國立臺灣大學理學院化學系
碩士論文
Department of Chemistry
College of Science
National Taiwan University
Master Thesis

光合作用能量傳輸網路分析
Theoretical Analysis of Energy Transfer Networks in Photosynthetic Systems

曾煒翔
Tseng, Wei-Hsiang

指導教授：鄭原忠博士
Advisor: Cheng, Yuan-Chung, Ph.D.

中華民國 106 年 7 月
July, 2017

doi:10.6342/NTU201704140

doi:10.6342/NTU201704140

doi:10.6342/NTU201704140

誌謝

感謝鄭原忠老師、金必耀老師在研究中給予的方向與指導。也感謝王昱夫、王佑仁、戴宏軒的討論與建議。最後感謝林冠宏、詹欣穆、孫敬對於程式的討論與改進。

doi:10.6342/NTU201704140

doi:10.6342/NTU201704140

摘要

爲了理解在光合系統中複雜激子的能量傳輸行爲，建立一個粗粒化模型是必要的。在本次工作中，我們開發了一個系統化的方法。藉由最小分割法來建構光合系統的粗粒化模型。我們使用這個方法處理三種不同的光合作用網路並且這些粗粒化模型可以很好的還原激子能量傳輸的動態演變。這些粗粒化模型可以給我們對這些光合作用系統有新的見解，也可以有效的讓我們理解複雜的動力學反應。

doi:10.6342/NTU201704140

doi:10.6342/NTU201704140

Abstract

To understand complex excitation energy transfer (EET) networks in photosynthetic systems, building a coarse-grained model is necessary to obtain a simplified representation. Here, we developed a systematic approach to produce coarse-grained models for photosynthetic systems by combining a minimum-cut method and a top-down clustering algorithm. The new approach was applied to investigate EET networks of three photosynthetic systems, and we demonstrate that our approach not only reproduces the population dynamics very well but also provides novel insights into the spatial-temporal EET dynamics in complex photosynthetic systems. The new approach could be a very powerful tool towards the elucidation of complex kinetic networks that is commonly encountered in Chemistry.

doi:10.6342/NTU201704140

doi:10.6342/NTU201704140

Contents

doi:10.6342/NTU201704140

doi:10.6342/NTU201704140

List of Figures

doi:10.6342/NTU201704140

doi:10.6342/NTU201704140

List of Tables

doi:10.6342/NTU201704140

doi:10.6342/NTU201704140

Chapter 1

Introduction

Photosynthesis, the process converts light energy to chemical energy, fuels most life on Earth. In the initial steps of photosynthesis, excitation energy transfer (EET) plays a crucial process. Light energy captured in the antenna moves between pigments and finally reaches the reaction center to trigger charge separation.

Photosynthetic systems utilize complex networks composed of chlorophyll molecules in protein-pigment antenna complexes to harvest sunlight and transfer the excitation energy to the reaction center for energy conversion. The excitation energy transfer (EET) networks of photosynthetic systems exhibit high quantum efficiency, and understanding the mechanism of energy-flow control in these networks may lead to novel design principles for light harvesting that can improve artificial photosynthetic systems.

Nevertheless, the size and complexity of the photosynthetic networks often prevent a clear understanding if the full details of the system is to be considered. To this end, many previous studies have simulated and constructed simplified network models for various photosynthetic systems. For instance, Ruban’s group have try to modeling the photosystem II based on the assumption that the excitation energy transfer within a pigment-protein complex is much faster than the intercomplex excitation energy transfer [14].

However, these simplified cluster model for photosynthesis networks are often constructed ad-hoc based on fitting to population dynamics or based on considerates of couplings between sites [16, 15] Therefore, a systematic method to yeild reduced model for EET dynamics in a photosynthetic system is desirable.

doi:10.6342/NTU201704140

In this study, we aim to take a graph theoretial perspective on photosynthetic EET networks. In this regard, we will apply graph clustering methods to investgate EET pathways and bottlenecks in photosynthetic systems. (Hereafter, ”network” and ”graph” are two terms that will be used interchangeably.)

The graph clustering is a central tool for the analysis of networks with applications ranging from the field of social sciences to biology and to the growing field of complex systems.[4]. The general aim of graph clustering is to separate graph into subgraphs, called clusters, that are only loosely connected to each other in the given network. The connections between the subgraphs would constitute the bottlenecks in the network. The reduce model of the network is called coarse-grained model (CGM).

In this work, we applied Ford-Fulkerson Algrithm (FFA) and developed some approaches to build the CGM of EET networks. Finally we do some benchmark and compare to the exact result from modified Redfield theory, and discussed the CGM we obtained.

doi:10.6342/NTU201704140

Chapter 2

Excitation Energy Transfer

In order to obtain the rate constant of EET in photosynthetic systems, we adopt the modified Redfield theory (MRT) [5]. In this section we describe the derivation of MRT rate constant and the constitution of EET networks.

2.1 Model Hamiltonian

In this work, we employ the Frenkel exciton Hamiltonian to describe photoexcitations of a pigment-protein complex aggregate with N pigments, which also called sites. Written in the electronic eigenbasis (the so called exciton basis), the Hamiltonian reads:

$$H = H_e + H_{ph} + H_{e-ph}, \tag{2.1}$$

$$H_e = \sum_{\alpha}^{N} \varepsilon_\alpha |\alpha\rangle\langle\alpha|, \tag{2.2}$$

$$H_{ph} = \sum_{i} \omega_i b_i^\dagger b_i, \tag{2.3}$$

$$\begin{aligned} H_{e-ph} &= \sum_{\alpha,\beta}^{N} |\alpha\rangle\langle\beta| \cdot \sum_{n=1}^{N} \sum_{i} C_{\alpha,n} C^*_{\beta,n} g_{ni} \omega_i (b_i + b_i^\dagger) \\ &\equiv \sum_{\alpha,\beta}^{N} |\alpha\rangle\langle\beta| \cdot (H_{e-ph})_{\alpha\beta}, \end{aligned} \tag{2.4}$$

where $|\alpha\rangle$ denotes the α-th exciton state, which is a linear combination of site excitation, $|\alpha\rangle = \sum_n^N C_{\alpha,n}|n\rangle$. In addition, ε_α denotes the excitation energy of $|\alpha\rangle$, $b_i^\dagger(b_i)$ is the

doi:10.6342/NTU201704140

creation (annihilation) operator of the i-th phonon mode, ω_i is the frequency of the phonon mode, and g_{ni} is the exciton-phonon coupling constant between the localized electronic excitation on site n and the i-th phonon mode. The exciton-phonon coupling constant is related to the displacement of the phonon coordinate in the excited state, which also defines the site reorganization energy $\lambda_n = \sum_i \sum_n g_{ni}^2 \omega_i$. Finally, basis transfermation from the site basis to the exciton basis yields the $\sum_n^N C_{\alpha,n} C_{\beta,n}$ factor, which can be considered as the overlap between exciton wavefunctions $|\alpha\rangle$ and $|\beta\rangle$.

The main idea behind the modified Redfield theory is to partition the Hamiltonian into a zeroth-order Hamiltonian including the diagonal fluctuations in the exciton basis:

$$H_0 = H_e + H_{ph} + \sum_{\alpha=1}^{N} |\alpha\rangle\langle\alpha| \cdot (H_{e-ph})_{\alpha\alpha} \tag{2.5}$$

and the perturbation part:

$$V = \sum_{\alpha\neq\beta}^{N} |\alpha\rangle\langle\beta| \cdot (H_{e-ph})_{\alpha\beta}. \tag{2.6}$$

2.2 Quantum master equation

We start from a general time-local quantum master equation that is derived using a second-order cumulant expansion respect to perturbation V. The quantum master equation written in the interaction picture of H_0 is

$$\dot{\sigma}_I(t) = -\int_0^t d\tau \mathrm{Tr_B} \left\{ [V_I(t), [V_I(\tau), \sigma_I(t) \otimes \rho_b^{eq}]] \right\}, \tag{2.7}$$

where $V_I(t)$ is defined in the interaction picture of H_0, i.e. $V_I(t) = e^{iH_0 t} V(t) e^{-iH_0 t}$. Note that in Eq. 2.7, we have assumed a product-state initial condition, $\rho(0) = \sigma(0)\rho_b^{eq}$, where $\rho_b^{eq} = \frac{e^{-\beta H_{ph}}}{Z}$ is the equilibrium density matrix of the bath. This assumption is justified with photoinduced EET process. Additionally, we also neglect the first-order average of the perturbation, i.e. $\langle V \rangle$ in perturbation. This assumption simplifies the equation of motion.

doi:10.6342/NTU201704140

In the Schrödinger picture, the reduced density matrix is calculated by

$$\sigma(t) = \mathrm{Tr_B}\left\{U_0(t)\sigma_I(t)\rho_b^{eq}U_0^\dagger(t)\right\}. \tag{2.8}$$

Note that $U_0(t)$ can not be simply separated into a system part and a bath part. Because the diagonal exciton-phonon coupling term contains both system and bath operator, the partial trace has to be taken after solving $U_0(t)\sigma_I(t)U_0^\dagger(t)$. Formally, the equation of motion for the reduced density matrix is

$$\dot{\sigma}(t) = -i\mathrm{Tr_B}\left\{[H_0, \sigma(t)\otimes\rho_b^{eq}]\right\} - \int_0^t d\tau \mathrm{Tr_B}\left\{[V,[V(-\tau),\sigma(t)\otimes\rho_b^{eq}]]\right\}. \tag{2.9}$$

The modified Redfield omits the coherence part of the density matrix and only considers the population trasfer between exciton states. The coherence part is the first term in Eq. 2.9. So the equation of motion in MRT is

$$\dot{\sigma}(t) = -\int_0^t d\tau \mathrm{Tr_B}\left\{[V,[V(-\tau),\sigma(t)\otimes\rho_b^{eq}]]\right\}, \tag{2.10}$$

which can be evaluated to yield

$$\dot{\sigma}_{\alpha\beta} = \sum_{\gamma\delta}\left(\Gamma_{\delta\beta,\alpha\gamma} + \Gamma^*_{\gamma\alpha,\beta\delta} - \delta_{\beta\delta}\sum_f \Gamma_{\alpha f,f\gamma} - \delta_{\alpha\gamma}\sum_f \Gamma^*_{\beta f,f,\delta}\right)\sigma_{\gamma\delta}(t), \tag{2.11}$$

where

$$\Gamma_{\alpha\beta,\gamma\delta} = \int_0^t d\tau \langle V_{\alpha\beta}V_{\gamma\delta}(-\tau)\rangle. \tag{2.12}$$

In Eq. 2.12, $V_{\alpha\beta}(-\tau) = \langle\alpha|V(-\tau)|\beta\rangle$. In general, from Eq. 2.10 the full Redfield tensor can be calculated in the modified Redfield representation. But in this work, we only consider the population transfer.

doi:10.6342/NTU201704140

The population transfer term ($\alpha = \beta, \gamma = \delta$) :

$$\dot{\sigma}_{\alpha\alpha}(t) = \sum_{\gamma}(R_{\alpha\gamma}\sigma_{\gamma\gamma}(t) - R_{\gamma\alpha}\sigma_{\alpha\alpha}(t)). \tag{2.13}$$

The transfer rate from $|\beta\rangle$ state to $|\alpha\rangle$ state is given by

$$R_{\alpha\beta}(t) = 2 \cdot \Re\left\{\int_0^t d\tau \langle V_{\beta\alpha} V_{\alpha\beta}(-\tau)\rangle\right\}. \tag{2.14}$$

Eq. 2.14 can be evaluated to yield:

$$R_{\alpha\beta}(t) = 2 \cdot \Re\left\{\int_0^t d\tau F^*_{\beta}(\tau) A_{\alpha}(\tau) \chi_{\alpha\beta}(\tau)\right\}, \tag{2.15}$$

where

$$A_{\alpha}(t) = e^{-i\varepsilon_{\alpha}t - g_{\alpha\alpha\alpha\alpha}(t)}, \tag{2.16}$$

$$F_{\alpha}(t) = e^{-i(\varepsilon_{\alpha} - 2\lambda_{\alpha\alpha\alpha\alpha}t - g^*_{\alpha\alpha\alpha\alpha}(t)} \tag{2.17}$$

and the perturbation induced dynamical term can be evaluated to yeild

$$\begin{aligned}\chi_{\alpha\beta}(t) = & e^{2(g_{\alpha\alpha\beta\beta}(t) + i\lambda_{\alpha\alpha\beta\beta}t)} \\ & \times (\ddot{g}_{\beta\alpha\alpha\beta}(t) - (\dot{g}_{\beta\alpha\alpha\alpha}(t) - \dot{g}_{\beta\alpha\beta\beta}(t) - 2i\lambda_{\beta\alpha\beta\beta}) \\ & \times (\dot{g}_{\alpha\beta\alpha\alpha}(t) - \dot{g}_{\alpha\beta\beta\beta}(t) - 2i\lambda_{\alpha\beta\beta\beta})),\end{aligned} \tag{2.18}$$

where $A(t)$ and $F(t)$ are related to the absorption and emission lineshape respecively, and the lineshape function $g(t)$ is defined as

$$g_{\alpha\beta\gamma\delta}(t) = \int_0^t dt_1 \int_0^{t_1} dt_2 \langle (H_{e-ph}(t_1))_{\alpha\beta} (H_{e-ph}(t_2))_{\gamma\delta}\rangle. \tag{2.19}$$

The time-correlation function in the integral depends on the exciton-phonon couplings. In condensed phased system, the number of phonon modes coupled to the exctions is so large that we should use spectral density funcitons to describe such interactions. Here we

doi:10.6342/NTU201704140

define general spectral density functions, $J_{nm}(\omega)$, as coupling-weighted density of states:

$$J_{nm}(\omega) = \sum_i g_{ni} g_{mi} \omega_i^2 \delta(\omega - \omega_i). \tag{2.20}$$

The spectral density is site-dependent and the definition generalizes the spectral density function to treat correlated-bath conditions. Diagonal $J_{nm}(\omega)$ describes cross correlation between site energy fluctuations on n and m. The lineshape function at exciton basis considering general bath spectral density functions $J_{nm}(\omega)$ can be evaluated from Eq. 2.20 to obtain

$$\begin{aligned} g_{\alpha\beta\gamma\delta}(t) &= \sum_{n,m}^{N} c_{\alpha,n} c^*_{\beta,n} c_{\gamma,m} c^*_{\delta,m} \int_0^\infty d\omega \frac{J_{nm}(\omega)}{\omega^2} \left\{ \coth(\frac{\omega}{2k_B T})[1 - \cos(\omega t)] + i[\sin(\omega t) - \omega t] \right\} \\ &= \sum_{n,m}^{N} c_{\alpha,n} c^*_{\beta,n} c_{\gamma,m} c^*_{\delta,m} \cdot g_{nm}(t) \end{aligned} \tag{2.21}$$

Therefore, after the spectral densities $J_{nm}(\omega)$ are specified, the lineshape funcitons can then be evaluated to calculate EET dynamics. To simplify the calculation, all the spectral densities are over-damped Brownian oscillator in this work. The spectral density funciton is

$$J(\omega) = \frac{2\lambda_0}{\pi} \frac{\omega \Gamma_0}{\omega^2 + \Gamma_0^2}, \tag{2.22}$$

and there are only two parameter, where λ_0 is reorganization energy and Γ_0 is the cut-off frequency.

2.3 Excitation energy transfer network

To simplify the calculation, we employ the Markovian limit in modified Redfield theory. The transfer ratefrom $|\beta\rangle$ state to $|\alpha\rangle$ state is given by

$$R_{\alpha\beta} = \lim_{t \to \infty} R_{\alpha\beta}(t) \tag{2.23}$$

doi:10.6342/NTU201704140

The master equation of the system is,

$$\frac{d}{dt}\mathbf{P} = \mathbf{RP} \tag{2.24}$$

, where P_α is the exciton populaiton in $|\alpha\rangle$ state.

We can convert the EET rate constant matrix **R** into a network. The nodes of the network is the exciton states of given system and the weight of each edges is the rate constant between two nodes. We define a weighted graph $G(V, E)$ to represent the network ,where V denotes all the nodes of the network, and E denotes all the edges between two nodes. A weighted graph is a network in which each edge is given a numerical weight, and in this work the weight is denote the rate constant between two exciton states(nodes).

doi:10.6342/NTU201704140

Chapter 3

Network analysis

In this work, the process of network analysis can devide into two part. The first part is constructing the minimum-cut binary tree (MBT). The second part is building the coarse-grained model.

3.1 Minimum-cut binary tree

In a simple chemical reaction, for instance,

$$\mathrm{A} \overset{\text{fast}}{\rightleftharpoons} \mathrm{B},$$
$$\mathrm{B} \xrightarrow{\text{slow}} \mathrm{C}.$$

The total rate of the reaction could be determined by the slowest step, $\mathrm{B} \longrightarrow \mathrm{C}$. In this case, we can say that A and B are in one cluster, and the reaction can be reduced,

$$[\mathrm{AB}] \xrightarrow{\text{slow}} \mathrm{C},$$

and this is the rate-determining step(RDS).

Following the same idea, for a complex photosynthetic network, we assume that the total dynamics can be reduced by such ”rate-determining step”. In network analysis, finding rate-determining step in a weighted graph is related to the minimum cut problem.

doi:10.6342/NTU201704140

We define a cut $C(S,T)$ to split a network into two subsets, S and T, which denote the source set and target(sink) subset, and the weight of the cut is the sum of the weights of the edges crossing the cut. And the minimum cut is a cut if the weight of the cut is not larger than any other cut. This concept is similar to the RDS in the chemical reaction.

But a graph with N vertices can at the most have $N(N-1)/2$ cut, it's expensive to find the minimum cut by comparing all the weight of cuts. In 1956, Ford and Fulkerson had proved that the minimum cut problem of the directed positive weight networks is equal to the maximum flow problem, called maximum-flow minimum-cut theorem. [3] Ford and Fulkerson also demonstrate a simple method to find the maximum flow [3], called Ford-Fulkerson Algorithm (FFA) (Alg. 1).

The idea of FFA is as follow: As there is a path from the source node to sink node, with available capacity on all edges in the path, we send flow along one of the paths. Then we find another path, and so on. When there is no path from source to sink. The subset S is the set of nodes reachable by source. And subset T is the set $(G - S)$

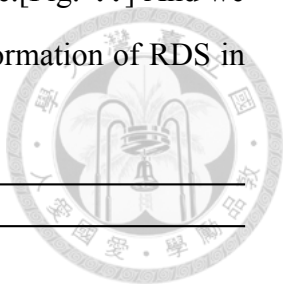
Algorithm 1: Ford-Fulkerson Algorithm (FFA)

input : Graph G and source s, sink t
output: Weight of minimum cut, F, residual graph, G_f
flow $F = 0$
for *each edge* $E(u,v) \in G.E$ **do**
    // initialize G_f
    $E(u,v).f = 0$
while *there exists a path* p *from* s *to* t *in the residual network* G_f **do**
    $c_f(p) = \min\{c_f(u,v) : E(u,v) \in p\}$
    $F = F + c_f(p)$
    for *each edge* $E(u,v) \in p$ **do**
        $E(u,v).f = E(u,v).f + c_f(p)$
        $E(v,u).f = E(v,u).f - c_f(p)$
return F, G_f

The complexity of Ford-Fulkerson Algorithm (FFA) is only $O(N(V)N(E)^2)$, where N denotes the number. But FFA is a greedy algorithm, it should fix the source node (s) and the target node (t), in our work, we choose the highest eigenenergy state as the source node and the lowest eigenenergy state as the sink node. By applying the FFA to the network, we can get the minimum cut of the network and two separate sub-networks. If we keep

doi:10.6342/NTU201704140

doing FFA to the sub-networks iteratively, we would get a binary tree.[Fig. **??**] And we call this minimum-cut binary tree (MBT), which contains all the information of RDS in the networks.



Algorithm 2: Construct Minimum-cut Binary Tree (CMBT)

input : Graph G
output: minimum cut tree
$s =$ heighest energy state
$t =$ lowest energy state
$root.data = G$
if $G_f.size > 1$ **then**
  $C, G_f = \text{FFA}(G, s, t)$
  // use Breadth-First-Search(BFS) to determine subgraph
  subgraph $S = \text{BFS}(G_f, s)$
  subgraph $T = G - S$
  $root.cut = C$
  $root.lt = \text{CMBT}(S)$
  $root.rt = \text{CMBT}(T)$
else
  return $root$ // leaf

3.2 Coarse-grained model

When we obtain the minimum-cut binary tree, we need a method to build the coarse-grained model. And we test if we should modified the MBT by normalizing the weight. Finally, we test three simple method and then identify which one is better.

3.2.1 MBT normalized

The weight of minimum cut would be larger if the subnetworks have more members. The weight has extensive property, and we also test if normalized the weight of minimum-cut could give the better result.

$$C_{norm} = \frac{C}{N_V(S) \times N_V(T)} \tag{3.1}$$

where C denotes the weight of minimum-cut.

doi:10.6342/NTU201704140

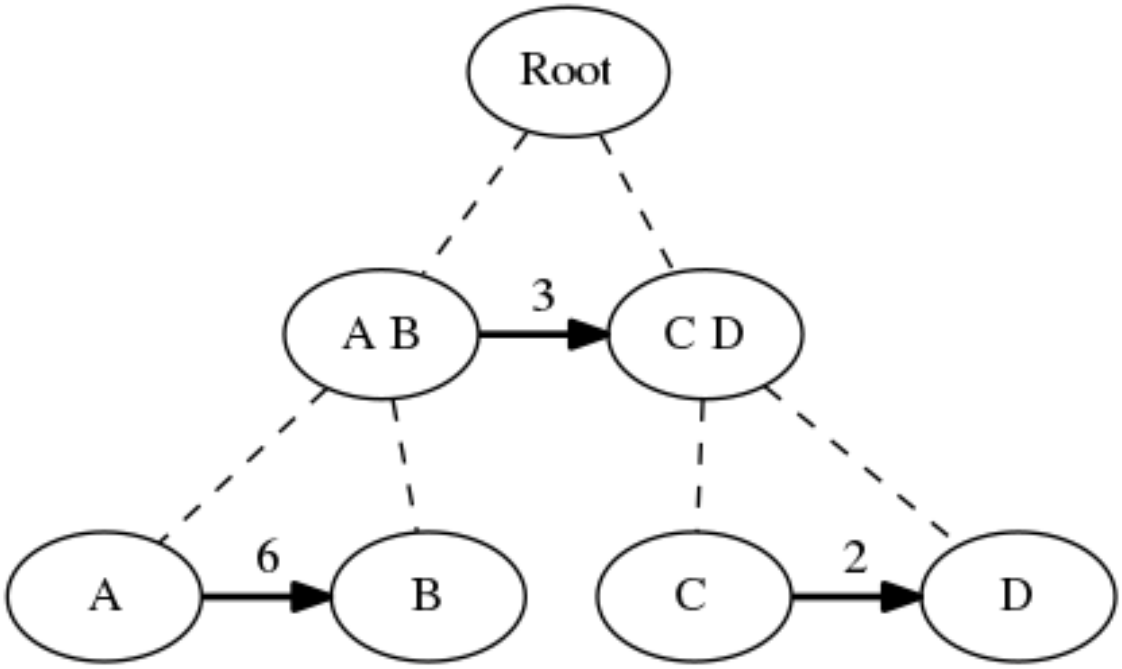


Figure 3.1: A simple minimum-cut binary tree, which original network has 4 nodes.

3.2.2 Simple cut-off method

The ideal is follow the FFA, the first cut should be the smallest minimum cut. So when we choose a cut-off, the subtree should be in the same cluster.

Setting a cut-off. From the top of the MBT, if the weight of minimum cut is large than the cut-off, then the sub-network is a cluster of the coarse-grained model. For example, if the cut-off is 5, the coarse-grained model of the simple MBT is $\{AB\}\{C\}\{D\}$. And if the smaller cut-off we choose, we would obtain the coarse-grained model with more clusters. The limit would be the same with the original networks. By tuning the cut-off, we can obtain any size of coarse-grained model. , Alg. 3

Algorithm 3: simple cut-off (SC) method

input : CGM, MBT, cut-off
output: coarse-grained model(CGM)
$root = MBT.root$
if $root.cut >$ *cut-off* **then**
  CGM.push(root)
else
  SC($CGM, root.lt$, cut-off)
  SC($CGM, root.rt$, cut-off)
return CGM

doi:10.6342/NTU201704140

3.2.3 Simple ratio cut-off method

This method is just do some modified of SC. We guess if that the nearest subtree has strong relationship. Setting a cut-off ratio. From the top of the MBT, if the weight of minimum cut divide by the weight of minimum cut of the top level is larger than the cut-off ratio, then the sub-network is a cluster of the coarse-grained model. For example, if the cut-off ratio is 3, the coarse-grained model is $\{AB\}\{C\}\{D\}$. Alg. 4 And we also can obtain differece size by tuning the cut-off.

Algorithm 4: simple ratio (SR) method

input : CGM, MBT, cut-off
output: coarse-grained model(CGM)
$root = MBT.root$
if $root.lt.cut/root.cut >$ *cut-off* **then**
  CGM.push(root.lt)
else
  SR($CGM, root.lt$, cut-off)
if $root.rt.cut/root.cut >$ *cut-off* **then**
  CGM.push(root.rt)
else
  SR($CGM, root.rt$, cut-off)
return CGM

3.2.4 Ascending cut-off method

Since each time we do the FFA, the source and target would be differece. So the simple-cut method may not be enough. The whole MBT's minimum-cuts may be equally important. This is another modified of SC. Consider all the nodes in one set, and then eliminate some elements into another set step by step, from smallest cut-off to largest cut-off. This method can give us 2 to N clusters CGM. For example, the smallest weight of the Fig. 3.1 is 2, then we can get 2 clusters model, $\{ABC\}\{D\}$. And the second weight is 3, we get $\{AB\}\{C\}\{D\}$. Finally the largest weight is 6, we get $\{A\}\{B\}\{C\}\{D\}$. Alg. 5

doi:10.6342/NTU201704140

Algorithm 5: ascending cut-off (AC) method

input : CGM, MBT, cut-off
output: coarse-grained model(CGM)
cut-off list is from 0 to cut-off, we can simply implentment by sorting all cut value of MBT
$root = MBT.root$
$CGM = \{1..N\}$
for *each cut-off* $\in$ *cut-off list* **do**
  $subtree =$
   find the maximum cut, which is smaller than cut-off, subtree of MBT
  $subtree = \max_{cut}\{subtree.lt, subtree.rt\}$
  $GCM = \{\{GCM - subtree\}\{subtree\}\}$
return CGM

3.3 Reduced dynamics

We test several approaches and need to identify which one would perform better. We follow the definition of the coarse-grained model, the model would reproduce complexity phenomena by simpler representation. So we try that if the coarse-grained model built in above can reproduce the propagation dynamics of the networks. We also need to know the transfer rates between cluster to cluster, we make an assumption that as the states were in the same cluster, they would reach thermal equilibrium quickly. In this quasi-thermal equilibrium approximation, we can define the transfer rate between clusters, Eq. 3.2

$$\mathbf{k}_{FI} = \sum_{\alpha \in F} \sum_{\beta \in I} \mathbf{R}_{\alpha\beta} \frac{e^{-\frac{\epsilon_\beta}{k_B T}}}{Z_I} \tag{3.2}$$

, where Z_I is the partition function of the cluster I.

And we can compare the reproduce population dynamics and the total dynamics to know which method is better in a simpler system. We define a cost function to show how similar of reduced dynamics and original dynamics. Eq. 3.4

$$cost = \frac{1}{N_c T} \sum_{I=1}^{N_c} \int_0^T dt (P^I_{original}(t) - P^I_{reduced}(t))^2 \tag{3.3}$$

$$P^I_{original}(t) = \sum_{state \in I} P^{state}_{original}(t) \tag{3.4}$$

,where N_c is number of clusters in coarse-grained model. T is a enough long time, in

doi:10.6342/NTU201704140

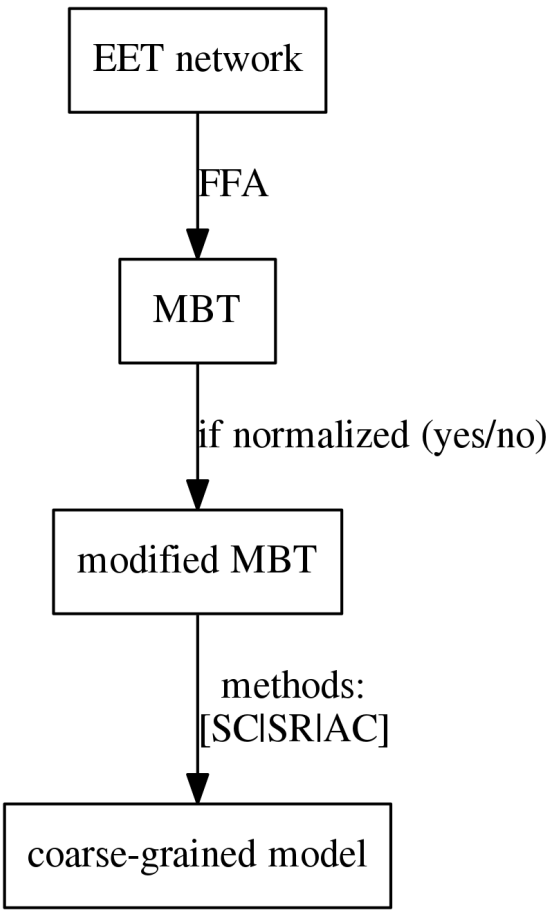


Figure 3.2: EET network ananlysis process

our work is 60 ps. And the using the better method to analysis more complex system. In additionally, we also check how the CGM. Because sometime the small cost is that the networks just separate by single node, and the population dynamics would not be change a lot. The total process is simple, shows in [Fig. 3.2].

We compare these method in following context, and the notation would be simplified as Tabel 3.1.

3.3.1 MBT rearrangement

In our study we also found that sometimes the MBT gives suspicious-looking tree that seems to lead to unreasonable clustering results. For example, in Fig. 3.3 we show a specific case. Nevertheless, we found that the MBT rearrangement does not improve the results when dealing with CGM of photosynthetic complexes. Therefore, the MBT rearangement results will not be presented here.

doi:10.6342/NTU201704140

Table 3.1: The notation of each method: In the context, we use the simple notation to denote each method.

MBT normalized	Methods	Notation
No	Simple Cut-off Method	SC
No	Simple Ration Cut-off Method	SR
No	Ascending Cut-off Method	AC
Yes	Simple Cut-off Method	norm-SC
Yes	Simple Ration Cut-off Method	norm-SR
Yes	Ascending Cut-off Method	norm-AC

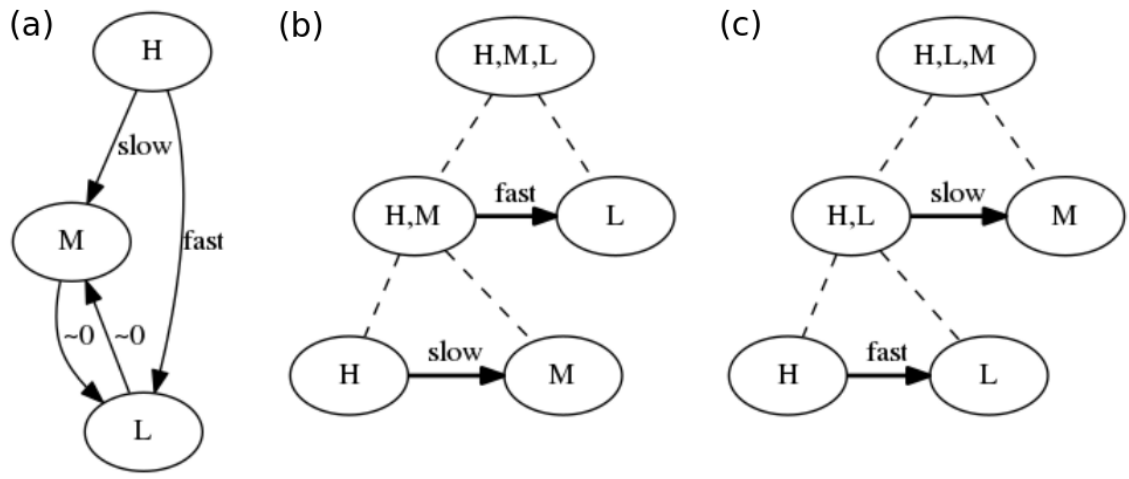


Figure 3.3: Reconstruct MBT by rearranging the subtree: (a) is the specific three nodes network. H denotes the highest exciton state. M denotes the second heigh exciton state. And L denotes the lowest exciton state. (b) shows the MBT construct by FFA. (c) shows the rearangement of MBT, the subtree with slower rate will switch with the faster one.

doi:10.6342/NTU201704140

Chapter 4

Fenna-Mattews-Olson complex

Fenna-Mattews-Olson (FMO) complex is a simple photosynthetic system in bacteria, we first test our methods on this system.

4.1 Effective Hamiltonian

To describe EET dynamics in FMO, we follow the simulations carried out by Wu et al. [16] and adopt the effective Hamiltonian proposed by Schmidt et al. [12]. Excitonic couplings in this Hamiltonian were calculated based on a 1.3 Å crystal structure [13]. This model Hamiltonian has been tested and shown to yield time-resolved transient-absorption and fluorescence spectra that are in excellent agreement with the experiments, hence the model should reproduce accurate population dynamics in the FMO system. (Tabel 4.1)

Table 4.1: The effective Hamiltonian of FMO is obtained from Schmidt et al. work[12] .(unit: cm^{-1})

BChl1	BChl2	BChl3	BChl4	BChl5	BChl6	BChl7	BChl8
12505	94.8	5.5	-5.9	7.1	-15.1	-12.2	39.5
94.8	12425	29.8	7.6	1.6	13.1	5.7	7.9
5.5	29.8	12195	-58.9	-1.2	-9.3	3.4	1.4
-5.9	7.6	-58.9	12375	-64.1	-17.4	-62.3	-1.6
7.1	1.6	-1.2	-64.1	12600	89.5	-4.6	4.4
-15.1	13.1	-9.3	-17.4	89.5	12515	35.1	-9.1
-12.2	5.7	3.4	-62.3	-4.6	35.1	12465	-11.1
39.5	7.9	1.4	-1.6	4.4	-9.1	-11.1	12700

doi:10.6342/NTU201704140

Table 4.2: The rate constant matrix of FMO is obtained from modified Redfield theory, and the parameters of spectral density function is from Wu et al. work[16].(unit: ps^{-1})

state 1	state 2	state 3	state 4	state 5	state 6	state 7	state 8
-0.373	1.159	0.191	0.293	0.128	0.092	0.084	0.004
0.309	-2.417	0.052	1.089	1.855	0.098	0.922	0.018
0.047	0.045	-1.342	0.249	0.138	3.632	0.179	0.275
0.022	0.485	0.127	-2.996	1.133	0.071	2.079	0.039
0.004	0.654	0.051	0.869	-3.82	0.062	1.563	0.064
-0.002	0.022	0.897	0.036	0.042	-4.414	0.019	1.082
-0.007	0.052	0.013	0.453	0.507	0.009	-4.872	0.031
0.0	0.0	0.011	0.007	0.017	0.45	0.026	-1.513

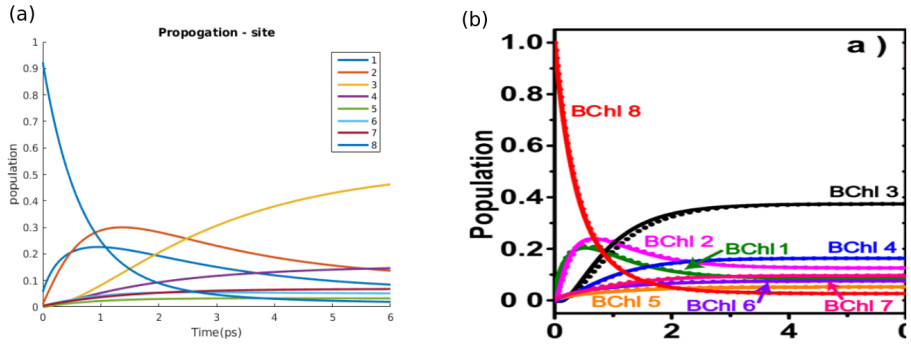


Figure 4.1: Check the rate constant matrix by comparing the population dynamics of MRT result(a) to Wu et al. work(b).

4.2 Rate constant matrix

The parameters of the spectral density function is obtained from Wu et al. work[16]. $T = 300.0\,K^{-1}$, $\lambda_0 = 35.0\,cm^{-1}$, $\Gamma = 666.7\,cm^{-1}$ (Tabel 4.2)

4.3 MRT population dynamics

To test if MRT [Fig. 4.2] could give the proper result of excitons population dynamics, we first check by comparing our result with Wu et al. work[16]. See Fig. 4.1 The transfer rate of MRT is smaller, but this difference would not impact on the coarse-grained. We can construct the MBT by FFA. Fig. 4.2

doi:10.6342/NTU201704140

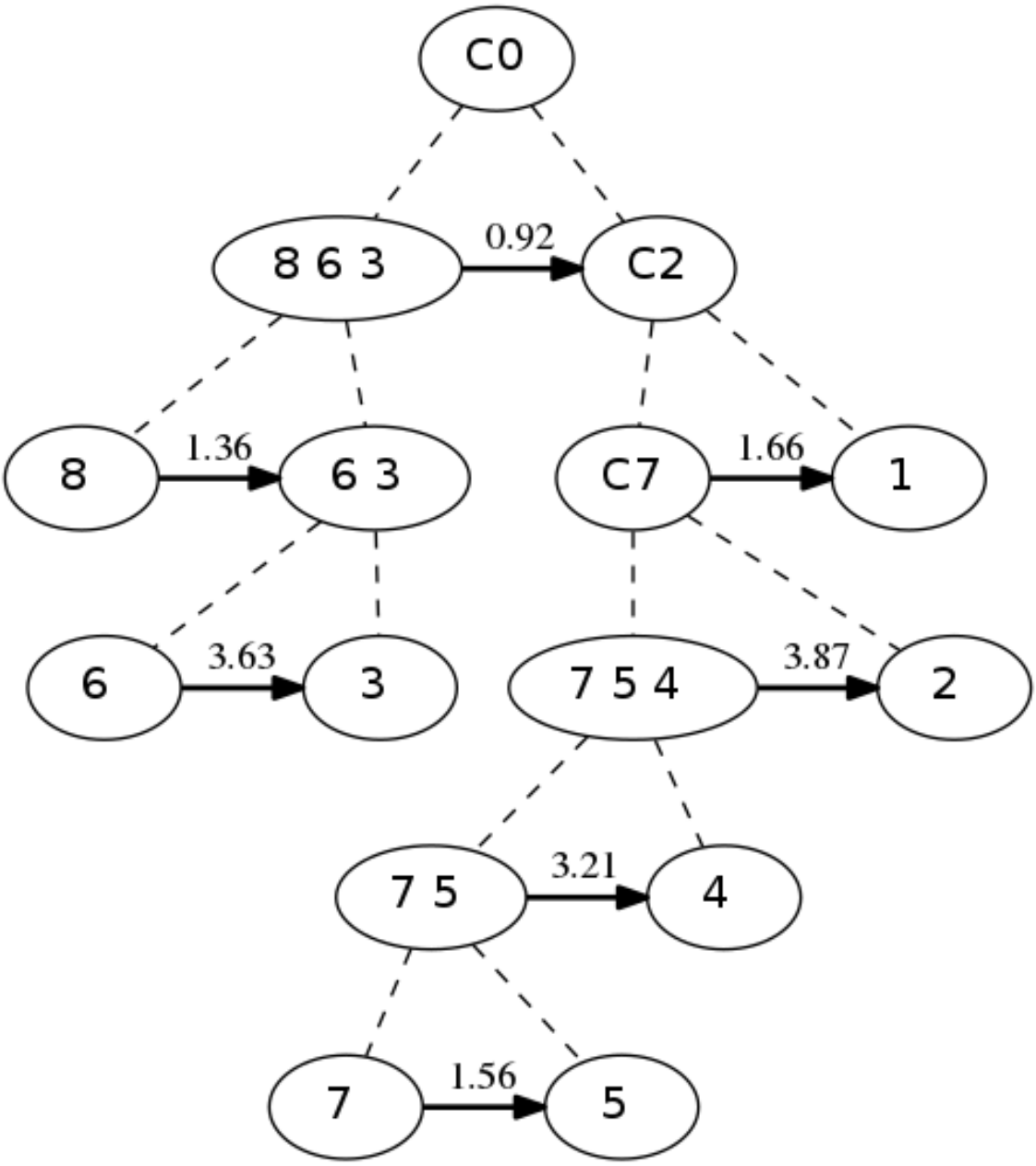


Figure 4.2: FMO Minimum-Cut Binary Tree (unit: ps^{-1}): The ellipses denote the subgraphs and each number α in the ellipse denotes the α-th exciton state.

doi:10.6342/NTU201704140

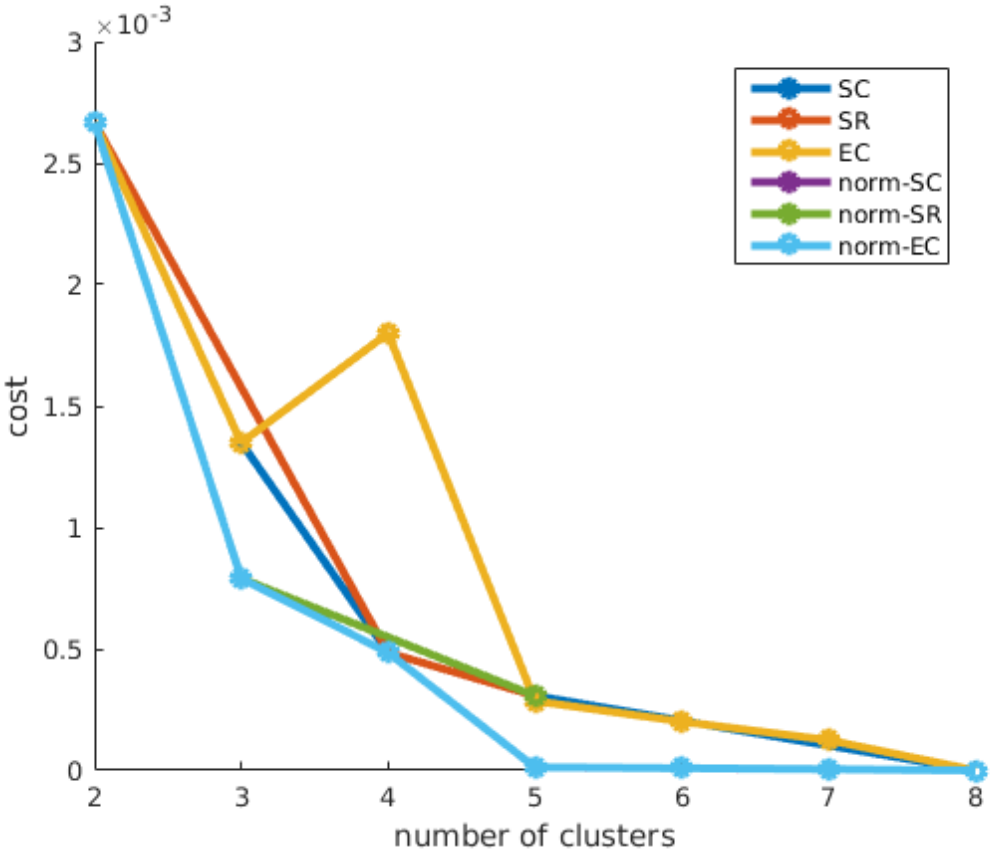


Figure 4.3: FMO cost comparing: Each color denotes a different method. In the case of FMO, the

4.4 Network analysis

In the chapter 3, we can build CGM at some options. First, if the MBT need to rearrange? Second if the weight of minimum-cut need to normalize? And after modifying the MBT, choosing a method to build the coarse-grained model. Each method would give us a difference CGM. Fig. 4.3

By the simplest strategy, we can choose the lowest cost method. In additionally, we also check what the order of the cluster that the given method cut out. But FMO networks is too small, so all these methods do not show obviously difference. Many CGM are the same although we use different method to build. And the SR method can only build three coarse-grained model whether we tune the cut-off ratio. So we would focus on the SC and AC methods.

We can choose any number of clustering. For instance we can test the population

doi:10.6342/NTU201704140

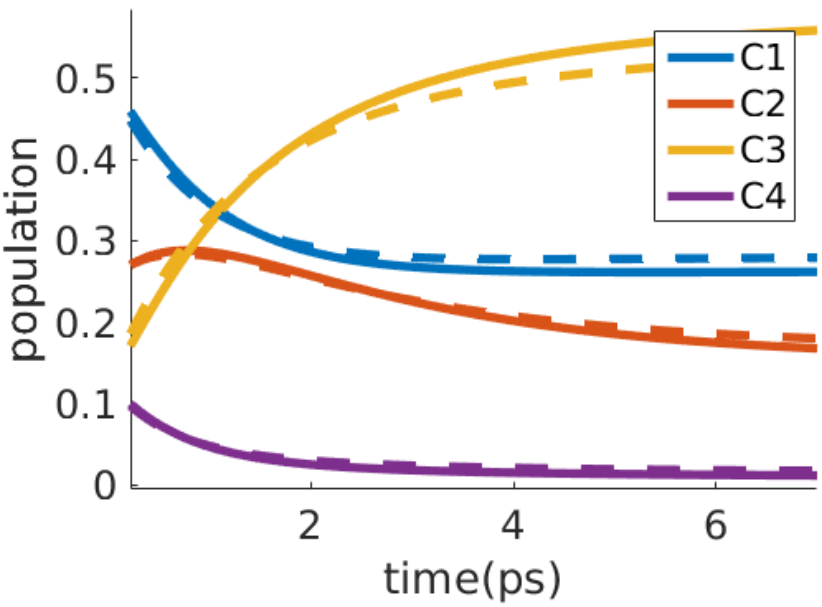


Figure 4.4: FMO 4 clusters population dynamics: The solid lines are the CGM dynamics, and the dash lines are the exact MRT dynamics.

dynamics with 4 clusters built by method SC, norm-SC, or norm-AC, see Fig. 4.4 The solid lines are the original population dynamics and the dashed lines are dynamics by the quasi-thermal equilibrium approximation. We can see that they are almost the same.

For a small network like FMO, all the clustering do not show obviously different. The coarse-grained model of FMO may be good in 4(or more) clusters. In Fig. 4.5, we show the simpler network of FMO. And in Fig. 4.6, we map the coarse-grained model into real space. We can see the energy flow can go through two pathways from highest energy exciton state to lowest energy exciton state.

doi:10.6342/NTU201704140

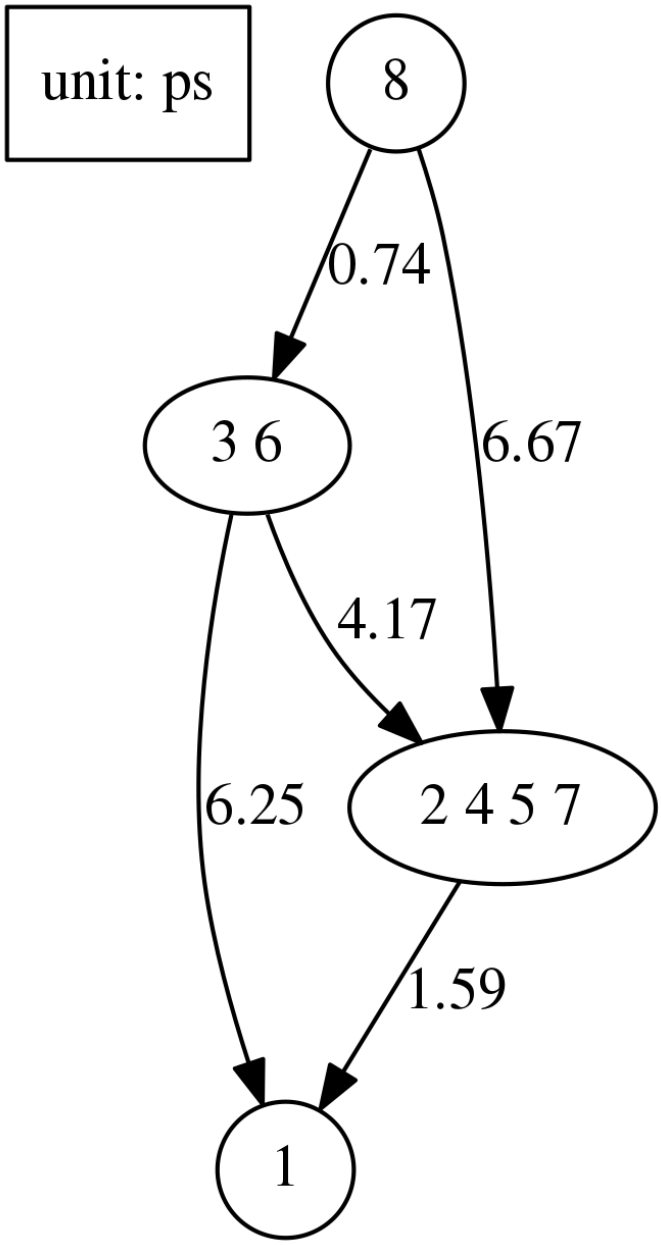


Figure 4.5: FMO 4 clusters coarse-grained model with abstract diagram, we can see the EET flow simply.

doi:10.6342/NTU201704140

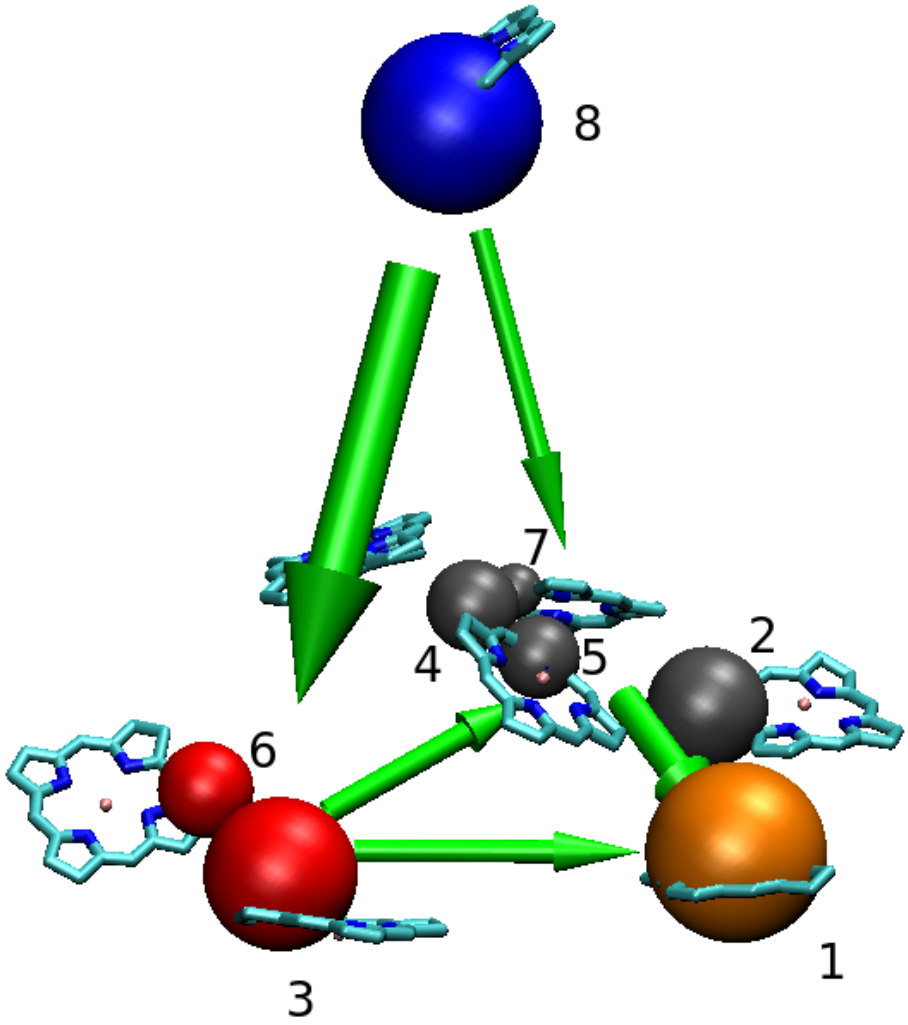


Figure 4.6: FMO 4 clusters coarse-grained model in real space: Each ball's centor denote the centor of each exciton state. And each color denote the cluster of CGM. The size of ball shows the relative population in the cluster base on the eigenenergies.

doi:10.6342/NTU201704140

doi:10.6342/NTU201704140

Chapter 5

Light Harvesting Complex II

LHCII complex is a membrane protein, it participates in the 1st steps of photosynthesis by harvesting sunlight and transfer excitation energy to the core complex. The structure of LHCII is a trimer. To simplified the system, we only consider the monomer of LHCII.

5.1 Effective Hamiltonian

To describe EET dynamics in a monomeric subunit of the LHCII complex, we follow the simulations carried out by Wu et al. [16] and adopt the effective Hamiltonian (Table 4.1) proposed by Novoderezhkin et al. [9]. Excitonic couplings in this Hamiltonian were calculated based on a 2.72 Å crystal structure [7] using a transition-dipole-transition-dipole interaction model for chlorophylls, and the site energies were determined by fitting to various experimental linear spectra of the LHCII systems. This model Hamiltonian has been tested and shown to yield time-resolved transient-absorption and fluorescence spectra that are in excellent agreement with the experiments, hence the model should reproduce accurate population dynamics in the LHCII monomer system. Note that although a more recent modification to the model Hamiltonian has bee proposed to improve the simulations of two-dimensional electronic spectra [1], the differences in the updated Hamiltonian are minor and should not affect the conclusion of our study.

doi:10.6342/NTU201704140

5.2 Rate constant matrix

For simplicity, we only consider the low-frequency Debye spectral density and ignore all the high-frequency components. This approximation neglects population dynamics from resonant electron-phonon couplings[9, 11, 15, 10, 6], which is acceptable considering our main purpose of calculating excitation energy transfer rate. The parameter of MRT is from previous work[16]. T = 300.0 K^{-1}, $\lambda_0 = 85.0\,cm^{-1}$, $\Gamma = 628.4\,cm^{-1}$ (Tabel 5.2)

5.3 MRT population dynamics

We also compare the MRT population dynamics with previous Wu et al. work[16]. In Fig. 5.1, we can see the transfer rate of MRT is smaller, but this difference would not impact on the coarse-grained.

5.4 Network analysis

The MBT of LHCII monomer is Fig. 5.2 And we also compare all the methods and check their costs. Fig. 5.4

As the CGM has more clusters, the reproduced population dynamics should close to the exact population dynamics, and has lower cost. But we can see that some methods has small cost when number of clusters is small, but get larger when the number of clusters is larger. Fig. 5.4 So we think that may not obtain suitable CGM. The method AC, norm-SC, norm-SR and norm-AC (Fig. 5.5) are apparently decrease as the numbers of cluster increase.

In additionally, comparing these methods, three methods are modified with normalized MBT. That maybe mean that the transfer rate constant has extensive property. Because SR and SC methods are not able to build the coarse-grained model with an arbitray number of clusters, so we focus on method 002 and method 012. The only different between them is if we normalize the MBT.

We can check the CGMs. In first cut, they both cut out exciton state 1, 2 and 5. And in

doi:10.6342/NTU201704140

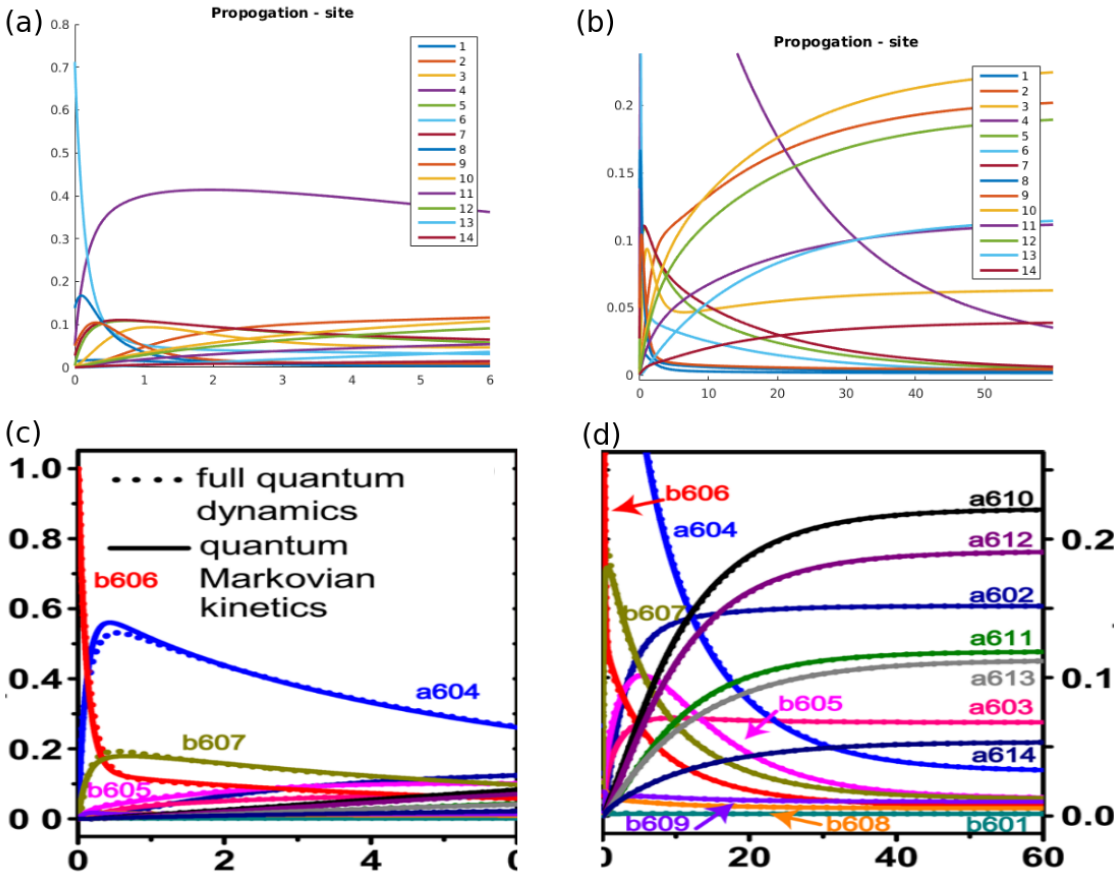


Figure 5.1: Check the rate constant matrix by comparing the population dynamics of MRT result(a,b) to Wu et al. work(c,d).

doi:10.6342/NTU201704140

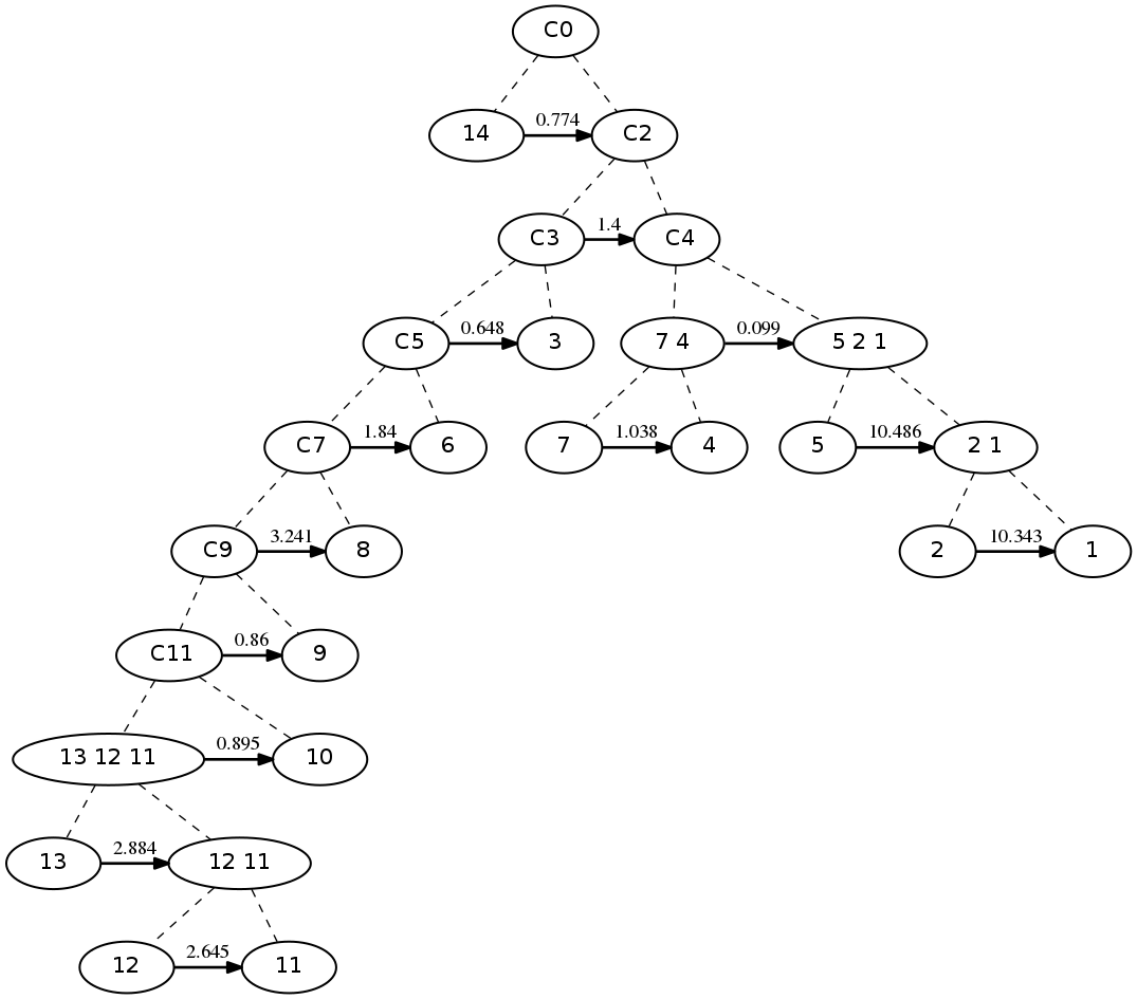


Figure 5.2: LHCII monomer MBT (unit: ps^{-1}): The ellipses denote the subgraphs and each number α in the ellipse denotes the α-th exciton state.

doi:10.6342/NTU201704140

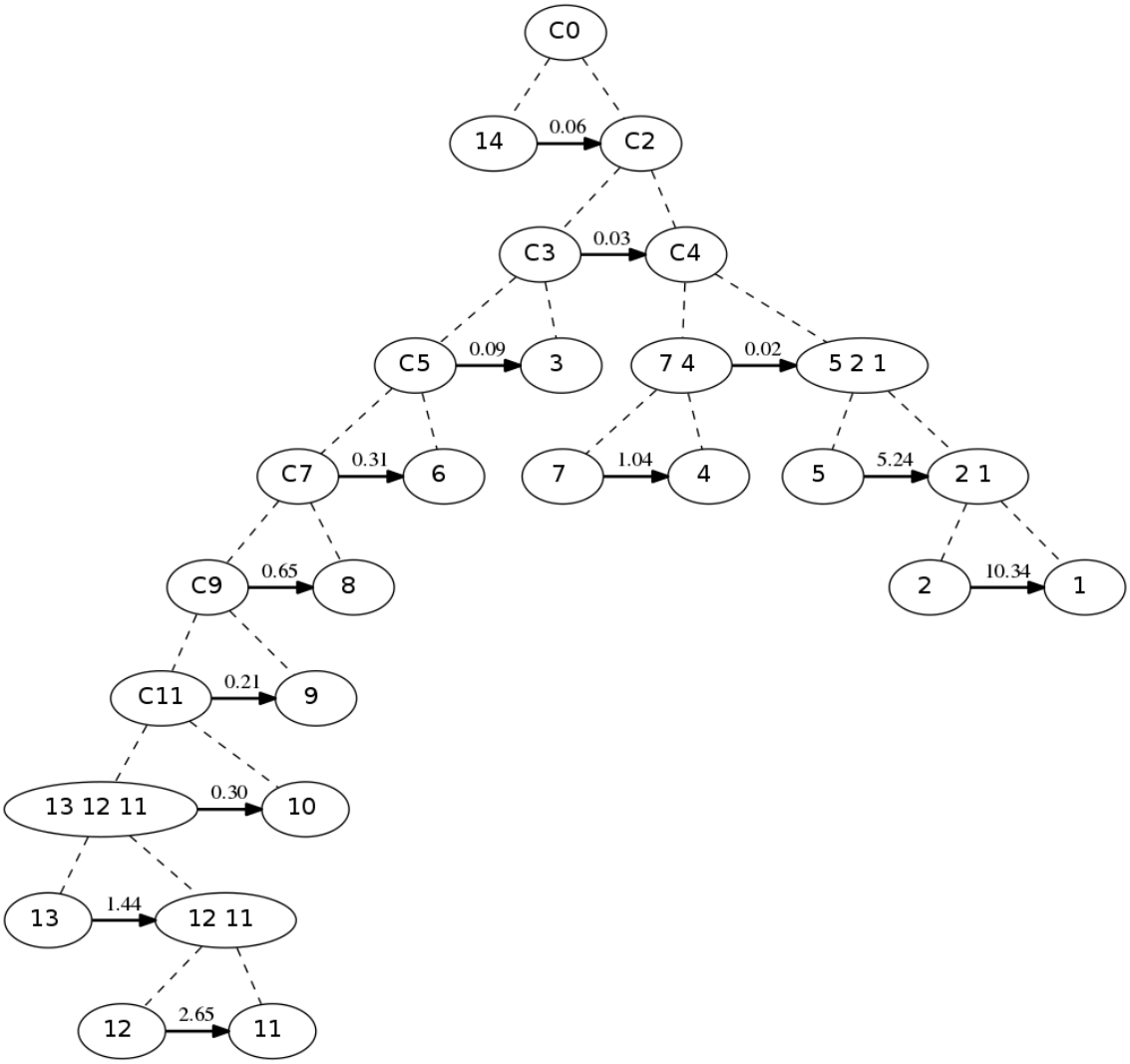


Figure 5.3: The modified MBT of LHCII: The capacity of minimum-cut is normalized.

doi:10.6342/NTU201704140

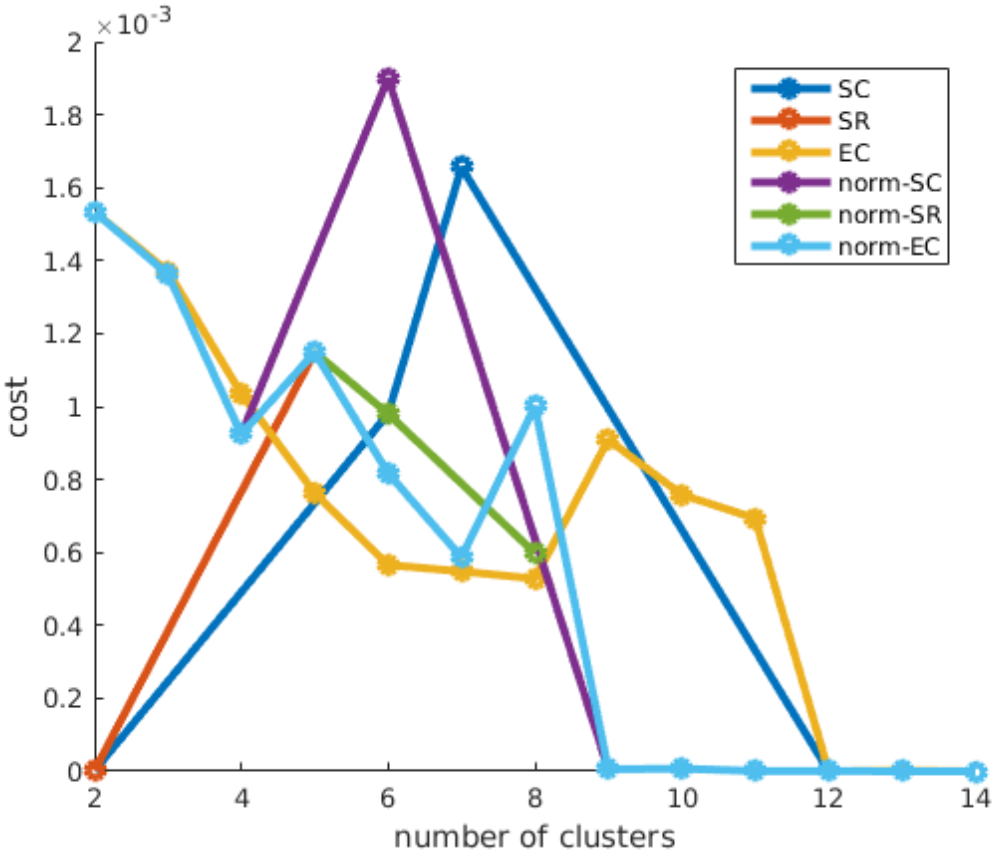


Figure 5.4: LHCII cost comparing: Each color denotes different methods of building coarse-grained model. And some methods like SC and norm-SC can not build coarse-grained model with an arbitrary number of clusters.

doi:10.6342/NTU201704140

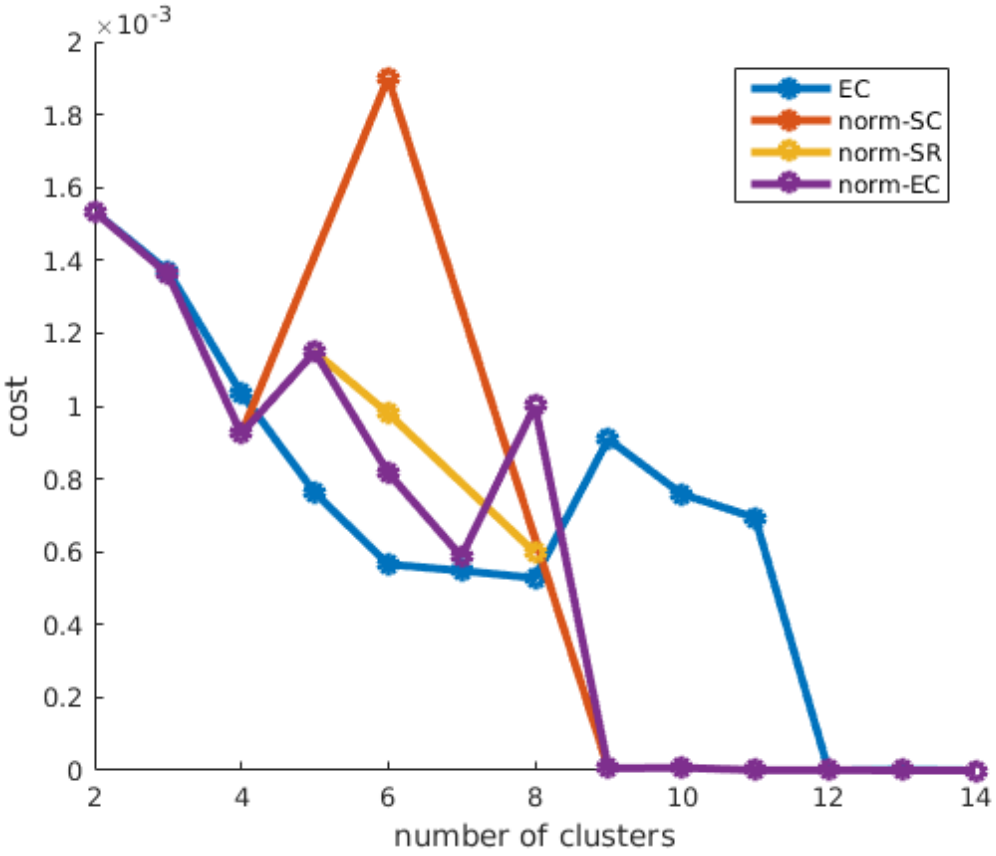


Figure 5.5: LHCII cost comparing: The methods have decreasing cost as increasing the number of clusters.

next cut, method 012 would cluster exciton state 4 and 7, but the method 002 would just cut out the node one-by-one. Exciton state 4 and 7 are localized at a613 and a614, and the EET transfer rate from state 7 to state 4 is 1.0 ps^{-1}. So it's reasonble to merge them together. Finally, we choose method 012. And this result also consists with Novoderezhkin et al. and Wu et al. works[8, 16].

5.5 Coarsed-grained model

In Fig. 5.5 and Fig. 5.6, we can see that the 9 clusters CGM fits population dynamics perfectly. And from small number of clusters, such as 5 clusters CGM, the network structure is simple and we can find that the 3rd exction state is the intermediate and may be important in this photosynthetic network. And in the 7 clusters model, the 9th and 10th exciton state is the terminal and could be neglect in the network. And Fig. 5.8 is the 9

doi:10.6342/NTU201704140

clusters CGM in the real space. In Fig. **??**, although the 9 clusters CGM has the best fitting coarse-grained dynamics. We can still obtain some graph information from 3-8 clusters CGMs. We can see that cluster {4,7} has departed early, and this result demonstrates that node 4 and 7 is the trap in the LHCII monomer EET networks. This can be compare with the result of Schlau-Cohen et al. work [11].

doi:10.6342/NTU201704140

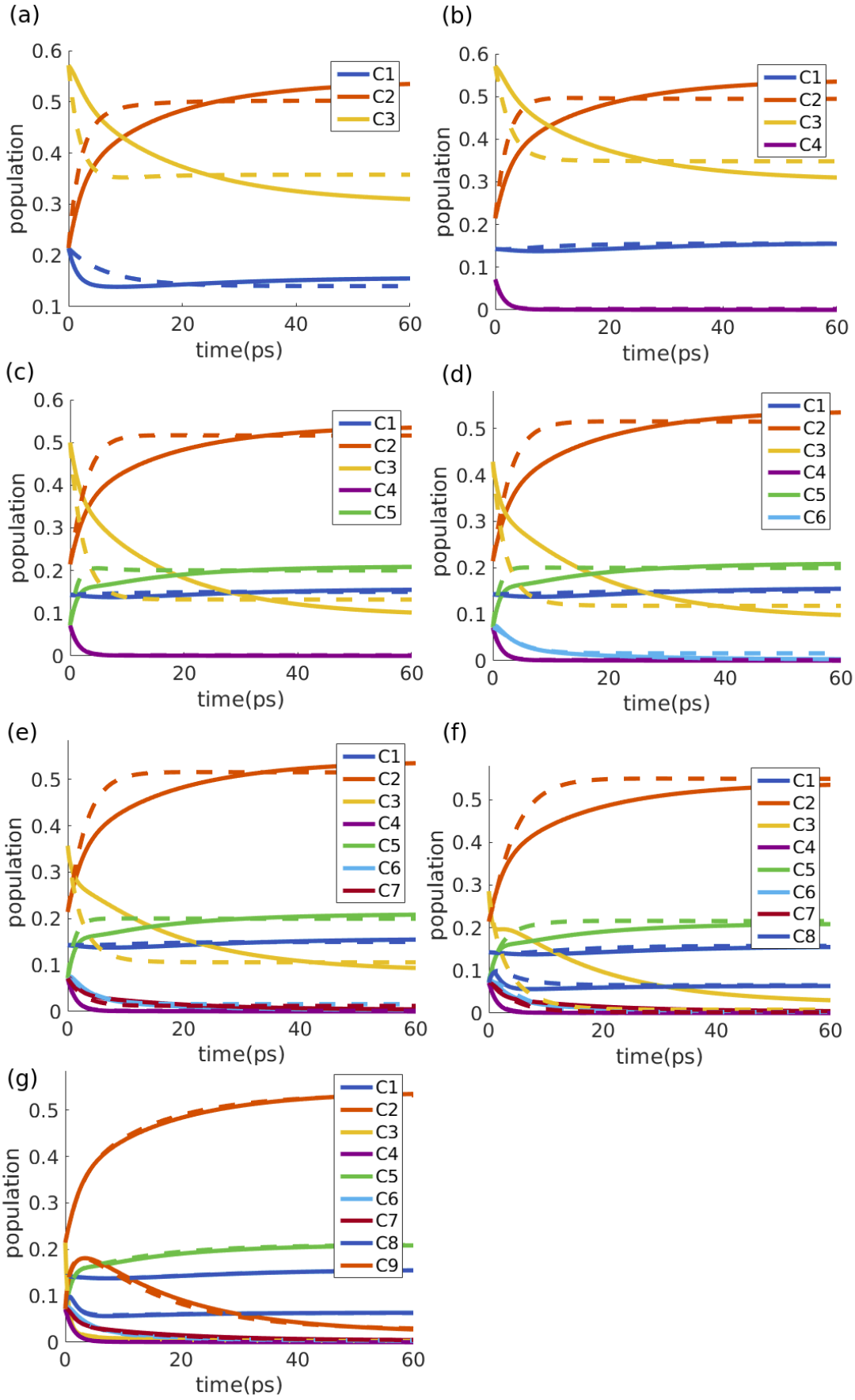


Figure 5.6: Coarse-grained dynamics. (a)-(g) each denotes the cluster population calculated by using the simplified model with 3-9 clusters respectively(solid line). For comparison, the clustered population calculated from the full 14-site model are also denote(dash line).

doi:10.6342/NTU201704140

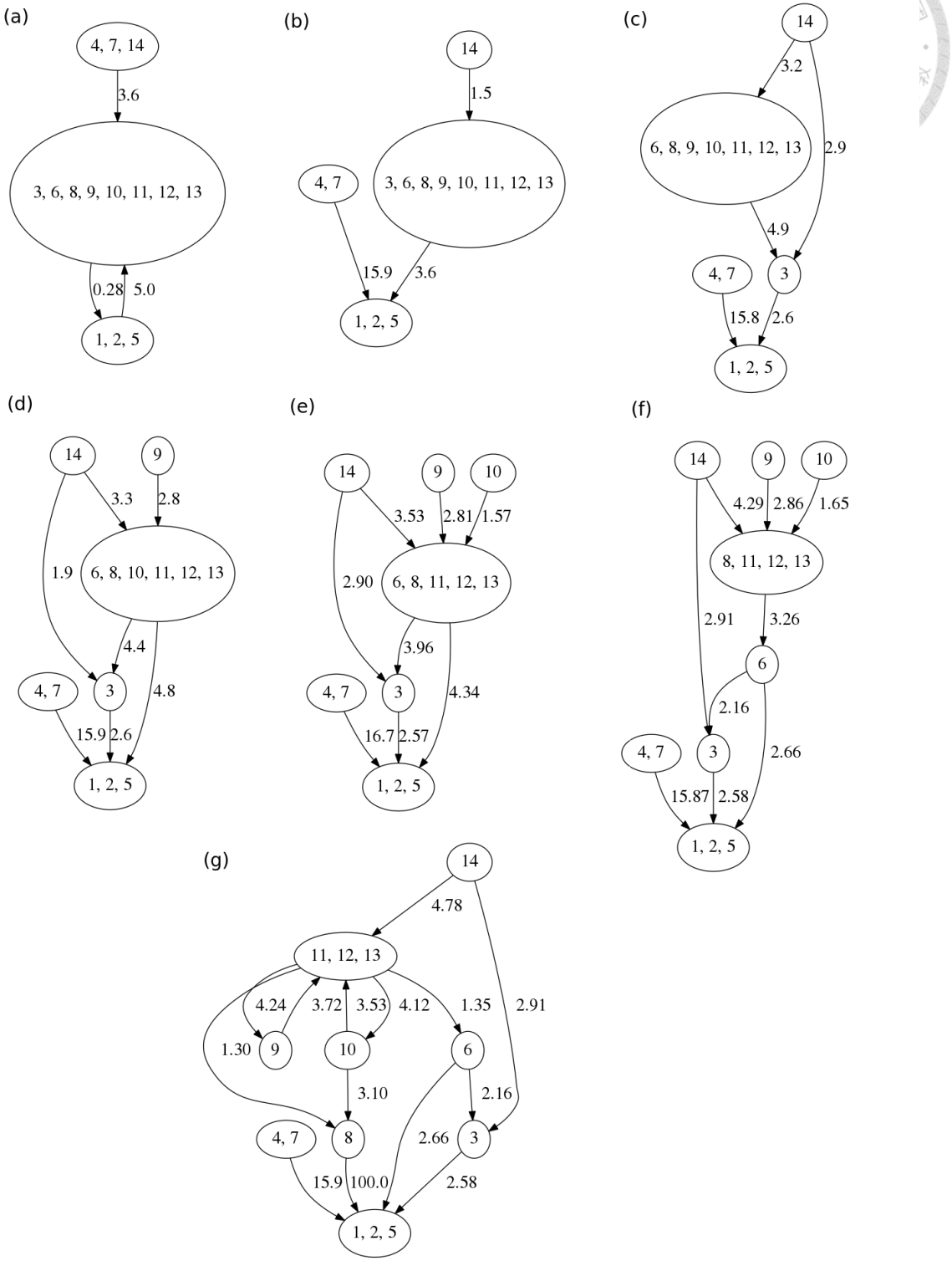


Figure 5.7: LHCII 3-9 clusters CGMs represent in abstract diagrams respectively. (unit: ps)

doi:10.6342/NTU201704140

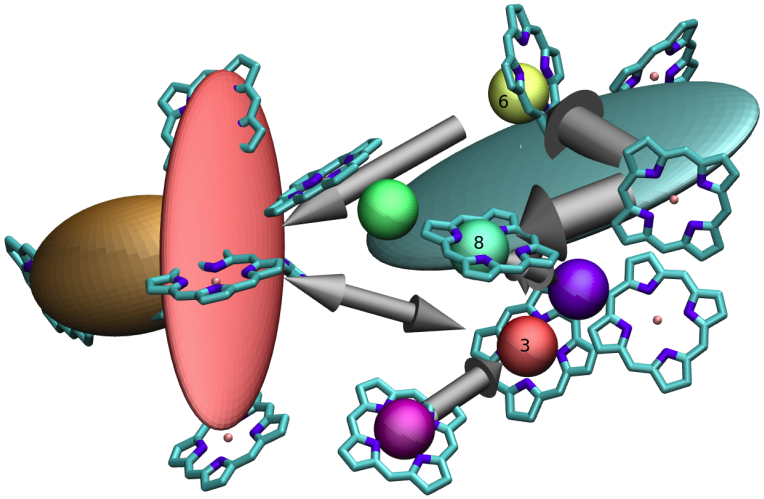


Figure 5.8: LHCII 9 clusters CGM in real space: Each color denotes a cluster. We can see the node 3(red ball) and node 6(yellow ball) are the intermediate. And the node 8(green ball) maybe the bottleneck state.

doi:10.6342/NTU201704140

Table 5.1: The effective Hamiltonian of LHCII is obtained from Novoderezhkin et al. work[9]. (unit: cm^{-1})

601	602	603	604	605	606	607	608	609	610	611	612	613	614
15764	-47.1	-6.1	-2.7	0.5	-2	-2.6	3.3	4.4	-4.5	24.5	2.3	-8.4	2.9
-47.1	15103	17.4	5.5	-0.2	4.9	6.2	-5.8	-21.9	-5.4	0.7	10.1	-1.9	0.1
-6.1	17.4	15262	-0.5	-0.2	-2.1	8.2	4.2	71.6	8.4	-0.7	-0.6	2.4	-5.7
-2.7	5.5	-0.5	15390	5.4	80.8	26	-5.7	-1.5	-0.2	-3.3	3.7	2.2	-2.8
0.5	-0.2	-0.2	5.4	15544	11.5	-5.2	-3.7	-0.1	0.8	1.1	-2.2	-1.2	0
-2	4.9	-2.1	80.8	11.5	15686	23.7	-6.7	-11.8	-0.6	-2	2.1	1.2	-1.8
-2.6	6.2	8.2	26	-5.2	23.7	15557	-3.5	-1.7	-0.4	-2.1	2.2	2.7	-2.4
3.3	-5.8	4.2	-5.7	-3.7	-6.7	-3.5	15678	26.1	57	4.8	-1.3	-2.2	1.4
4.4	-21.9	71.6	-1.5	-0.1	-11.8	-1.7	26.1	15616	1.1	3.3	-0.2	-2.5	2
-4.5	-5.4	8.4	-0.2	0.8	-0.6	-0.4	57	1.1	15038	-26.4	12.4	6	-1.2
24.5	0.7	-0.7	-3.3	1.1	-2	-2.1	4.8	3.3	-26.4	15180	105	-0.8	0.6
2.3	10.1	-0.6	3.7	-2.2	2.1	2.2	-1.3	-0.2	12.4	105	15082	-1	-0.2
-8.4	-1.9	2.4	2.2	-1.2	1.2	2.7	-2.2	-2.5	6	-0.8	-1	15160	-28
2.9	0.1	-5.7	-2.8	0	-1.8	-2.4	1.4	2	-1.2	0.6	-0.2	-28	15309

doi:10.6342/NTU201704140

Table 5.2: The rate constant matrix of LHCII is obtained from modified Redfield theory, and the parameters of spectral density function is from Wu's work[16]. (unit: ps^{-1})

state 1	state 2	state 3	state 4	state 5	state 6	state 7	state 8	state 9	state 10	state 11	state 12	state 13	state 14
-10.147	10.343	0.315	0.038	7.493	0.234	0.002	0.004	0.001	0.001	0.021	0.122	0.021	0.014
8.273	-11.264	0.037	0.045	2.993	0.142	0.003	0.003	0.001	0	0.039	0.229	0.038	0.01
0.236	0.035	-0.532	0.009	0.168	0.464	0.005	0.012	0.001	0.006	0.135	0.017	0.013	0.344
0.02	0.028	0.001	-0.425	0.006	0.009	1.038	0.003	0	0.002	0.001	0.001	0.001	0.008
1.565	0.812	0.036	0.003	-10.671	0	0.008	0.001	0	0.001	0.066	0.007	0.004	0.082
0.053	0.045	0.14	0.005	0	-0.915	0.062	0.001	0.005	0.029	1.588	0.14	0.077	0.05
0	0	0.001	0.325	0.006	0.027	-1.12	0.005	0.001	0.004	0.014	0.002	0.005	0.008
0	0.001	0.002	0	0.001	0	0	-0.166	0.081	0.323	0.019	0.416	2.402	0.024
-0	-0	-0	0	0	0	0	0.021	-0.361	0	0.022	0.122	0.716	0.007
-0	-0	-0	-0	0	0.002	0.001	0.082	0	-0.654	0.009	0.117	0.769	0.018
0	0	0	-0	0.004	0.037	0.001	0.001	0.012	0.004	-3.675	2.645	0.613	0.032
0	0	0	-0	-0	0	-0	0.002	0.042	0.041	1.524	-5.627	2.271	0.037
0	0	0	-0	-0	0	-0	0.031	0.215	0.238	0.227	1.785	-7.055	0.14
0	0	0	0	0	0	-0	0	0.002	0.005	0.01	0.024	0.125	-0.774

doi:10.6342/NTU201704140

doi:10.6342/NTU201704140

Chapter 6

Photosystem I

Photosystem I(PSI) is an integral membrane protein, and the lowest excited state is located at the reaction center. The reaction center is actually the sink of the excitation energy transfer network.

6.1 Effective Hamiltonian

The effective Hamiltonian for PSI constructed by Fleming and coworkers [2].

6.2 Rate constant matrix

This rate constant also obtained from MRT. This work is calculated by Ai, Qing. The spectral density function is the same, the Debye spectral density.T = 300.0 K^{-1}, $\lambda_0 = 100.0\,\mathrm{cm}^{-1}$, $\Gamma = 100.4\,\mathrm{cm}^{-1}$

6.3 Network analysis

The MBT is large so we do not plot here, and we the methods compare is Fig. 6.1. And we also choose the method norm-AC because it success in LHCII. Fig. 6.2

Refer to the Fig. 6.2, the CGM of PSI may be good in 12 clusters or more. We can see that the CGM has two big clusters, which is opsite from reaction center, and we can find

doi:10.6342/NTU201704140

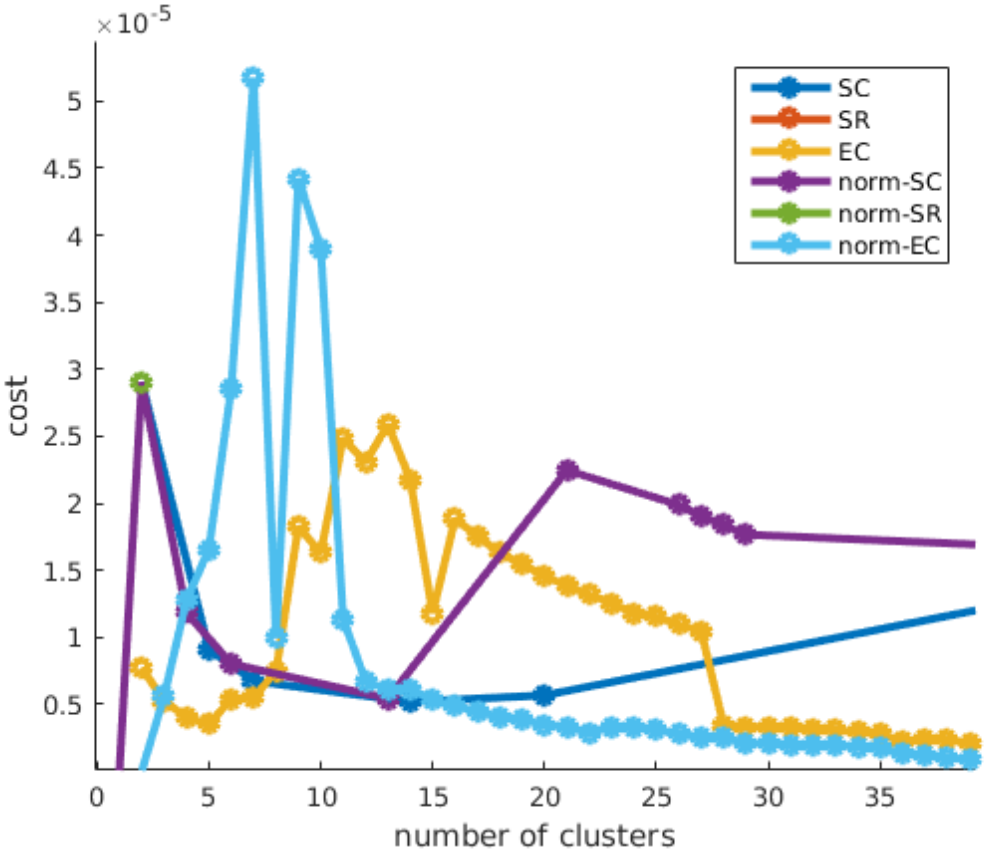


Figure 6.1: PSI cost comparing

doi:10.6342/NTU201704140

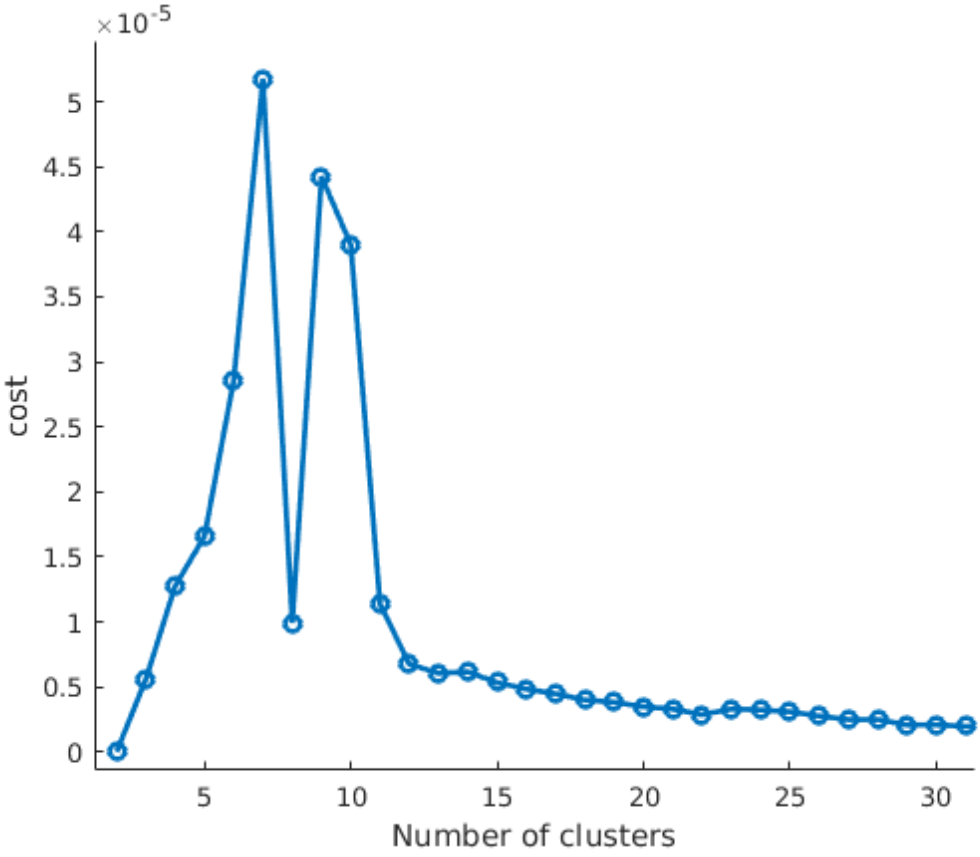


Figure 6.2: PSI norm-AC

doi:10.6342/NTU201704140

that some states can comute these two big clusters. The rate to reaction center is slower and the critical clusters would be the cluster {3, 65}, {4} and {5}.

doi:10.6342/NTU201704140

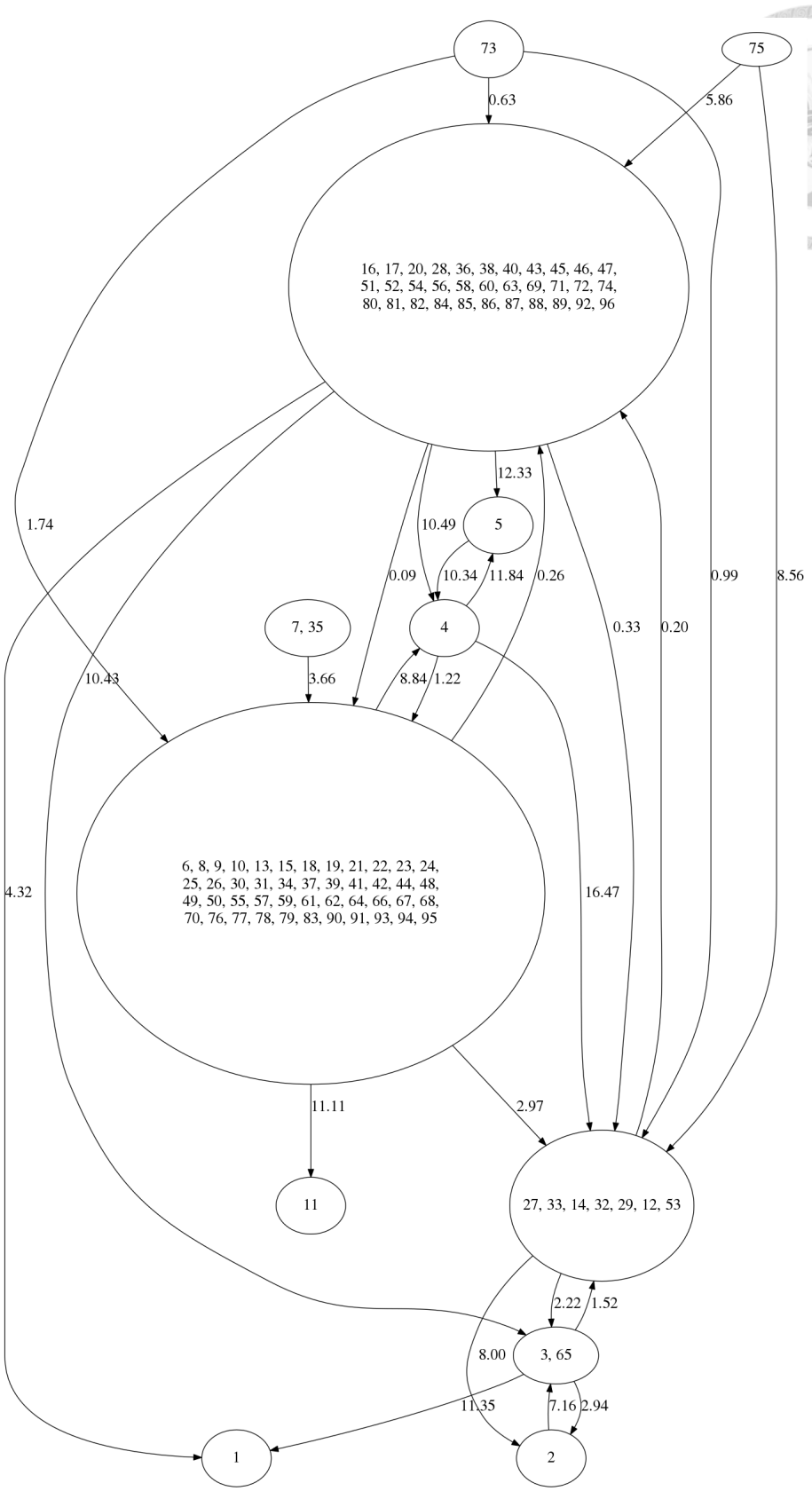


Figure 6.3: PSI 12 clusters CGM (unit: ps): We can devide the PSI into two big cluster and some isolated node.

doi:10.6342/NTU201704140

doi:10.6342/NTU201704140

Chapter 7

Conclusions

In this work, we demonstrate that CGM of EET networks can reporduce population dynamics correctly. We test FMO, LHCII monomer and PSI EET networks and demonstrate we develop a systematic process for clustering photosynthetic networks.

We demonstrate the better options of clustering graph. First, rearanging the subtree of the MBT is not necessary, and this operation would destroy the information of bottlenecks in the given networks. Second, the capacity of minimum-cut is extensive. So the MBT should normalize the capacity. Third, the simple cut-off method and ratio cut-off method do not consider the total information of MBT. These two method usually stop at the top of MBT. So the elimination cut-off method is the best.

We can finally build the CGM of networks and obtain the pathways by the process we develop.

doi:10.6342/NTU201704140

doi:10.6342/NTU201704140

Bibliography

[1] T. R. Calhoun, N. S. Ginsberg, G. S. Schlau-Cohen, Y.-C. Cheng, M. Ballottari, R. Bassi, and G. R. Fleming. Quantum Coherence Enabled Determination of the Energy Landscape in Light-Harvesting Complex II. *J. Phys. Chem. B*, 113(51):16291–16295, dec 2009.

[2] A. Damjanović, H. M. Vaswani, P. Fromme, and G. R. Fleming. Chlorophyll excitations in photosystem I of Synechococcus elongatus. *J. Phys. Chem. B*, 106(39):10251–10262, 2002.

[3] L. R. Ford and D. R. Fulkerson. Maximal flow through a network. *J. Can. mathématiques*, 8:399–404, 1956.

[4] R. Görke, T. Hartmann, and D. Wagner. Dynamic Graph Clustering Using Minimum-Cut Trees. *J. Graph Algorithms Appl.*, 16(2):411–446, 2012.

[5] Y. H. Hwang-Fu, W. Chen, and Y. C. Cheng. A coherent modified Redfield theory for excitation energy transfer in molecular aggregates. *Chem. Phys.*, 447:46–53, 2015.

[6] C. Kreisbeck, T. Kramer, and A. Aspuru-Guzik. Scalable high-performance algorithm for the simulation of exciton dynamics. Application to the light-harvesting complex II in the presence of resonant vibrational modes. *J. Chem. Theory Comput.*, 10(9):4045–4054, 2014.

doi:10.6342/NTU201704140

[7] Z. Liu, H. Yan, K. Wang, T. Kuang, J. Zhang, L. Gui, X. An, and W. Chang. Crystal structure of spinach major light-harvesting complex at 2.72 A resolution. *Nature*, 428(6980):287–292, 2004.

[8] V. Novoderezhkin, A. Marin, and R. van Grondelle. Intra- and inter-monomeric transfers in the light harvesting LHCII complex: the Redfield—Förster picture. *Phys. Chem. Chem. Phys.*, 13(38):17093, 2011.

[9] V. I. Novoderezhkin, M. A. Palacios, H. Van Amerongen, and R. Van Grondelle. Excitation dynamics in the LHCII complex of higher plants: Modeling based on the 2.72 a crystal structure. *J. Phys. Chem. B*, 109(20):10493–10504, 2005.

[10] T. Renger, M. E. Madjet, A. Knorr, and F. Müh. How the molecular structure determines the flow of excitation energy in plant light-harvesting complex II. *J. Plant Physiol.*, 168(12):1497–1509, 2011.

[11] G. S. Schlau-Cohen, T. R. Calhoun, N. S. Ginsberg, E. L. Read, M. Ballottari, R. Bassi, R. Van Grondelle, and G. R. Fleming. Pathways of energy flow in LHCII from two-dimensional electronic spectroscopy. *J. Phys. Chem. B*, 113(46):15352–15363, 2009.

[12] M. Schmidt Am Busch, F. Müh, M. El-Amine Madjet, and T. Renger. The eighth bacteriochlorophyll completes the excitation energy funnel in the FMO protein. *J. Phys. Chem. Lett.*, 2(2):93–98, 2011.

[13] D. E. Tronrud, J. Wen, L. Gay, and R. E. Blankenship. The structural basis for the difference in absorbance spectra for the FMO antenna protein from various green sulfur bacteria. *Photosynth. Res.*, 100(2):79–87, may 2009.

[14] L. Valkunas, G. Trinkunas, J. Chmeliov, and A. V. Ruban. Modeling of exciton quenching in photosystem II. *Phys. Chem. Chem. Phys.*, 11(35):7576–7584, 2009.

[15] R. van Grondelle and V. I. Novoderezhkin. Energy transfer in photosynthesis: experimental insights and quantitative models. *Phys Chem Chem Phys*, 8(7):793–807, 2006.

doi:10.6342/NTU201704140

[16] J. Wu, Z. Tang, Z. Gong, J. Cao, and S. Mukamel. Minimal Model of Quantum Kinetic Clusters for the Energy-Transfer Network of a Light-Harvesting Protein Complex. *J. Phys. Chem. Lett.*, pages 1240–1245, 2015.

doi:10.6342/NTU201704140