Graduate School of Life Science
Hokkaido University

# Development on Methods to Decipher Timescale Hierarchy in Reaction Networks

（化学反応ネットワークに埋め込まれた時間階層構造の解読方法の開発）

*Yutaka Nagahata*

supervised by
Dr. Tamiki Komatsuzaki

March 2016

# Preface

In 1969, a research workshop named "Mőssbauer Spectroscopy in Biological Systems" was held in Robert Allerton Park in Illinois. In the proceeding of the workshop, Cyrus Levinthal pointed out the following paradox[1]. In protein conformations, "*there are indeed certain structural restrictions in a polypeptide structure.*" For example, some types of bonds restrict adjacent atoms to trigonal planar, tetrahedral or others. Under the restrictions on bond angles,

> *Considerations of such factors allow us to predict only 450 degrees of freedom in a protein structure containing 150 amino acids, for example. Of these 450 degrees of freedom, 300 would be due to rotations and 150 would be due to relative bond angles of the side chains*

> *Even if we knew these angles to better than a tenth of a radian, there would be* $10^{300}$ *possible configurations in our theoretical protein. In nature, proteins apparently do not sample all of these possible configurations.*

This is a well-known paradox named Levinthal's paradox. On the other hand, he also mentions the possible energy profiles in a folding process as follows.

> *We feel that protein folding is speeded and guided by the rapid formation of local interactions which then determine the further folding of the peptide. This suggests local amino acid sequences which form stable interactions and serve as nucleation points in the folding process. Then, is the final conformation necessarily the one of lowest free energy? We do not feel that it has to be. It obviously must be a metastable state which is in a sufficiently deep energy well to survive possible perturbations in a biological system. If it is the lowest energy state, we feel it must be the result of biological evolution; i.e., the first deep metastable trough reached during evolution happened to be the lowest energy state.*

The concrete hypothesis for "*the result of biological evolution*" was mentioned in 1980s by Go's "consistency principle" [2] and sophisticated by Bryngelson and Wolynes's "the principle of minimum frustration" [3]. Levinthal also mentioned about the relation between folding and amino acid sequence.

> *You may then ask the question, "Is a unique folding necessary for any random 150-amino acid sequence?" and I would answer, "Probably not." Some experimental support for this statement comes from the difficulty many of me are all too aware of in attempting to crystallize peptides."*

In 1990s, protein energy landscape resulting from "biological evolution" is elucidated in model systems. In 1992, Onuchic and his coworkers eluci-

dated "folding funnel" in a lattice protein model, which consists of 27 bead chain expected to form a $3 \times 3 \times 3$ lattice. Onuchic and his coworkers randomly gave contact energies, conducted Monte Carlo simulation, and saw funnel-like energy profile along folding process. In 1991 and 1994 [4, 5] Shakhnovich, Karplus and their coworkers systematically investigated the model and its folding processes in Monte Carlo simulation. In the paper published in 1994[4], they investigate 200 model proteins with randomly given contact energies. 30 models were folded within $5 \times 10^7$ Monte Carlo steps and each of folded model had a pronounced energy minimum as a native ($3 \times 3 \times 3$ cubic) state. They also introduced a reaction coordinate of the folding process as numbers of native contacts. Along the reaction coordinate, the funnel shape appears again. In 1995, Onuchic, Wolynes, and their coworkers have discussed the relationship between folding funnel and configuration (Levinthal's) entropy $S_{\mathrm{L}} = k_{\mathrm{B}} \ln(N(E))$ [6], where $k_{\mathrm{B}}$, and $N(E)$ are Boltzmann constant, and the number of states to be searched, respectively. They visualized a schematic folding funnel for a more realistic 60 amino acid three letter codes model (i.e., a protein model consist from three kinds of beads) corresponding to a helical protein. The funnel expresses the molten globule (i.e. compact and partially folded conformations) region, the transition state ensemble (a transition state region representing an ensemble of structures that acts as boundaries of folding/unfolding reactions), and the locally glassy (i.e., no pronounced energy minimum) region.

The studies of folding funnel help us to understand how protein-like model folds. However in reality, a protein has an energy surface in space with gigantic degrees of freedom (dimension), so that one needs to use a method

to project the energy landscape in very high dimensional space into one or two dimensional space.

In 1997, Becker and Kerplus introduced, so called, Disconnectivity graph(DG) [7] to visualize the characteristic topography of a potential energy surfaces. Since the potential energy surface can be derived by *ab initio* (quantum chemical, electronic state) computation, the graph can reflect the energy profiles of molecular conformations. The method was applied to various systems, such as not only small proteins (e.g. GB1: B1 domain of protein G[8], beta3s: the 20-residue miniprotein[9], three-color 46-bead model (BLN))[10], but also molecular clusters[11], cluster molecules[11], and some mesoscopic systems consisting of rigid building blocks[12, 13].

The story of this thesis starts from describing the details of DG and its variants. I propose another DG, which shed light on another feature of an energy landscape of molecule(s).

# Acknowledgement

I would like to express my special appreciation and thanks to my advisor Professor Dr. Tamiki Komatsuzaki, you have been a tremendous mentor for me. I would like to thank you for encouraging my research and for allowing me to grow as a research scientist. Your advice on both research as well as on my career have been priceless.

I would also like to thank my collaborators professor Tetsuya Taketsugu, associate professor Shinnosuke Kawai, associate professor Satoshi Maeda, associate professor Takashi Horiyama, associate professor Chun-Biu Li, assistant professor Hiroshi Teramoto, and Mr. Yousuke Sumiya, for fruitful discussion with them and their kind support. I would also like to thank my co-workers professor Shin-ichi Minato, associate professor Yoshinobu Kawahara, research assistant professor James N. Taylor, assistant professor Masaru Kobayashi, and Dr. Akinori Baba for their variable comments.

A special thanks to my family. Words cannot express how grateful I am to my mother, and my late father for all of the sacrifices that you've made on my behalf. Your prayer for me was what sustained me thus far. I would also like to thank all of my friends who supported me.

# Author's Declaration

I, *Yutaka Nagahata*, declare that the PhD thesis entitled "*Development on Methods to Decipher Timescale Hierarchy in Reaction Networks*" has been compiled by me for *Degree of Philosophy in Life Science* under the supervision of *Dr. Tamiki Komatsuzaki, the graduate school of Life Science, Hokkaido University.*

This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work.

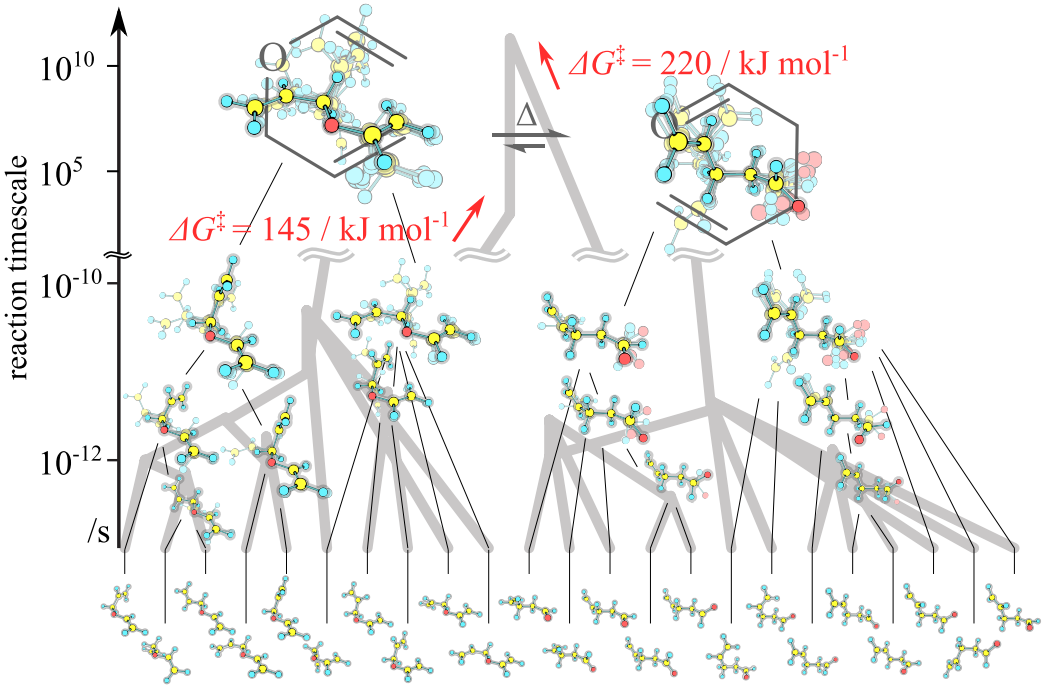
# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction




Markovian dynamics on complex reaction networks [14] are one of the most intriguing subjects in a wide range of scientific and engineering research

topics such as chemical reaction, protein folding and functions, metabolic networks, pharmacokinetics, epidemics, ecology, social networks, social networks, neural networks, evolutionary game theory networks, and petri nets as listed in a review 14. In addition to these topics, various dynamical systems has been approximated by Markov chain such as standard dynamical systems [15–17] (e.g., Henon map [15], Chua circuit [15], Lorentz attractor [16], Ikeda attractor [16], and standard map [17]), molecular dynamics [18–21], celestial mechanics and mission design [22, 23], epidemics (dynamical systems point of view) [24–27], neural dynamics [28, 29] by discretizing (i.e., approximating with discrete spectrum and its eigenfunctions) Perron-Frobenius Operator (e.g., Forkker-Plank Operator, Liouville Operator).

Especially in molecular dynamics, recent developments of experimental technologies and drastic improvement of computer performance elucidate more complex but more accurate nature of large molecules. These developments and improvements allow me to obtain detailed descriptions of rate equations[30, 31] or an energy surface which dynamics of a molecule follows [32–35]. However, its gigantic degree-of-freedoms prevent me from making a bridge between experimental data and theories of molecular dynamics. Energy landscape[33, 36] is one of the versatile concepts to understand feature of complex, global energy profile of molecule(s) and has been discussed in wide range of model systems: proteins [8–10], molecular clusters[11], atomic clusters[11], and some mesoscopic systems consisting of rigid building blocks [12, 13]. One of the most popular techniques to visualize the topographical features underlying energy landscapes is the disconnectivity graph (DG) [7, 11, 33, 37–40].

The original DG is based on the information of the sequence of minimum-saddle-minimum-$\cdots$ on the underlying multi-dimensional potential energy surface, and captures "disconnected" components (energetically inaccessible regions at a given total energy $E$ [7]). The component can be extracted by the list of saddle point energies $E_{ij}^{\ddagger}$ between basin $i$ and basin $j$. If $E_{ij}^{\ddagger}$ is less than $E$, then basins $i$ and $j$ are considered to be connected (energetically mutually accessible). The algorithm assigns basins $i$ and $j$ to the same superbasin, if $E$ is larger than $E_{ij}^{\ddagger}$. Free energy DGs were also proposed for free energy landscapes in which energy basins are grouped together if the free energies of activation are less than either thermal energy [7] or a certain threshold for free energies of activation of a pair of basins to be grouped or unified [9, 41].

To incorporate multiple pathways, the transition disconnectivity graph (TRDG) was developed to transform a reaction network with the equilibrium distribution into an energy landscape, which is based on *rates* across the network. [39] TRDG searches for a set of dividing surfaces (termed "cuts" in graph theory [42]) that form boundaries of two disjoint subnetworks over a given network. Each dividing surface is a bottleneck of gross rate (equilibrium flow) between a pair of nodes [43, 44](i.e., a node in one subnetwork and the other node in another subnetwork). The TRDG visualizes hierarchical relation (Lemma 8.35 in Ref. 44) of the dividing surfaces defined for all pairs of nodes. There exist cases in which the dividing surface minimizing the gross rate partitions a single node out of all nodes.[45, 46] The TRDG procedure includes a protocol to overcome this issue [39, 45, 46] but this protocol is only effective under certain limited conditions.

In this thesis, an another DG including a new clustering method (incor-

porating network transition state) and visualization technique is introduced. The method visualizes hierarchical organization of reaction timescales which provides more direct intuition in experiment than profiles of an energy surface, which has been visualized by DGs. The method is demonstrated by using a reaction network of an organic reaction derived from *ab initio* (quantum chemical, electric state) computation: the reaction network of Claisen rearrangement of allyl vinyl ether (AVE) consisting of 23 nodes and 66 links (saddles on the energy landscape) connecting them. The method obtains accurate overall rate constant between 10 elementary states (reactant superbasin) and 13 elementary states (product superbasin). I compare the kinetic properties of my disconnectivity graph to those of the transition matrix of the rate equations. My graph can properly reveal the hierarchical organization of timescales in a network and indicate existence of stiff substructures within each superbasin.

In chapter 2, the method proposed in this thesis will briefly explained. In chapter 3, basic theories, methods, and mathematics for the developed method are explained in detail. In chapter 4, the method I developed is explained. In chapter 5, the developed method is demonstrated on five state and three state illustrative Markov chains. In chapter 6, the developed method is applied to a reaction network derives from *ab initio* computation, and how the method deciphers timescale nature of reaction networks is shown. Finally, I conclude in chapter 7.

# Chapter 2

# Abstract of Proposed Methods

Let me first explain my approach briefly. Suppose that a reaction network is composed of nine nodes $V = \{1, 2, \ldots, 9\}$ corresponding to 9 different chemical species, with each reaction associated with a rate constant $k_{ij}$ (i.e., $k_{ii} = 0$)between the nodes $i, j \in V$ (see Fig. 2.1(a)). For each $\forall i \in V$, one can define first-order rate equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{X}_i] = \sum_{j \in V} \left( k_{ij}[\mathrm{X}_j] - k_{ji}[\mathrm{X}_i] \right),$$

where $[\mathrm{X}_i]$ denotes the density of $i$th species $(i = 1, \ldots, n)$.

By using the solution of the equations, I will introduce Timescale of Reaction, Network Transition States. Then, I propose Hierarchical Organization of Network TS and visualization method of it.

## 2.1 Timescale of Reaction

Let me show relation between rate constants and timescale to react. The escaping process from a species $i$ is described by

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{X}_i] = \sum_{j \in V} -k_{ji}[\mathrm{X}_i],$$

The solution of this equation is $[\mathrm{X}_i](\Delta t) = [\mathrm{X}_i](0)\exp[-\Delta t \sum_{j\in V} k_{ji}]$. From the solution one can derive probability to be remain in the spices $i$ at $\Delta t$ is $e^{-k_i \Delta t}$, where $k_i := \sum_{j\in V} k_{ji}$. Hence, the escaping probability within time $\Delta t$ is

$$F_{\mathrm{escape}}(t) := 1 - e^{-k_i \Delta t},$$

so that the expectation time to escape is

$$\int_0^\infty t \, \frac{\mathrm{d}}{\mathrm{d}t} F_{\mathrm{escape}}(t) dt = 1/k_i.$$

Therefore $1/k_i$ is expectation time to react from the species $i$.

## 2.2 Network Transition State

When one is only capable of observing longer timescales than elementary reactions, it is expected that one observes coarse-grained reaction network. To visualize hierarchical relation of coarse-grained networks over a observation timescale, one should find sets of species which are indistinguishable in a timescale. Such coarse-graining will be achieved when overall reactions have slower timescales than that of equilibration in coarse-grained states. It seems

that overall rate constant relate to such coarse-graining, however, there are unstable species that are difficult to visit but easy to return back.

Let me introduce overall rate constants of the forward $k_{\mathrm{f}}$ and backward $k_{\mathrm{b}}$ reactions. Here, I seek the cut across the network satisfying that the larger overall rate constant of forward and backward reactions is the minimum over all the possible cuts: $\min \max\{k_{\mathrm{f}}, k_{\mathrm{b}}\}$. Let me say this is the longest reaction timescale of reaction over all the possible cut.

## 2.3 Hierarchical Organization of Network TS

Let me show the cut schematically in Fig. 2.1. Let me say that the resultant disjoint subnetworks by the cut: R and P (i.e., R $= \{1, 2, 3, 4, 5\}$ and P $= \{6, 7, 8, 9\}$). Hereafter I use the term 'superbasin' S to refer to a subnetwork (i.e. a set of nodes and links connecting them) whose total number of nodes is more than one.

Next, I seek to further divide the superbasins R and P into smaller superbasins individually. Let the resulting superbasins in R be $\mathrm{R}_\alpha (= \{1, 2, 3\})$ and $\mathrm{R}_\beta (= \{4, 5\})$; let the timescale of reactions from $\mathrm{R}_\alpha$ or $\mathrm{R}_\beta$ to P are still $1/k_{\mathrm{R}}$. Until the resulting superbasins cannot be divided further into smaller basins (i.e., single basins), I proceed to divide each superbasin into smaller ones hierarchically. I can visualize such hierarchical organization of timescales existing in networks as in Fig. 2.1-(b).

When I refer timescales of reactions, i.e., the inverse of overall rate constants, I instead have two values between R and P, that is, $1/k_{\mathrm{R}}$ for the forward reaction and $1/k_{\mathrm{P}}$ for the backward reaction. Here I introduce 'a

pair of points' to represent $1/k_{\mathrm{R}}$ and $1/k_{\mathrm{P}}$ that are connected with each other vertically so that the high (low) point is $\max\{1/k_{\mathrm{R}}, 1/k_{\mathrm{P}}\}$ ($\min\{1/k_{\mathrm{R}}, 1/k_{\mathrm{P}}\}$) at the same position along the horizontal axis as in Fig. 2.1-(b). I term such a pair of points as 'vertex pair' or 'TS' hereinafter.

(Super)basins with shorter (longer) reaction timescales are placed to the left (right) of the vertex pair along the horizontal axis in Fig. 2.1-(b). In this example, because $1/k_{\mathrm{R}} < 1/k_{\mathrm{P}}$, the superbasin R (P) is placed to the left (right) from the TS with the slowest timescale. For shorter timescale, the superbasins are further decomposed such that, e.g., R is decomposed into $\mathrm{R}_\alpha = \{1, 2, 3\}$ and $\mathrm{R}_\beta = \{4, 5\}$. Similarly, $\mathrm{R}_\alpha$ ($\mathrm{R}_\beta$) are placed to the left (right) from the new TS, because $1/k_{\mathrm{R}_\alpha} < 1/k_{\mathrm{R}_\beta}$. To clarify from which "parent" superbasin "children" (super)basins originate, children are connected to their parent so that one point in the parent vertex pair of $\min\{1/k_{\mathrm{R}_{\mathrm{parent}}}, 1/k_{\mathrm{P}_{\mathrm{parent}}}\}$ (in this example, $1/k_{\mathrm{R}}$) is connected to another point in the child vertex pair of $\min\{1/k_{\mathrm{R}_{\mathrm{child}}}, 1/k_{\mathrm{P}_{\mathrm{child}}}\}$ (in this example, $1/k_{\mathrm{R}_\alpha}$). Likewise, this procedure is continued until each of the resulting superbasins reach to a single basin. One can see that the order (or index) of basins (i.e., nodes) is, hence, determined so as to reflect such hierarchical nature of timescales in the network. Importantly, the tree graph structure thus constructed can tell us in what timescale each reaction occurs.

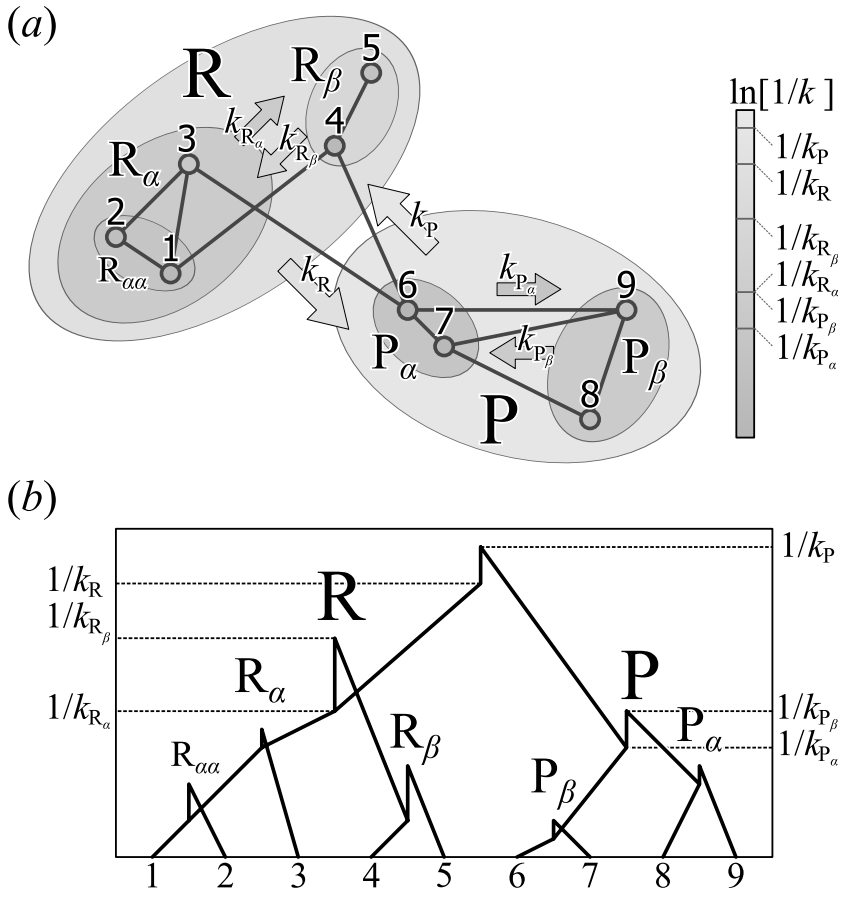



Figure 2.1: A schematic illustration of transformation from a network to a tree of reaction timescale hierarchy. (a) a reaction network composed of nine nodes with rate constants $k$ between them (links). (b) a tree of reaction timescale hierarchy representing the hierarchical organization of reaction timescales from (super)basins in the network. In (a), each gray ellipse represents each (super)basin $\mathrm{R}, \mathrm{R}_\alpha, \mathrm{R}_\beta, \mathrm{R}_{\alpha\alpha}, \mathrm{P}_\alpha$, and $\mathrm{P}_\beta$. Each large arrow represents transition between the subnetworks with the overall rate constants, e.g., $k_\mathrm{R}$. The color grade of ellipses, and large arrows represent the extent of timescale of a reaction: the darker the color, the faster the timescale.

# Chapter 3

# Backgrounds and Reviews

In this chapter, basic theories, algorithms, and mathematics relating to the developed methods are explained in detail.

In the first two sections: Sec. 3.1 Basics of Graph and Sec. 3.2 Non-negative Matrix, I introduce basic notations and mathematics for network theory.

Then, in the following three sections: Sec. 3.3 Transition State Theory, Sec. 3.4 Discrete Time Markov Chain, and Sec. 3.5 Disconnectivity Graphs, I will introduce basic theories and mathematics for relation between a energy landscape and a reaction network.

Finally, in the last two sections: Sec. 3.6 Clusterings for Disconnectivity Graphs, and Sec. 3.7 Combinatorial Search on Graph, I will introduce clustering methods and algorithms for visualizations of a reaction network.

## 3.1 Basics of Graph

Let me introduce definitions for a graph [42]. For a set of *nodes (vertices)* $V$ and set of *links (edges)* $E(= V \times V)$ between them, a *graph* $\mathcal{G} = (V, E)$ is defined. $\mathcal{G}$ is *directed graph (digraph)* when its edges $(i, j), (j, i) \in E(i, j \in V)$ are not equal $(i, j) \neq (j, i)$, or else $\mathcal{G}$ is called *undirected graph*. One can define a subgraph having a subset of nodes $S \subset V$ and all the links between them. The subgraph is called *vertex-induced subgraph*.

$w : E \to = [0, \infty)$ is called *edge weight*. To simplify later explanation, let me introduce the following definition:

$$\forall S, T \subset V, \quad w[S, T] := \sum_{i \in S, j \in T} w[i, j]. \tag{3.1}$$

If $w[i_1, i_2], \ldots, w[i_{k-1}, i_k]$ are positive, then $i_1, (i_1, i_2), i_2, \ldots, i_{k-1}, (i_{k-1}, i_k), i_k$ is called a *path* of length $k - 1$ from $i_1$ to $i_k$. For the later simplicity, let me call a path when the length of the path is not specified.

If any two of its vertices are linked by a path in G then the graph is called *connected*. If there are paths from $i$ to $j$ and vice versa for all $i, j \in V$, then the graph (subgraph) is called *strongly connected (strongly connected component)*. For the later simplicity, let me call a network as a connected graph.

Later I will explain that the reaction network having the equilibrium should be strongly connected.

## 3.2 Non-negative Matrix

In this section, I will introduce the mathematical property of the non-negative matrices. In Sec. 3.4 I will introduce relation between rate equation Eq. (3.7) and (discrete time) Markov chain Eq. (3.11). The properties of the non-negative matrices gives existence of unique steady distribution of Markov chain, ergodicity (see Sec. 3.4.2 for more detail).

### 3.2.1 Irreducibility

Non-negative matrix $A$ is *irreducible*, iif (if and only if) there exists $k$ for all $i, j$ such that $a_{ij}^{(k)}$ is positive, where $a_{ij}^{(k)} := (A^k)_{ij}$ is $i, j$th element of matrix $A^k$. In addition, a non-negative matrix is called *reducible*, if it is not irreducible.

For example,

$$A = \begin{pmatrix} A_{11} & * \\ O & A_{22} \end{pmatrix}$$

is reducible. Because

$$A^k = \begin{pmatrix} A_{11}^k & * \\ O & A_{22}^k \end{pmatrix}.$$

Hence, there are no $k$ for $i > j$ such that $a_{ij}^{(k)}$.

Notice that it is trivial by definition that matrix $W = (w[i, j])$ is irreducible iif graph $\mathcal{G}$ is strongly connected.

### 3.2.2 Perron-Frobenius Theorem

Let me introduce *spectral radius* of a matrix $A$ as $\rho(A) := \max_k |\lambda_k|$, where $\lambda_k$ is an eigenvalue of $A$. For irreducible, non-negative matrix, the following strong theorem holds.

Notice that, a vector, all the elements of which is positive, is called positive vector.

**Theorem 3.2.1** (Perron-Frobenius)

*If a Matrix $A$ is irreducible and non-negative, the followings hold.*

- *$\rho(A)$ is an eigenvalue of $A$, and the eigenvalue is simple.*
- *$\rho(A) > 0$ and there is a positive eigenvector associated with $\rho(A)$.*

**proof.** (See textbook 66) ∎

The theorem provides a necessary condition of existence of equilibrium distribution $\pi$ on reaction network.

If $A$ is positive and irreducible, the theorem become simplified as follows:

**Theorem 3.2.2** (Perron)

*If a matrix $A$ is irreducible and positive, the followings hold.*

- *$\lambda_1 = \rho(A)$ is the only eigenvalue of $A$ satisfying $|\lambda_1| \geq \rho(A)$.*
- *$\rho(A) > 0$ and there is a positive eigenvector associated with $\rho(A)$.*

**proof.** (See textbook 66) ∎

Later, I will sate that a $\Delta t(>0)$ time step transition-probability matrix derived from a rate equations satisfying detailed balance is positive, so that we use the latter theorem. Notice that, the latter theorem is not holds for all the Markov chains, since Markov chain is able to have zero transition probability.

### 3.2.3 Period of Matrix

Some non-negative matrices have eigenvalues $\lambda_1, \ldots, \lambda_\sigma$ satisfying $\rho(A) = |\lambda_k|$ $(k = 1, \ldots, \sigma)$. The eigenvalues can be written as $\lambda_k = |\rho(A)| \exp[2\pi i \frac{k}{\sigma}]$. $\sigma$ is called *period* of the matrix $A$. Note that, if a matrix $A$ has $\sigma > 1 (\sigma = 1)$ then $A$ is *periodic* (*aperiodic*).

It is trivial that if a transition-probability matrix is periodic, then the matrix does not have a unique steady distribution.

## 3.3 Transition State Theory

The *transition state theory* (TST)[47–53] is a theory to estimate a rate constant. In this section, I will introduce background, form of rate constant, hypothesise of the theory, and exceptional cases.

### 3.3.1 Background

One of the oldest empirical law of rate constant is found in 1889 by Svante Arrhenius known as Arrhenius plot(equation).

$$k = Ae^{-E_\mathrm{a}/RT}, \tag{3.2}$$

where $A$ is pre-exponential factor, $E_{\mathrm{a}}$ is activation energy, $R$ is gas constant, and $T$ is Temperature. After his finding, it was found that many reaction follow the law, but it was also found non-Arrhenius type systems.

### 3.3.2 Rate Constant

In 1935, H. Eyring [54], and M. Evans and M. Polanyi [55] introduce *Transition State Theory* (TST) independently to estimate *rate constant* statistical mechanically. The form of rate constant they introduce is

$$k = (k_{\mathrm{B}}T/h)\exp[-\Delta G^{\ddagger}/RT] \tag{3.3}$$

$$= (k_{\mathrm{B}}T/h)e^{\Delta S^{\ddagger}/R}e^{-\Delta H^{\ddagger}/RT}, \tag{3.4}$$

where $k_{\mathrm{B}}$ is Boltzmann constant, $h$ is Planck constant, $\Delta G^{\ddagger}$ is free energy (Gibbs energy) of activation, $\Delta H^{\ddagger}$ is enthalpy of activation, and $\Delta S^{\ddagger}$ is entropy of activation. Note that the last equality is from relation $\Delta G = -T\Delta S + \Delta H$. Hence, roughly $A \approx (k_{\mathrm{B}}T/h)e^{\Delta S^{\ddagger}/R}$ and $E_a \approx \Delta H^{\ddagger}$.

### 3.3.3 Hypothesise

To derive the relation, Eyring assume three things[53]: Transition State(TS), Quasi-equilibrium, existence of Reaction Coordinate.

**Transition State**

To assume rate constant, he introduce transition state (TS) between reactant (Re) and product (Pr)

$$\mathrm{Re} \rightleftarrows \mathrm{TS} \longrightarrow \mathrm{Pr}.$$

He assumes all the molecule passing through the TS should not return back to the reactant. From this formulation, the rate constant is able to be estimated by

$$k_{\mathrm{Re}\to\mathrm{Pr}} = \frac{k_{\mathrm{TS}\to\mathrm{Pr}}[\mathrm{TS}]_{\mathrm{eq}}}{[\mathrm{Re}]_{\mathrm{eq}}}. \tag{3.5}$$

**Quasi-equilibrium**

To compute the rate constant, $k_{\mathrm{TS}\to\mathrm{Pr}}[\mathrm{TS}]_{\mathrm{eq}}$ and $[\mathrm{Re}]_{\mathrm{eq}}$ should be derived. To derive these, Eyring assumes equilibrium between TS and Re so that $[\mathrm{TS}]_{\mathrm{eq}}$ and $[\mathrm{Re}]_{\mathrm{eq}}$ attains equilibrium distribution. This condition is called quasi-equilibrium.

In order to estimate these, one should use statistical mechanics. If the system is under equilibrium, density of molecular spices $[\mathrm{X}]_{\mathrm{eq}}$ is in proportion to probability to be observed as $X$:

$$[\mathrm{X}]_{\mathrm{eq}} \propto \int_{\mathrm{X}} \rho \mathrm{d}\Omega$$

where $\rho$ is equilibrium distribution, $\Omega$ is sample space, and X is area of spices $X$.

### Reaction Coordinate

The numerator of Eq.(3.5) is flux from TS to Pr. In order to derive these statistical-mechanically, one should introduce *reaction coordinate* $q^\ddagger$ and *dividing surface* $\{(p,q)|s(p,q)=0\}$.

The flux from TS to Pr can be derived

$$k_{\mathrm{TS}\to\mathrm{Pr}}[\mathrm{TS}]_{\mathrm{eq}} \propto \int_{\mathrm{TS}} \chi_{\mathrm{R}} J \rho_{\mathrm{R}} \mathrm{d}\Omega$$

where $J(p,q)$ is flux function, $\chi_{\mathrm{R}}(p,q)$ is a *simple function* which takes 1 at reactive area, where trajectories are reactive in phase space, and TS is area of TS.

The flux function $J(p,q)$ should defined on a dividing surface $\{(p,q)|s(p,q)=0\}$, a (codimension one) surface dividing phase space into Re and Pr. Typically, the dividing surface is defined on the origin of reaction coordinate, i.e., $q^\ddagger = 0$, so that $J = \delta(s(q))\nabla_q s(q) dq/dt = \delta(q^\ddagger)\mathrm{d}q^\ddagger/\mathrm{d}t$. If $q^\ddagger = 0$ is a saddle point of potential energy surface, the reactive area on the dividing surface is approximately area satisfying $p^\ddagger > 0$. [1] Therefore, the simple function can be rewritten as $\chi_{\mathrm{R}} := \Theta(p^\ddagger)$. The error of this estimation has been corrected by *transmission coefficient* $\kappa$, a empirical parameter.

Typically area of Re is a potential energy *basin* and that of TS is *ridge* given by $s(q)(=q^\ddagger) = 0$ which is boundary of Re and Pr basins; and the former is approximated from the local minimum and the latter is from a

[1] This assumption is known as *point of no return*. The assumption is not always true, because some trajectories recross the surface (overestimation) and/or quantum wave packet has tunnelling effect (underestimation). To correct these effect empirically, *transmission coefficient* $\kappa$ has been introduced.

*saddle* point (local minimum on the ridge but not the minimum of a basin). By using relation $dq_j/dt = m_j p_j$,

$$k_{\mathrm{Re}\to\mathrm{Pr}} = \frac{Z_{\mathrm{TS}}}{Z_{\mathrm{Re}}} e^{-\frac{-\Delta E^\ddagger}{RT}} \int \frac{p^\ddagger}{m^\ddagger} e^{-\frac{p^{\ddagger 2}/2m^\ddagger}{k_{\mathrm{B}}T}} \Theta(p^\ddagger) \frac{dp^\ddagger}{h}$$

where $\Delta E^\ddagger$ is difference between the saddle point energy and the minimum point energy of basin Re, and $Z_{\mathrm{X}} := \int_{\mathrm{X}} \rho d\Omega$ is partition function of X. Let I introduce free energy of X: $F_{\mathrm{X}}$ by

$$F_{\mathrm{X}} = -RT \ln[Z_{\mathrm{X}}].$$

Typically, chemical reactions proceed under constant pressure, so that the equilibrium distribution is estimated by *isothermal-isobaric ensemble* ($NPT$ ensemble: number of molecule $N$, pressure $P$, and temperature $P$ is constant). Under the condition, the free energy (invariant) of the system is Gibbs energy $F_{\mathrm{X}} = G_{\mathrm{X}}$. By using

$$\int \frac{p^\ddagger}{m^\ddagger} e^{-\frac{p^{\ddagger 2}/2m^\ddagger}{k_{\mathrm{B}}T}} \Theta(p^\ddagger) \frac{dp^\ddagger}{h} = \frac{k_{\mathrm{B}}T}{h},$$

one derive (3.3) for free energy of activation of a reaction Re $\longrightarrow$ Pr: $\Delta G^\ddagger_{\mathrm{Pr,Re}} := G^\ddagger_{\mathrm{Pr,Re:TS}} - G_{\mathrm{Re}}$, where $G^\ddagger_{\mathrm{Pr,Re:TS}}$ is the free energy of the transition state (the ridge being boundary of Re and Pr).

### 3.3.4 Exceptional Cases

Although the theoretical formulation has not been achieved yet, there exist exceptional cases, e.g. barrierless reaction[56, 57], reaction associated with potential energy ridge accompanied with higher index saddles[58–60], and dynamical effect in the reaction process under high energy condition[61–65].

## 3.4 Discrete Time Markov Chain

*Markov chain* (or more precisely *discrete-time Markov chain*: DTMC) is defined on set of states $V = \{1, 2, \dots\}$ and by *transition probabilities* between them (i.e., $i, j \in V,\ p_{\Delta t}[i|j]$).

Some reaction network can be translated into a special class of Markov chain known as ergodic and reversible Markov chain. However, There are some Markov chains that do not satisfy such properties. For example, some of the Markov chains do not have a unique steady distribution even after infinite iteration. In this section I will introduce necessary conditions to achieve them.

### 3.4.1 DTMC Derived from Rate Equation

#### 1st Order Rate Equations

Let me consider a reaction network that consists of $n$ *chemical species* $\mathrm{X}_1, \dots, \mathrm{X}_n$ and *first-order rate equations* between them:

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{X}_i] = \sum_{j \neq i} \left( k_{ij}[\mathrm{X}_j] - k_{ji}[\mathrm{X}_i] \right),$$

where $[\mathrm{X}_i]$ denotes *density* of the species $\mathrm{X}_i$ $(i = 1, \ldots, n)$.

$$k_{ij} := \sum_{k=1}^{m_{ij}} \frac{k_\mathrm{B} T}{h} \exp[-\Delta G_{ij}^{(k)\ddagger}/RT],$$

is a reaction rate constant for a reaction from $\mathrm{X}_j$ to $\mathrm{X}_i$. $\Delta G_{ij}^{(k)\ddagger} = G_{ij}^{(k)\ddagger} - G_j$ is a free energy of activation for the reactions between the $i$th and the $j$th chemical species and the free energy of the $j$th chemical species $G_j$. $G_{ij}^{(k)\ddagger}$ is the free energy of the $k$th reaction channel, out of $m_{ij}$ reaction channels. Note that $G_{ij}^{(k)\ddagger} = G_{ji}^{(k)\ddagger}$ holds.

In addition, let the reaction network be strongly connected. Under the setting, the reaction network has a unique steady distribution (see Sec.3.4.1 and Sec.3.2.2); and it satisfies *detailed balance*:

$$k_{ij}\pi_j = k_{ji}\pi_i \tag{3.6}$$

because of $k_{ij}\pi_j = k_{ji}\pi_i \propto \exp[G_{ij}^{(k)\ddagger}/RT]$. The steady distribution $\pi$ satisfying detailed balance is called the *equilibrium distribution*.

Let me use another expression for these rate equations

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{X}] = M[\mathrm{X}], \tag{3.7}$$

where $[\mathrm{X}] := ([\mathrm{X}_1], \ldots, [\mathrm{X}_n])^\mathrm{T}$ denotes the density (column) vector, $[\mathrm{X}]^\mathrm{T}$ its transpose of the vector $[\mathrm{X}]$, and $M := K - \sum_{j=1}^{n} \mathrm{diag}[(k_{j1}, \ldots, k_{jn})^\mathrm{T}]$ a $n \times n$ matrix with $K := (k_{ij})$ where $\mathrm{diag}[v]$ is a diagonal matrix whose $i$th diagonal element is the $i$th element of vector $v$. For the expression, the detailed

balance can express as

$$MD = (MD)^{\mathrm{T}}. \tag{3.8}$$

### Solution of the Equation

Since the equations are just linear ordinary differential equations, one can derive the solution of the equations especially when $M$ is diagonalizable: $M = U\Lambda U^{-1}$, where $U$ is regular matrix, $\Lambda(= \mathrm{diag}[\lambda_1, \ldots, \lambda_n])$ is a diagonal matrix having eigenvectors of $M$ in its entries. In case, (3.7) becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}U^{-1}[\mathrm{X}] = \Lambda U^{-1}[\mathrm{X}].$$

Hence, by introducing $U^{-1}[\mathrm{X}] = (y_1, \ldots, y_n)^{\mathrm{T}}$, the equations consist of

$$\frac{\mathrm{d}}{\mathrm{d}t}y_i = \lambda_i y_i,$$

for $i \in V$. For each $i$, the corresponding equation has a solution, i.e. $y_i(t) = e^{\lambda_i t}y_i(0)$.

By using *matrix exponential*:

$$\exp[A] := \sum_{k=0}^{\infty} \frac{A^k}{k!}, \tag{3.9}$$

the solution of the equation (3.7) can be rewritten as

$$[\mathrm{X}](t + \Delta t) = \exp[\Delta t M][\mathrm{X}](t). \tag{3.10}$$

### Corresponding Markov Chain

By normalizing the density vector [X], one can derive the equivalent Markov chain (see Ref. 67 or see in next section)

$$p_{t+\Delta t} = P_{\Delta t} p_t, \tag{3.11}$$

where $p_t := [\mathrm{X}](t)/\sum_{i=1}^{n}[\mathrm{X}_i](t)$ is a *marginal probability vector* and $P_{\Delta t} = (p_{\Delta t}[i|j]) := \exp[\Delta t M]$ is a *transition (probability) matrix*.

This transformation can be justified if $\exp[\Delta t M]$ satisfies the following two conditions:

$$\begin{aligned} &\text{(i)} \quad \mathbf{1}^{\mathrm{T}}\exp[\Delta t M] = \mathbf{1}^{\mathrm{T}}, \\ &\text{(ii)} \quad \text{for all } i, j, (\exp[\Delta t M])_{ij} \geq 0, \end{aligned}$$

where $\mathbf{1} = (1, \ldots, 1)^{\mathrm{T}}$.

The statement (i) can be proved[67] by the conservation property of the rate equation, Eq. (3.7): $\mathbf{1}^{\mathrm{T}} M = \mathbf{0} \Rightarrow \mathbf{1}^{\mathrm{T}} \mathrm{d}[\mathrm{X}]/\mathrm{d}t = 0$. In other words, for all $\Delta t \geq 0$ and for all [X],

$$\begin{aligned} &\mathbf{1}^{\mathrm{T}} \cdot [\mathrm{X}](t) = \mathbf{1}^{\mathrm{T}} \cdot [\mathrm{X}](t + \Delta t) \\ \Leftrightarrow \quad &(\mathbf{1}^{\mathrm{T}} - \mathbf{1}^{\mathrm{T}}\exp[\Delta t M]) \cdot [\mathrm{X}](t) = 0. \end{aligned}$$

The statement (ii) can be proved, since $M$ is a *Metzler matrix*[68]: a $n \times n$ real square matrix $A$ whose all the off-diagonal entries are non-negative.

For any Metzler matrix $A$, there exists a real positive scalar constant $a$ ($> 0$) such that $(A + aI)$ is non-negative (i.e., all the elements of $(A + aI)$ are non-negative) where $I$ is an identity matrix. In addition, $\exp[\Delta tA] = \exp[\Delta t(A + aI)]e^{-\Delta ta}$ holds, because $A$ and $I$ commute with each other, i.e. $AI = IA$. Hence for all $\Delta t > 0$, $\exp[\Delta tA]$ is non-negative, since $e^{-\Delta ta} > 0$.

**Positivity of $P_{\Delta t}$**

If the matrix $(A + aI)$ is irreducible, then for all $i, j(= 1, \ldots n)$, at least one of non-negative matrices $(A + aI)^1, (A + aI)^2, \ldots, (A + aI)^n$ has a positive $i, j$ element. Hence for all $\Delta t > 0$,

$$\exp[\Delta t(M + aI)] = \sum_{k=0}^{\infty} \frac{(\Delta t(M + aI))^k}{k!} \tag{3.12}$$

is positive. Therefore $\exp[\Delta tM]$ is positive, since $\exp[\Delta tM] = \exp[\Delta t(M + aI)]\mathrm{e}^{-\Delta ta}$. Note that $(A + aI)$ is irreducible iif $A$ is irreducible. It is because self loops (i.e., link that connects a node to itself) is irrelevant to strong connectedness of a network. Therefore, if the matrix $M$ is irreducible, $\exp[\Delta tM]$ is positive.

Therefore, Perron's theorem holds for a reaction networks defined by (3.3) and (3.7); and the network has the unique equilibrium distribution.

### 3.4.2 Ergodicity

If a transition probability matrix of a finite Markov chain is aperiodic and irreducible, then the matrix is called *ergodic*. In case, there exists a unique *steady distribution* $\pi$, because of Perron-Frobenius theorem 3.2.2.

In addition, some Markov chain has a relation known as *reversibility*:

$$P_{\Delta t} D = (P_{\Delta t} D)^{\mathrm{T}}, \tag{3.13}$$

where $D := \mathrm{diag}[\pi]$, and $X^{\mathrm{T}}$ is transpose of matrix $X$. This relation is corresponding to detailed balance which is expected for systems under equilibrium.

## 3.5 Disconnectivity Graphs

In this section, I will introduce some visualization methods for energy landscapes, i.e. *disconnectivity graph*s (DSs). Let me briefly summarize the DG and the variants of the DG. The original DG was proposed in 1997 by O. M. Becker and M. Karplus [7]. In addition to the DG, another DG, called *canonical DG*, was proposed in the same paper. The DG was applied to free energy surface instead of potential energy surface by D. Evans and D. Wales, and called *free energy DG*[37]. These DGs were applied to various systems, such as not only small proteins (e.g. GB1: B1 domain of protein G[8], beta3s: the 20-residue miniprotein[9], three-color 46-bead model (BLN)[10]), but also molecular clusters[11], cluster molecules[11], and some mesoscopic systems consisting of rigid building blocks[12, 13]. In 2002, the first top-down type DG was proposed by S. Krivov and M. Karplus[38, 39] and it called *transition DG*.

In the following subsections, I will explain about the DG, Canonical DG, and Transition DG.

Table 3.1: Variants of Disconnectivity graphs(DSs) and their classification based on their methods

| | TS's (free) energy | Free energy of activation |
|---|---|---|
| Bottom-up | DG[7], Free energy DG[37] | Canonical DG[7] |
| Top-down | Transition DG[38, 39] | Proposed method |

### 3.5.1 Disconnectivity Graph, and Free Energy DG

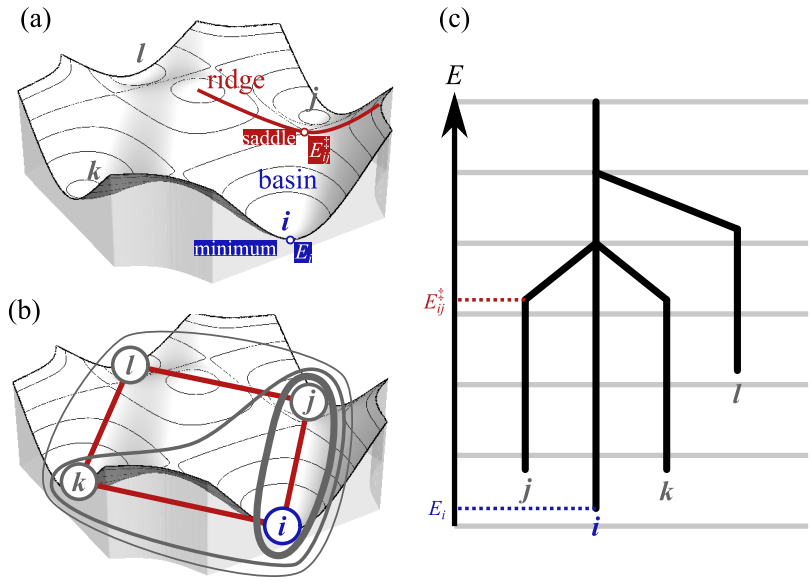



Figure 3.1: Potential energy landscape and its disconnectivity graph. (a) Potential energy surface, a basin and its local minimum energy $E_i$, a ridge and its saddle point energy $E_{ij}^{\ddagger}$. (b) Corresponding network which edges have saddle point energies and vertices have local minimum energies of basins. (c) Disconnectivity graph of the potential energy surface.

DG[7] is a visualization method for multi-dimensional potential energy surface. DG expresses energetically inaccessible ('disconnected') subspaces parametrized by total energy $E$. This disconnectivity is described by saddle points (the local minimums of ridges) energy of potential energy surface, because if the total energy is larger(smaller) than a saddle point energy e.g.,

$E^{\ddagger}_{ij}$, adjacent areas are energetically accessible (inaccessible), as shown in Fig. 3.1-(a). Specifically, a saddle point connects potential energy basin (adjacent area of a local minimum $E_i$ and that of $E_j$), e.g., basin $i$ and $j$ or set of basins.

The change of disconnectivity by energy $E$ increase is similar to pouring water into surface. If you pour more water into a basin of the surface, water pool on the basin is getting larger. Then an adjacent basin is flooded with water. The two pools on the basins connect each other and become a single pool. Finally, all the pools on the basins connect each other and become a single pool.

Therefore, by using lits of saddle points energies, e.g., $E^{\ddagger}_{ij}$ and their adjacent basins, e.g., $i$th and $j$th basins, or by using corresponding network which edges have saddle point energies and vertices have local minimum energies of basins as shown in Fig. 3.1-(b), DG can be constructed. For example, as shown in Fig. 3.1-(c), at first the basin $i$ and $j$ are connected at $E = E^{\ddagger}_{ij}$. Then the basins $i, j$ and the basin $k$ are connected at $E = E^{\ddagger}_{ik}$. Finally, the basins $i, j, k$ and the basin $l$ are all together at $E = E^{\ddagger}_{kl}$

Free energy DG[37] is simple extension of DG. The free energy DG uses list of free energy of ridges, e.g., $G^{\ddagger}_{ij}$ instead of saddle point energies, e.g., $E^{\ddagger}_{ij}$. Since transition state (TS) is typically assigned at a adjacent ridge of a saddle, one can say that both method use TS's (free) energy.

The DG and free energy DG are well established projection method of morphological features of multi-dimensional surface. These DGs were applied to various systems, such as not only small proteins (e.g. GB1: B1 domain of protein G[8], beta3s: the 20-residue miniprotein[9], three-color 46-bead

model (BLN)[10]), but also molecular clusters[11], cluster molecules[11], and some mesoscopic systems consisting of rigid building blocks[12, 13].

### 3.5.2 Canonical DG

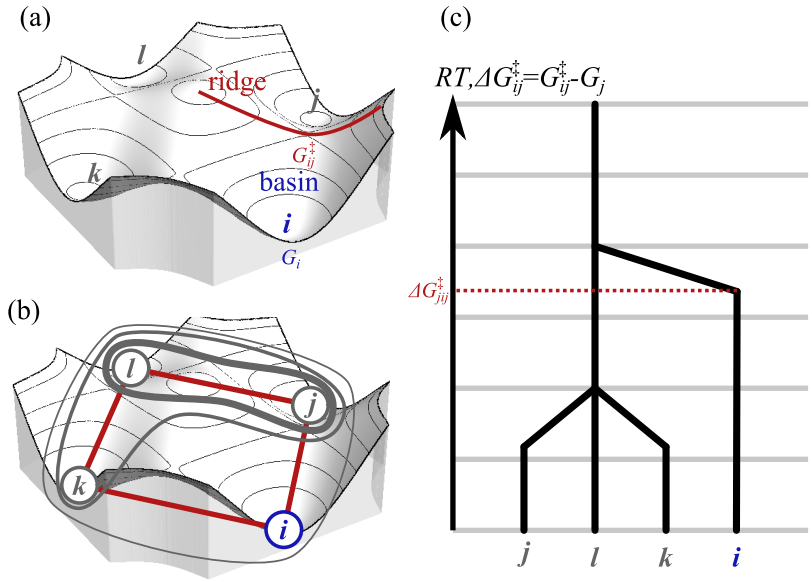



Figure 3.2: Potential energy landscape and its canonical disconnectivity graph. (a) Potential energy surface, a basin and its free energy $G_i$, a ridge and its free energy $G_{ij}^{\ddagger}$. (b) Corresponding network which edges have free energies of activation $\max\{\Delta G_{ij}^{\ddagger}, \Delta G_{ji}^{\ddagger}\}$ and vertices have free energy of basins $G_i$. (c) Canonical disconnectivity graph of the potential energy surface.

Canonical DG visualize observable set of reactions for given temperature[7] or timescale[41]. Canonical DG uses list of free energy of activation $\max\{\Delta G_{ij}^{\ddagger}, \Delta G_{ji}^{\ddagger}\}$ instead of TS's (free) energy. The disconnectivity (originally it considered as kinetic connectivity) is parametrized by thermal energy $RT$ (or $k_{\mathrm{B}}T$; it depends on the physical unit of free energy)[7], or by reaction timescale $1/k_{ij}$ [41], where $T$ is temperature, $R$ gas constant, $k_{\mathrm{B}}$ Boltzmann constant, $k_{ij}$

rate constant of reaction from $j$ to $i$. Since this observability is not directly reflect global dynamics of reaction process, the method have not been applied possibly due to difficulty of its interpretation.

### 3.5.3 Transition DG

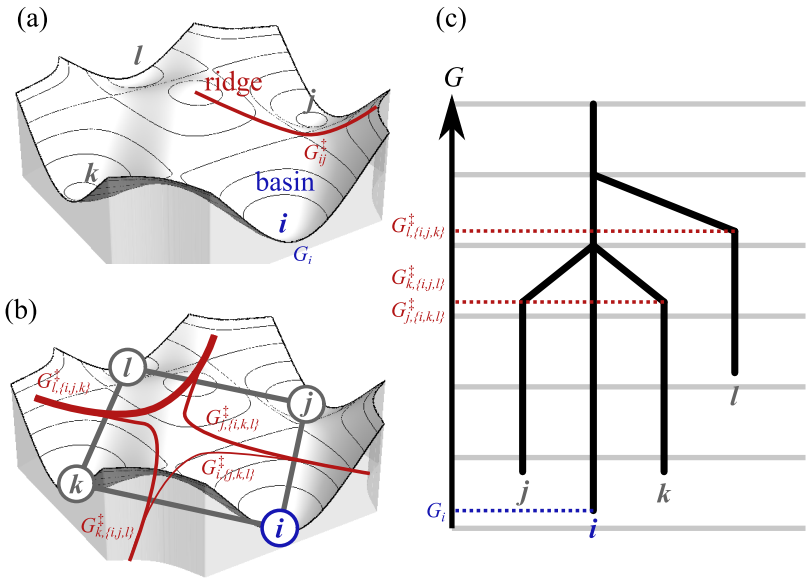



Figure 3.3: Potential energy landscape and its transition disconnectivity graph. (a) Potential energy surface, a basin and its free energy $G_i$, a ridge and its free energy $G^{\ddagger}_{ij}$. (b) Corresponding network which edges have TS's free energies $G^{\ddagger}_{ij}$ and vertices have free energy of basins $G_i$. (c) transition disconnectivity graph of the potential energy surface.

Transition DG express hierarchies of reaction bottle-necks[38, 39]. Transition DG uses partitions which have the 1st, the 2nd, the 3rd,..., minimum gross equilibrium reaction rate. The y-axis of the graph represent absolute reaction barrier hight of partition boundary. It is because, reaction rate has relation $k_{ij}[X_j]_{\mathrm{eq}} \propto \exp[-G^{\ddagger}_{ij}/RT]$, so that gross reaction rate

$\sum_{i\in I, j\in J} k_{ij}[X_j] \propto \exp[-G^{\ddagger}_{IJ}/RT]$, where $[X_j]_{\mathrm{eq}}$ is equilibrium distribution of $j$th chemical species and $I, J$ are disjoint set of nodes.

It was pointed out[39] that the procedure of transition DG conduct partitions which are physically unimportant. It is because the gross equilibrium reaction rate only depend on boundary of partition, but does not depend on size or depth of partitioned subspace. Therefore, entropic effect is not taken into account. For some trivial cases, correction method is proposed[39]. However not only the correction does not improve for the worst case, but also physical interpretation and when the correction should be made are unclear.

## 3.6 Clusterings for Disconnectivity Graphs

In this section, I review some clustering methods relating to DG and its variants. There are two typical approaches to make a cluster from data points: top-down and bottom-up. *Agglomerative clustering*[69]: merging a pair of data points iteratively, is subgroup of the bottom-up clustering methods. *Divisive clustering*[69]; partitioning data points into two or more is subgroup of the top-down clustering methods. The former is used in the DG[7] and the free energy DG[37], and the latter is used in the transition DG[38, 39]. My method proposed in Sec.4 uses another Divisive clustering method.

### 3.6.1 Agglomerative Clustering

Agglomerative clustering is a bottom-up type clustering method. For given data points and similarity (e.g., distance between a pair of data points),

the method merge a pair of data points iteratively. There are several fast algorithms for agglomerative clusterings, but interpretation is less clear in relationship between the resultant clusters. This is because that the clustering is achieved only by similarity of a pair of data points. However, when the purpose of employing agglomerative clustering is not concern in the relationships between the resultant clusters, the agglomerative clustering gives you rigorous result with light computation cost, e.g. DG.

#### Single Linkage

*Single linkage*[70] is a typical agglomerative method. The DG, free energy DG, and canonical DG use the method[7, 37].

The method uses list of similarity for each pair of data point (e.g. list of saddle point energies). The list is sorted in descending order. By looking the list from the top, the method merges pair of data points having the largest similarity. By keeping the remaining list, the method merges points having the next largest similarities.

### 3.6.2 Divisive Clustering

*Divisive clustering* is top-down method achieving optimal value of *cost function*, so that resultant clusters have clear relationships. In many cases, to find such clusters, algorithms need to see vast number of combinations: *combinatorial explosion problem.*

Minimum Cut is a problem which has fast algorithms which are free from combinatorial explosion problem. However in many cases, the combinatorial

explosion problem is inevitable. Hence, *heuristics* (approximation method) have been proposed to approximate optimal clusters with relatively light computation cost.

### Minimum Cut

*Minimum (capacity) Cut problem* [44] is a problem to find a bipartition which has the smallest in sum of similarities (capacities) of the boundary edges from all the other bipartitions. The method need *source (start)* node $s$ and *sink (target)* node $t$ as input, thus the resultant partition is called $s,t$-*Minimum Cut*. For example, if one uses similarity of data points $i, j$ as $w[i, j]$, Minimum Cut provides

$$\min_{S \subset V} w[S, S^{\mathrm{c}}],$$

where $s \in S, t \in S^{\mathrm{c}}$. The boundary of the $s,t$-Minimum Cut will change only when new source $s'$ and sink $t'$ is on the same side in previous partition (i.e. $s,t \in S$ or $s,t \in S^{\mathrm{c}}$); and the resultant new boundary never crosses the boundary already obtained (Lemma 8.35 in Ref. 44). Because of this two properties, by recursively applying the $s,t$-Minimum Cut, one obtains succinct data structure expressing all the possible $s,t$-Minimum Cuts. The data structure is known as *Minimum Cut (Gomory-Hu) tree.*[43, 44]

However, because Minimum Cut only refer the boundary of cluster, the cut sometimes partition a network into a single node and the other. This is know as *trivial(unbalanced) cut problem* [39, 45, 46]: a typical problem of Minimum Cut in clustering. There is a method to correct the cases known as *balanced Minimum Cut problem.*[39, 45, 46] The correction is made in order

to balance Minimum Cut. The balanced Minimum Cut balance number of nodes:

$$\min_{S \subset V}(c|S| + w[S, S^{\mathrm{c}}]),$$

where $|S|$ is number of node in $S$, and $c$ is parameter. This balancing makes relationships between the resultant clusters ambiguous. In addition, if the $S$ consists of a single node, no correction made.

Transition DG[38, 39] uses Minimum Cut tree to visualize hierarchical relation of equilibrium flows. The correction is also introduced for transition DG [39] to find entropically stabilized superbasin. They chose $s$ having the largest free energy of basin $G$ (corresponding to largest weighted edge degree) and they chose $c$ variationally (i.e. starting from $c = 0$, by increasing $c$ the first transition of boundary is taken.). However, justification of the correction only by one of the resultant clusters and by addition of the term $c|S|$ has not been clarified. In addition, $c|S|$ can relate to entropic stabilization in $S$ when $c|S| = \pi[S](\propto \exp[-G/RT])$, but it is not clear for the other case. Thus, interpretation of the correction is physically ambiguous.

### Spectral Clustering

There are several measures to quantify quality of clustering known as cluster fitness measures[69]. One of the well know measure is called *graph conductance*[69, 71–77]:

$$\phi(S) := \min\left\{\frac{w[S, S^{\mathrm{c}}]}{w[V, S^{\mathrm{c}}]}\frac{w[S^{\mathrm{c}}, S]}{w[V, S]}.\right\} \tag{3.14}$$

The graph conductance provides us density of similarities $w$ on boundary between $S$ and $S^{\mathrm{c}}$.

If one uses similarity as *equilibrium joint probability*:

$$\pi_{\Delta t}[i,j] := p_{\Delta t}[i,j]\pi_{[}j], \tag{3.15}$$

the conductance becomes,

$$\phi_p(S) := \max\left\{p_{\Delta t}[S|S^{\mathrm{c}}], p_{\Delta t}[S^{\mathrm{c}}|S]\right\}, \tag{3.16}$$

where

$$p_{\Delta t}[S|S^{\mathrm{c}}] := \frac{\pi_{\Delta t}[S,S^{\mathrm{c}}]}{\pi[S^{\mathrm{c}}]}. \tag{3.17}$$

The meaning of the conductance then is physically clear; transition probability from an equilibrated superbasin $S$ to the other $S^{\mathrm{c}}$. This is similar to quasi-equilibrium approximation in transition state theory.

It is interesting to obtain the cluster attaining the minimum in conductance: the partition $(S, S^{\mathrm{c}})$ satisfying $\phi_p(S) = \min_{T\subset V}\phi_p(T)$. Finding the partition is known as the *minimum conductance cut* (MCC) problem. The problem is belong to NP-hard problem [71]; the solution of the problem with large number of nodes and links is not able to be found by any algorithm in practically manageable time. Therefore, another algorithms which approximate the solution and find solution quickly are used in practice. However by following reason, one cannot use approximation algorithms (heuristics) to obtain the MCC.

To obtain cluster attaining small $\phi$, *spectral clustering* methods are devel-

oped [69, 75–78]. For a heuristic of the MCC, the normalized graph Laplacian $L$ defined by

$$L = I - D^{-1/2} P_{\Delta t} D^{1/2}, \tag{3.18}$$

was used, where $I$ is an $n \times n$ identity matrix, $D^{1/2} = \mathrm{diag}[(\pi[1]^{1/2}, \ldots, \pi[n]^{1/2})^{\mathrm{T}}]$, and $D^{-1/2}$ is the inverse matrix of $D^{1/2}$. The method conducts one dimensional search based on the eigenvector $v = (v[1], \ldots, v[n])^{\mathrm{T}}$ of the second smallest eigenvalue of $L$. The search is conducted on sorted vertices $\{1, \ldots, n\}$ to $\{j_1, \ldots, j_n\}$ in order that the coefficients $v[j_1], \ldots, v[j_n]$ in the eigenvector are ordered in ascent order as $v[j_1] \geq \cdots \geq v[j_n]$ (ordering in descent order also works). Then, by classifying vertices $V_k = \{j_1, \ldots, j_k\}$ into $S$ and the rest $S^c$ consequently up to finding the minimum value of $\phi_p(V_k)$ among all $1 \leq k \leq n$. It is proved that the obtained value $\phi_p$ is within a certain bound [69, 72–77].

Finally, it should be noted that although one might imagine that the eigenvector with the second smallest eigenvalue of $L$ provides the smallest overall rate constant over the network, i.e., MCC, between two disjoint subnetworks, it is not necessarily so. One may read its mathematical discussions in, e.g., Ref. 72–75. In some cases, the MCC and the second minimum conductance cut are very different partitions although their conductance is close each other. This is a problem when one applies the MCC recursively to obtain hierarchical organization, because large difference in early partition possibly makes large difference in the resultant hierarchical organization.

## 3.7 Combinatorial Search on Graph

In this thesis, I introduce algorithm to obtain the MCC and MCC-based TST for network TS. Although the algorithm cannot find solution of problems with large number of nodes and links, the algorithm is faster than an algorithm which enumerate of all possible candidates.

Given a set of all the nodes $V$ of a network, let me consider all the possible bi-partitions of $V$, i.e., partitions of $V$ into two disjoint subsets $S$ and $S^{\mathrm{c}}$. For the number of nodes $n$, the number of all the possible bi-partitions is $(2^n - 2)/2 = 2^{n-1} - 1$, where $-2$ is to remove cases that either $S$ or $S^{\mathrm{c}}$ is empty and the multiplication by $1/2$ is due to symmetry $(S, S^{\mathrm{c}}) = (S^{\mathrm{c}}, S)$. In order to obtain the MCC, I use binomial decision tree, which is a naive but very general graphical tool to enumerate combination of partitions.

### 3.7.1 Binomial Decision Tree

The *binomial decision tree* (schematically depicted in Fig. 3.4) forms data structure for combinatorial enumeration of all the possible series of binomial decisions. Each series of binomial decisions starts from a single vertex in the tree (the top rectangle vertex of Fig. 3.4). Each vertex in the tree has two output branches representing decisions "true" and "false" (the arrows from each rectangle vertex in Fig. 3.4), respectively. A series of decisions or each path in the tree represents a *bit vector* $b = (b[1], \ldots, b[n])^{\mathrm{T}} \in \{0, 1\}^n$ where 1 represents "true" and 0 represents "false", for example. *bit vector* To use the binomial decision tree to enumerate all the possible bi-partitions, I relate each series of binomial decisions to each bi-partition. To do that,

it is sufficient to construct a unique bi-partition for each bit vector $b$, since each series of binomial decisions can be represented by a unique bit vector $b$.

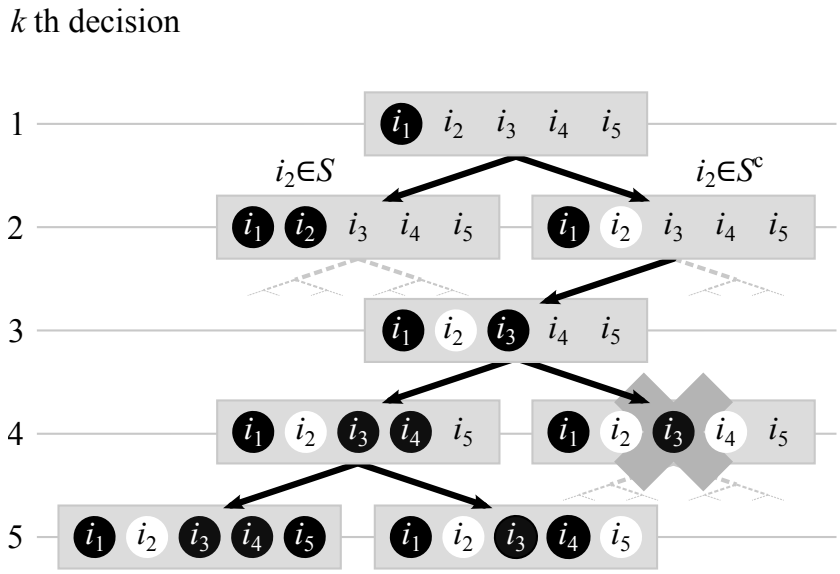



Figure 3.4: Schematic picture of part of a binomial decision tree for five elements with the branch and bound method. The five elements $\{i_1, \ldots, i_5\}$ are to be decided whether $i_k \in S$ or $i_k \in S^{\mathrm{c}}$. Each computation step corresponds to *decision* and each gray-rectangle vertex gives the (temporal) classification of the elements at each decision step. If $i_k$ is allocated to $S_k$ ($S_k^{\mathrm{c}}$), I use a black (white) disk, and if $i_k$ has not yet been allocated, $i_k$ is left as it has been. Each black arrow corresponds to a possible choice of bi-partitioning at each decision step. Gray dashed lines are some of such choices omitted for simplicity. The $\times$-mark on the lower right at fourth decision corresponds to a situation that one does not need to search any longer along the path. See Text for more details.

Let $\{i_1, \ldots, i_n\}$ be a permutation of the indices of the nodes $V = \{1, \ldots, n\}$. I recursively construct subsets of $V_k = \{i_1, \ldots, i_k\}$ i.e., $S_k$ and $S_k^{\mathrm{c}} (= V_k \setminus S_k)$ $(k = 1, \cdots, n)$ along each path of the decision tree as follows: Let $S_1 = \{i_1\}$ and $S_1^{\mathrm{c}} = \emptyset$ be a pair of subsets of $V_1 (= \{i_1\})$. Starting from such a pair of $S_1$ and $S_1^{\mathrm{c}}$, I construct a series of $V_k$, $S_k$, and $S_k^{\mathrm{c}}$ (Note that these fixed subsets do not lose any generality). Suppose that I have constructed $S_{k-1}$ and $S_{k-1}^{\mathrm{c}}$

up to $(k-1)$. Then, I construct $S_k$ and $S_k^{\rm c}$ such that $S_k = S_{k-1} \cup \{i_k\}$ and $S_k^{\rm c} = S_{k-1}^{\rm c}$ if $b[i_k] = 1$ and $S_k = S_{k-1}$ and $S_k^{\rm c} = S_{k-1}^{\rm c} \cup \{i_k\}$ if $b[i_k] = 0$. If I continue this recursive construction up to $k = n$, all the vertices in $V$ are allocated either $S_n$ or $S_n^{\rm c}$ and $(S_n, S_n^{\rm c})$ becomes a bi-partition of $V$. By this procedure, I can construct a unique bi-partition for a given bit vector $b$.

An example of a decision tree with five elements is shown schematically in Fig. 3.4. In the tree, the first decision for $i_1$ is unimportant due to the symmetry $(S, S^{\rm c}) = (S^{\rm c}, S)$, and, thus, $i_1$ can be allocated to $S_1$ without any loss of generality. The second decision for $i_2$ has two possible allocations: $S_2 = S_1 \cup \{i_2\}$, $S_2^{\rm c} = S_1^{\rm c}$ and $S_2^{\rm c} = S_1^{\rm c} \cup \{i_2\}$, $S_2 = S_1$. The former corresponds to the left branch and the latter corresponds to the right branch in the figure. In Fig. 3.4 if $i_k$ is allocated to $S_k$ ($S_k^{\rm c}$), I draw a black (white) disk, and if $i_k$ has not yet been allocated, $i_k$ is left as it has been. Each path from the top to the bottom of the tree corresponds to a series of decisions; hence the number of such paths in the tree is equal to the number of possible combinations of bi-partitions. The number of computation (decision) in the tree is $\sum_{k=2}^{n} 2^{k-1} = 2^{n+1} - 2$. Hence the computational cost increases exponentially when the number of elements increases.

### 3.7.2 Branch and Bound Method

However in some cases, one can limit amounts of computational cost by using the data structure. Let me consider an optimization problem to find a partition $(S, S^{\rm c})$ yielding the minimum value of some function $f(S)$ among all the other partitions. For example, suppose that one knows a function $g(S_k)$

which monotonically increases as a function of the number of decisions in the decision tree ($k \to n$) converging into the objective function $f(S)(= g(S_n))$, and that one knows a certain bound $\alpha$ such that the optimal answer $f(S)$ is smaller than that value. Then one does not need to continue to search all the rest beyond some branch which has a larger value of $g(S_k)$ than the bound $\alpha$. Hence, the method to stop searching beyond such branches is called branch-and-bound method. For example in Fig. 3.4, if one allocates $i_2, i_4$ to $S^{\mathrm{c}}$ and $i_1$, $i_3$ to $S$, then $g(S_4) > \alpha$, implying that, in the case, regardless of allocation of $i_5$, $f(S) > \alpha$.

### 3.7.3 Searching Order

In general, it is difficult to find the permutation or the order $\{i_1, \ldots, i_n\}$ that best minimizes the steps of the search (decision) even with the relation $g(S_k) \leq f(S)$ and the minimum value of $\alpha$.

The order $\{i_1, \ldots, i_n\}$ such that it satisfies $\deg[i_1] \geq \cdots \geq \deg[i_n]$ is commonly used, where $\deg[i] := w[V, i]$ is edge degree (i.e. $\pi[i]$). This choice was found to be more efficient to accelerate finding combinations to exceed the bound $\alpha$.

# Chapter 4

# Method

Transition state theories (TST) assume quasi-equilibrium between the reactant state and the transition state (TS), which divides the entire space into those of the reactant and the product. It needs the existence of timescale separation between the equilibration within the reactant and transition through the TS to the product. It is thus natural to define the network TS for reaction networks on the basis of the timescale separation. Network TS: a surface dividing the network into two disjoint subnetworks should have the longest reaction timescale. Compared to referring gross rates [38, 39], there exist two overall rate constants from one subnetwork to another and vice versa. I refer the larger overall rate constant and search the cut having the minimum over all the possible cuts. This cut is known to be minimum conductance cut(MCC) obtained by the minimization of graph conductance $\phi_p$ [69, 72, 73, 75–78] defined by Eq. (3.16):

$$\phi_p(S) := \max\left\{p_{\Delta t}[S^{\mathrm{c}}|S], p_{\Delta t}[S|S^{\mathrm{c}}]\right\},$$

where the set of nodes in a given network $V$ is partitioned into two disjoint subsets $S$ and its compliment $S^{\mathrm{c}}(S^{\mathrm{c}} = V \setminus S)$, $p_{\Delta t}[S^{\mathrm{c}}|S]$ and $p_{\Delta t}[S|S^{\mathrm{c}}]$ are transition probabilities from $S$ to $S^{\mathrm{c}}$ and from $S^{\mathrm{c}}$ to $S$ within timescale $\Delta t$, respectively. In the following, the notation $(S, S^{\mathrm{c}})$ denotes the partition for the sake of simplicity.

## 4.1 Deriving Markov Chain from Rate Equation

The Markov chain was derived from 1st order rate equation defined in Eq. (3.7)

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathrm{X}] = M[\mathrm{X}].$$

Form the solution (Eq. (3.10)), one can derive the equivalent Markov chain(Eq.(3.11))

$$p_{t+\Delta t} = P_{\Delta t} p_t,$$

where $p_t := [\mathrm{X}](t)/\sum_{i=1}^{n}[\mathrm{X}_i](t)$ is a marginal probability vector and $P_{\Delta t} = (p_{\Delta t}[i|j]) := \exp[\Delta t M]$ is a transition (probability) matrix.

The values of $p_{\Delta t}[S^{\mathrm{c}}|S]$ and $p_{\Delta t}[S|S^{\mathrm{c}}]$ are defined in Eq. (3.17):

$$p_{\Delta t}[S|S^{\mathrm{c}}] := \frac{\pi_{\Delta t}[S, S^{\mathrm{c}}]}{\pi[S^{\mathrm{c}}]}.$$

The partition $(S, S^{\mathrm{c}})$ satisfying $\phi_p(S) = \min_{T \subset V} \phi_p(T)$ is called the minimum conductance cut (MCC) [69, 72, 73, 75–78].

Notice that after finding the MCC, I can translate the transition probabilities into overall rate constants by using the relation

$$M = \ln[P_{\Delta t}]/\Delta t. \tag{4.1}$$

More specifically, If one can assume equilibration within each resultant superbasin of the MCC $S, S^{\mathrm{c}}$ before the reaction takes place, the transition probabilities between them is simply written as $p_{\Delta t}[S^{\mathrm{c}}|S]$ and $p_{\Delta t}[S|S^{\mathrm{c}}]$. For the cases, Eq. (4.1) is

$$\begin{pmatrix} -k_{S^{\mathrm{c}},S} & k_{S,S^{\mathrm{c}}} \\ k_{S^{\mathrm{c}},S} & -k_{S,S^{\mathrm{c}}} \end{pmatrix} = -\frac{\ln(1 - p_{\Delta t}[S|S^{\mathrm{c}}] - p_{\Delta t}[S^{\mathrm{c}}|S])}{\Delta t(p_{\Delta t}[S|S^{\mathrm{c}}] + p_{\Delta t}[S^{\mathrm{c}}|S])} \begin{pmatrix} -p_{\Delta t}[S^{\mathrm{c}}|S] & p_{\Delta t}[S|S^{\mathrm{c}}] \\ p_{\Delta t}[S^{\mathrm{c}}|S] & -p_{\Delta t}[S|S^{\mathrm{c}}] \end{pmatrix}. \tag{4.2}$$

## 4.2 Function for Branch and Bound Method

By using branch and bound method on a decision tree and $\psi$ defined below, one searches the MCC. The function $\psi(S_k; V_k)$ (corresponding to the function $g(S_k)$) is another form of the graph conductance $\phi_p(S)$:

$$\psi(S) := \max\left\{ \frac{\pi_{\Delta t}[S^{\mathrm{c}}, S]}{\min\{\frac{1}{2}, 1 - \pi[S^{\mathrm{c}}]\}}, \frac{\pi_{\Delta t}[S, S^{\mathrm{c}}]}{\min\{\frac{1}{2}, 1 - \pi[S]\}} \right\}. \tag{4.3}$$

The advantage of this form $\psi(S)$ is to provide a function $\psi(S_k; V_k)$

$$\psi(S_k; V_k) := \max\left\{ \frac{\pi_{\Delta t}[S_k^{\mathrm{c}}, S_k]}{\min\{\frac{1}{2}, 1 - \pi[S_k^{\mathrm{c}}]\}}, \frac{\pi_{\Delta t}[S_k, S_k^{\mathrm{c}}]}{\min\{\frac{1}{2}, 1 - \pi[S_k]\}} \right\}. \tag{4.4}$$

By using this form, one can prove that $\psi(S_k; V_k)$ monotonically increases as $k$ increases: $\psi(S_1; V_1) \leq \cdots \leq \psi(S_n; V_n) = \phi_p(S; V) = \phi_p(S)$.

**Theorem 4.2.1** *Suppose $S_n \neq \emptyset$ or $S_n \neq V_n$.*

$$\psi(S_k; V_k) := \max \left\{ \frac{\pi_{\Delta t}[S_k^{\mathrm{c}}, S_k]}{\min\{\frac{1}{2}, 1 - \pi[S_k^{\mathrm{c}}]\}}, \frac{\pi_{\Delta t}[S_k, S_k^{\mathrm{c}}]}{\min\{\frac{1}{2}, 1 - \pi[S_k]\}} \right\}$$

*satisfy $\psi(S_1; V_k) \leq \cdots \leq \psi(S_n; V_n) = \phi_p(S_n; V_n)$ under the setting described in chapter 3.*

**proof.**

First I transform the conductance $\phi_p$ to $\psi$, where $S \neq \emptyset$ or $S \neq V$.

$$\begin{aligned}
\phi_p(S; V) &= \max \{p_{\Delta t}[S^{\mathrm{c}}|S], p_{\Delta t}[S|S^{\mathrm{c}}]\} \\
&= \max \left\{ \frac{\pi_{\Delta t}[S^{\mathrm{c}}, S]}{\pi[S]}, \frac{\pi_{\Delta t}[S, S^{\mathrm{c}}]}{\pi[S^{\mathrm{c}}]} \right\} \\
&= \max \left\{ \frac{\pi_{\Delta t}[S^{\mathrm{c}}, S]}{\min\{\frac{1}{2}, \pi[S]\}}, \frac{\pi_{\Delta t}[S, S^{\mathrm{c}}]}{\min\{\frac{1}{2}, \pi[S^{\mathrm{c}}]\}} \right\} \\
&= \max \left\{ \frac{\pi_{\Delta t}[S^{\mathrm{c}}, S]}{\min\{\frac{1}{2}, 1 - \pi[S^{\mathrm{c}}]\}}, \frac{\pi_{\Delta t}[S, S^{\mathrm{c}}]}{\min\{\frac{1}{2}, 1 - \pi[S]\}} \right\} = \psi(S; V).
\end{aligned}$$

Because of reversibility (corresponding to detailed balance): $\pi_{\Delta t}[S^{\mathrm{c}}, S] = \pi_{\Delta t}[S, S^{\mathrm{c}}]$, in the second equality, $\phi_p(S; V)$ becomes $\pi_{\Delta t}[S^{\mathrm{c}}, S]$ divided by the smaller divisor, either $\pi[S]$ or $\pi[S^{\mathrm{c}}]$. Then, the third equality is due to $\min\{\pi[S], \pi[S^{\mathrm{c}}]\} \leq 1/2 \leq \max\{\pi[S], \pi[S^{\mathrm{c}}]\}$. The fourth equality is simply because $\pi[S] + \pi[S^{\mathrm{c}}] = 1$. Since $S_n = S$ and both $\pi[S_k]$ and $\pi_{\Delta t}[S_k, S_k^{\mathrm{c}}]$ monotonously increase as $k$ increases, the numerator of $\psi(S_k; V_k)$ monotonously increases and the denominator of that monotonously decreases as $k$ increases.

Thus the inequality $\psi(S_1; V_1) \leq \cdots \leq \psi(S_n; V_n) = \phi_p(S; V)$ holds. ∎

## 4.3 Recursive Partition Scheme with Variable Divisions

The (first) MCC is naturally regarded as the network TS that dominates the "slowest" gross reaction over a reaction network. The second procedure in my algorithm is to further decompose the acquired subnetworks $S$ and $S^c$ recursively to a set of smaller subnetworks. This provides me with the hierarchical organization of timescale of reactions buried in a complexity of networks.

To make the tree graph facilitate that the the lower the hierarchy of the (super)basins on energy landscape the faster the timescale of reactions escaping from the (super)basins, I introduce a modified recursive partition scheme below. To avoid confusion, I use the term 'vertex' to indicate both vertex pairs and points without any children (called 'leaf' in graph theory[42]) by which the tree is defined, instead of using the term 'node' in the reaction network. For example, let me consider a simple three-states Markov chain given in Fig. 4.1 and focus only on the gray colored tree graph for the time being. Each of the three leaves of the gray tree graph in Fig. 4.1 (b) corresponds to each node in the reaction network in Fig. 4.1 (a).

A set of these leaves is defined by $V_{\text{Leaf}} := \{1, \cdots, n\}$. The rest of all the vertexes corresponds to the vertex pairs, that is, the MCCs of a hierarchical bi-partition. In similar to Fig. 5.3, the transition probabilities across the

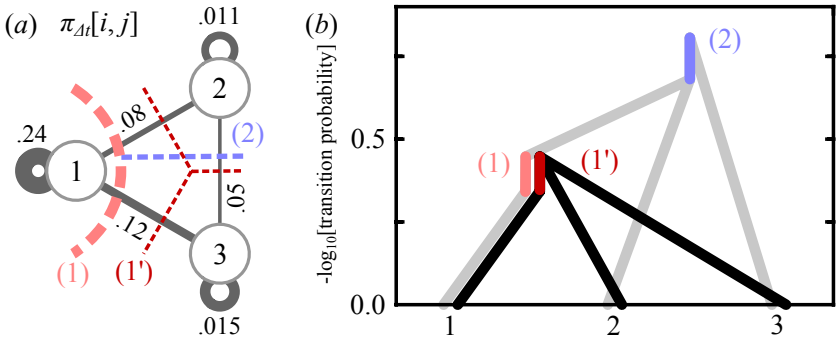



Figure 4.1: An illustrative example of a three-states Markov chain that is inconsistent with desired ordering of timescales by bi-partitioning. (a) the network of the Markov chain, and (b) its tree graphs. The transition probabilities $p_{\Delta t}[2|3] \approx .27, p_{\Delta t}[3|2] \approx .35$ are smaller than those $p_{\Delta t}[\{1\}|\{2,3\}] \approx .61, p_{\Delta t}[\{2,3\}|\{1\}] \approx .45$. In this case, the partition $(\{2\},\{3\})$ cannot be regarded as the appropriate partition compared to the partition $(\{2,3\},\{1\})$. See the Text in detail.

children's MCC is smaller than that of the parental one (i.e., yielding a larger free energy of activation in the children vertex). Since there must be $n-1$ such vertexes for the hierarchical bi-partition, I denote a set of such vertexes as $V_{\text{Fork}} := \{n+1, \cdots, 2n-1\}$. In addition, I denote the set of all the vertexes of the tree as $V_{\text{Tree}} := V_{\text{Leaf}} \cup V_{\text{Fork}}$.

I check whether the following condition holds or not for all pairs between a vertex $(T, T^{\text{c}})$ and its parent vertex $(S, S^{\text{c}})$ where $T^{\text{c}} := S \backslash T$ is a complement of $T$ relative to $S$:

$$\phi_p(S;V) = \max\{p_{\Delta t}[S|S^{\text{c}}], p_{\Delta t}[S^{\text{c}}|S]\} < \min\{p_{\Delta t}[T|T^{\text{c}}], p_{\Delta t}[T^{\text{c}}|T]\}. \quad (4.5)$$

This means that reactions across the cut $(T, T^{\text{c}})$ has shorter timescales than the timescales of reactions across the cut $(S, S^{\text{c}})$, which is consistent to the desired hierarchical ordering of timescales. If the condition Eq. (4.5) does

not hold, such vertex $(T, T^{\rm c})$ is discarded, and the children of the vertex (accompanied by a direct descendant) under $(T, T^{\rm c})$ (i.e., a part of "grand children" of $(S, S^{\rm c})$) are directly connected to the parent $(S, S^{\rm c})$. One can illustrate this modification by looking into the black tree graph in Fig. 4.1 (b). That is, the vertex (2) has smaller transition probabilities than those of the parent vertex (1), i.e., violating Eq. (4.5), and, hence, the vertex (2) is discarded and its children vertexes, that is, the leaves of the tree $\{2\}$ and $\{3\}$, become directly connected to the vertex (1) (namely, adopted as 'children' of the vertex (1)). Then, the new vertex (1') is interpreted as 'tri-partition' as in Fig. 4.1 (a).

Likewise, starting from vertexes whose children are all leaves toward the root vertex[42] (i.e., the 'ancestor' vertex of all vertexes), I first focus on vertexes whose children are all leaves and check whether Eq. (4.5) holds between each of these vertexes and its parent. I discard some vertexes such that Eq. (4.5) does not hold with their parents, and adopt the children of the discarded vertexes as new children of the grand parent simultaneously at the generation of the vertexes whose children are all leaves. Next I check and adopt recursively from the generation to the next generation (i.e., the parents and the grand parents,. . ., after modifiying the relation between vertexes whose children are all leaves and their parents) until Eq. (4.5) becomes satisfied for all remaining (child-parent relationship's) pairs of vertexes or vertexes in question along the algorithm are adopted to the root vertex.

# Chapter 5

# Demonstration in Model Systems

## 5.1 A Five State Markov Chain and Partition Methods

In this section, I demonstrate my algorithm by using a five-states Markov chain showing in Fig. 5.1-(a). I briefly explain a drawback of defining a network TS by minimizing flux over networks. The stationary flow $\pi_{\Delta t}$ is given in Table 5.1 and in Fig. 5.1-(a).

Table 5.1: $\pi_{\Delta t}[i, j]$ and $\pi[i]$.

| $i$ | $j=1$ | $j=2$ | $j=3$ | $j=4$ | $j=5$ | $\pi[i]$ |
|---|---|---|---|---|---|---|
| 1 | 0.096 | 0.032 | 0.041 | 0.001 | 0.022 | 0.192 |
| 2 | 0.032 | 0.050 | 0.016 | 0.001 | 0.001 | 0.100 |
| 3 | 0.041 | 0.016 | 0.122 | 0.064 | 0.001 | 0.244 |
| 4 | 0.001 | 0.001 | 0.064 | 0.137 | 0.071 | 0.274 |
| 5 | 0.022 | 0.001 | 0.001 | 0.071 | 0.095 | 0.190 |

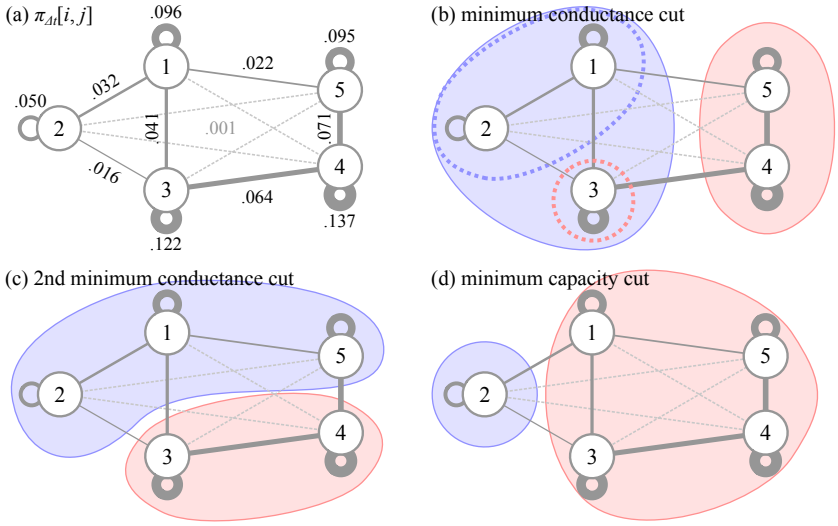



Figure 5.1: A five-states Markov chain. The thickness of links (including self-transitions, i.e., self-loops) reflects the size of $\pi_{\Delta t}[i,j]$. (a) $\pi_{\Delta t}[i,j]$, (b) The first MCC to separate the whole network into two disjoint subnetworks $\{1,2,3\}$ and $\{4,5\}$, (c) The second MCC to separate the whole network into two disjoint subnetworks $\{1,2,5\}$ and $\{3,4\}$, (d) A cut to most minimize the flux over the entire network, that is, $(\{2\},\{1,3,4,5\})$.

Table 5.2: The list of all the possible partitions and their related quantities.

| $S$ | $S^c$ | $\pi_{\Delta t}[S,S^c]$ | $\pi[S]$ | $\pi[S^c]$ | $\phi_p(S;V)$ | $p_{\Delta t}[S^c\|S]$ | $p_{\Delta t}[S\|S^c]$ |
|---|---|---|---|---|---|---|---|
| 1 | 2, 3, 4, 5 | 0.096 | 0.192 | 0.808 | 0.500 | 0.500 | 0.119 |
| 2 | 1, 3, 4, 5 | **0.050** | 0.100 | 0.900 | **0.500** | 0.500 | 0.056 |
| 3 | 1, 2, 4, 5 | 0.122 | 0.244 | 0.756 | 0.500 | 0.500 | 0.161 |
| 4 | 1, 2, 3, 5 | 0.137 | 0.274 | 0.726 | 0.500 | 0.500 | 0.189 |
| 5 | 1, 2, 3, 4 | 0.095 | 0.190 | 0.810 | 0.500 | 0.500 | 0.117 |
| 1, 2 | 3, 4, 5 | 0.082 | 0.292 | 0.708 | 0.281 | 0.281 | 0.116 |
| 1, 3 | 2, 4, 5 | 0.136 | 0.436 | 0.564 | 0.312 | 0.312 | 0.241 |
| 1, 4 | 2, 3, 5 | 0.231 | 0.466 | 0.534 | 0.496 | 0.496 | 0.433 |
| 1, 5 | 2, 3, 4 | 0.147 | 0.382 | 0.618 | 0.385 | 0.385 | 0.238 |
| 2, 3 | 1, 4, 5 | 0.140 | 0.344 | 0.656 | 0.407 | 0.407 | 0.213 |
| 2, 4 | 1, 3, 5 | 0.185 | 0.374 | 0.626 | 0.495 | 0.495 | 0.296 |
| 2, 5 | 1, 3, 4 | 0.143 | 0.290 | 0.710 | 0.493 | 0.493 | 0.201 |
| 3, 4 | 1, 2, 5 | 0.131 | 0.518 | 0.482 | **0.272** | 0.253 | 0.272 |
| 3, 5 | 1, 2, 4 | 0.215 | 0.434 | 0.566 | 0.495 | 0.495 | 0.380 |
| 4, 5 | 1, 2, 3 | **0.090** | 0.464 | 0.536 | **0.194** | 0.194 | 0.168 |

Let me compare the (first) MCC and the second MCC, and a cut to minimize the flux over the network. Note that the cut to minimize the flux over the network could be regarded as a generalization of variational transition state theories (TST) that have been successfully applied to molecules [52, 79] to complex reaction networks. The list of all the possible partitions is shown in Table 5.2. It is found that the MCC is $(\{1,2,3\},\{4,5\})$ (see also Fig. 5.1-(b)).

If the bound $\alpha$ estimated from the spectral clustering is less than the second minimum conductance value, then only the possible cut whose conductance value is less than the value of $\alpha$ is the MCC to be acquired. However in this model, the bound of the spectral clustering (see Theorem 1 in Ref. 72) is 1.080 ($= 2\sqrt{\lambda}$ , where the $\lambda$ is the second smallest eigenvalue of the normalized graph Laplacian $L$), so that there is no guarantee for finding the (first) MCC in terms of spectral clustering. As demonstrated in the figure, the cut yielding the second minimum conductance value, i.e., $(\{1,2,5\},\{3,4\})$ with $\phi_p = 0.272$ (Fig. 5.1-($c$)) is not necessarily close (e.g., in the sense of Hamming distance) to the MCC $(\{1,2,3\},\{4,5\})$ with $\phi_p = 0.194$ (Fig. 5.1-($b$)) even if the conductance values are relatively close to each other.

The cut across which the flux (=gross rate) is minimum, called minimum capacity cut in graph theory[44], (i.e., corresponding to variational TST) is given by a cut $(\{2\},\{1,3,4,5\})$, that is, $\pi(\{2\},\{1,3,4,5\}) = 0.050$ is the smallest over those of the other cuts. Note however that this cut has a the largest gross rate constant in this example (i.e., $\phi_p(\{2\},\{1,3,4,5\}) = 0.500$), that is, faster timescale across the cut than the timescales across the MCC. This implies that the cut to minimize the flux over the network does not

necessarily serve as a reaction barrier. The well-established variational TST in molecules [52, 79] requires *a priori* knowledge of an approximate location of TS, that is, near a saddle linking a pair of basins, and it variationally optimizes the location of TS so as to minimize the flux locally starting from the approximate location of the TS. If the variational TST would search the location of TS more globally over the whole region on a potential energy surface, without any *a priori* knowledge on the location of saddle, then the surface across which the flux is minimized may not be proper for defining the TS that should dominate the reaction. Referring potential energy landscapes, a surface across which the flux is minimized can be located near an edge of potential energy surface. In case, the shallow basin at high energy will partitioned by the surface. To such a high potential energy edge, the system hardly reaches (i.e., a very small transition probability) but easily goes down to the basin (i.e., a very high transition probability). This is probably the present situation: in Table 5.2 one can see in this example that $\pi_{\Delta t}[\{1,3,4,5\}|\{2\}]$ is the largest (0.500) but $\pi_{\Delta t}[\{2\}|\{1,3,4,5\}]$ is the smallest (0.056). One possible prescription proposed previously [39] is to use the so-called balanced Minimum Cut, that is, roughly corresponding to finding a cut to minimize the flux under such condition that two disjoint subnetworks, say, $R$ and $P$, divided by the cut, have smaller $|P|$ (i.e. number of nodes in $P$) than the minimum capacity cut. This prescription may work for cases that actual reaction barrier to dominate the timescale of reactions divides the whole space into two subspaces of similar size $|S| \approx |S^{\mathrm{c}}|$. However, in general, one cannot know *a priori* that this condition is met for a given network in question.

## 5.2 Demonstration of Algorithm

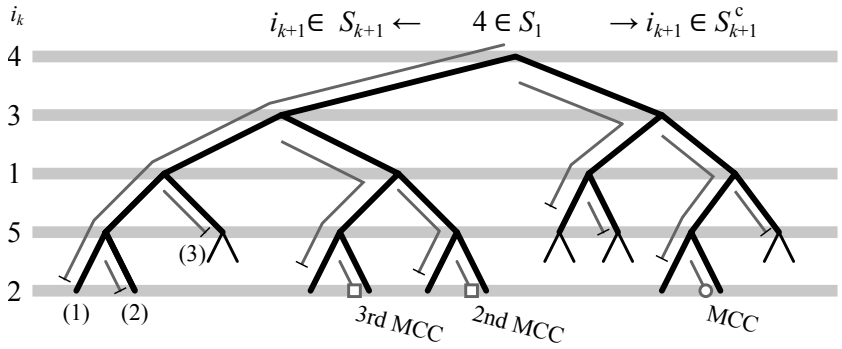



Figure 5.2: Binomial decision tree for a greedy search of the MCC of the five-states Markov chain in Fig. 5.1 with initial $\alpha = 0.3$.

In the following, I demonstrate an algorithm in terms of a binomial decision tree as shown in Fig. 5.2. First, I set a permutation of the indexes of the nodes to $\{4, 3, 1, 5, 2\}$ in descending order of their equilibrium distributions $\pi$, $\pi[4] \geq \pi[3] \geq \pi[1] \geq \pi[5] \geq \pi[2]$. Let me assign $4 \in S_1 \subset V$ and set $\alpha = 0.3$ in the beginning ($\alpha$ was determined using a heuristics described in the previous chapter). The thick black lines are branches of the binomial decision tree to enumerate all the possible partitions of the five-states reaction network in Fig. 5.1. The thick gray horizontal lines indicate the first, second, ..., the fifth decisions. The MCC-searches are performed along the branches of the decision tree, and the thin gray lines along the binomial decision tree indicate "trajectories" of the searches. The stop signs $\dashv$ on the end of the lines indicate either that searches for the rest of all the branches below are discarded due to $\psi(S_k; V_k) > \alpha$ at the $k$th decision ($k < n(= 5)$) or that searches ending up with the $n$th decision have $\phi_p(S_n; V_n)$ whose value is larger than $\alpha$. The other signs $\bigcirc$ and $\square$ in the figure indicate the partitions of the

MCC, and those of the second and third MCCs whose conductance value are smaller than $\alpha$, respectively. The search is started from assigning the most populated node (i.e., node 4 in this example) to a subset $S_1$ and leaving all the others unassigned (cf. Fig. 2.1). If a trajectory (a thin line) goes to the left (right) side from the $k$th decision (e.g., at $k = 1$ and $i_1(= 4) \in S_1$), the next most populated node $i_{k+1}$ is assigned to the subset $S_{k+1}$ ($S^c_{k+1}$) (e.g., at $k = 2$, $\{i_1(= 4), i_2(= 3)\} \in S_2$ in the left branch from the top).

For example, trajectory (1) (the thin line denoted by (1)), all the way down to the left from the top, reaches the partition $(\{1, 2, 3, 4, 5\}, \emptyset)$. There exists no actual partition. Hence, I discard such partition. Trajectory (2) reaching $(\{1, 3, 4, 5\}, \{2\})$ is turned out to have $\phi_p(\{2\}) = 0.5 > \alpha$. Trajectory (3) discards any further down-search beyond $(\{1, 3, 4\}, \{5\})$ because the $\psi$ value exceeds $\alpha$. This implies that these three trajectories do not provide any candidate of the MCC.

When the algorithm finds partitions that satisfy the condition of $\phi_p(S) < \alpha$ (indicated as '3rd MCC' and '2nd MCC' in the figure), the value of $\alpha$ is updated with a smaller conductance value than the smallest one found so far during the search process. Likewise, my current algorithm searches trajectories on the binomial decision tree sequentially with updating the value of $\alpha$ and discarding any further down-search when the $\psi$ value exceeds the (updated) value of $\alpha$. This is how my branch-and-bound method works. In this example, my algorithm reaches the MCC (a circle in Fig. 5.2) and, in all the rest of searches or trajectories, the condition of $\phi_p(S) < \alpha$ was not satisfied and thus the MCC and its conductance value $\alpha$ were kept as they are up to the end of all the searches. Note that the MCC to give the *smallest*

conductance value over *all* the possible cuts are only identified, in principle, for completing all searches on the binomial decision tree.

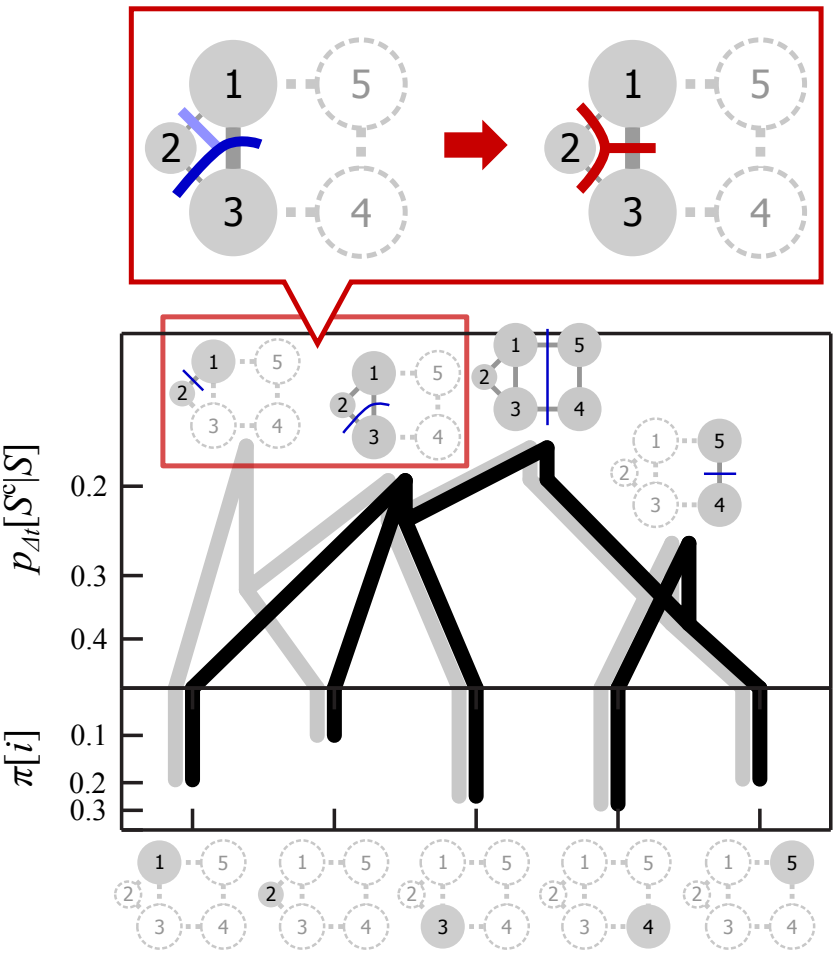



Figure 5.3: Timescale hierarchical tree of the five-states Markov chain in Fig. 5.1. The gray colored tree graph is based on recursive bi-partitions in which each resulting subnetwork is divided into two disjoint (sub)networks in each hierarchical decomposition. The black colored tree graph is based on a modified recursive partition scheme whose number of divisions changes adaptively. Since the cut $(\{1\},\{2\})$ has a longer reaction timescale than that the cut $(\{1,2\},\{3\})$ has, the bi-partition $(\{1,2\},\{3\})$ is replaced by the tri-partition $(\{1\},\{2\},\{3\})$ as the right partition. See the Text in more detail.

The (first) MCC is naturally regarded as the network TS that dominates

the "slowest" gross reaction over a reaction network. The second procedure in my algorithm is to further decompose the acquired subnetworks $S$ and $S^c$ recursively to a set of smaller subnetworks. This provides me with the hierarchical organization of timescale of reactions buried in a complexity of networks. For example if the first partition to yield the MCC is achieved by $(S, S^c)$ the algorithm is continued to search the MCC for the resulting subnetworks $S$ and $S^{\mathrm{c}}$. For instance, the MCC in this five-states Markov chain is $(\{1,2,3\},\{4,5\})$. Then I further search the MCCs for $\{1,2,3\}$ and $\{4,5\}$ individually. Here, I consider to search the MCC only for $\{1,2,3\}$ because it is trivial that the only meaningful partition of $\{4,5\}$ is $(\{4\},\{5\})$. All the possible partitions of $\{1,2,3\}$ are listed in Table 5.3. According to Table 5.3, the MCC of $\{1,2,3\}$ is found to be $(\{1,2\},\{3\})$ whose value of the conductance $\phi_p(S)$ is 0.234. Then the MCC of $\{1,2\}$ is expected to be $(\{1\},\{2\})$. Fig. 5.3 gives the corresponding tree graph (gray colored) that is based on the recursive bi-partitions in which the obtained subnetworks are imposed to be divided into two disjoint (sub)networks. One can see that $\min\{p_{\Delta t}[\{1\}|\{2\}], p_{\Delta t}[\{2\}|\{1\}]\}(= p_{\Delta t}[\{2\}|\{1\}])$ of the partition $(\{1\},\{2\})$ is larger than $\max\{p_{\Delta t}[\{1,2\}|\{3\}], p_{\Delta t}[\{3\}|\{1,2\}]\}(= p_{\Delta t}[\{1,2\}|\{3\}])$ of the partition $(\{1,2\},\{3\})$, i.e., the (longer) timescale of the reaction $\{1\} \to \{2\}$ exceeds the (shorter) timescale of the reaction $\{3\} \to \{1,2\}$. This is inconsistent with expectation that the lower the hierarchy of the (super)basins on energy landscape, i.e., from "a parent" superbasin to the "children" (super)basins, the faster the timescale of reactions escaping from the children (super)basins. This is due to the fact that my algorithm is based on bi-partitions irrespective of which subnetworks are divided. In what follows,

I present a modified recursive partition scheme whose number of divisions changes adaptively.

Table 5.3: The list of all the possible cuts of $\{1, 2, 3\}$ and their related quantities

| $S$ | $S^{\mathrm{c}}$ | $\pi_{\Delta t}[S, S^{\mathrm{c}}]$ | $\pi[S]$ | $\pi[S^{\mathrm{c}}]$ | $\phi_p(S)$ | $p_{\Delta t}[S^{\mathrm{c}}\|S]$ | $p_{\Delta t}[S\|S^{\mathrm{c}}]$ |
|---|---|---|---|---|---|---|---|
| 1 | 2, 3 | 0.073 | 0.192 | 0.344 | 0.380 | 0.380 | 0.212 |
| 2 | 1, 3 | 0.048 | 0.100 | 0.436 | 0.480 | 0.480 | 0.110 |
| 3 | 1, 2 | 0.057 | 0.244 | 0.292 | **0.234** | 0.234 | 0.195 |

# Chapter 6

# Demonstration in Chemical System

The Claisen rearrangement of AVE [80, 81] proceeds from allyl vinyl ether (3-(ethenyloxy)prop-1-ene) into 4-Pentenal (pent-4-enal), taking several hours at room temperature:

$$\mathrm{CH_2{=}CH{-}CH_2{-}O{-}CH{=}CH_2}$$
$$\overset{\Delta}{\rightleftharpoons} \mathrm{CH_2{=}CH{-}CH_2{-}CH_2{-}CH{=}O.}$$

## 6.1 *Ab Initio* Calculation and Rate Constants

The original reaction network was obtained previously by single component artificial force induced reaction method at B3LYP/6-31G level [80]. In this study, the reaction network was recalculated by a more accurate computational method CCSD(T)/jun-cc-pVTZ//M062X/6-311+G(2d,p).

The method searches minima and index-one saddles connecting them. I here assume that reaction events are dominated mainly through the index-one saddles as transition state theories assume. Note that my method proposed in this thesis can be applied to first-order rate equations or a discrete-time Markov chain, and the approximation based on the transition state theory is not necessarily required and only for illustrative demonstrations.

## 6.2 Derived Reaction Network

An *ab initio* electronic structure calculation of the Claisen rearrangement of AVE reveals a complex potential energy surface that consists of 23 minima and 66 saddles linking them[80]. Figure 6.1 presents the reaction network of the Claisen rearrangement of AVE. Each numbered node in the figure corresponds to each potential energy minimum (i.e., conformer) with the corresponding energy basin. Each edge/link connecting two nodes corresponds to the saddle connecting between the two basins. No edge, e.g., between node 1 and node 6, implies that no saddle connecting between the two was found by the single component artificial force induced reaction method. The MCC search and its recursive applications are carried out at temperature $T = 473.15$ K($= 200$ °C) that reflects rate constants between nodes with $\Delta t \approx 13$ fs. This timescale was chosen such that it is faster than the reaction timescale between nodes 14 and 15, which is the fastest reaction timescale in the network, i.e., $1/k_{\mathrm{Max}} (= 1/k_{14,15})$, such that $1 - e^{-k_{\mathrm{Max}}\Delta t} = 0.05$.

The dashed line cutting links between nodes 1 and 18 and nodes 7 and 17 is the network TS assigned by the first application of the MCC that has

the minimum value of graph conductance $\phi_p$ defined by Eq. (3.16). This TS divides the entire network to the two subnetworks, i.e., a network consisting of the reactant superbasin (right, 1-10) (AVE) and that of the product superbasin (left, 11-23) (4-petenal). Each arc or ring represents a superbasin in which the equilibration over a subnetwork occurs for a timescale represented by the corresponding color-grade (the detailed computation of arc/ring will be presented in the section of Fig. 6.1). For example, the system is equilibrated over a subnetwork composed of nodes 14 and 15 within the product superbasin at the fastest timescale $\sim 10^{-12.5}$ s although the equilibration over a subnetwork composed of nodes 11-15 and that of nodes 16-23 takes a much longer timescale.

## 6.3 Computation Cost

For problems which are NP-hard or NP-complete, it is known that the branch-and-bound method does not improve the worst case computation cost. However, in practice, the method does improve its cost (e.g. deceleration of exponential increase of the cost). In case of my algorithm applying to the reaction network associated with Claisen rearrangement of AVE, while the maximum number of trajectories to be searched on the binomial decision tree is $\sum_{k=2}^{23} 2^{k-1} = 2(2^{23-1} - 1) = 8\ 388\ 606$ (where 23 is the total number of nodes (i.e., conformers) and the minus one comes from the fact that the first decision is unimportant due to the symmetry $(S, S^{\mathrm{c}}) = (S^{\mathrm{c}}, S)$), the total number of trajectories actually searched on the binomial decision tree by setting an initial value of $\alpha$ to an approximate minimum conductance value

estimated by the spectral clustering (see Preliminary chapter) was found to be about three times of the depth of the binomial decision tree (i.e., total number of nodes minus one) $\approx (23-1)\times 3(=66)$. It is noted worthy that to single out or identify *the* trajectory ending up with the desired MCC requires, at least, $(23-1)\times 2(=44)$, that is, the depth of the tree multiplied by the factor of two, left and right branches along the trajectory at each decision.

The corresponding tree graph is shown in Fig. 6.2. The vertical axis gives the inverse of the rate constant $k_{S,T}$ across the network TS from the (super)basin $S$ to $T$, and the free-energy-of-activation $\Delta G^{\ddagger}_{S,T}$ from the (super)basin $S$ to $T$. The overall rate constants are derived from Eq. (4.2) by using the two transition probabilities $P_{\Delta t}[S|T]$ and $P_{\Delta t}[T|S]$ calculated by Eq. (3.17), and each rate constant is translated into the free-energy-of-activation by Eq. (3.3) assuming a single reaction channel:

$$\Delta G^{\ddagger}_{ij} = -RT \ln \left[ \sum_{k=1}^{m_{ij}} \exp[-\Delta G^{(k)\ddagger}_{ij}/RT] \right].$$

One can easily see the existence of two large superbasins separated by a large free-energy-of-activation, and more importantly the timescale hierarchy of reactions, e.g., nodes 14 and 15 are separated by the lowest free-energy-of-activation over which the system reacts at the fastest timescale.

## 6.4 Quasi-equilibrium Over Superbasins

Here I scrutinize the validation of the overall rate constant of the Claisen rearrangement of AVE $k_{\mathrm{Re}\to\mathrm{Pr}}(T)(=k_{\mathrm{Pr,Re}})$ whose network TS, as defined

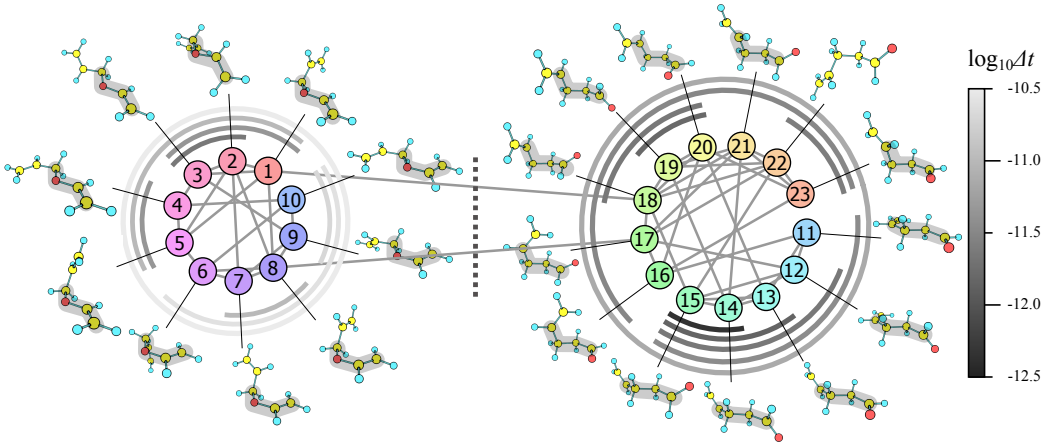



Figure 6.1: The reaction network of the Claisen rearrangement of AVE. See the main text for full detail.

by the first application of the MCC, exhibits Arrhenius-type temperature dependence.

Any form of TS theories assumes quasi-equilibrium before crossing the TS from the state of the reactant. The quasi-equilibrium results from the existence of timescale separation between the timescale of relaxation within the reactant state before transition to the TS and then to the product. Even for complex reaction networks where several timescales may coexist, it may be natural to postulate quasi-equilibrium over a subnetwork from which the reaction occurs, as a generalization of the original TS theories. This corresponds to the situation that any initial condition prepared in the subnetwork (i.e., the node from which system starts) will be lost, thus providing a unique overall rate constant from one subnetwork to another.

Whether quasi-equilibrium is approximately attained over the extracted, network TS can be examined by checking similarity of column vectors of the $n \times n$ transition matrix $P_{\Delta t}$ as a function of $\Delta t$ ($n$: the number of nodes).

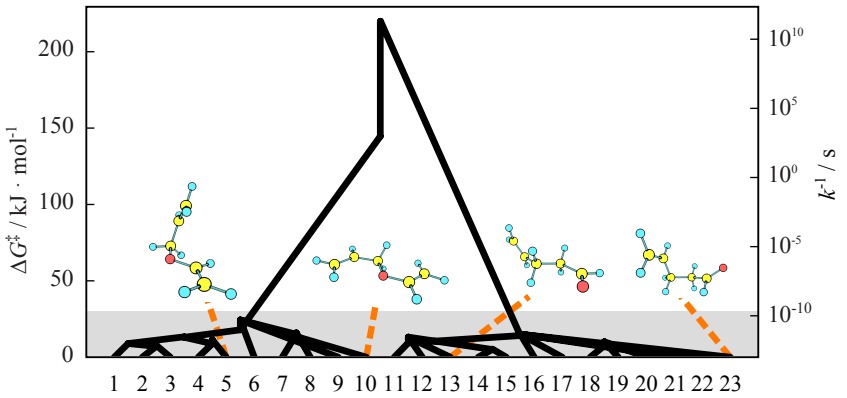



Figure 6.2: The tree of reaction timescale hierarchy of the Claisen rearrangement of AVE at $437.15K(= 200°\mathrm{C})$. The tree is derived from the recursive application of MCC (see Fig. 2.1). The vertical axis gives the inverse of the rate constant $k_{T,S}$ across the network TS from the (super)basin S to T (right) and the free-energy-of-activation $\Delta G^{\ddagger}_{T,S}$ from the (super)basin S to T (left). Because the reactant and the product superbasins of AVE are well separated by a large barrier, the vertical $\Delta t$-axis is log-scaled to clarify the timescale relationship among conformers of AVE. The gray area $(0 - 30\ \mathrm{kJ \cdot mol^{-1}})$ is magnified in Fig. 6.5-(a). The horizontal axis gives the indices of nodes assigned in Fig. 6.1. The most stable structure in each of superbasins $\{1, 2, \ldots, 5\}, \{6, \ldots, 10\}, \{11, \ldots, 15\}, \{16, \ldots, 23\}$ is depicted, and each orange dashed line indicates to which node the conformer belongs. Each of four superbasins was found to have a stiff structural motif which does not change (except structure 22) in the superbasin as shown in a gray shade on each structure in Fig. 6.1.

Suppose that the system starts from the node $i$ at time $t = 0$ and it propagates in time through the network: the $j$th component of $n$-dimensional column vector of the initial distribution $p_{t=0}$ is given by $p_{t=0}[j] = \delta_{ij}$ ($\delta_{ij}$: Kronecker delta), and the time propagation initiated at the node $i$ is represented by $P_{\Delta t} p_{t=0} = (p_{\Delta t}[1|i], \ldots, p_{\Delta t}[n|i])^{\mathrm{T}}$ which is equal to the $i$th column vector of the transition matrix $P_{\Delta t}$. That is, the $i$th column vector of the matrix $P_{\Delta t}$ is interpreted as the $\Delta t$-time propagated probability distribution starting from the node $i$ at time $t = 0$. In this article I quantify the similarity

of any pair of two column vectors of $P_{\Delta t}$ in terms of Hellinger distance $d_H$, which is defined by $d_H(x,y) := (\sum_i(\sqrt{x_i}-\sqrt{y_i})^2/2)^{1/2}$ satisfying $0 \leq d_H \leq 1$ when the vectors $x$, and $y$ are probability distributions. I regard two column vectors, e.g., the $i$th and $j$th column vectors, $(p_{\Delta t}[1|i],\ldots,p_{\Delta t}[n|i])^{\mathrm{T}}$ and $(p_{\Delta t}[1|j],\ldots,p_{\Delta t}[n|j])^{\mathrm{T}}$ as the same when $d_H$ between the two is smaller than a certain threshold (I set the value to be 0.05). When column vectors become regarded as the same at all $\Delta t$ longer than $\tau$, $\tau$ is identified as the timescale to merge the two column vectors.

It is found for the Claisen rearrangement of AVE at $T = 473.15$ K that the quasi-equilibrium is attained in the reactant subnetwork at the timescale of $\Delta t \geq 10^{-10.50}$ s $\approx$ 32 ps. The column values of the transition matrix $P_{\Delta t}$ are shown with $\Delta t = 10^{-12.35}$ s and $\Delta t = 10^{-10.50}$ s in Fig. 6.3. As seen in Fig. 6.3-(a), at $\Delta t = 10^{-12.35}$ s, two column vectors

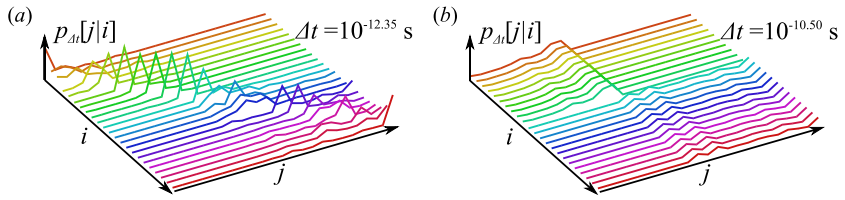



Figure 6.3: The column vectors $(p_{\Delta t}[1|i],\ldots,p_{\Delta t}[n|i])^{\mathrm{T}}$ of the transition matrix $P_{\Delta t}$ with timescale $\Delta t$ of (a) $10^{-12.35}$ s and (b) $10^{-10.50}$ s. The index $i$ is the same as the node $i$ and its color shown in Fig. 6.1. The equivalence of the two types of distributions observed at $10^{-10.50}$ s was quantified by the Hellinger distance $d_H$ between any pair of the column vectors having $i$ belonging to either of subsets $\{1,\ldots,10\}$ and $\{11,\ldots,23\}$, which was smaller than the threshold, $d_H < 0.05$.

14 and 15 yield a similar probability distribution though the rest do not. As the timescale $\Delta t$ increases, at $\Delta t = 10^{-10.50}$ s only two types of probability distributions remain as seen in Fig. 6.3-(b), which result from the

initial conditions starting from any node in the subset of nodes 1-13 (i.e., reactant subnetwork) and those from any node in the subset of nodes 14-23 (i.e., product subnetwork). This implies that the detailed information from which node the system starts in each subnetwork was lost. Because of the similarity, $\sum_{j\in \mathrm{Pr}} p_{\Delta t}[j|i_1] \approx \sum_{j\in \mathrm{Pr}} p_{\Delta t}[j|i_2]$ for any pair of $i_1$ and $i_2$ belonging to the reactant subnetwork Re and $\sum_{j\in \mathrm{Re}} p_{\Delta t}[j|i_1] \approx \sum_{j\in \mathrm{Re}} p_{\Delta t}[j|i_1]$ for any pair of $i_1$ and $i_2$ belonging to the product subnetwork Pr are satisfied. By using this relation, one can approximate $p_{\Delta t}[\mathrm{Pr}|\mathrm{Re}]$ and $p_{\Delta t}[\mathrm{Re}|\mathrm{Pr}]$ from $\sum_{j\in \mathrm{Pr}} p_{\Delta t}[j|i], i \in \mathrm{Re}$, and $\sum_{j\in \mathrm{Re}} p_{\Delta t}[j|i], i \in \mathrm{Pr}$, respectively. Then one can derive rate constants of the rearrangement through Eq. (4.2) by using a pair of approximated transition probabilities. By checking the convergence, I obtained converged rate constants (with less than 0.001% error) at $\Delta t = 10$ $\mu$s, and I used these rate constants to validate the MCC-based rate constants (see "direct estimation from $P_{\Delta t}$" in Fig. 6.4 and Table 6.1).

## 6.5 Arrhenius Plot

Now let me examine temperature dependency of rate constants of the Claisen rearrangement of AVE among several approximations; namely, the direct solution of the overall rate constant from the transition matrix $P_{\Delta t}$ with $\Delta t = 10$ $\mu$s, the MCC-based rate constant (regarded as MCC-based transition state theory (TST)) assuming $\Delta t \approx 13$ fs and the equilibrium distribution $\pi$, and the conventional rate-limiting step approximations.[82] Fig. 6.4 and Table 6.1 present the Arrhenius-plot and the estimated pre-exponential factor $A$ and activation energy $E_a$, respectively. One can see that the direct

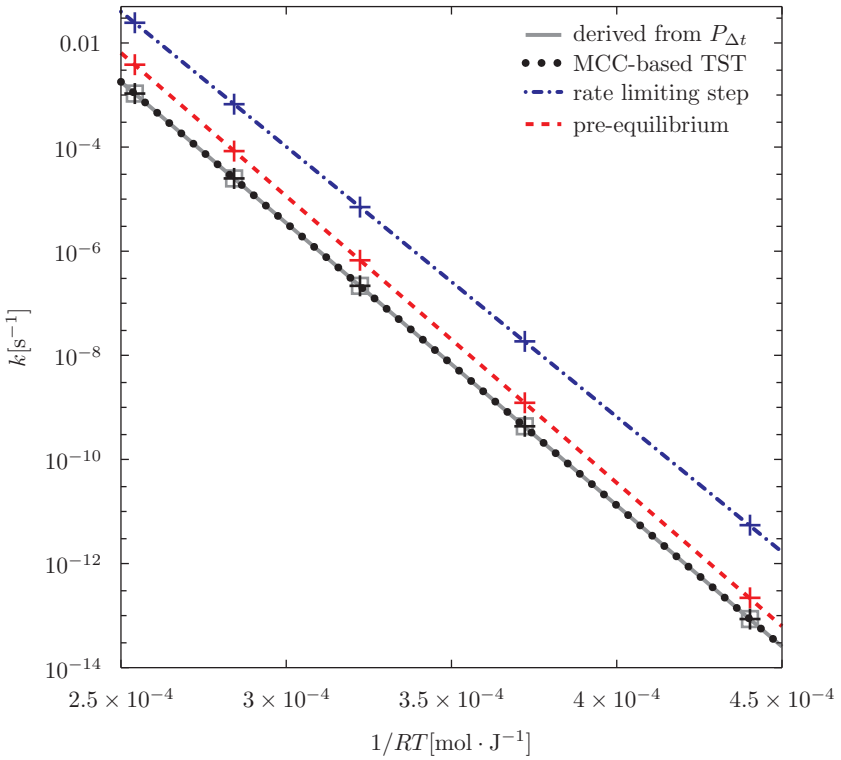



Figure 6.4: Temperature dependency of the rate constant of the Claisen rearrangement of AVE at $T = 273.15$ K, 323.15 K, 373.15 K, 423.15 K, 473.15 K, and 523.15 K. Each point is calculated by the direct estimation from the transition matrix $P_{\Delta t}$ (gray square), the MCC-based TST (black cross), the rate constant of rate-limiting step (blue cross), the rate constant of pre-equilibrium assumption (red cross).

Table 6.1: The estimated pre-exponential factor $A$ and activation energy $E_a$ from the Arrhenius-type temperature dependency ($k = A\exp[-E_a/RT]$) of the Claisen rearrangement of AVE. The values $A$ and $E_a$ are estimated by using least squares method.

| | $A[\mathrm{s}^{-1}]$ | $E_a$ $[\mathrm{kJ} \cdot \mathrm{mol}^{-1}]$ |
|---|---|---|
| Direct estimation from $P_{\Delta t}$ | $6.53 \times 10^{10}$ | 125 |
| MCC-based TST | $6.53 \times 10^{10}$ | 125 |
| rate-limiting step | $3.78 \times 10^{11}$ | 120 |
| Pre-equilibrium assumption | $3.78 \times 10^{11}$ | 127 |

estimation from $P_{\Delta t}$ and the MCC-based TST provide the same quantities of the pre-exponential factor $A = 6.53 \times 10^{10}$ $\mathrm{s}^{-1}$ and the activation energy

$E_a = 125$ kJ $\cdot$ mol$^{-1}$. Note that my MCC-based TST is based on a dividing surface over a given network with equilibrium distribution across which the passage is the slowest among all the possible cuts, which is the reasonable definition of network TS of reaction networks. However, if quasi-equilibrium is not attained in the reactant (and also product) subnetwork before the reaction takes place, or if the timescale of passage over the second slowest MCC would be comparable to that over the slowest MCC, the MCC-based TST is not expected to coincide with the direct solution from $P_{\Delta t}$. It should be noted that the approximation based on the rate-limiting step that is from node 1 to 18 presents a faster rate constant over a range of temperatures (shifted almost constantly from $k$ evaluated from the direct solution from $P_{\Delta t}$ and the MCC-based TST) (Fig. 6.4). The rate constant of the rate-limiting step has a lower activation energy and a larger pre-exponential factor than those of the direct solution from $P_{\Delta t}$. The pre-equilibrium assumption (classified as one of the rate-limiting step approximations in this thesis) relies on a single sequential reaction path from the most stable node (basin) to a saddle associated with the rate-limiting step reaction through several nodes (basins), in which thermal equilibria are assumed to attain between any consecutive reactions.[82] The pre-equilibrium assumption provides an estimation of the free-energy-of-activation that is derived from the difference between the saddle point free energy for the rate-limiting step reaction and the free energy of the most stable basin among basins in the reactant superbasin. Table 6.1 tells me that the pre-equilibrium assumption correctly captures the activation energy but does not capture the pre-exponential factor. The discrepancy observed in pre-exponential factor $A$ (corresponding to the constant vertical

shift in Fig. 6.4) is interpreted as follows: $A$ corresponds to $\sim \exp(\Delta S^{\ddagger}/R)$ ($\Delta S^{\ddagger}$: the entropy change from that of "TS" to that of the reactant state) if the temperature dependence of the entropy change is negligible.

The larger $A$ in the rate-limiting step approximations suggests that the entropy loss from the reactant superbasin to the "TS" (note that the definition of the TS is different from the MCC) is estimated to be smaller than that by the direct solution of $P_{\Delta t}$ and the MCC-based TST, both of which take into account multiple pathways to reach the TS of the reaction.

The activation energy of the MCC-based TST 125 kJ $\cdot$ mol$^{-1}$ fairly coincides with that estimated by a gas phase experiment 127 kJ $\cdot$ mol$^{-1}$ in Ref. 83, although the pre-exponential factor $A$ does not. The deviation observed in $A$ may arise, for example, from the level of the *ab initio* calculation I used, the experimental accuracy, and/or the assumption of TST such as no-recrossing assumption and no significant quantum effect.

## 6.6 Hierarchical Organization of Reaction Timescales

Here I investigate whether not only the first application of the MCC but also the second, third, ... application of the MCCs capture multiple timescales of the underlying complex kinetics by using the direct solution of transition matrix $P_{\Delta t}$ with various $\Delta t$. Note that, in some cases, it is more natural to partition a network into three or more subnetworks compared to cases that MCC always partitions a network into two subnetworks. I generalize my scheme by introducing an algorithmic correction (see previous Chapter for more detail) and call the resultant tree 'a tree of reaction timescale hierarchy'.

### 6.6.1 Hierarchies of Merging

Fig. 6.5 (b) shows the transition matrix of $\Delta t$ propagation $p_{\Delta t}[j|i]$. The system is interpreted as initiating from the node $i$ (at $\Delta t = 0$ s) and propagates through the network (from $\Delta t = 10^{-12.35}$ s to $10^{-10.50}$ s). In Fig. 6.5 (b), each colored distribution gives the column vector $p_{\Delta t}[\cdot|i]$ of the transition matrix at a given $\Delta t$ whose color indicates from which node the system starts. As time $\Delta t$ goes by some of the column vectors start to merge. I monitored time to merge by checking whether the Hellinger distance among a set of column vectors is less than the threshold of 0.05. The short and long horizontal color bars in Fig. 6.5 (b) represent a set of initial nodes whose column vectors merge into some single distribution (see also the list of time to merge in Fig. 6.6).

As an overall trend, the observed order of time to merge in Fig. 6.5 (b) qualitatively agrees with the information of the tree of reaction timescale hierarchy in Fig. 6.5 (a). The tree not only reproduces the overall rate constant over the first application of the MCC in the network but also captures the hierarchical nature of timescales in the network.

### 6.6.2 Hierarchies of Conformers

The gray circles and arcs in Fig. 6.1 represent a series of short and long bars appearing in Fig. 6.5(b)-2, ..., Fig. 6.5 (b)-8. Referring to the conformers associated with individual nodes, one can infer the origin of stability of conformations. For example, the conformations of the carbon chain $=CH-CH_2-CH_2-CHO$ (indicated by a gray shade in Fig. 6.1) are different

between the conformers 11-15 and 16-23 in the product subnetwork. The conformers 11-15 and 16-23 differ with each other in the C-C-C-C dihedral angle for the $=CH-CH_2-CH_2-CHO$ part. The former 11-15 have the C-C-C-C angle of 174°-180° but the latter 16-23 have that of 60°-75°. The rate constants reveal that the rotation of this dihedral angle is relatively slow. Likewise, one can also consider the origin of stability in the reactant subnetwork. The conformations of $-CH_2-O-CH=CH_2$ are different between the conformers 1-5 and 6-10. The MCC-based tree shows that the rotation around the C-O-C-C dihedral angle for the $-CH_2-O-CH=CH_2$ part is slow. This is consistent to my chemical knowledge that there may be a $n-\pi$ conjugation between the lone-pairs of the O atom and the out-of-plane $\pi$-orbital of the C=C double bonds. This $n-\pi$ conjugation would explain the slow C-O-C-C rotation observed in the present analysis. My MCC-based approach can shed light on the stability of conformations by referring to the classification of their kinetic properties.

## 6.7 The List of Time to Merge

In this section, I show a detailed list of time to merge as a function of $\Delta t$ in the reaction network of Claisen rearrangement of AVE. Each black or pink horizontal bar in Fig. 6.6 indicates a set of nodes in a subnetwork $S$ satisfying $d_H(p_{\Delta t}[\cdot|i], p_{\Delta t}[\cdot|j]) < 0.05$ for all $i, j \in S$. For example at $10^{-11.85}$s, the black horizontal bar links nodes 13, 14, and 15 satisfying $d_H(p_{\Delta t}[\cdot|i], p_{\Delta t}[\cdot|j]) < 0.05$ for all the combinations of $i, j \in \{13, 14, 15\}$. The pink horizontal bars and their timescales are transient timescales at which some of the states merge

with sharing some nodes. For example, at $10^{-10.55}$s nodes $1, \ldots, 7$ are grouped together and at the same time nodes $5, \ldots, 10$ are also grouped together. This can be interpreted as the two subnetworks (corresponding to superbasins) at the timescale overlapping with each other by sharing nodes $5, 6$, and 7.

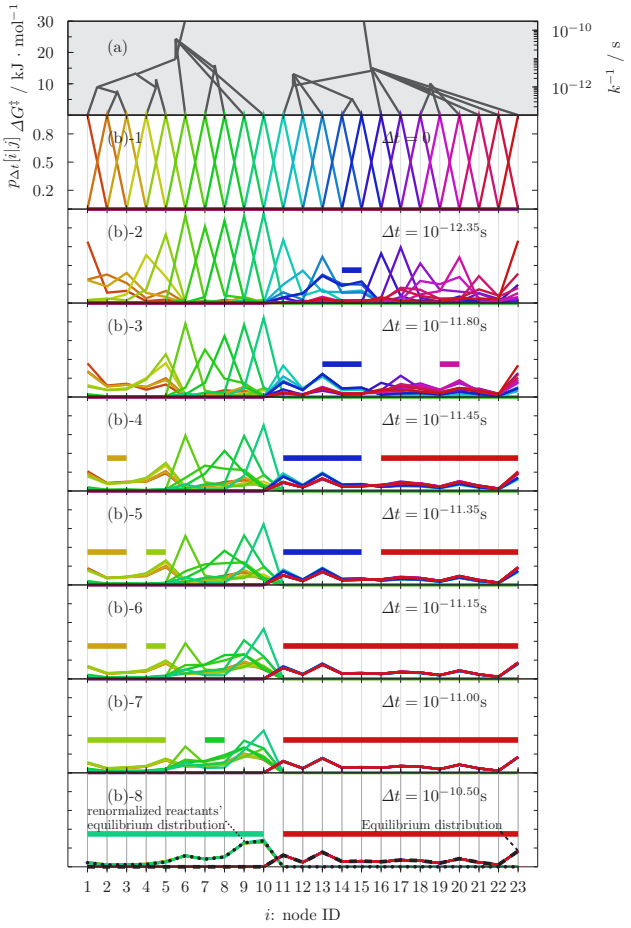



Figure 6.5: Comparison of timescales between the MCC-based TST and the direct observation of $P_{\Delta t}$ of the isomerization reaction network of Claisen rearrangement of AVE. A part of Fig. 6.2 whose range of the free-energy-of-activation is $0 - 30$ kJ $\cdot$ mol$^{-1}$ is taken on the top row figure. The seven other row figures show column vectors of $P_{\Delta t}$ with $\Delta t = 0$ s, $10^{-12.35}$ s $\approx$ 0.45 ps, $10^{-11.80}$ s $\approx$ 1.6 ps, $10^{-11.45}$ s $\approx$ 3.5 ps, $10^{-11.35}$ s $\approx$ 4.5 ps, $10^{-11.15}$ s $\approx$ 7.1 ps, $10^{-11.00}$ s $\approx$ 10 ps, $10^{-10.50}$ s $\approx$ 32 ps. Initially, i.e., $\Delta t = 0$, the system starts from the node $i$ (with the same color) indicated in Fig. 6.1. Instead of showing delta functions located at the node $i$, line plots, "triangle shapes", are drawn for the sake of visual clarity at $\Delta t = 0$ s. The horizontal axis gives the index of node. The short and long color bars in the seven figures represent a set of initial nodes (the system starts from) whose column fall into almost the same in the Hellinger distance less than 0.05, along $\Delta t = 10^{-13+0.05m}$ s where $m = 1, 2, \ldots, 60$. For all the pair of nodes $(i, j)$ within the reactant and product superbasins I identified the time to satisfy $d_H(p_{\Delta t}[\cdot|i], p_{\Delta t}[\cdot|j]) < 0.05$ (see the full list in Fig. 6.6).

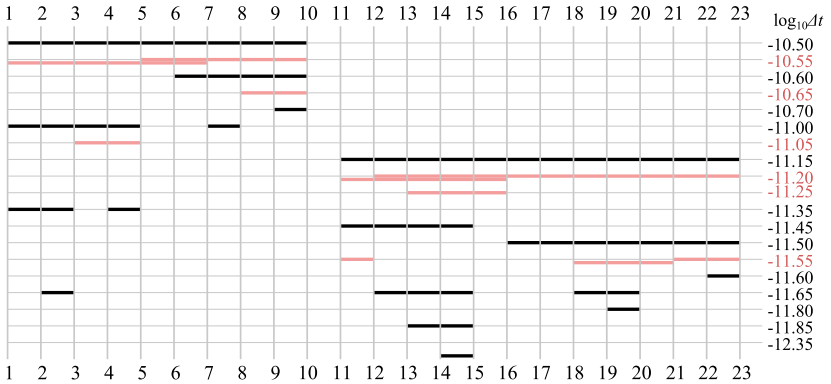



Figure 6.6: The complete list of the hierarchical organization of timescales $P_{\Delta t}$

# Chapter 7

# Conclusion

In this thesis, I have newly defined a series of network TSs over a reaction network on the basis of overall rate constants by which one can construct a tree of reaction timescale hierarchy that can be regarded as a variant of disconnectivity graph. The tree can reveal the hierarchical nature of timescales of reactions/structural transitions buried in the network, e.g., in what timescale each reaction among subnetwork takes place, and what actually serves as the TS for each conformational change. I have applied my method to a reaction network of Claisen rearrangement of AVE. I compared the kinetic properties of my disconnectivity graph to those of the transition matrix of the rate equations. It was shown that the MCC-based TST presents a reliable estimation of the overall rate constant by taking into account the entropic contribution but also reproduces how each subnetwork starts to merge into a larger one; in the other words, how many states should be considered at a given timescale.

The reason why my MCC-based TST can estimate the overall rate constants reliably is that the MCC properly assigns an equilibrated subnet-

work. As can be seen in Fig. 6.5 (b)-8, the distributions starting from any of nodes belonging to the either of the reactant or the product subnetwork almost converge into the equilibrium distribution within the subnetwork at $\Delta t = 10^{-10.50}$ s.

It should be noted that this is the case that the generalized pre-equilibrium assumption[84] works. The equilibration is due to a very high free-energy-of-activation separating the network into the reactant and product subnetworks compared to activation energies over TSs within each subnetwork. Then, the MCC should be the boundary between them so that the transition probabilities of reaction between the subnetworks become very small and this equilibration is fairly similar to the assumption of my MCC-based TST which is based on equilibrium distribution. Hence the transition probabilities between the subnetworks are almost exactly transformed into overall rate constants by Eq.(4.2). For this reason, the MCC-based TST derives a proper overall rate constant.

The striking consequence of this present study is to provide a rigorous definition of network TSs over complex reaction networks on the basis of MCC. This method avoids a drawback of the spectral clustering method, corresponding to an eigenvalue problem of the transition matrix, in that the second smallest eigenvalue is not necessarily present in the MCC.[72–75] However, it is known that the combinatorial explosion problem is inevitable for finding two disjoint subnetworks induced by $S$ and $S^{\mathrm{c}}$ to satisfy a clear criteria such as MCC.[71] Therefore, less rigorous but practical approaches would also be important for some applications. We have just recently developed another approach that a node in a network (from which the system

escapes at the maximum rate constant to another node) is recursively contracted to the neighboring nodes in the network to yield the slowest overall rate constant over the entire network. [85]

Complex reaction networks[14] have been widely studied to reveal complex kinetics in both computer simulations[7, 11, 21, 32, 37–41, 84, 86, 87] and single molecule time series analysis [30, 31, 88–91]. One of the most intriguing subjects is how to take into account multiple pathways to connect any pair of basins (nodes) and to construct an effective energy landscape free from any choice of the coordinate into which the energy landscape is projected. The MCC approach provides me with a natural scheme to represent the organization of timescales or free-energies-of-activation. Characterizing timescale hierarchies [11, 41, 86, 87] in molecular systems can tell me the relationship between (overall) reactions of long timescale and (elementary) reactions of short timescale (e.g., non-exponential kinetics), and also the relationship between the underlying molecular structures (when available like in computer simulations) such as secondary and tertiary structures of proteins and the underlying complex kinetics, e.g., folding-unfolding. My method is expected to provide a new framework to enable me to explore such relationship in terms of the network.

# Bibliography

[1] C. Levinthal, in *Mossbauer Spectroscopy in Biological Systems*, edited by J. T. P. DeBrunner and E. Munck (University of Illinois Press, Monticello, Illinois., 1969), pp. 22–24, URL [website].

[2] N. Go, "The Consistency principle in protein structure and pathways of folding," Advance Biophysics **18**, 149 (1984), ISSN 0065-227X, URL [website].

[3] J. D. Bryngelson and P. G. Wolynes, "Spin glasses and the statistical mechanics of protein folding.," Proceedings of the National Academy of Sciences of the United States of America **84**, 7524 (1987), ISSN 0027-8424, URL [website].

[4] A. Šali, E. Shakhnovich, and M. Karplus, "How does a protein fold?," Nature **369**, 248 (1994), ISSN 0028-0836, URL [website].

[5] E. Shakhnovich, G. Farztdinov, A. M. Gutin, and M. Karplus, "Protein folding bottlenecks: A lattice Monte Carlo simulation," Physical Review Letters **67**, 1665 (1991), ISSN 0031-9007, URL [website].

[6] J. N. Onuchic, P. G. Wolynes, Z. Luthey-Schulten, and N. D. Socci, "To-

ward an outline of the topography of a realistic protein-folding funnel.," Proceedings of the National Academy of Sciences of the United States of America **92**, 3626 (1995), ISSN 0027-8424, URL [website].

[7] O. M. Becker and M. Karplus, "The topology of multidimensional potential energy surfaces: Theory and application to peptide structure and kinetics," The Journal of Chemical Physics **106**, 1495 (1997), ISSN 00219606, URL [website].

[8] J. M. Carr and D. J. Wales, "Refined kinetic transition networks for the GB1 hairpin peptide," Physical Chemistry Chemical Physics **11**, 3341 (2009), ISSN 1463-9076, URL [website].

[9] J. M. Carr and D. J. Wales, "Folding Pathways and Rates for the Three-Stranded $\beta$-Sheet Peptide Beta3s using Discrete Path Sampling," The Journal of Physical Chemistry B **112**, 8760 (2008), ISSN 1520-6106, URL [website].

[10] M. A. Miller and D. J. Wales, "Energy landscape of a model protein," The Journal of Chemical Physics **111**, 6610 (1999), ISSN 00219606, [website], URL [website].

[11] D. J. Wales, M. A. Miller, and T. R. Walsh, "Archetypal energy landscapes," Nature **394**, 758 (1998), ISSN 00280836, URL [website].

[12] D. J. Wales, "The energy landscape as a unifying theme in molecular science," Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **363**, 357 (2005), ISSN 1364-503X, URL [website].

[13] D. Chakrabarti, S. N. Fejer, and D. J. Wales, "Rational design of helical architectures," Proceedings of the National Academy of Sciences **106**, 20164 (2009), ISSN 0027-8424, URL [website].

[14] J. Goutsias and G. Jenkinson, "Markovian dynamics on complex reaction networks," Physics Reports **529**, 199 (2013), ISSN 03701573, URL [website].

[15] M. Dellnitz and O. Junge, "On the Approximation of Complicated Dynamical Behavior," SIAM Journal on Numerical Analysis **36**, 491 (1999), ISSN 0036-1429, URL [website].

[16] G. Froyland and M. Dellnitz, "Detecting and Locating Near-Optimal Almost-Invariant Sets and Cycles," SIAM Journal on Scientific Computing **24**, 1839 (2003), ISSN 1064-8275, URL [website].

[17] G. Froyland, "Statistically optimal almost-invariant sets," Physica D: Nonlinear Phenomena **200**, 205 (2005), ISSN 01672789, URL [website].

[18] P. Deuflhard, M. Dellnitz, O. Junge, and C. Schütte, in *Computational Molecular Dynamics: Challenges, Methods, Ideas*, edited by P. Deuflhard, J. Hermans, B. Leimkuhler, A. E. Mark, S. Reich, and R. D. Skeel (Springer-Verlag, Berlin Heidelberg, 1999), vol. 45 of *Lecture Notes in Computational Science and Engineering*, pp. 98–115, URL [website].

[19] P. Deuflhard and M. Weber, "Robust Perron cluster analysis in conformation dynamics," Linear Algebra and its Applications **398**, 161 (2005), ISSN 00243795, URL [website].

[20] V. S. Pande, K. Beauchamp, and G. R. Bowman, "Everything you wanted to know about Markov State Models but were afraid to ask.," Methods (San Diego, Calif.) **52**, 99 (2010), ISSN 1095-9130, URL [website].

[21] G. R. Bowman, V. S. Pande, and F. Noé, eds., *An Introduction to Markov State Models and Their Application to Long Timescale Molecular Simulation*, Advances in Experimental Medicine and Biology (Springer, Dordrecht, The Netherlands, 2014), ISBN 978-94-007-7605-0, URL [website].

[22] M. Dellnitz, O. Junge, W.-S. Koon, F. Lekien, M. W. Lo, J. E. Marsden, K. Padberg, R. Preis, S. D. Ross, and B. Thiere, "Transport in dynamical astronomy and multibody problems," International Journal of Bifurcation and Chaos **15**, 699 (2005), ISSN 0218-1274, URL [website].

[23] J. E. Marsden and S. D. Ross, "New methods in celestial mechanics and mission design," Bulletin of the American Mathematical Society **43**, 43 (2005), ISSN 0273-0979, URL [website].

[24] E. M. Bollt, L. Billings, and I. B. Schwartz, "A manifold independent approach to understanding transport in stochastic dynamical systems," Physica D: Nonlinear Phenomena **173**, 153 (2002), ISSN 01672789, URL [website].

[25] L. Billings, E. M. Bollt, and I. Schwartz, "Phase-Space Transport of

Stochastic Chaos in Population Dynamics of Virus Spread," Physical Review Letters **88**, 234101 (2002), ISSN 0031-9007, URL [website].

[26] L. Billings and I. B. Schwartz, "Exciting chaos with noise: unexpected dynamics in epidemic outbreaks," Journal of Mathematical Biology **44**, 31 (2002), ISSN 0303-6812, URL [website].

[27] L. Billings and I. B. Schwartz, "Identifying almost invariant sets in stochastic dynamical systems.," Chaos **18**, 023122 (2008), ISSN 1089-7682, URL [website].

[28] T. Yamanobe, "Global dynamics of a stochastic neuronal oscillator," Physical Review E **88**, 052709 (2013), ISSN 1539-3755, URL [website].

[29] T. Yamanobe, "Stochastic phase transition operator," Physical Review E **84**, 011924 (2011), ISSN 1539-3755, URL [website].

[30] C.-B. Li, H. Yang, and T. Komatsuzaki, "Multiscale complex network of protein conformational fluctuations in single-molecule time series.," Proceedings of the National Academy of Sciences of the United States of America **105**, 536 (2008), ISSN 1091-6490, URL [website].

[31] T. Sultana, H. Takagi, M. Morimatsu, H. Teramoto, C.-B. Li, Y. Sako, and T. Komatsuzaki, "Non-Markovian properties and multiscale hidden Markovian network buried in single molecule time series.," The Journal of chemical physics **139**, 245101 (2013), ISSN 1089-7690, URL [website].

[32] N. Hori, G. Chikenji, R. S. Berry, and S. Takada, "Folding energy landscape and network dynamics of small globular proteins," Proceedings of

the National Academy of Sciences of the United States of America **106**, 73 (2009), ISSN 0027-8424, URL [website].

[33] D. J. Wales, *Energy Landscapes: Applications to Clusters, Biomolecules and Glasses*, 51 (Cambridge University Press, 2003), ISBN 0521814154, URL [website].

[34] S. Maeda, K. Ohno, and K. Morokuma, "Systematic exploration of the mechanism of chemical reactions: the global reaction route mapping (GRRM) strategy using the ADDF and AFIR methods," Physical Chemistry Chemical Physics **15**, 3683 (2013), ISSN 1463-9076, URL [website].

[35] S. Kawai and T. Komatsuzaki, "Effect of timescale on energy landscape: Distinction between free-energy landscape and potential of mean force," Physical Review E **87**, 030803 (2013), ISSN 1539-3755, URL [website].

[36] J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, "THEORY OF PROTEIN FOLDING: The Energy Landscape Perspective," Annual Review of Physical Chemistry **48**, 545 (1997), ISSN 0066-426X, URL [website].

[37] D. A. Evans and D. J. Wales, "Free energy landscapes of model peptides and proteins," The Journal of Chemical Physics **118**, 3891 (2003), ISSN 00219606, URL [website].

[38] S. V. Krivov and M. Karplus, "Free energy disconnectivity graphs: Application to peptide models," The Journal of Chemical Physics **117**, 10894 (2002), URL [website].

[39] S. V. Krivov and M. Karplus, "Hidden complexity of free energy surfaces for peptide (protein) folding.," Proceedings of the National Academy of Sciences of the United States of America **101**, 14766 (2004), ISSN 0027-8424, URL [website].

[40] D. J. Wales, "Energy landscapes: some new horizons," Current Opinion in Structural Biology **20**, 3 (2010), ISSN 0959440X, URL [website].

[41] D. J. Wales and P. Salamon, "Observation Time Scale, Free-Energy Landscapes, and Molecular Symmetry," Proceedings of the National Academy of Sciences of the United States of America **111**, 617 (2014), ISSN 0027-8424, URL [website].

[42] R. Diestel, *Graph Theory*, Graduate Texts in Mathematics (Springer-Verlag, New York, 2000), electric ed., ISBN 0-387-95014-1, URL [website].

[43] R. E. Gomory and T. C. Hu, "Multi-Terminal Network Flows," Journal of the Society for Industrial and Applied Mathematics **9**, 551 (1961), ISSN 0368-4245, URL [website].

[44] B. Korte and J. Vygen, in *Combinatorial Optimization* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012), Algorithms and Combinatorics, chap. 8, pp. 173–209, 5th ed., ISBN 978-3-642-24487-2, URL [website].

[45] G. W. Flake, S. Lawrence, and C. L. Giles, in *Proceedings of the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM Press, Boston, MA, USA, 2000), pp. 150–160, ISBN 1581132336, URL [website].

[46] G. W. Flake, R. E. Tarjan, and K. Tsioutsiouliklis, "Graph Clustering and Minimum Cut Trees," Internet Mathematics **1**, 385 (2004), ISSN 1542-7951, URL [website].

[47] A. Fernández-Ramos, J. A. Miller, S. J. Klippenstein, and D. G. Truhlar, "Modeling the kinetics of bimolecular reactions.," Chemical Reviews **106**, 4518 (2006), ISSN 0009-2665, URL [website].

[48] C. Jaffé, J. Palacián, S. Kawai, P. Yanguas, and T. Uzer, "A New Look at the Transition State: Wigner's Dynamical Perspective Revisited," Advances in Chemical Physics **130A**, 171 (2005), URL [website].

[49] K. J. Laidler and M. C. King, "Development of transition-state theory," Journal of Physical Chemistry **87**, 2657 (1983), ISSN 0022-3654, URL [website].

[50] P. Pechukas, "Transition State Theory," Annual Review of Physical Chemistry **32**, 159 (1981), ISSN 0066-426X, URL [website].

[51] D. G. Truhlar, W. L. Hase, and J. T. Hynes, "Current status of transition-state theory," Journal of Physical Chemistry **87**, 2664 (1983), ISSN 0022-3654, URL [website].

[52] D. G. Truhlar, B. C. Garrett, and S. J. Klippenstein, "Current Status of Transition-State Theory," Journal of Physical Chemistry **100**, 12771 (1996), ISSN 0022-3654, URL [website].

[53] S. Glasstone, K. J. Laidler, and H. Eyring, *The theory of rate processes : the kinetics of chemical reactions, viscosity, diffusion and electrochemi-*

*cal phenomena*, International chemical series (McGraw-Hill Book Company, inc., New York and London, 1941), URL [website].

[54] H. Eyring, "The Activated Complex in Chemical Reactions," The Journal of Chemical Physics **3**, 107 (1935), ISSN 00219606, URL [website].

[55] M. G. Evans and M. Polanyi, "Some applications of the transition state method to the calculation of reaction velocities, especially in solution," Transactions of the Faraday Society **31**, 875 (1935), ISSN 0014-7672, URL [website].

[56] A. Junginger and R. Hernandez, "Uncovering the Geometry of Barrierless Reactions Using Lagrangian Descriptors," The Journal of Physical Chemistry B p. acs.jpcb.5b09003 (2015), ISSN 1520-6106, URL [website].

[57] F. A. L. Mauguière, P. Collins, Z. C. Kramer, B. K. Carpenter, G. S. Ezra, S. C. Farantos, and S. Wiggins, "Phase space barriers and dividing surfaces in the absence of critical points of the potential energy: Application to roaming in ozone," J. Chem. Phys. **144**, 054107 (2016), ISSN 0021-9606.

[58] P. Collins, G. S. Ezra, and S. Wiggins, "Index k saddles and dividing surfaces in phase space with applications to isomerization dynamics.," The Journal of Chemical Physics **134**, 244105 (2011), ISSN 1089-7690, URL [website].

[59] Y. Nagahata, H. Teramoto, C.-B. Li, S. Kawai, and T. Komatsuzaki, "Reactivity Boundaries to Separate the Fate of a Chemical Reaction

Associated with an Index-two saddle," Physical Review E **87**, 062817 (2013), URL [website].

[60] Y. Nagahata, H. Teramoto, C.-B. Li, S. Kawai, and T. Komatsuzaki, "Reactivity boundaries for chemical reactions associated with higher-index and multiple saddles," Physical Review E **88**, 042923 (2013), ISSN 1539-3755, URL [website].

[61] H. Teramoto, M. Toda, and T. Komatsuzaki, "Dynamical Switching of a Reaction Coordinate to Carry the System through to a Different Product State at High Energies," Physical Review Letters **106**, 054101 (2011), ISSN 0031-9007, URL [website].

[62] H. Teramoto, M. Toda, M. Takahashi, H. Kono, and T. Komatsuzaki, "Mechanism and Experimental Observability of Global Switching Between Reactive and Nonreactive Coordinates at High Total Energies," Physical Review Letters **115**, 093003 (2015), ISSN 0031-9007, URL [website].

[63] C.-B. Li, A. Shoujiguchi, M. Toda, and T. Komatsuzaki, "Definability of No-Return Transition States in the High-Energy Regime above the Reaction Threshold," Physical Review Letters **97**, 028302 (2006), ISSN 0031-9007, URL [website].

[64] C.-B. Li, M. Toda, and T. Komatsuzaki, "Bifurcation of no-return transition states in many-body chemical reactions.," The Journal of Chemical Physics **130**, 124116 (2009), ISSN 1089-7690, URL [website].

[65] F. Mauguiere, V. Tyuterev, and S. C. Farantos, "Bifurcation effects and patterns in the vibrational excited states of isotopically substituted water," Chemical Physics Letters **494**, 163 (2010), ISSN 00092614, URL [website].

[66] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra* (SIAM, 2000), ISBN 9780898719512, URL [website].

[67] N. G. V. Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, North-holland, 2007), 3rd ed., ISBN 978-0-444-52965-7, URL [website].

[68] T. Kaczorek, *Positive 1D and 2D Systems*, Communications and Control Engineering (Springer London, London, 2002), ISBN 978-1-4471-1097-2, URL [website].

[69] S. E. Schaeffer, "Graph clustering," Computer Science Review **1**, 27 (2007), ISSN 15740137, URL [website].

[70] R. Sibson, "SLINK: an optimally efficient algorithm for the single-link cluster method," (1973), URL [website].

[71] J. Šíma and S. E. Schaeffer, in *SOFSEM 2006*, edited by J. Wiedermann, G. Tel, J. Pokorný, M. Bieliková, and J. Štuller (Springer, Berlin/Heidelberg, 2006), vol. 3831 of *Lecture Notes in Computer Science*, pp. 530–537, ISBN 978-3-540-31198-0, URL [website].

[72] F. Chung, in *Combinatorics, Paul Erdős is Eighty*, edited by D. Miklós,

V. T. Sós, and T. Szönyi (János Bolyai Mathematical Society, Budapest, 1996), vol. 2, pp. 157–172.

[73] F. Chung, "Laplacians and the Cheeger Inequality for Directed Graphs," Annals of Combinatorics **9**, 1 (2005), ISSN 0218-0006, URL [website].

[74] R. Kannan, S. Vempala, and A. Vetta, "On Clusterings: Good, Bad and Spectral," Journal of the ACM **51**, 497 (2004), ISSN 00045411, URL [website].

[75] U. von Luxburg, "A tutorial on spectral clustering," Statistics and Computing **17**, 395 (2007), ISSN 0960-3174, URL [website].

[76] D. A. Spielman and S.-H. Teng, "Spectral partitioning works: Planar graphs and finite element meshes," Linear Algebra and its Applications **421**, 284 (2007), ISSN 00243795, URL [website].

[77] T. Bühler and M. Hein, in *Proceedings of the 26th Annual International Conference on Machine Learning* (ACM Press, Montreal, Canada, 2009), pp. 1–8, ISBN 9781605585161, URL [website].

[78] Y. Kawahara, K. Nagano, and Y. Okamoto, "Submodular fractional programming for balanced clustering," Pattern Recognition Letters **32**, 235 (2011), ISSN 01678655, URL [website].

[79] J. C. Keck, "Variational Theory of Chemical Reaction Rates Applied to Three-Body Recombinations," The Journal of Chemical Physics **32**, 1035 (1960), ISSN 00219606, URL [website].

[80] S. Maeda, T. Taketsugu, and K. Morokuma, "Exploring transition state structures for intramolecular pathways by the artificial force induced reaction method," Journal of Computational Chemistry **35**, 166 (2014), ISSN 01928651, URL [website].

[81] A. M. Martín Castro, "Claisen Rearrangement over the Past Nine Decades," Chemical Reviews **104**, 2939 (2004), ISSN 0009-2665, URL [website].

[82] P. Atkins and J. de Paula, *Atkins' Physical Chemistry* (OUP Oxford, 2014), 10th ed., ISBN 9780199697403.

[83] F. W. Schuler and G. W. Murphy, "The Kinetics of the Rearrangement of Vinyl Allyl Ether," Journal of the American Chemical Society **72**, 3155 (1950), ISSN 0002-7863, URL [website].

[84] M. Rae and M. N. Berberan-Santos, "A Generalized Pre-Equilibrium Approximation in Chemical and Photophysical Kinetics," Journal of Chemical Education **81**, 436 (2004), ISSN 0021-9584, URL [website].

[85] Y. Sumiya, Y. Nagahata, T. Komatsuzaki, T. Taketsugu, and S. Maeda, "Kinetic Analysis for the Multistep Profiles of Organic Reactions: Significance of the Conformational Entropy on the Rate Constants of the Claisen Rearrangement," The Journal of Physical Chemistry A **119**, 11641 (2015), ISSN 1089-5639, URL [website].

[86] J. P. K. Doye, M. A. Miller, and D. J. Wales, "The double-funnel energy landscape of the 38-atom Lennard-Jones cluster," The Journal of Chemical Physics **110**, 6896 (1999), ISSN 00219606, URL [website].

[87] F. Calvo, "Free-energy landscapes from adaptively biased methods: Application to quantum systems," Physical Review E **82**, 046703 (2010), ISSN 1539-3755, URL [website].

[88] A. Baba and T. Komatsuzaki, "Construction of effective free energy landscape from single-molecule time series.," Proceedings of the National Academy of Sciences of the United States of America **104**, 19297 (2007), ISSN 1091-6490, URL [website].

[89] J. N. Taylor, C.-B. Li, D. R. Cooper, C. F. Landes, and T. Komatsuzaki, "Error-based Extraction of States and Energy Landscapes from Experimental Single-Molecule Time-Series," Scientific Reports **5**, 9174 (2015), ISSN 2045-2322, URL [website].

[90] S. Chaudhury, J. Cao, and N. A. Sinitsyn, "Universality of Poisson Indicator and Fano Factor of Transport Event Statistics in Ion Channels and Enzyme Kinetics," The Journal of Physical Chemistry B **117**, 503 (2013), ISSN 1520-6106, URL [website].

[91] J. Cao and R. J. Silbey, "Generic Schemes for Single-Molecule Kinetics. 1: Self-Consistent Pathway Solutions for Renewal Processes," The Journal of Physical Chemistry B **112**, 12867 (2008), ISSN 1520-6106, URL [website].

# Index