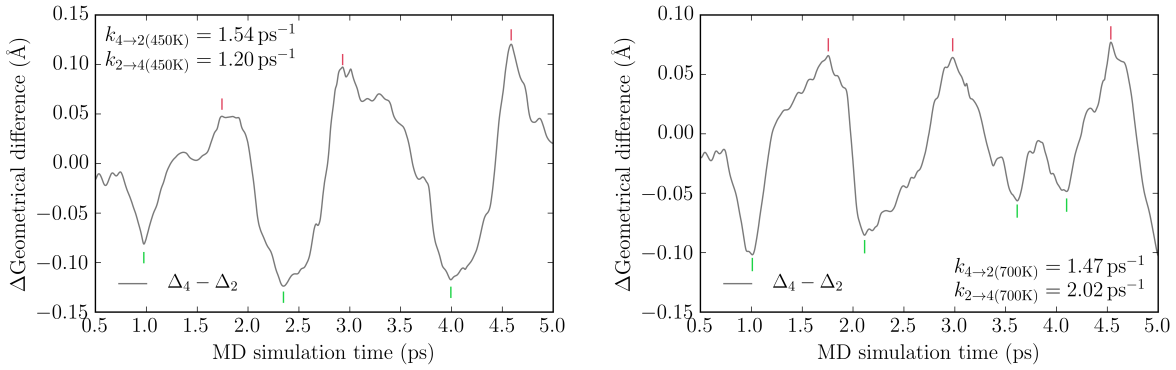


Figure 4.9 The difference between geometrical differences measured from isomers #4 and #2 based on MD trajectories starting from isomer #2 simulated at (a, left) 450 K and (b, right) 700 K.

Table 4.2 The electronic energies, HTST predicted rate constants and rate constants approximately estimated from MD trajectories for the fastest isomerization processes, namely, transition between isomers #1 and #3 and between isomers #2 and #4, in MD simulation starting from isomer #1 and #2, respectively.

$\#_{init}$	$\#_{final}$	E_{init} (eV)	E_{TS} (eV)	T = 450 K		T = 700 K	
				k_{HTST} (1/ps)	k_{MD} (1/ps)	k_{HTST} (1/ps)	k_{MD} (1/ps)
3	1	0.226	0.227	10.6	2.47	10.7	2.01
1	3	0.000	0.227	0.0337	2.33	0.272	3.08
4	2	0.249	0.267	14.7	1.54	17.2	1.47
2	4	0.073	0.267	3.56	1.20	21.3	2.02

Conclusions

The method development for efficient exploration of catalytic cluster PES, is of critical importance in the computational study of these systems. The general guide for studying these systems is first determining the important features of the system under consideration (such as the bond length distribution, point group symmetry, space group symmetry, and number of identical atoms) and then utilizing some of these features to speed up computation. Based on this simple idea we have proposed several different approaches in this dissertation.

We reported a new method, AFFCK, for finding the global and local minima of gas phase clusters. Based on traditional CK, AFFCK utilizes an intermediate step where all candidate structures are pre-optimized using a classical FF. Despite the skepticism toward FF in general when applied to clusters, it works surprisingly well, because FF in AFFCK is learned on-the-fly, for every given system under consideration. As a result, FF energies, energy rankings, and geometries of optimized species are shown be in good agreement with DFT results. The number of steps required for the final DFT geometry optimization of all minima is tremendously reduced, and thus, overall, the method is much more efficient than the traditional CK. We illustrate its performance on the platinum clusters. For Pt_8, we identified the global minimum structure that is much lower in energy than the previously reported one. We also note that for clusters that possess a mix of covalent and delocalized bonding, such as clusters of boron, AFFCK is less efficient than for all-metal clusters that possess only

delocalized bonding. We would recommend AFFCK for clusters of transition metals, such as those used in catalysis.

We proposed another parallel global optimization scheme based on DNN fitting, called NN-PGOPT, and demonstrated that this new scheme is able to successfully find the global minima of gas phase Pt_9 and Pt_{13}, which are in agreement with some other literatures, as well as low energy optima, which are important at realistic temperature. Using the S-BLDA structure generation method, the configuration space can be randomly sampled in an efficient and natural way to generate initial structures. With the help of NN based local optimization combined with limited-step DFT optimization, the global optimization is proved to be faster than the traditional full-step DFT optimization embedded way. In addition, the training of DNN can be greatly accelerated by modern GPU accelerators. An efficient DFS based structure similarity measurement algorithm has also been proposed and duplicates can be excluded at different stages of the global search. The tricapped octahedron structure is found to be the putative global minimum of Pt_9 using TPSSh functional, but a planar structure has even lower energy when PBE functional is used. The tricapped pentagonal prism structure is found to be the putative global minimum of Pt_{13} using PBE functional, while TPSSh functional favors a square pyramid structure. However, at catalysis relevant temperature, the low symmetry and high multiplicity structures are predicted to be more populated, which is found to be a functional independent fact, for the system investigated in this work. Particularly, the structure transitions for Pt_9 and Pt_{13} clusters can be identified, at relatively low and high temperatures, respectively,

based on the ensemble-average representation of local minima. The energy separation of isomers can also have a significant influence on the properties observed at finite temperature.

For surface supported clusters, we utilized the standard BH approach for global optimization, and proposed a new approach for constructing isomerization graph of Pt_7 isomerization on α-alumina surface. A MST derived from this graph shows all MEPs of transition between any two Pt_7 isomers whose energies are below 0.6 eV, which are represented as vertices in the MST. We also introduced a structure difference measurement based on mean distance in cluster geometries and surface space group symmetry, and a fast polynomial algorithm based on bipartite model for efficient evaluation. This measurement is used to generate initial NEB paths and analyze MD trajectories. From both the HTST isomerization graph and MD trajectories, we find that the shape of surface supported Pt_7 isomers can be highly fluxional at finite temperature, with energy barriers of direct paths along MEP varying from ca. 0.00 eV to 0.78 eV. Based on our local fluxionality picture, the Pt_7 isomerization on α-alumina surface can be understood in two different timescales. The short time (less than 1 ns when T = 700 K) behavior is frequent transition among similar shapes within local region, and the relatively longer time (greater than 1 ns when T = 700 K) behavior is thermal equilibrium across regions. At lower temperature, the time span for local fluxionality can be much larger. We note that this locally fluxional picture may have important implications for explaining catalytic mechanisms. On the one hand, the structurally similar isomers within each region share some common binding sites for reactant or product molecules. These common binding sites will exist for a relatively long time

when the system only fluctuates locally (namely, within the region) but not globally. On the other hand, after considerably longer time the thermal equilibrium among all low-energy isomers will be achieved, which brings to question the reactivity on such clusters being considered at nearly fixed geometries (usually the global minimum).[117] An ensemble representation consisting of low energy isomers may be a better computational model for such occasions.[87,118]

Finally, we proposed a revised BH global optimization approach, with core-shell separation, to address the challenging task of global optimization of surface-supported clusters with high coverage of adsorbates. Using this revised method, we show that the high hydrogen coverage changes the preferred geometry of alumina-supported Pt_7 clusters. This observation implies that the isomers obtained from global optimization of the bare Pt_7 cluster on a support surface (or in the gas phase) may not be used as a "reference state" for the system with large number of adsorbates. For H-covered Pt_7 on α-Al_2O_3, a rich spectrum of isomers thermally-accessible at 700-900 K is found (700-900 K is a typical temperature range for catalytic dehydrogenation). Among the isomers, many have a common structure of the Pt core beneath the bound adsorbates, and thus these groups of isomers differ only by the adsorbate positions. Several typical core types, labelled A-F, have been identified. Their presence in the statistical ensemble of the catalyst states changes as a function of temperature: The structures of the type A (which includes the global minimum) prevail below 200 K. Above that temperature, the contributions of the structures B-F grow, and at 700 K and above all of them become nearly equal with the population of A. Hence, all these states may play a role in a catalytic process on this cluster catalyst. In addition, for these H-

covered systems, the H atoms and the Pt atoms have qualitatively different mobilities. This is seen not only from the appearance of the typical Pt core structures, A-F, but also from the ab initio MD simulations at 700 K, which show that Hs can quite easily move and change binding site on the cluster, while the Pt atoms are less mobile, and do not move in the time scale of the MD (10 ps). This further supports the idea that the general global optimization approach with a unified treatment for all atoms (those of the cluster and of the adsorbates) may not be very efficient for modelling catalytic clusters in realistic conditions, and further validates the method with the core-shell separation presented in this study.

In summary, the computational study of catalytic clusters is a very attractive and active research field. On one hand, the non-scalable feature, surface effect, finite temperature and realistic coverage all together make the computational modeling of the system a very difficult task. However, it is also believed that the true understanding of the reaction catalyzed by these clusters deeply relies on the suitable treatment of these effects. On the other hand, these clusters are also very "simple" objects. They may have point group symmetry, atomic permutation symmetry and space group symmetry. Their bond length may obey simple normal distribution. The mobility of coverage atoms and cluster atoms may at quite different scale. Whether we can find and incorporate these properties in any PES exploration method is then essential for the performance of the method. In view of this, we think that there are still plenty of opportunities in the development of methods in this field and we believe that more comprehensive and efficient methods can be proposed in the near future.

Appendix

Parallel global optimization code PGOPT

The AFFCK method discussed in this dissertation has been implemented in an open source package called AFFCK (available from https://github.com/hczhai/AFFCK).

For most of remaining methods proposed in this dissertation, their implementation is integrated in a single open source package called PGOPT (available from https://github.com/hczhai/PGOPT).

This appendix will give a short introduction of the installation and usage of PGOPT code.

Installation

This code contains three sub-packages. *ACNN* contains all core algorithms. *PGOPT* includes settings for supercomputer environment. *STMOLE* defines the interface to *VASP* and *TURBOMOLE*.

If you want to test the code in a non-supercomputer environment, *PGOPT* and *STMOLE* are not needed.

Here we will explain installation steps using a *docker* (https://docs.docker.com/) image of *anaconda2* (https://hub.docker.com/r/continuumio/anaconda). Please first install the docker application in your system, then open a terminal to continue. (You

can also use anaconda with *python2* installed in your local environment without docker.) The following commands will download an image of anaconda2:

```
docker pull continuumio/anaconda
docker run –it –-rm continuumio/anaconda /bin/bash
```

Now you are inside the docker container. We need to add the following python packages and several packages needed for compiling fortran code:

```
pip install theano reportlab dill
conda install pygpu
apt-get update
apt-get -y install gfortran g++ make vim
```

Now get the copy of PGOPT:

```
cd ~
git clone https://github.com/hczhai/PGOPT.git PGOPT-PROGRAMS
```

Then we can compile the fortran code. This will generate some warnings but the process should not produce any error.

```
cd ~/PGOPT-PROGRAMS/ACNN/formod
make
```

Now add the following to `~/.bashrc` (you may need `vim ~/.bashrc` first):

```
BASE=~/PGOPT-PROGRAMS
export STMOLE_HOME=$BASE/STMOLE
export ACNNHOME=$BASE/ACNN
export PGOPTHOME=$BASE/PGOPT
export PATH=$STMOLE_HOME:$PATH
export PATH=$PGOPTHOME:$ACNNHOME:$PATH
```

Then apply the these environment changes:

```
source ~/.bashrc
```

Gas phase cluster generation

The main program is called `acnnmain` which can be found under `$ACNNHOME` defined previously. You need to prepare an input file for generating structures. There are some example input files under ACNN/tests. Here as an example, we will try to generate some gas phase Pt_7 structures using S-BLDA.

```
cd $ACNNHOME/tests/structure-generation
acnnmain pt7-gas.json
```

Note that the output directory is indicated in the input file. Here we can find the results in `./OUT-pt7-gas/fil_structs.xyz.0`. The structures are written in XYZ format. A single file will contains more than one structures. Visualization software such as *jmol* is useful for examine all structures within only one file. For example, if *jmol* is installed, you can type:

```
jmol ./OUT-pt7-gas/fil_structs.xyz.0
```

to look at the structures.

The PGOPT code also has its own visualization implementation. It will generate a PDF file containing images of all structures in the given input file. To generate the PDF, use the following input file (this only works after you run `acnnmain pt7-gas.json`):

```
acnnmain pt7-gas-draw.json
```

Then you will find the PDF in `./OUT-pt7-gas/report.pdf`.

Surface supported cluster generation

The following command will generate Pt_7 structures on α-Al_2O_3 surface. The surface is described by a XYZ file. There are some example surface files under `$ACNNHOME/tests/surfaces`. The computational cell information is written in the comment line of the XYZ file. It can be either 3 numbers or 5 numbers. The program always assume the Z direction is normal to the surface plane. If the cell size is described by 3 numbers "n_1 n_2 n_3", then the XYZ components of cell axes are a = (n_1, 0, 0), b = (0, n_2, 0), c = (0, 0, n_3). If the cell size is described by 5 numbers "n_1 n_2 n_3 n_4 n_5", then the XYZ components of cell axes are a = (n_1, n_2, 0), b = (n_3, n_4, 0), c = (0, 0, n_5). Usually n_5 is larger than the actual height of the surface because of the added vacuum gap. So in the end of comment line there is an additional number in parenthesis, indicating the unit cell height along Z without vacuum gap.

```
acnnmain pt7-alpha.json
```

The surface group and unit cell (the minimal unit cell, not the computational unit cell) information, indicated in the input file may help determining structure duplicates. If these information is unavailable, use the computational cell for unit cell and "P 1" for space group, and [0.0, 0.0] for space group transformation reference point.

Then we can find the results in `./OUT-pt7-alpha/fil_structs.xyz.0`.

Structure filtering

The following command will try to find unique structures from a given example XYZ file containing some local minima (`$ACNNHOME/tests/data/pt4b4-local.xyz`).

```
cd $ACNNHOME/tests/filtering
acnnmain pt4b4-local.json
```

The unique structures will be in `./OUT-pt4b4-filter/fil_structs.xyz.0`. The additional file `./OUT-pt4b4-filter/fil_list.txt.0` shows how many duplicates of each unique structure appear in the original input XYZ file (the "multi" column). The other additional file `./OUT-pt4b4-filter/fil_corr.txt.0` lists the structural difference data. In this file, if one line ends with "*", then the structure is selected as unique structure, because its structural difference to all previous unique structures is higher than the threshold. The second last column "mindm" shows the minimal value over structural difference to all previous structures.

If the structure filtering should be performed on surface support clusters, the "creation-surface" section should be given in input file, which contains the same information as that in the input file for creation.

Neural network fitting

Note that NN fitting is only implemented for gas phase clusters containing only one type of element.

The following command will try to fit a neural network based on an example Pt_9 data (`$ACNNHOME/tests/data/pt9-structs.xyz`). Note that for realistic results, we need to set "sample_number" parameter to [200000, 20000, 20000] and "epochs" to

2000. This calculation normally requires large memory or GPU. If there is no GPU to use, it will automatically switch to CPU.

```
cd $ACNNHOME/tests/nn_fitting
export OMP_NUM_THREADS=40
acnnmain pt9-fit.json
```

After the fitting is finished, the fitted network will be stored in `./OUT-pt9-nn/fit_network.dill.0`. We have a reference result from executing "`acnnmain pt9-fit.json`" stored under `$ACNNHOME/tests/nn_fitting_ref`.

Next we need to create some new structures, then use the network to optimize them. The file `./OUT-pt9-nn/fit_network.dill.0` is required for optimization.

```
acnnmain pt9-create.json
acnnmain pt9-opt.json
```

The optimized structures and energies are in `./OUT-pt9-nn/opt_structs.xyz.0`.

Parallel global optimization

For large scale global optimization (such as BH global optimization of Pt_7 supported on α-Al_2O_3 surface), we typically need hundreds of thousands of CPU hours. This can only be possible if the calculation is performed in a supercomputer and highly parallelized. To perform this kind of calculation, PGOPT has to know the format of the submission script and command of a certain supercomputer environment. PGOPT uses the command `hostname` to detect the supercomputer environment. The submission command names and number of cores in each node are then specified in the script `PGOPT/pg/hosts.py`. The templates for the submission script of different types of

jobs are in folder `PGOPT/scripts.templates`. The environment dependent data for each supercomputer is given in folder `PGOPT/scripts.spec`.

PGOPT will also need to invoke electronic structure software (such as VASP) for energy evaluation. The VASP binaries should appear in `$VASPHOME`. The uncompressed pseudopotential files should be stored in `$VASP_PP_PATH`, organized as follows

```
$VASP_PP_PATH/potpaw_GGA/<element-name>/POTCAR
$VASP_PP_PATH/potpaw_PBE/<element-name>/POTCAR
```

In parallel environment, command `pgopt` will be used to initialize calculation (`pgopt init`), edit input file (`pgopt set`), generating job scripts (`pgopt relax/pgopt torun`) submit jobs (`pgopt submit`), checking results (`pgopt log/pgopt mclog`), and finalize results (`pgopt report`). A list of pgopt commands can also be used as a way for tracking the settings for a global optimization (which is computational enbironment independent) and repeating at a different time or different environment.

For example, the following pgopt recipe will perform global optimization of Pt_7 gas phase clusters in parallel (the default local optimization method is DFT using VASP with PBE functional)

```
pgopt init Pt7 500
pgopt set relax
pgopt set opts+ "encut=300;lwave=T;cell=18"
pgopt set creation order 3
pgopt set creation 2d 0.2
pgopt relax 1
pgopt torun para 0 24
pgopt submit relax 1
pgopt submit torun para
```

With this setup, 500 initial structures will be generated and optimized independently (in parallel) to obtain around 500 local minima. Then these local minima will be checked for duplicates and sorted. When the calculation finished, the results can be found in XYZ format in `master/1.0/par_local.xyz.0`.

Specifically, "`pgopt init Pt7 500`" will initialize a calculation for Pt_7 with 500 initial structures. But no structures will be created after this command. Instead, a template input file (in *json* format, which can be read by `acnnmain`) named "para-template.json" will be created. This input file contains many default settings. Then several "`pgopt set`" commands are used to edit the input file. The input file can also be edited directly, but with "`pgopt set`" it will be easier to track all the changes in a future time. "`pgopt set relax`" adds a VASP section in the input file, with the default method DFT and functional PBE. The next command changes some VASP parameters. Then the two "`pgopt set creation`" commands changes the "creation" section in the input file.

Then we have to create submission scripts. There are two types of scripts/jobs. Master job typically uses only one CPU node, which is responsible for generating initial structures, filtering structures, making BH perturbations and monitoring worker jobs. Worker jobs are responsible for performing the time-consuming electronic structure computation. Here "`pgopt relax 1`" will generate a master job script "tomater/relax-1.0/run-master.sh" (in scratch directory). Here "`pgopt torun para 0 24`" will generate 24 worker job scripts numbered from zero under "torun" directory (in scratch

directory). Command "`cd `pgopt show``" can be used to switch back and forth between main and scratch directory. The last two "`pgopt submit`" commands submit the scripts into the job queue.

When the calculation is running, "`pgopt log`" can be used to check the status of the calculation. For BH calculation, "`pgopt mclog`" can be used to check the status of individual MC walkers. After the calculation is finished, "`pgopt report`" can be used to generate a PDF report.

Basin hopping global optimization

The following pgopt recipe can be used to perform a BH global optimization for gas phase Pt_7 cluster. In the "pgopt init" command we have specified (optional) that two nodes will be used for each worker job.

```
pgopt init Pt7 10 "--nodes=2"
pgopt set mc
pgopt set mc temperature 1500
pgopt set mc short-distance-factor 0.9
pgopt set mc max-iter 200
pgopt set relax
pgopt set do-monte-carlo T
pgopt set opts+ "encut=300;lwave=T;cell=18"
pgopt set opts^ "scf(iter=100)"
pgopt set args step 50
pgopt set creation order 3
pgopt set creation 2d 0.2
pgopt relax 1
pgopt torun para 0 10
pgopt submit relax 1
pgopt submit torun para
```

With this setup, 10 BH walkers, each with 200 moves will be used. The temperature for BH will be set to 1,500 K.

Basin hopping global optimization with coverage

The following pgopt recipe can be used to perform a BH global optimization for gas phase $Pt_7H_{10}CH_3$ system, with structure generation and BH parameters set similar to those discussed in Section 2.4.

```
pgopt init "Pt7|H10(CH3)" 1 "--surface=alpha4rz" "--nodes=8"
pgopt set molecules+ CH3
pgopt set creation ~lowpos 0.1
pgopt set creation ~loworderelems H
pgopt set creation ~nomoleelems -CH3,H
pgopt set creation ~nocluselems -H,Al,H,O,C,-CH3,Al,H,O
pgopt set creation number 10
pgopt set mc
pgopt set mc temperature 1500
pgopt set mc short-distance-factor 0.85
pgopt set mc max-iter 300
pgopt set mc detailed-balance F
pgopt set mc swap-site T
pgopt set mc swap-site-make-space CH3
pgopt set mc swap-site-rate 0.25
pgopt set mc keep-ch3 T
pgopt set mc solid-move CH3
pgopt set mc light-shell T
pgopt set 0 relax
pgopt set 0 do-monte-carlo T
pgopt set 0 opts^ "sigma=0.1;encut=300;ediff=1E-4;ediffg=1E-3;"
pgopt set 0 opts^ "lreal=A;scf(iter=300)"
pgopt set 0 opts^ "lwave=F;algo=Fast;npar=16"
pgopt set 0 args step 50
pgopt set 0 args max_step 750
pgopt relax 1
pgopt torun para 0 10 "--time=12:00:00"
pgopt submit relax 1
pgopt submit torun para 0 10
```

In order to create surface supported clusters, the surface files “alpha4rz.xyz” and “alpha4rz.json” must present in the main directory before executing the commands. Also the “CH3.xyz” should be provided before “`pgopt set molecules+ CH3`” command. These files can be found in “ACNN/tests/surfaces” directory of the package.

Isomerization pathway optimization

As a first step, type "`pgopt filter 1`" in a converged global optimization calculation will try to find all unique local minima using a stricter condition for structure similarity. The results can be found in a XYZ file in "`master/1.0/par_filter.xyz.0`".

As a second step, type "`pgopt connect 1`" in a converged global optimization calculation will generate the images for NEB pathway calculation for the isomerization between unique isomers in "`master/1.0/par_filter.xyz.0`". Detailed parameters can be set by modifying input file "`tomaster/relax-1.0/connect.json`" before issuing this command. The images will be stored in "`master/1.0/conn_images.xyz.0`". This file is required for starting the NEB calculation. Copy and rename this file (for example, to "images.xyz") in a new clean directory.

In the new directory, use the following pgopt recipe to perform the pathway optimization (Assuming the system is Pt_7 and the surface file name starts with "alphaz").

```
pgopt init Pt7 0 "--surface=alphaz" "--nodes=20"
pgopt set 0 neb
pgopt set 0 sources^ read,images.xyz
pgopt set 0 opts^ "sigma=0.1;encut=400"
pgopt set 0 opts^ "ediff=5E-6;ediffg=-0.05;lclimb=F;"
pgopt set 0 opts^ "lreal=A;scf(iter=300);lwave=T;algo=Fast"
pgopt neb 1 "--time=24:00:00"
pgopt torun para 0 10 "--time=24:00:00"
pgopt submit neb 1
pgopt submit torun para 0 10
```

Note that the node number in first line must be a multiple of number of images. After the calculation is finished, the command "`pgopt report`" can be used to generate pathway curves and the isomerization graph.

References

[1] Alexandrova, A. N.; Boldyrev, A. I. Search for the Li n 0/+1/-1 (n = 5−7) Lowest-Energy Structures Using the Ab Initio Gradient Embedded Genetic Algorithm (GEGA). Elucidation of the Chemical Bonding in the Lithium Clusters. *J. Chem. Theory Comput.* **2005**, *1*, 566–580.

[2] Alexandrova, A. N. H ·(H2O)n Clusters: Microsolvation of the Hydrogen Atom via Molecular Ab Initio Gradient Embedded Genetic Algorithm (GEGA). *J. Phys. Chem. A* **2010**, *114*, 12591–12599.

[3] Kanters, R. P. F.; Donald, K. J. CLUSTER : Searching for Unique Low Energy Minima of Structures Using a Novel Implementation of a Genetic Algorithm. *J. Chem. Theory Comput.* **2014**, *10*, 5729–5737.

[4] Davis, J. B. A.; Shayeghi, A.; Horswell, S. L.; Johnston, R. L. The Birmingham Parallel Genetic Algorithm and Its Application to the Direct DFT Global Optimisation of Ir N (N = 10–20) Clusters. *Nanoscale* **2015**, *7*, 14032–14038.

[5] Bandow, B.; Hartke, B. Larger Water Clusters with Edges and Corners on Their Way to Ice: Structural Trends Elucidated with an Improved Parallel Evolutionary Algorithm. *J. Phys. Chem. A* **2006**, *110*, 5809–5822.

[6] Call, S. T.; Zubarev, D. Y.; Boldyrev, A. I. Global Minimum Structure Searches via Particle Swarm Optimization. *J. Comput. Chem.* **2007**, *28*, 1177–1186.

[7] Avendaño-Franco, G.; Romero, A. H. Firefly Algorithm for Structural Search. *J. Chem. Theory Comput.* **2016**, *12*, 3416–3428.

[8] Wang, J.; Ma, L.; Zhao, J.; Jackson, K. A. Structural Growth Behavior and Polarizability of Cd(n)Te(n) (N=1-14) Clusters. *J. Chem. Phys.* **2009**, *130*, 214307.

[9] Wales, D. J.; Doye, J. P. K. Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *J. Phys. Chem. A* **1997**, *101*, 5111–5116.

[10] Zhai, H.-J.; Zhao, Y.-F.; Li, W.-L.; Chen, Q.; Bai, H.; Hu, H.-S.; Piazza, Z. A.; Tian, W.-J.; Lu, H.-G.; Wu, Y.-B.; et al. Observation of an All-Boron Fullerene. *Nat. Chem.* **2014**, *6*, 727–731.

[11] Oganov, A. R. *Modern Methods of Crystal Structure Prediction*; Oganov, A. R., Ed.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2010.

[12] Sergeeva, A. P.; Averkiev, B. B.; Zhai, H.-J.; Boldyrev, A. I.; Wang, L.-S. All-Boron Analogues of Aromatic Hydrocarbons: B 17 − and B 18 −. *J. Chem. Phys.* **2011**, *134*, 224304.

[13] Vilhelmsen, L. B.; Hammer, B. A Genetic Algorithm for First Principles Global Structure Optimization of Supported Nano Structures. *J. Chem. Phys.* **2014**, *141*.

[14] Hjorth Larsen, A.; Jørgen Mortensen, J.; Blomqvist, J.; Castelli, I. E.; Christensen, R.; Dułak, M.; Friis, J.; Groves, M. N.; Hammer, B.; Hargus, C.; et al. The Atomic Simulation Environment—a Python Library for Working with Atoms. *J. Phys. Condens. Matter* **2017**, *29*, 273002.

[15] Zhai, H.; Ha, M.-A.; Alexandrova, A. N. AFFCK: Adaptive Force-Field-Assisted Ab Initio Coalescence Kick Method for Global Minimum Search. *J. Chem. Theory Comput.* **2015**, *11*, 2385–2393.

[16] Saunders, M. Stochastic Search for Isomers on a Quantum Mechanical Surface. *J. Comput. Chem.* **2004**, *25*, 621–626.

[17] Weisstein, E. W. Sphere Point Picking http://mathworld.wolfram.com/SpherePointPicking.html.

[18] Jmol: An Open-Source Java Viewer for Chemical Structures in 3D http://www.jmol.org/.

[19] Sumpter, B. G.; Noid, D. W. Potential Energy Surfaces for Macromolecules. A Neural Network Technique. *Chem. Phys. Lett.* **1992**, *192*, 455–462.

[20] Malshe, M.; Narulkar, R.; Raff, L. M.; Hagan, M.; Bukkapatnam, S.; Komanduri, R. Parametrization of Analytic Interatomic Potential Functions Using Neural Networks. *J. Chem. Phys.* **2008**, *129*.

[21] Le, H. M.; Raff, L. M. Molecular Dynamics Investigation of the Bimolecular Reaction BeH + H 2 → BeH 2 + H on an Ab Initio Potential-Energy Surface Obtained Using Neural Network Methods with Both Potential and Gradient Accuracy Determination. *J. Phys. Chem. A* **2010**, *114*, 45–53.

[22] Li, J.; Jiang, B.; Guo, H. Permutation Invariant Polynomial Neural Network Approach to Fitting Potential Energy Surfaces. II. Four-Atom Systems. *J. Chem. Phys.* **2013**, *139*, 1–8.

[23] Behler, J. Atom-Centered Symmetry Functions for Constructing High-Dimensional Neural Network Potentials. *J. Chem. Phys.* **2011**, *134*, 074106.

[24] Behler, J. Constructing High-Dimensional Neural Network Potentials: A Tutorial Review. *Int. J. Quantum Chem.* **2015**, *115*, 1032–1050.

[25] Artrith, N.; Morawietz, T.; Behler, J. High-Dimensional Neural-Network Potentials for Multicomponent Systems: Applications to Zinc Oxide. *Phys. Rev. B* **2011**, *83*, 153101.

[26] Ouyang, R.; Xie, Y.; Jiang, D. Global Minimization of Gold Clusters by Combining Neural Network Potentials and the Basin-Hopping Method. *Nanoscale* **2015**, *7*, 14817–14821.

[27] Chiriki, S.; Bulusu, S. S. Modeling of DFT Quality Neural Network Potential

for Sodium Clusters: Application to Melting of Sodium Clusters (Na20 to Na40). *Chem. Phys. Lett.* **2016**, *652*, 130–135.

[28] Manzhos, S.; Carrington, T. A Random-Sampling High Dimensional Model Representation Neural Network for Building Potential Energy Surfaces. *J. Chem. Phys.* **2006**, *125*, 084109.

[29] Manzhos, S.; Carrington, T. Using Neural Networks, Optimized Coordinates, and High-Dimensional Model Representations to Obtain a Vinyl Bromide Potential Surface. *J. Chem. Phys.* **2008**, *129*, 224104.

[30] Bengio, Y. Learning Deep Architectures for AI. *Found. Trends® Mach. Learn.* **2009**, *2*, 1–127.

[31] Krizhevsky, A.; Sutskever, I.; Hinton, G. E. Imagenet Classification with Deep Convolutional Neural Networks. *Neural Inf. Process. Syst. 2012* **2012**.

[32] Bengio, Y.; Lamblin, P.; Popovici, D.; Larochelle, H. Greedy Layer-Wise Training of Deep Networks Yoshua. *Adv. Neural Inf. Process. Syst.* **2007**, *19*, 153.

[33] Chetlur, S.; Woolley, C.; Vandermersch, P.; Cohen, J.; Tran, J.; Catanzaro, B.; Shelhamer, E. CuDNN: Efficient Primitives for Deep Learning. *CoRR* **2014**, *abs/1410.0*, 1–9.

[34] Murrell, J. N.; Carter, S. Approximate Single-Valued Representations of Multivalued Potential Energy Surfaces. *J. Phys. Chem.* **1984**, *88*, 4887–4891.

[35] Tersoff, J. Modeling Solid-State Chemistry: Interatomic Potentials for Multicomponent Systems. *Phys. Rev. B* **1989**, *39*, 5566–5568.

[36] Jarrett, K.; Kavukcuoglu, K.; Ranzato, M.; LeCun, Y. What Is the Best Multi-Stage Architecture for Object Recognition? *2009 IEEE 12th Int. Conf. Comput. Vision. IEEE* **2009**, 2146–2153.

[37] Glorot, X.; Bordes, A.; Bengio, Y. Deep Sparse Rectifier Neural Networks. *Int. Conf. Artif. Intell. Stat.* **2011**, *15*, 315–323.

[38] Hochreiter, S.; Bengio, Y.; Frasconi, P.; Schmidhuber, J. Gradient Flow in Recurrent Nets: The Difficulty of Learning LongTerm Dependencies. In *A Field Guide to Dynamical Recurrent Networks*; IEEE, 2009; pp 237–243.

[39] The Theano Development Team; Al-Rfou, R.; Alain, G.; Almahairi, A.; Angermueller, C.; Bahdanau, D.; Ballas, N.; Bastien, F.; Bayer, J.; Belikov, A.; et al. Theano: A Python framework for fast computation of mathematical expressions http://arxiv.org/abs/1605.02688.

[40] Sutskever, I.; Martens, J.; Dahl, G.; Hinton, G. On the Importance of Initialization and Momentum in Deep Learning. In *Proceedings of the 30th International Conference on Machine Learning, Atlanta, Georgia, USA*; 2013; Vol. 3, pp 1139–1147.

[41] Zhu, C.; Byrd, R. H.; Lu, P.; Nocedal, J. Algorithm 778: L-BFGS-B: Fortran Subroutines for Large-Scale Bound-Constrained Optimization. *ACM Trans. Math. Softw.* **1997**, *23*, 550–560.

[42] Ballard, A. J.; Stevenson, J. D.; Das, R.; Wales, D. J. Energy Landscapes for a Machine Learning Application to Series Data. *J. Chem. Phys.* **2016**, *144*, 124119.

[43] Jensen, F. *Introduction to Computational Chemistry*; John wiley & sons: Chichester, 2007.

[44] Kresse, G.; Hafner, J. Ab Initio Molecular Dynamics for Liquid Metals. *Phys. Rev. B* **1993**, *47*, 558–561.

[45] Kresse, G.; Hafner, J. Ab Initio Molecular-Dynamics Simulation of the Liquid-Metal–Amorphous-Semiconductor Transition in Germanium. *Phys. Rev. B* **1994**, *49*, 14251–14269.

[46] Kresse, G.; Furthmüller, J. Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Comput. Mater. Sci.* **1996**, *6*, 15–50.

[47] Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54*, 11169–11186.

[48] TURBOMOLE V6.3 2011, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007. http://www.turbomole.com.

[49] Perdew, J. P.; Wang, Y. Accurate and Simple Analytic Representation of the Electron-Gas Correlation Energy. *Phys. Rev. B* **1992**, *45*, 13244–13249.

[50] Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[51] Perdew, J. P.; Ernzerhof, M.; Burke, K. Rationale for Mixing Exact Exchange with Density Functional Approximations. *J. Chem. Phys.* **1996**, *105*, 9982–9985.

[52] Alexandrova, A. N.; Boldyrev, A. I.; Zhai, H.-J.; Wang, L.-S. All-Boron Aromatic Clusters as Potential New Inorganic Ligands and Building Blocks in Chemistry. *Coord. Chem. Rev.* **2006**, *250*, 2811–2866.

[53] Shen, L.; Dadras, J.; Alexandrova, A. N. Pure and Zn-Doped Pt Clusters Go Flat and Upright on MgO(100). *Phys. Chem. Chem. Phys.* **2014**, *16*, 26436–26442.

[54] Ha, M.-A.; Dadras, J.; Alexandrova, A. Rutile-Deposited Pt–Pd Clusters: A Hypothesis Regarding the Stability at 50/50 Ratio. *ACS Catal.* **2014**, *4*, 3570–3580.

[55] Weigend, F. Accurate Coulomb-Fitting Basis Sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057.

[56] Xiao, L.; Wang, L. Structures of Platinum Clusters: Planar or Spherical? *J. Phys. Chem. A* **2004**, *108*, 8605–8614.

[57] Nie, A.; Wu, J.; Zhou, C.; Yao, S.; Luo, C.; Forrey, R. C.; Cheng, H. Structural Evolution of Subnano Platinum Clusters. *Int. J. Quantum Chem.* **2007**, *107*, 219–224.

[58] TURBOMOLE V6.6 2014, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007. http://www.turbomole.com.

[59] Treutler, O.; Ahlrichs, R. Efficient Molecular Numerical Integration Schemes. *J. Chem. Phys.* **1995**, *102*, 346–354.

[60] Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. Comparative Assessment of a New Nonempirical Density Functional: Molecules and Hydrogen-Bonded Complexes. *J. Chem. Phys.* **2003**, *119*, 12129–12137.

[61] Soini, T. M.; Genest, A.; Nikodem, A.; Rösch, N. Hybrid Density Functionals for Clusters of Late Transition Metals: Assessing Energetic and Structural Properties. *J. Chem. Theory Comput.* **2014**, *10*, 4408–4416.

[62] Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.

[63] Singh, N. B.; Sarkar, U. Structure, Vibrational, and Optical Properties of Platinum Cluster: A Density Functional Theory Approach. *J. Mol. Model.* **2014**, *20*, 2537.

[64] Kumar, V.; Kawazoe, Y. Evolution of Atomic and Electronic Structure of Pt Clusters: Planar, Layered, Pyramidal, Cage, Cubic, and Octahedral Growth. *Phys. Rev. B* **2008**, *77*, 205418.

[65] Chaves, A. S.; Rondina, G. G.; Piotrowski, M. J.; Tereshchuk, P.; Da Silva, J. L. F. The Role of Charge States in the Atomic Structure of Cu n and Pt n (n = 2–14 Atoms) Clusters: A DFT Investigation. *J. Phys. Chem. A* **2014**, *118*, 10813–10821.

[66] Winczewski, S.; Rybicki, J. Structure of Small Platinum Clusters Revised. *Comput. Methods Sci. Technol.* **2011**, *17*, 75–85.

[67] Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50*, 17953–17979.

[68] Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. MOLPRO, version 2015.1, a package of ab initio programs http://www.molpro.net.

[69] Chou, J. P.; Hsing, C. R.; Wei, C. M.; Cheng, C.; Chang, C. M. Ab Initio Random Structure Search for 13-Atom Clusters of Fcc Elements. *J. Phys. Condens. Matter* **2013**, *25*, 125305.

[70] Bunău, O.; Bartolomé, J.; Bartolomé, F.; Garcia, L.-M. Large Orbital Magnetic Moment in Pt 13 Clusters. *J. Phys. Condens. Matter* **2014**, *26*, 196006.

[71] Sun, Y.; Zhang, M.; Fournier, R. Periodic Trends in the Geometric Structures of 13-Atom Metal Clusters. *Phys. Rev. B* **2008**, *77*, 075435.

[72] Zhang, M.; Fournier, R. Density-Functional-Theory Study of 13-Atom Metal Clusters M 13, M=Ta-Pt. *Phys. Rev. A* **2009**, *79*, 043203.

[73] Sinfelt, J. H. . In *Handbook of Heterogeneous Catalysis*; Ertl, G., Knözinger, E., Weitkamp, J., Eds.; Wiley: Weinheim, 1997; pp 1939–1955.

[74] Sattler, J. J. H. B.; Ruiz-Martinez, J.; Santillan-Jimenez, E.; Weckhuysen, B. M. Catalytic Dehydrogenation of Light Alkanes on Metals and Metal Oxides. *Chem. Rev.* **2014**, *114*, 10613–10653.

[75] Barron, Y.; Cornet, D.; Maire, G.; Gault, F. G. The Mechanism of Isomerization of Hexanes on Platinum Catalysts. *J. Catal.* **1963**, *2*, 152–155.

[76] Raybaud, P.; Chizallet, C.; Mager-Maury, C.; Digne, M.; Toulhoat, H.; Sautet, P. From γ-Alumina to Supported Platinum Nanoclusters in Reforming Conditions: 10years of DFT Modeling and Beyond. *J. Catal.* **2013**, *308*, 328–340.

[77] Rötzer, M. D.; Crampton, A. S.; Krause, M.; Thanner, K.; Schweinberger, F. F.; Heiz, U. Controlling Hydrogenation Selectivity by Size: 3-Hexyne on Supported Pt Clusters. *J. Phys. Chem. C* **2019**, *123*, 5518–5524.

[78] Jimenez-Izal, E.; Gates, B. C.; Alexandrova, A. N. Designing Clusters for Heterogeneous Catalysis. *Phys. Today* **2019**, *72*, 38–43.

[79] Baxter, E. T.; Ha, M.-A.; Cass, A. C.; Alexandrova, A. N.; Anderson, S. L. Ethylene Dehydrogenation on Pt 4,7,8 Clusters on Al 2 O 3 : Strong Cluster Size Dependence Linked to Preferred Catalyst Morphologies. *ACS Catal.* **2017**, *7*, 3322–3335.

[80] Jahel, A.; Avenier, P.; Lacombe, S.; Olivier-Fourcade, J.; Jumas, J.-C. Effect of Indium in Trimetallic Pt/Al2O3SnIn–Cl Naphtha-Reforming Catalysts. *J. Catal.* **2010**, *272*, 275–286.

[81] Gao, M.; Horita, D.; Ono, Y.; Lyalin, A.; Maeda, S.; Taketsugu, T. Isomerization in Gold Clusters upon O 2 Adsorption. *J. Phys. Chem. C* **2017**, *121*, 2661–2668.

[82] Lim, D.-H.; Wilcox, J. DFT-Based Study on Oxygen Adsorption on Defective Graphene-Supported Pt Nanoparticles. *J. Phys. Chem. C* **2011**, *115*, 22742–22747.

[83] Liu, J.-C.; Tang, Y.; Chang, C.-R.; Wang, Y.-G.; Li, J. Mechanistic Insights into Propene Epoxidation with O2 –H2O Mixture on Au7/α-Al2O3 : A Hydroproxyl Pathway from Ab Initio Molecular Dynamics Simulations. *ACS Catal.* **2016**, *6*,

2525–2535.

[84] Mager-Maury, C.; Bonnard, G.; Chizallet, C.; Sautet, P.; Raybaud, P. H2-Induced Reconstruction of Supported Pt Clusters: Metal-Support Interaction versus Surface Hydride. *ChemCatChem* **2011**, *3*, 200–207.

[85] Gorczyca, A.; Raybaud, P.; Moizan, V.; Joly, Y.; Chizallet, C. Atomistic Models for Highly-Dispersed PtSn/Γ-Al 2 O 3 Catalysts: Ductility and Dilution Affect the Affinity for Hydrogen. *ChemCatChem* **2019**, *11*, 3941–3951.

[86] Chen, H.-Y. T.; Tosoni, S.; Pacchioni, G. Hydrogen Adsorption, Dissociation, and Spillover on Ru 10 Clusters Supported on Anatase TiO 2 and Tetragonal ZrO 2 (101) Surfaces. *ACS Catal.* **2015**, *5*, 5486–5495.

[87] Sun, G.; Sautet, P. Metastable Structures in Cluster Catalysis from First-Principles: Structural Ensemble in Reaction Conditions and Metastability Triggered Reactivity. *J. Am. Chem. Soc.* **2018**, *140*, 2812–2820.

[88] Steinmann, S. N.; Corminboeuf, C. A Generalized-Gradient Approximation Exchange Hole Model for Dispersion Coefficients. *J. Chem. Phys.* **2011**, *134*, 044117.

[89] Gautier, S.; Steinmann, S. N.; Michel, C.; Fleurat-Lessard, P.; Sautet, P. Molecular Adsorption at Pt(111). How Accurate Are DFT Functionals? *Phys. Chem. Chem. Phys.* **2015**, *17*, 28921–28930.

[90] Zhang, Z.; Jimenez-Izal, E.; Hermans, I.; Alexandrova, A. N. Dynamic Phase Diagram of Catalytic Surface of Hexagonal Boron Nitride under Conditions of Oxidative Dehydrogenation of Propane. *J. Phys. Chem. Lett.* **2019**, *10*, 20–25.

[91] Zhai, H.; Alexandrova, A. N. Fluxionality of Catalytic Clusters: When It Matters and How to Address It. *ACS Catal.* **2017**, *7*, 1905–1911.

[92] Ha, M. A.; Baxter, E. T.; Cass, A. C.; Anderson, S. L.; Alexandrova, A. N. Boron

Switch for Selectivity of Catalytic Dehydrogenation on Size-Selected Pt Clusters on Al2O3. *J. Am. Chem. Soc.* **2017**, *139*, 11568–11575.

[93] Jimenez-Izal, E.; Zhai, H.; Liu, J.-Y.; Alexandrova, A. N. Nanoalloying MgO-Deposited Pt Clusters with Si To Control the Selectivity of Alkane Dehydrogenation. *ACS Catal.* **2018**, *8*, 8346–8356.

[94] Jimenez-Izal, E.; Liu, J.-Y.; Alexandrova, A. N. Germanium as Key Dopant to Boost the Catalytic Performance of Small Platinum Clusters for Alkane Dehydrogenation. *J. Catal.* **2019**, *374*, 93–100.

[95] Ponce, Y. M. Total and Local Quadratic Indices of the Molecular Pseudograph's Atom Adjacency Matrix: Applications to the Prediction of Physical Properties of Organic Compounds. *Molecules* **2003**, *8*, 687–726.

[96] Lv, J.; Wang, Y.; Zhu, L.; Ma, Y. Particle-Swarm Structure Prediction on Clusters. *J. Chem. Phys.* **2012**, *137*, 084104.

[97] Abraham, N. L.; Probert, M. I. J. Improved Real-Space Genetic Algorithm for Crystal Structure and Polymorph Prediction. *Phys. Rev. B* **2008**, *77*, 134117.

[98] Even, S. . In *Graph Algorithms*; Even, G., Ed.; Cambridge University Press: Cambridge, 2011; pp 46–48.

[99] Kuhn, H. W. The Hungarian Method for the Assignment Problem. *Nav. Res. Logist. Q.* **1955**, *7*, 83–97.

[100] Fukuda, K.; Matsui, T. Finding All the Perfect Matchings in Bipartite Graphs. *Appl. Math Lett.* **1994**, *7*, 15–18.

[101] Griffiths, M.; Niblett, S. P.; Wales, D. J. Optimal Alignment of Structures for Finite and Periodic Systems. *J. Chem. Theory Comput.* **2017**, *13*, 4914–4931.

[102] Helmich, B.; Sierka, M. Similarity Recognition of Molecular Structures by

Optimal Atomic Matching and Rotational Superposition. *J. Comput. Chem.* **2012**, *33*, 134–140.

[103] Allen, W. J.; Rizzo, R. C. Implementation of the Hungarian Algorithm to Account for Ligand Symmetry and Similarity in Structure-Based Design. *J. Chem. Inf. Model.* **2014**, *54*, 518–529.

[104] Wagner, A.; Himmel, H. J. ARMSD: A Comprehensive Tool for Structural Analysis. *J. Chem. Inf. Model.* **2017**, *57*, 428–438.

[105] Temelso, B.; Mabey, J. M.; Kubota, T.; Appiah-Padi, N.; Shields, G. C. ArbAlign: A Tool for Optimal Alignment of Arbitrarily Ordered Isomers Using the Kuhn-Munkres Algorithm. *J. Chem. Inf. Model.* **2017**, *57*, 1045–1054.

[106] Vajda, S.; Pellin, M. J.; Greeley, J. P.; Marshall, C. L.; Curtiss, L. A.; Ballentine, G. A.; Elam, J. W.; Catillon-Mucherie, S.; Redfern, P. C.; Mehmood, F.; et al. Subnanometre Platinum Clusters as Highly Active and Selective Catalysts for the Oxidative Dehydrogenation of Propane. *Nat. Mater.* **2009**, *8*, 213–216.

[107] Häkkinen, H.; Abbet, S.; Sanchez, A.; Heiz, U.; Landman, U. Structural, Electronic, and Impurity-Doping Effects in Nanoscale Chemistry: Supported Gold Nanoclusters. *Angew. Chemie Int. Ed.* **2003**, *42*, 1297–1300.

[108] Xing, X.; Li, X.; Yoon, B.; Landman, U.; Parks, J. H. Dynamic Fluxionality and Enhanced CO Adsorption in the Presence of Coadsorbed H2O on Free Gold Cluster Cations. *Int. J. Mass Spectrom.* **2015**, *377*, 393–402.

[109] Satoh, H.; Oda, T.; Nakakoji, K.; Uno, T.; Tanaka, H.; Iwata, S.; Ohno, K. Potential Energy Surface-Based Automatic Deduction of Conformational Transition Networks and Its Application on Quantum Mechanical Landscapes of D-Glucose Conformers. *J. Chem. Theory Comput.* **2016**, *12*, 5293–5308.

[110] Negreiros, F. R.; Aprà, E.; Barcaro, G.; Sementa, L.; Vajda, S.; Fortunelli, A. A

First-Principles Theoretical Approach to Heterogeneous Nanocatalysis. *Nanoscale* **2012**, *4*, 1208–1219.

[111] Henkelman, G. Atomistic Simulations of Activated Processes in Materials. *Annu. Rev. Mater. Res.* **2017**, *47*, 199–216.

[112] Koehl, P. Minimum Action Transition Paths Connecting Minima on an Energy Surface. *J. Chem. Phys.* **2016**, *145*, 184111.

[113] Wales, D. J. Energy Landscapes: Calculating Pathways and Rates. *Int. Rev. Phys. Chem.* **2006**, *25*, 237–282.

[114] Jorgensen, W. L.; Thomas, L. L. Perspective on Free-Energy Perturbation Calculations for Chemical Equilibria. *J. Chem. Theory Comput.* **2008**, *4*, 869–876.

[115] Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187–199.

[116] Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **2008**, *100*, 1–4.

[117] Campbell, C.; van Santen, R.; Stamatakis, M.; Collis, N.; Freund, H.-J.; Plaisance, C.; Sauer, J.; Garrett, B.; Gross, E.; Kotarba, A.; et al. Catalyst Design from Theory to Practice: General Discussion. *Faraday Discuss.* **2016**, *188*, 279–307.

[118] Zhai, H.; Alexandrova, A. N. Ensemble-Average Representation of Pt Clusters in Conditions of Catalysis Accessed through GPU Accelerated Deep Neural Network Fitting Global Optimization. *J. Chem. Theory Comput.* **2016**, *12*, 6213–6226.