Examining Saddle Point Searches in the Context of Off-Lattice Kinetic Monte Carlo

J. Hicks[1,*] and T. P. Schulze[2]

[1] *Department of Mathematics and Statistics, East Tennessee State University, Johnson City, TN 37614 - 0661, USA*
[2] *Department of Mathematics, University of Tennessee, Knoxville, TN 37996 - 1300, USA*

Abstract. In calculating the time evolution of an atomic system on diffusive timescales, off-lattice kinetic Monte Carlo (OLKMC) can sometimes be used to overcome the limitations of Molecular Dynamics. OLKMC relies on the harmonic approximation to Transition State Theory, in which the rate of rare transitions from one energy minimum to a neighboring minimum scales exponentially with an energy barrier on the potential energy surface. This requires locating the index-1 saddle point, commonly referred to as a transition state, that separates two neighboring energy minima. In modeling the evolution of an atomic system, it is desirable to find all the relevant transitions surrounding the current minimum. Due to the large number of minima on the potential energy surface, exhaustively searching the landscape for these saddle points is a challenging task. In examining the particular case of isolated Lennard-Jones clusters of around 50 particles, we observe very slow convergence of the total number of saddle points found as a function of successful searches. We seek to understand this behavior by modeling the distribution of successful searches and sampling this distribution to create a stochastic process that mimics this behavior. Finally, we will discuss an improvement to a rejection scheme for OLKMC where we terminate searches that appear to be failing early in the search process.

AMS subject classifications: 52B10, 65D18, 68U05, 68U07

Key words: Off-lattice kinetic Monte Carlo, Energy Landscape, Saddle Point Search, Dimer Method, Lennard-Jones clusters

1 Introduction

Simulating the time evolution of an atomic scale system in which a chemical reaction or diffusion occurs is an essential task in the study of condensed matter physics and material science. While molecular dynamics (MD) is often the preferred approach, it suffers from a severe timescale limitation due to the need to integrate the classical equations of motion for all the particles in the system using a time step on the order of a femtosecond. This restricts the timescale on which MD

---

*Corresponding author. *Email addresses:* `hicksj2@etsu.edu` (J. Hicks), `t.p.schulze@gmail.com` (T. P. Schulze)

can simulate events to mere nanoseconds, while diffusion and chemical reactions can take much longer. Off-lattice kinetic Monte Carlo (OLKMC), first introduced by Henkelman and Jónsson [4], is aimed at overcoming this limitation. As explained further below, the essential challenge with OLKMC lies in repeatedly building catalogs of saddle points/transition states. In this paper, we explore several related aspects of these saddle point searches.

OLKMC relies on the observation that the system will spend the majority of its time oscillating within the N_p-particle configuration space about a local minimum of the potential energy function, with rare transitions from one basin of attraction to another. The energy landscape typically features an enormous number of local minima, each of which is connected to a large number of neighboring minima that can be reached by crossing a single saddle point. In view of this, OLKMC seeks to replace the Newtonian dynamics of MD with a Markov chain model, with dynamics given by jumps between discrete states, represented by the local minima. The rates for these transition processes are estimated using the harmonic approximation to Transition State Theory (TST) [1], also known as Vineyard Theory [13]:

$$R_{ij} = K \exp^{-\Delta\phi_{ij}/k_B T}, \tag{1.1}$$

where $\Delta\phi_{ij}$ is the energy barrier that must be overcome in moving from basin i to a neighboring basin j, k_B is Boltzmann's constant, T is the system temperature, and K is a prefactor which we will take to be constant for the purpose of our discussion.

Determining the set of accessible states and the corresponding rates at each time step requires searching the potential energy surface (PES) for index-1 saddle points, where the gradient of the potential is zero and all but one of the principal curvatures is positive. The energy barrier is the difference between the saddle point (*transition state*) energy and the minimum energy, also known as the *binding state* energy:

$$\Delta\phi_{ij} = \phi(X_{ij}) - \phi(X_i), \tag{1.2}$$

where X_i is the current minimum configuration and X_{ij} is an index-1 saddle configuration. Ideally—for the method to yield a faithful representation of the dynamics—one would need to build an exhaustive catalog of *connected* index-1 saddle points at each discrete time step, where "connected" indicates that the current binding site can be reached from the saddle point/transition state by a path that is strictly descending. The difficulty in building such a catalog stems from the large number of individual searches that must be performed in order to explore the PES.

For the most part, we focus on two issues associated with OLKMC. The first issue is relevant to all OLKMC methods and concerns the *global search* for all of the connected saddle points as outlined above, while the second issue deals with an inefficiency in doing a certain type of *local search* to be described below. The issue associated with the global search is illustrated by considering what we will refer to as the *accumulation plot*, the top curve shown in Fig. 1. The lower curve is an analogous result associated with the local search. The accumulation plot shows the number of distinct, connected saddle points found as a function of the number of successful saddle point searches. Each of these searches is initiated from an initial guess, which is generated by what we will refer to as a *perturbation scheme*. Much of what we discuss is independent of the specific perturbation scheme chosen, but our specific scheme will be introduced in Section 2. This particular

accumulation plot was generated using a cluster of 55 particles arranged in what is thought to be the global minimum configuration for a Lennard-Jones potential [15]. Notice that the number of connected saddle points continues to grow, with no clear upper bound in sight. In Section 5, we will examine this behavior using a stochastic model, with an eye toward understanding why this happens.

While OLKMC is faster than the corresponding MD simulation, it is still computationally challenging and there is a great need for finding ways to accelerate the method. Ruzayqat and Schulze [12] introduce a scheme that relies on a standard Monte Carlo technique known as *rejection* , and this is where the local search mentioned above comes into play. In such a scheme, one samples a majorizing distribution with approximate rates, $\hat{r} \geq r$, rejecting an appropriate fraction of selected events, $\frac{\hat{r}-r}{\hat{r}}$, so that one is effectively sampling the original distribution. In Ref. [12], the set of connected saddle points is partitioned into disjoint subsets associated with individual atoms. The atom associated with a particular subset is referred to as the *key atom*, and is defined as the atom whose position changes by the greatest magnitude in going from the local minimum to a connected saddle point. The success of the rejection scheme relies on choosing a perturbation scheme that can target a specific atom, so that a high percentage of the searches initiated with a given target yield a saddle point where the targeted atom is the key atom. If this is the case, the resulting saddle point will be referred to as a *key connected* saddle. The most straight-forward way to accomplish this is to simply perturb the system so that the targeted atom is initially the atom whose position changes by the greatest magnitude in the sense identified above. This procedure is what we refer to as a local search. For the rejection scheme introduced in Ref. [12], rate estimates $\hat{r}_J \geq r_J$ are provided for each of these subsets, so that one can select a candidate subset J before performing a local search to evaluate the true rate r_J. This allows one to select an event without doing a much more costly global saddle point search.

In Fig. 1, we present the accumulation plot for both the local and the global searches. The two curves are generated from the same set of data, consisting of 36,849 successful dimer searches out of a total of 107,100 searches, with the data for non-key connected saddles having been discarded for the local search. This is the correct way to define the accumulation plot in the context of the rejection scheme, as this method discards the non-key connected saddles and only uses information learned about key connected saddles. This reveals a significant inefficiency in the local searches in that they locate many non-key connected saddles in the process of finding the set of key connected saddles. In a perfectly efficient scheme, these curves would coincide. In Section 6, we will introduce an *early termination* scheme that aims to overcome this problem. Early termination will monitor the individual searches in a way that will ensure that the atom selected retains key-atom status, terminating the search if this is not the case. Early termination will be shown to reduce the cost in computational time, allowing one to reinvest the time saved into more searches.

In Section 2, we discuss the particular system and algorithm choices that we use for our examples in later sections. In Section 3, we look at the *basins of attraction* for individual saddle points in 2D cross sections, illustrating the fractal and complex nature of the basins. In Section 4 we discuss the distribution of saddle points that is the result of the coupling of these basins with a distribution of initial guesses. In Section 5, we use these distributions to model the behaviour of accumulation plots. In Section 6, we introduce the early termination scheme for the Rejection
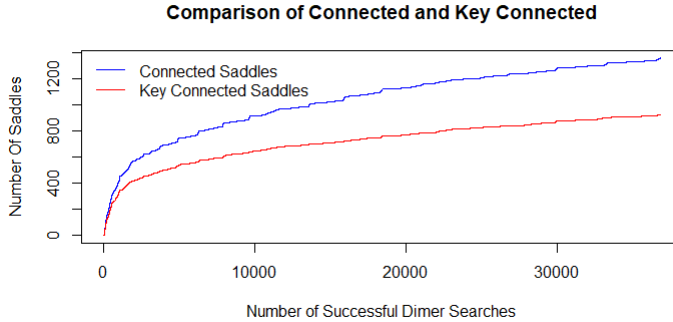
Comparison of Connected and Key Connected

Figure 1: The accumulation plot for a 55 particle configuration (upper, blue curve) and the corresponding accumulation plot for the local saddle point search (lower, red curve).

OLKMC discussed above. We recap our main observations and summarize our conclusions in the final section.

2 Implementation

This section describes various implementation details, such as choice of potential, the way constraints are placed on the system, and the method by which saddle points are located, including the way individual searches start from a distribution of initial guesses. While the observations, analysis and methods we present in later sections are largely independent of the choices described in this section, the data we present is specific to these choices.

All of our data is for isolated clusters ranging in size from 8-55 particles. These configurations interact via a Lennard-Jones potential:

$$\phi(X) = \sum_{i<j} \phi_{ij}(r_{ij}) \text{ for } X \in \mathbb{R}^{3N_p}, \tag{2.1}$$

where the sum is over all pairs of N_p particles, r_{ij} is the distance between particles i and j, and the pair potential ϕ_{ij} is given by

$$\phi_{ij}(r_{ij}) = 4\varepsilon[(\frac{\sigma}{r_{ij}})^{12} - (\frac{\sigma}{r_{ij}})^6]. \tag{2.2}$$

The parameter σ is the distance at which the pair potential $\phi_{ij}(\sigma)$ is zero, and is related to the bond spacing that minimizes the pair potential, found at a distance of $2^{\frac{1}{6}}\sigma$. The parameter ε is the well-depth for the pair potential, i.e. $\phi_{ij}(2^{\frac{1}{6}}\sigma) = -\varepsilon$.

Building the catalog of transition states requires the repeated application of two basic algorithms: a means to *relax* the system to a local minimum along a path that stays in the current basin of attraction, i.e. by a strictly descending path, and an algorithm for finding a saddle point. For relaxation we will use the Polak-Ribie're Non-Linear Conjugate Gradient method [11] . To locate

saddle points, we use the well established and competitive Dimer Method introduced by Henkelman and Jonsson [2]. There exists a number of competing methods, including some improvements on the original Dimer Method that could also be used [3, 5, 7, 9, 10, 16–19]. Whereas steepest descent follows the force, $-\nabla\phi$, to a local minimum, the Dimer Method attempts to follow what is known as the *Householder Vector* to a saddle point. The Householder Vector,

$$F^\dagger = F - 2(F \cdot \hat{\eta})\hat{\eta}, \tag{2.3}$$

is constructed by first identifying the eigenvector, $\hat{\eta}$, corresponding to the smallest eigenvalue of the Hessian, and then reversing this component of the force. Intuitively, this guides the system to a saddle point by maximizing in the lowest curvature direction and minimizing in all other directions. The eigenvector $\hat{\eta}$ is also known as the *min-mode*, and, when at a critical point, corresponds to the direction of least curvature. More generally, it is the direction with the smallest second derivative. Unlike other methods that use the Householder Vector, the Dimer Method identifies the min-mode with the aid of a simple center difference approximation along a line in configuration space:

$$\hat{\eta} = \underset{\hat{\eta}}{\text{argmin}}\, C(\hat{\eta}), C(\hat{\eta}) \approx [\phi(X_2) - 2\phi(X_0) + \phi(X_1)]/\Delta x^2, \tag{2.4}$$

where (X_2, X_1) are the endpoints, referred to as the *dimer*, and X_0 is the center of the finite difference stencil.

Potentials are generally invariant under any translation or rotation, with the result that the Hessian will have six zero eigenvalues. We can eliminate this degeneracy by considering a constrained system that pins a particle to the origin, a second particle to a line passing through the origin, and a third particle to a plane containing this line. This has the additional advantage of reducing the total number of degrees of freedom by 6, so that $X \in \mathbb{R}^{3N_p-6}$. Wth these constraints, minima are stationary points where the reduced Hessian, $H \in \mathbb{R}^{(3N_p-6)\times(3N_p-6)}$, has strictly positive eigenvalues. In addition to making the computation slightly faster, this makes it easier to identify saddle configurations. For a large system, it is common to constrain a larger number of particles that lie outside an active region [6], relying on the assumption that these particles move a negligible amount between the binding state and the transition state. For smaller system sizes, twenty or fewer particles, we use minimal constraints, and for larger systems we use what we refer to as a *constrained core* consisting of a four-particle tetrahedral structure at the center of the cluster.

We assume that each initial guess determines a particular outcome, i.e. a specific saddle point (connected or otherwise) or a failure to converge within certain specified tolerances. Within the context of the Dimer Method this means a "guess" refers to both the initial configuration and the initial dimer orientation. These choices are made at random according to some some probability distribution, which we refer to as a *perturbation scheme*. We follow the scheme proposed in Ref. [12], where each atom in the initial minimum $X = \{\mathbf{x}_i \in \mathbb{R}^3\}_{i=1}^{N_p}$ is perturbed by a magnitude that decreases with increasing distance from a central atom j. This reflects the assumption that most transitions are localized around a single atom or group of atoms. The particular scheme is given by

$$x_i^k = x_i + \frac{k\sigma}{N}\hat{n}_i^k \exp\{-\|x_i - x_j\|_2\} \text{ for all } i = 1, 2, \dots N_p, \tag{2.5}$$

where N is the number of initial guesses, $k = 1, 2, \ldots, N$ is the the guess number, $\hat{n}_i^k \in \mathbb{R}^3$ is a random unit vector, and σ is the length scale parameter in the Lennard Jones potential.

Upon finding a saddle point, the configuration is requenched to ensure that it is connected to the initial minimum. The Hessian is then utilized to ensure that the saddle point is index-1, and higher indexed saddles are discarded. The saddle point is cataloged for future reference so we may check for duplicate saddle points.

The duplicate check is performed by calculating the distance between a newly found saddle point and previously cataloged saddle points. The saddle point is considered a duplicate if the magnitude of the difference in position falls below a specified threshold. Care is taken to ensure the threshold is not too small or too large. In the case where the threshold is too small, two saddle points that have nearly identical configurations would be counted as distinct, and therefore incorrectly inflate the number of saddle points. Whereas in the case where the threshold is too large, two saddle points that have significantly different configurations would be counted as the same saddle point, incorrectly yielding a smaller number of total saddle points. We find setting the threshold on the magnitude of the difference in the position coordinates to 0.05 works well. In Fig. 2, we demonstrate this behavior for a small number of searches for a nine-particle configuration. Notice that there is a range of values for which the criteria could be set to ensure that we do not inflate or deflate the number of connected saddles.
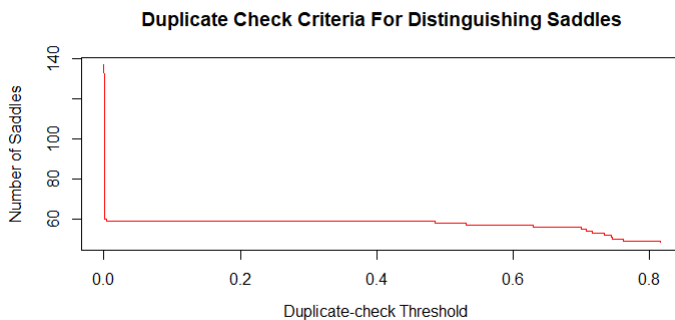

Figure 2: A plot of the total number of saddles found using a fixed number of searches and a nine-particle configuration as a function of the duplicate-check threshold. Note that there is a range of acceptable threshold values, including the 0.05 that we use for our calculations.

3 Basins of Attraction

In this section, we will briefly examine the basins of attraction for saddle points. While we are working within the full configuration space, $\mathbb{R}^{3N_p}$, we will be examining two dimensional cross sections of this space to gain a sense of what the basins look like. Wales [14] employed a similar approach to visualize the basins of attraction of transition states and minima on PES's. Massen and Doye [8] find that the basins of attraction provide a fractal-like tiling of the energy surface.

Our focus is specifically on the basins for the saddle points. For a fixed search algorithm, such as an implementation of the Dimer Method, each initial guess (defined to include the dimer orientation) produces a unique saddle point. For a given basin plot, we fix the initial dimer orientation, so that the plot will reveal how the starting configurations partition that slice of configuration space into domains of attraction for individual saddle points. Our aim is to understand why it is difficult to complete an exhaustive search for saddle points and to gain insight into potential improvements in the perturbation scheme. Ultimately, the distribution of initial guesses (perturbation scheme) combines with the search algorithm to produce a distribution of saddle points. We will use the distribution of the saddle points in Section 5 to gain an understanding of the slow convergence observed in the accumulation plot.

For illustrative purposes, we consider an eight-particle configuration that initially minimizes the Lennard-Jones potential. This will allow us to build up what we believe to be an exhaustive catalog of the connected saddle points, a task that is difficult for larger clusters. The accumulation plot for this configuration is presented in Fig. 3, where the horizontal axis represents the number of successful Dimer searches. We continue to search the PES until no new saddle points are found. Typically one does not know if all the saddle points have been found, and in most systems with larger than eight particles, complete convergence was elusive. For this configuration we were able to find 48 saddles connected to the starting configuration.
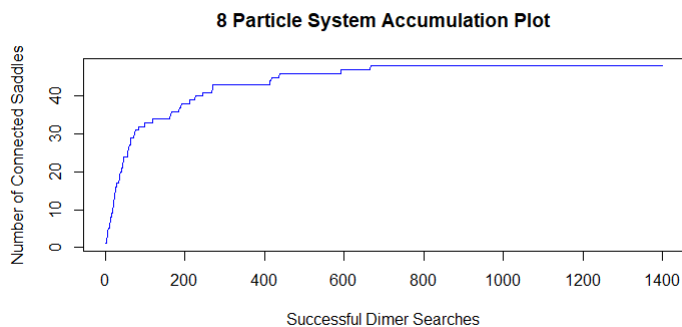

Figure 3: The number of connected saddles found as a function of the number of successful dimer searches for the eight-particle configuration described in the text. We were able to find 48 connected saddles for this configuration.

In order to illustrate the basins of attraction, a plane is chosen for the 2D cross section, and small steps within the plane are taken to create a grid (Fig. 4a). While each starting position is within this fixed plane and each dimer starts with the same orientation $\hat{\eta}$, the subsequent Dimer search takes place in the entire $3N_p-$dimensional configuration space. We outline this procedure in Algorithm 1. If the Dimer Method completes the search and converges to a connected index-1 saddle point, the saddle point is either assigned a new saddle index number i or assigned an index number based upon the catalog of previously found saddle points. Each saddle index is assigned a color to distinguish the distinct connected saddle points. Also, if the Dimer Method fails to

---

Algorithm 1: Exploring a 2D Cross Section

1: Choose a particle j and an initial configuration $X_0 \in \mathbb{R}^{3N_p}$ according to the perturbation scheme outlined in Section 2.

2: Choose a plane passing through X_0. In practice, we choose two of the three coordinates corresponding to particle j, labeling these coordinates x_j^1 and x_j^2.

3: Choose the starting dimer orientation $\hat{\eta}$, which is to be used at each initiation of the Dimer Method.

4: We explore the chosen 2D cross section by defining a grid of initial coordinates using multiples $\{m,n\}$ of a small distance δx, keeping all of the other coordinates fixed to the values of X_0:

$$x_j^1 = x_j^1 + m\delta x$$
$$x_j^2 = x_j^2 + n\delta x.$$

5: Initiate the Dimer Method for each configuration in the grid.

---

converge the point is assigned the color black, or, if it converges to a nonconnected saddle point, it is assigned the color green, as presented in Table 1.

Table 1: Legend to Signify Dimer Results.

Dimer Result	Color
Failed to Converge	⬛
Converged to Nonconnected Saddle Point	🟩
Converged to Connected Saddle Point	48 other colors

Twenty-four distinct saddle points were located using this cross section. Although we have assigned every saddle point a specific color, this cannot be seen due to the relatively large percentage of nonconnected saddles (green) and the relatively coarse resolution. In Fig. 4b, we show the basin map without the nonconnected saddles to better view the basins for the connected saddle points. We get an even better view in Figs. 4c and 4d, which are zoomed in versions of Figs. 4a and 4b. Examining these figures, we learn a few things. First, as is to be expected, perturbations that place the configuration a large distance away from the minimizing configuration are increasingly likely to fail to converge or, at best, converge to a nonconnected saddle point. Second, the boundaries of individual saddle basins are fractal in nature and not entirely connected, consistent with the observations of others [14]. This can be traced to the nonlinear, iterative process introduced by Newton's method, which is used to do the line searches needed to minimize various quantities. To the extent that there is a coherent pattern to the structures seen, this pattern will frequently be disjoint as it is interrupted by various abrupt changes in the search algorithm, e.g. a change
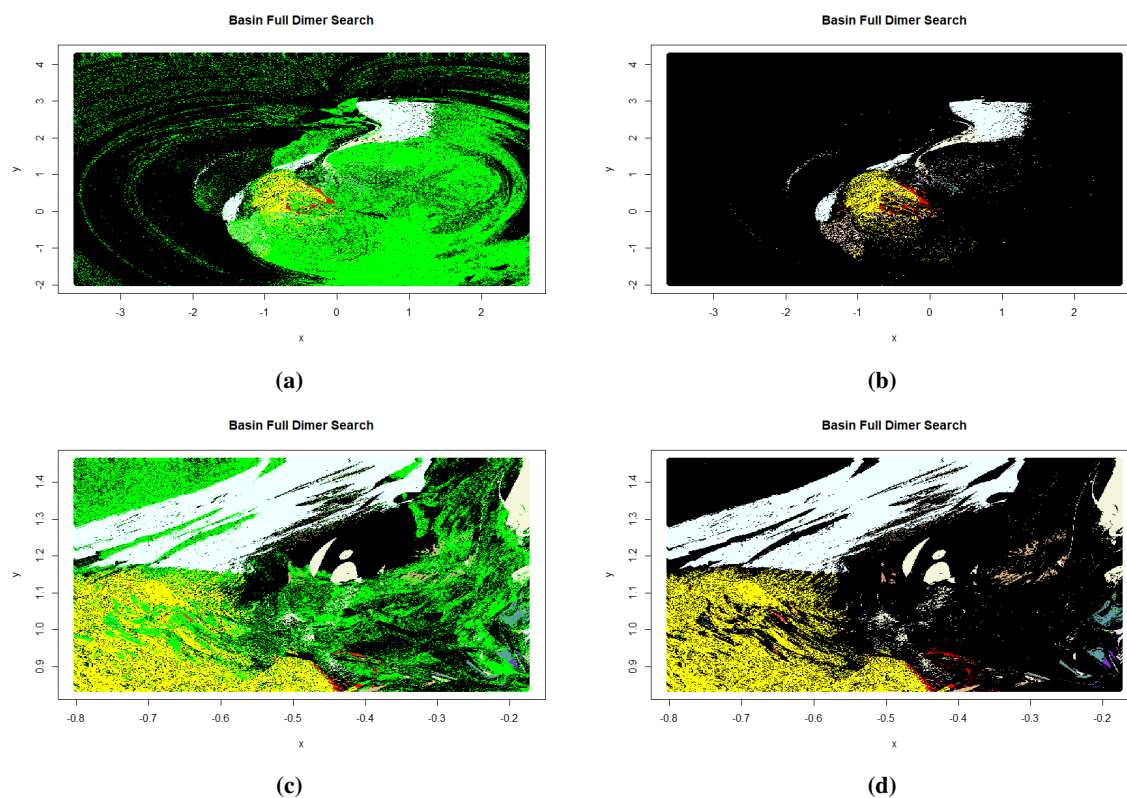
Figure 4: a) Basin plot for a two dimensional cross section of configuration space with initial configurations colored according to Table 2. b) The same plot with the non-connected saddles removed so that the connected saddles can be seen more clearly. c) A zoomed in version of figure (a). d) A zoomed in version of figure (b). There are 24 distinct connected saddles within the plane.

in search direction. Overall, for the dimer method, we find that perturbations that are especially close to the local minimum work well in that this region tends to be a chaotic blend of very small basins of attraction leading to many different saddle points.

While the details vary dramatically depending on the plane and particle chosen, the qualitative features just described are similar, as can be seen by examining the additional cross section shown in Fig 5. In Fig. 5b, we examine a zoomed in portion of this same cross section that is near the minimum, and again see the random nature for small perturbations about the minimum. This portion of this particular 2D cross section is especially rich in distinct saddle points, and we are able to find 41 of the 48 distinct saddle points observed in our exhaustive search. While this is somewhat unusual, it is nevertheless remarkable that it can occur at all. Placing constraints on the configuration to reduce the dimensions of the configuration space will also alter these plots, but, again, retaining qualitatively similar features.
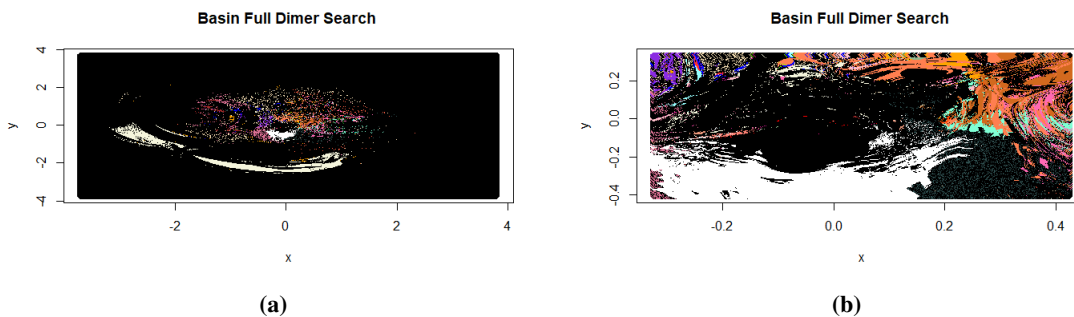

(a)   (b)

Figure 5: a) Basin plot for an additional two dimensional cross section of configuration space with initial configurations with the connected saddle points removed. b) A zoomed in version that contains 41 of the 48 saddle points found in the exhaustive search.

Table 2: Fraction of Successful Dimer Searches. The first two columns are for the portions of the 2D cross section examined in Figs. 4a and 4c,and the third column is for data collected on the full Configuration Space that is presented in Fig. 3.

Dimer Result	Fig. 4a	Fig. 4c	Full Space
Failed to Converge	0.5926	0.3550	0.2258
Converged to Nonconnected Saddle Point	0.3491	0.2911	0.4898
Converged to Connected Saddle Point	0.0583	0.3539	0.2845

In Table 2 we present the fraction of successful Dimer searches. The first two columns are for the portions of the 2D cross section examined in Figs. 4a and 4c, and the third column is for data collected on the full configuration space that is presented in Fig. 3. When searching the full configuration space, we perturbed each particle of our eight-particle system 500 times using the scheme outlined in Section 2 to generate 4000 total guesses. We tallied the number of connected and nonconnected saddle points to generate the data listed in Table 2. We note that for searches

within the plane, as we confine the search to be closer to the minimum, the portion of connected saddles increases. This is another indication that small perturbations about the minimum are best when searching for connected saddle points.

4 Distributions

When performing the global saddle point search of the PES for the eight-particle configuration in the last section we were able to obtain what we believe to be an exhaustive catalog of 48 connected saddle points. We increased the number of searches beyond that illustrated in Fig. 3 and kept track of the number of times each saddle point on the surface was found to produce the distribution found in Fig. 6a. In this figure, the index for each saddle point is based on the order in which it was encountered. This is why the distribution tends to be decreasing in magnitude, as we are much more likely to encounter an easily found saddle point early in the search process. It would be nice
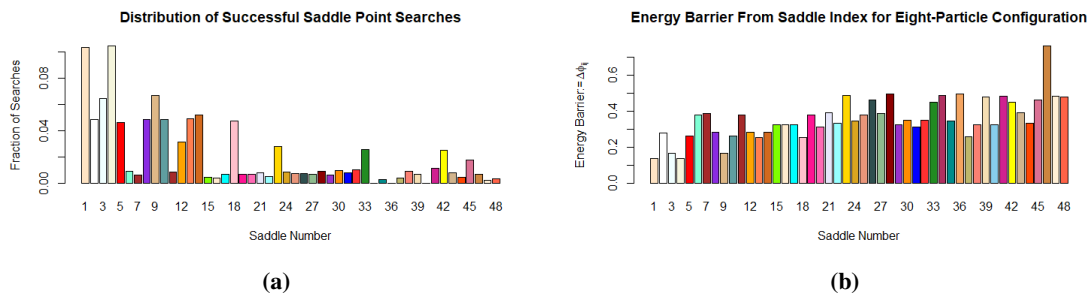

Figure 6: The distribution of successful saddle point searches for the 48 distinct connected saddles found for the eight-particle configuration, along with a plot of the corresponding energy barriers. The colors for the saddle index correspond to the color assigned in the previous section.

if the saddle points with large energy barriers, corresponding to rare and possibly less important events, were the difficult to find saddle points, but examining the corresponding plot of energy barriers presented in Fig. 6b, we see that this is not entirely the case. The extent to which this is true depends on the potential, the saddle-searching algorithm and the details of its implementation, including the perturbation scheme. Nevertheless, one cannot expect a strong correlation between these two plots unless one performs costly computations that locate saddle points via something closely imitating MD. Instead, we will try to gain a better understanding of what these distributions look like and how they affect the accumulation of an exhaustive list of connected saddle points.

Consider again Fig. 6a, and let M_i with $i = 1, 2, \ldots, N$ be the number of times saddle point i was found using g guesses. We will refer to the fraction of searches that yield saddle i as the *empirical distribution*:

$$p_i = \frac{M_i}{\sum_{i=1}^{N} M_i}, \tag{4.1}$$

$i = 1, 2, \ldots, N$ and note that $\sum_{i=1}^{N} p_i = 1$. From this we can imagine defining an ideal distribution

$$\bar{p}_i = \lim_{g \to \infty} p_i. \tag{4.2}$$

We now present data for a larger system of 55 particles, shown in Fig. 7. The four constrained particles near the center of the cluster are shown in red. As discussed earlier, these constraints are introduced to decrease the computational cost of performing the Dimer Method. This is an approximation that relies on the assumption that these interior particles move a negligible amount between the binding state and the transition state.
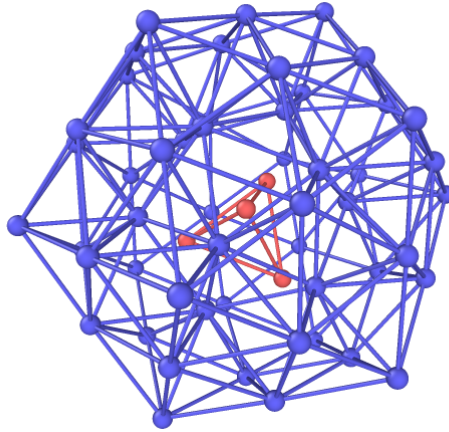

Figure 7: The 55 particle quenched system with a constrained core represented in red, and the 51 particles that are free to move during the Dimer Method searches shown in blue.

In Fig. 8, we present the saddle point distribution p_i for the 1358 connected saddle points that were found using 36,849 successful dimer searches. In Fig. 8a, the distribution is shown in the order in which the saddle point was encountered during the dimer searches, while in Fig. 8b, the distribution is sorted from greatest to lowest frequency.

We find that the data in Fig. 8b, the sorted distribution, is well represented by a Geometric Distribution:

$$p_i = \left(\frac{1-p}{1-p^N} \right) p^{i-1}, \tag{4.3}$$

where i is the saddle index, and we have normalized so that the sum over the first N values of p_i is one. Notice that this is a two parameter (p and N) family of normalized distributions. Keeping N fixed to the 1358 saddle points that were found, we find that a nonlinear least-squares fit of Eq. (4.3) to the data in Fig. 8b gives $p = 0.9949$. This is shown in Fig. 9. Note that p is close to one, a fact that can be understood by realizing that the p^N in the denominator of Eq. (4.3) is going to be negligible. As a result, the $1 - p$ in the numerator must be approximately equal to the first data point in the graph, which is a relatively small fraction as the number of saddle points grows large.
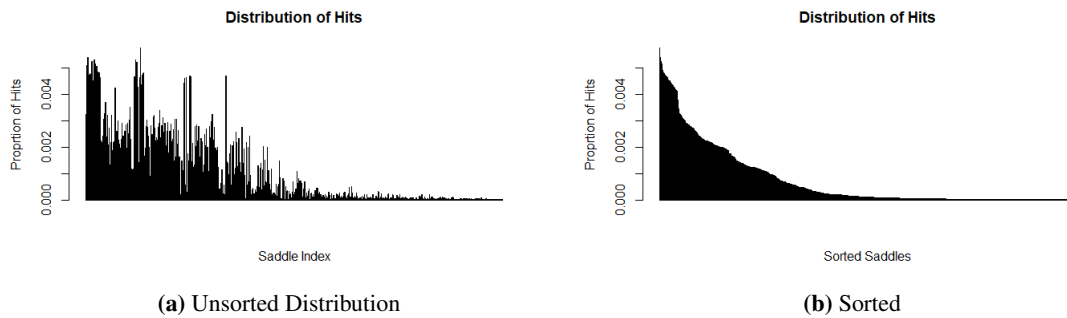
(a) Unsorted Distribution

(b) Sorted

Figure 8: The unsorted distribution of hits is presented in Fig. 8a for which the 1358 distinct connected saddles were found with 36,849 successful Dimer searches. These are presented in the order in which the saddles were encountered. Fig. 8b presents the distribution sorted from greatest frequency to lowest frequency.
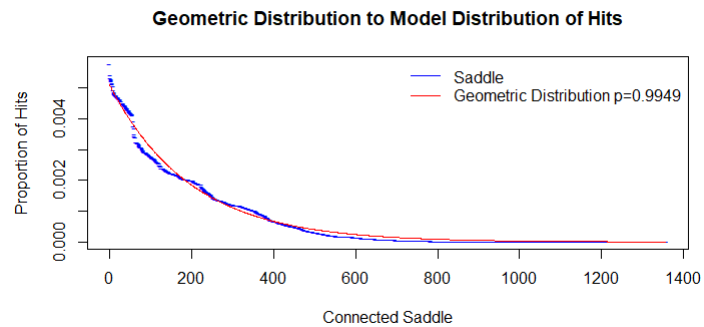

Figure 9: Nonlinear least-squares fit of the geometric distribution, Eq. (4.3), to the sorted empirical distribution shown in Fig. 8b. For this calculation, we kept $N = 1358$ and found $p = 0.9949$.

5   Accumulation Plots

Recall that the accumulation plot represents the number of connected saddles found for a given initial configuration as a function of the number of successful searches. Examining Fig. 1, it is not clear how close we are to converging to the total number of connected saddle points. We will model this behavior in an aim to help us understand why the growth rate of the number of connected saddles is so slow.

Let $[i_1, i_2, \ldots, i_g]$ be a sequence of saddle points recovered from g independent searches, where i_j is the saddle index assigned to the *jth* successful search. Note that the same saddle point could appear in this sequence multiple times. Let $c(n, g; N)$ represent the number of sequences that can yield n distinct saddles from g guesses when a total of N saddles are available to choose from. We have

$$\sum_{n=1}^{N} C(n, g; N) = N^g. \tag{5.1}$$

Let $P(n, g; N)$ be the probability of getting n distinct saddles using g guesses. For an arbitrary distribution $\{\bar{p}_i\}$, we have $\mathbb{P}([i_1, i_2, \ldots, i_g]) = \prod_{j=1}^{g} \bar{p}_{i_j}$ and

$$P(n, g; N) = \sum_{k=1}^{C(n, g; N)} \prod_{j=1}^{g} \bar{p}_{i_j}. \tag{5.2}$$

A recursive formula for computing the number of sequences that yield n distinct saddle points for a number of guesses g and a fixed N is

$$C(n, g+1; N) = nC(n, g; N) + (N - n + 1)C(n - 1, g; N). \tag{5.3}$$

This recursion formula reflects the fact that there are n ways to add one more saddle to each of the $C(n, g; N)$ sequences resulting from g guesses that yielded n distinct saddles, and $N - n + 1$ ways to add one more saddle to each of the $C(n - 1, g : N)$ sequences resulting from g guesses that yielded $n - 1$ distinct saddles.

The recursion formula Eq. (5.3) can be solved by introducing a related sequence $a_g^{(n)}$ in the form

$$C(n, g; N) = \frac{N!}{(N - n)!} a_g^{(n)}. \tag{5.4}$$

Directly substituting Eq. (5.4) into Eq. (5.3) yields

$$a_{g+1}^{(n)} = a_g^{(n)} \cdot n + a_g^{(n-1)}, \tag{5.5}$$

with the initialization of $a_n^{(n)} = 1$.

For $g \geq n$, the solution to Eq. (5.5) can be represented in the form :

$$a_g^{(n)} = \sum_{k=1}^{n} b_{n,k} k^g, \text{ with } a_n^{(n)} = 1, \tag{5.6}$$

where

$$b_{n,k} = \frac{(-1)^{n+k}}{k!(n-k)!}, \tag{5.7}$$

for $n, k \geq 1$.

Our main focus in this section is modeling the accumulation plot using the expected number of saddles of the distribution as

$$<n> = \sum_{n=1}^{N} nP(n,g;N), \tag{5.8}$$

for each fixed number of guesses g and fixed N.

We start by examining the special case of a uniform distribution $\{\bar{p}_i \equiv \bar{p} = \frac{1}{N}\}$. While this is not especially realistic, it can be analyzed more completely and provides insight into the general case.

Now, as we are taking independent guesses,

$$\mathbb{P}([i_1, i_2, ..., i_g]) = \prod_{i=1}^{g} \bar{p}_i = \bar{p}^g = \frac{1}{N^g}, \tag{5.9}$$

and

$$P(n,g;N) = \frac{C(n,g;N)}{N^g}. \tag{5.10}$$

The corresponding result for the Geometric Distribution Eq. (4.3) is

$$
\begin{aligned}
P(n,g;N) &= \sum_{k=1}^{C(n,g;N)} \prod_{j=1}^{g} \left(\frac{1-p}{1-p^N}\right) p^{i_j - 1} \\
&= \left[\left(\frac{1-p}{1-p^N}\right)^g\right] \sum_{k=1}^{C(n,g;N)} \sum_{j=1}^{g} p^{i_j - 1}.
\end{aligned}
\tag{5.11}
$$

In Fig. 10a, we use Eqs. (5.8) and (5.10) to compute the expected value of an accumulation curve for a uniform distribution of $N = 100$ objects/saddle points. The analogous calculation using Eq. (5.11) for the geometric distribution is uncomputable due to the need to enumerate and store the extremely large number $C(n,g;N)$ of distinct sequences. Similarly, it is not possible to carry out the calculation for the uniform case if the total number of objects N gets too large. For these cases, we turn to a Monte Carlo sampling of the distributions to simulate an accumulation process analogous to that in the actual saddle point search. Some curves comparing the outcomes for the uniform distribution and several cases of the geometric distribution are shown in Fig. 10b.

As a final way of examining the behavior of accumulation plots, we will return to the 55 particle configuration used in Fig 1. This accumulation plot is again shown (in blue) in Fig. 11. Recall that we were able to find $N = 1358$ connected saddles using just over $30,000$ independent searches. The red curves are realizations generated using a Monte Carlo sampling of the distribution Eq. (4.3) with the value of $p = 0.9949$ calculated for Fig. 9 and varying the value of N to obtain a best least squares fit. The resulting $N = 4660$ is still smaller than we believe the actual
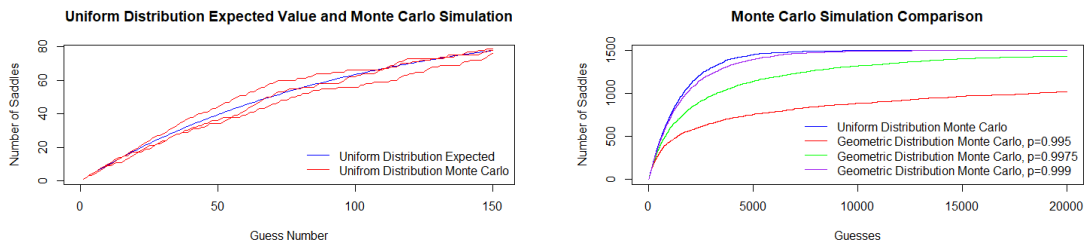
(a) Expected Value and Monte Carlo Simulations for the Uniform Distribution.

(b) Monte Carlo Simulations for the Uniform Distribution and several cases for the Geometric Distribution.

Figure 10: In Fig. 10a the blue curve is the expected number of saddles found as a function of successful search attempts, as modeled using Eq. (5.8) with $N = 100$. The corresponding red curves are generated using Monte Carlo sampling of the uniform distribution $\bar{p}_i = \frac{1}{N}$. In Fig. 10b, the blue curve is a Monte Carlo simulation with $N = 1,500$ saddle points for the uniform distribution (blue), and the remaining curves are simulations for several values of p in the geometric distribution Eq. (4.3).

total number of saddle points to be. For example, extending our search to 3,518,985 successful guesses, we have found 7,068 distinct connected saddles connected to the initial configuration. Our point here is not to estimate the ultimate number of saddle points, though that would be desirable, but merely to demonstrate that the slow accumulation behavior is to be expected, especially for a distribution that tails off quickly.

6 Early Termination

Recall that in the rejection scheme introduced by Ruzayqat and Schulze [12], transitions from the current minimum configuration are partitioned into N_p mutually exclusive subsets by associating each saddle point with a particular particle. This is done by assigning a saddle point to the particle that moves the most when the system is moved from the minimum to the transition state. The perturbation scheme discussed in Section 2 allows us to control the particle that initially moves the most. Once a transition state is found, it is requenched to check that it is a connected saddle. For the connected saddle, if the particle that was initially perturbed is the one that has moved the greatest distance from the initial configuration we then classify the state as a key connected saddle. If it were possible to ensure that the perturbation along with the Dimer search led us to a key connected saddle, a lot of valuable time could be saved and reinvested into more Dimer searches to obtain a more comprehensive list of saddles. To this aim, we will implement an early termination technique, where we monitor the distances particles have traveled from the initial minimum as the dimer algorithm works its way to a saddle point. If the initially perturbed particle no longer holds the status of being the particle that has moved the greatest distance, the Dimer search is terminated, and the next search begins.

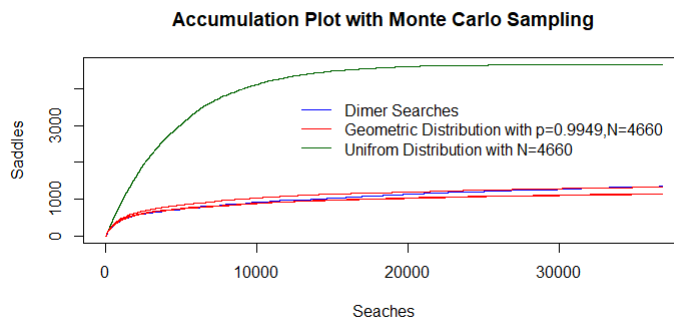We will illustrate this using the same 55 particle system used in previous sections. In our

Accumulation Plot with Monte Carlo Sampling

Figure 11: The blue curve is the accumulation plot originally shown in Fig. 1 , where 1,358 saddle points had been located using over 30,000 successful searches. The red curves are a couple of realizations of the Monte Carlo simulations using the geometric distribution with $p = 0.9948$ fit to the empirical distribution shown in Fig. 9 , with a value of $N = 4,660$ chosen to give the best fit. The green curve, shown for comparison, is Monte Carlo simulation with the uniform distribution and this same value of N.

examination of the basins of attraction for the saddle points, we concluded that small perturbations about the minima were the most effective for exploring the PES for the saddles. This was due to the chaotic nature of the basins of attraction. In Fig. 12, we present the distance each unconstrained particle in the configuration moved as a function of the Dimer iteration until convergence to a saddle point was observed. Recall that the perturbation scheme gives control over the particle that has initially moved the greatest distance, and a key connected saddle is a transition state where the initially perturbed particle retains the status of being the particle that has traveled the greatest distance from the initial configuration. In Fig. 12a and Fig. 12b, we present cases where the particle that was initially perturbed the most, remained the particle that moved the most, resulting in a key connected saddle point. It is often the case that the initial perturbation coincides with the key atom, but this is not always the case, as seen in Fig. 12c and Fig. 12d.

In view of these observations, we introduce an early termination procedure, where we monitor the progress of the search every 100 iterations to see if the initially perturbed particle remains the particle that moved the most. If this is the case, the Dimer search continues. If not, we exit the current search and start again. If one chooses to use more or less iterations for the early termination check, then one finds somewhat less or more saddle points, respectively.

In Table 3, we tabulate the number of searches per particle, connected saddles found, key connected saddles found, and key connected saddles found with the early termination procedure. These are independent runs as we increase the number of searches per unconstrained particle for the Dimer Method. It is clear that we are able to get close to the number of key connected saddles found using a full, non-terminated search, but we lose some as the number of searches increases. This loss is more than made up for by the computation time that is saved, as this can be reinvested into additional searches.
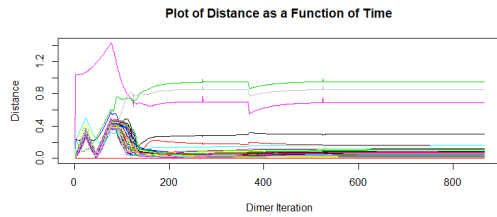
In Table 4, we give the ratio of key connected saddle points to connected saddle points and that
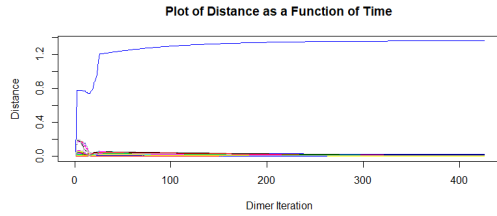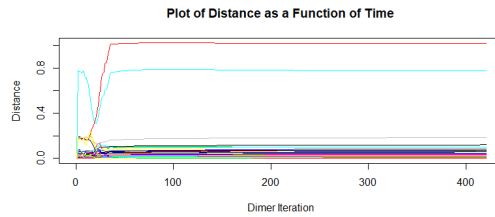
(a) Particle 33 is the initially perturbed particle.
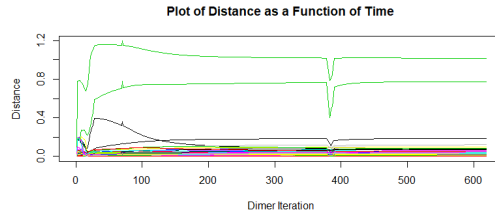
(b) Particle 26 is the initially perturbed particle.

(c) Particle 27 initially perturbed.

(d) Particle 46 initially perturbed.

Figure 12: In Fig. 12a and Figure 12b, we present cases in which the resulting connected saddle is a key connected saddle, while in Fig. 12c and Fig. 12d the resulting connected saddle did not maintain the key connected status.

Table 3: Table to compare the counts of Key connected saddles to connected saddles and the Key connected saddles caught with early termination to the total number of Connected saddles. Results are given for several, successively larger, searches, as indicated by the number of guesses per particle.

Searches per Unconstrained Particle	Connected Saddles	Keyed (Regular)	Early Termination
10	159	127	127
50	416	327	320
250	712	516	505
500	890	611	588

of key connected saddle points caught by early termination to the connected saddle points. Again, notice that we lose some key connected saddles as early termination is implemented, but the ratio does not differ substantially.

Table 4: Table to compare the proportion of Key connected saddles to connected saddles and the Key connected saddles caught with early termination to the total number of Connected saddles.

Searches/Particle	Proportion Key/connected	ET/connected
10	.7874	.7874
50	.7861	.7692
250	.7247	.7093
500	.6865	.6607

Finally, we compare the performance of the rejection-based OLKMC with and without early termination. For this comparison we used 100 guesses per particle to perform saddle point searches focussed on each of the 51 unconstrained particles, and found a CPU time of 2,661 seconds for the scheme without early termination and 946 seconds for the scheme with early termination. For these simulations, the early termination procedure was approximately 2.5 times faster.

7 Conclusions

This paper is broadly motivated by OLKMC and the challenging task of finding all of the saddle points connected to a given local minimum of a high-dimensional energy landscape. We began with the observation, illustrated in Fig. 1 by what we refer to as an accumulation plot, that this goal often seems unattainable—there are too many saddle points and progress in finding them diminishes as more saddle points are found. Throughout the paper, we used freely suspended Lennard-Jones clusters with between 8 and 55 particles. It may be that this is a somewhat more challenging configuration than a free-surface environment with a constrained underlayer, which is often considered in the OLKMC literature. Nevertheless it seems that this slow accumulation and explosive growth of the number of located saddle points is a fundamental challenge that must be overcome if OLKMC is to become more broadly applicable.

While a variety of saddle point finding methods may be used, they share the common feature that one must equip them with a set of initial guesses for the configurations. The guesses are crafted so that 1) they yield a high percentage of connected saddle points, and 2) fully explore the energy landscape in the vicinity of a given local minimum. Usually this features a random choice that can be thought of as a probability density $\rho(X)$ of points in configuration space. In Section 2, we refer to this as the perturbation scheme, and discuss a specific example of such a scheme that we use throughout the paper [12].

In Section 3 we map some two-dimensional cross sections of these initial guesses, identifying the basins of attraction for specific saddle points. While these observations are not new [14], our aim is to emphasize a few features of these searches. First, the way the located saddle point is connected to the initial guess is chaotic with a fractal-like structure outside of a relatively small

region surrounding each saddle point. This behavior can, for the most part, be traced back to the nonlinear iterative process inherent to the Newton's method that is employed in the various line searches needed to locate saddle points. Our observation was that this was especially true in the immediate vicinity of the local minimum being explored, and that this was actually a good region to probe for saddle points as it satisfies the two goals identified in the previous paragraph.

In Section 4 we discuss the distribution p_i of saddle points that is found due to the combination of search algorithm and perturbation scheme. We observe that for the system we explored here and the algorithm choices we made, this distribution is exponentially decreasing, so that it is well approximated by a geometric distribution. The rapidly decaying tail of this distribution contributes to the slow progress observed in the accumulation plots. Further study is required to determine how general this behavior is. For example, it may be that one could get a flatter distribution by combining multiple saddle search methods and perturbation schemes.

The distribution $\bar{p}_i$ discussed in Section 4 can be related, in principle, to the basins of attraction introduced in Section 3 and the perturbation scheme discussed in Section 2. Let $\Omega_i = \{$perturbations $X \in \mathbb{R}^{3N_p}$ that leads to saddle $i\}$ with $\Omega = \bigcup_{i=1}^{N} \Omega_i$ being the set of initial guesses for which the search converges to a connected saddle point. For any two distinct Ω_i's, notice $\Omega_i \cap \Omega_j = \{\varnothing\}$, that is they are mutually exclusive and partition Ω. Conceptually, the perturbation scheme $\rho(X)$ and the basins of attraction Ω_i combine to give the probability, $\bar{p}_i$, that an initial guess leads to saddle point i:

$$\bar{p}_i = \frac{\int_{\Omega_i} \rho(X) dX}{\int_{\Omega} \rho(X) dX}. \tag{7.1}$$

In section 5 we examine the effect of the distribution p_i on the behavior of the accumulation plot. While we were able to make some analytic progress in the special case of a uniform distribution, this was mostly done using Monte Carlo sampling. Our main objective in this section was to demonstrate that the slow accumulation behavior observed in Fig. 1 is an inherent feature of the search process, especially in cases where the distribution $\bar{p}_i$ has an exponentially diminishing tail. While it appears difficult to get an accurate result, this procedure can be used to get a crude estimate of the number of saddles left to be located, and this may be a promising avenue for further study.

In Section 6 we consider a type of local saddle point search introduced by Ruzayqat and Schulze[12] aimed at finding just the subset of saddle points associated with a particular particle, e.g. the particle that moves the most in the transition from the binding to the transition state. If such a search could be done efficiently, it allows a significant improvement of the OLKMC algorithm through the use of a rejection scheme. However, as illustrated in Fig. 1, the local search suffers from the slow accumulation problem to an even greater extent than the global search, presenting a serious challenge to this type of algorithm. Our final observation is that this can be substantially mitigated through the use of what we refer to as an early termination procedure, where one simply monitors the motion of the key particle as the search proceeds and ends the search if another particle has at some point moved further. We found that this more than doubled the efficiency of the rejection OLKMC algorithm.

Acknowledgments

References

[1] Eyring, Henry. The Activated Complex in Chemical Reactions." *J. Chem. Phys*, 3.2 (1935): 107–115. 1935.

[2] Graeme Henkelman and Hannes Jónsson. A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. *J. Chem. Phys*, 111(15):7010–7022, 1999.

[3] Graeme Henkelman, Gísli Jóhannesson, and Hannes Jónsson. Methods for finding saddle points and minimum energy paths. In *Theoretical methods in condensed phase chemistry*, pages 269–302. Springer, 2002.

[4] Graeme Henkelman and Hannes Jónsson. Long time scale kinetic monte carlo simulations without lattice approximation and predefined event table. *J. Chem. Phys*, 115(21):9657–9666, 2001.

[5] Andreas Heyden, Alexis T. Bell, and Frerich J. Keil. Efficient methods for finding transition states in chemical reactions: Comparison of improved dimer method and partitioned rational function optimization method. *J. Chem. Phys*, 123(22), 2005.

[6] S. Kadkhodaei and A. van de Walle. A simple local expression for the prefactor in transition state theory. *J. Chem. Phys*, 150(14):144105, 2019.

[7] Johannes Kästner and Paul Sherwood. Superlinearly converging dimer method for transition state search. *J. Chem. Phys*, 128(1):014106, 2008.

[8] Claire P Massen and Jonathan PK Doye. Power-law distributions for the areas of the basins of attraction on a potential energy landscape. *Phys. Rev. E*, 75(3):037101, 2007.

[9] RA Olsen, GJ Kroes, G Henkelman, A Arnaldsson, and H Jónsson. Comparison of methods for finding saddle points without knowledge of the final states. *J. Chem. Phys*, 121(20):9776–9792, 2004.

[10] Vanden-Eijnden Ren. A Climbing String Method for Saddle Point Search. *J. Chem. Phys*, 138.13: 134105–134105, 2013.

[11] William H Press, Brian P Flannery, Saul A Teukolsky, William T Vetterling, et al. Numerical recipes, 1989.

[12] Hamza M Ruzayqat and Tim P Schulze. A rejection scheme for off-lattice kinetic monte carlo simulation. *J. Chem. Theory Comput.*, 14(1):48–54, 2017.

[13] George H Vineyard. Frequency factors and isotope effects in solid state rate processes. *J Phys Chem Solids*, 3(1-2):121–127, 1957.

[14] David J Wales. Basins of attraction for stationary points on a potential-energy surface. *J. Chem. Soc., Faraday Trans.*, 88(5):653–657, 1992.

[15] David J Wales and Jonathan PK Doye. Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms. *J. Phys. Chem. A*, 101(28):5111–5116, 1997.

[16] Penglishhao Xiao, Qiliang Wu, and Graeme Henkelman. Basin constrained κ-dimer method for saddle point finding. *J. Chem. Phys*, 141(16):164111, 2014.

[17] Bing Yu and Lei Zhang Global optimization-based dimer method for finding saddle points. *Discrete & Continuous Dynamical Systems-B* 22.11, 2011

[18] Yi Zeng, Penghao Xiao, and Graeme Henkelman. Unification of algorithms for minimum mode optimization. *J. Chem. Phys*, 140(4):044115, 2014.

[19] E. Zhou The Gentlest Ascent Dynamics. *Nonlinearity*, 24.6: 1831-1842, 2011