Circulation theory of enzyme kinetics

Yuhao Jiang

Date: August 3, 2021

1 Introduction

2 Monocyclic markov chains and cycle fluxes

We consider the following m-step($n \geq 2$) enzyme kinetics model:

$$\boldsymbol{E+S \rightleftharpoons ES \rightleftharpoons EP_1 \rightleftharpoons EP_2 \rightleftharpoons \cdots EP_{m-2} \rightleftharpoons E+P} \tag{1}$$

where $\boldsymbol{E}$ is an enzyme turning the substrate $\boldsymbol{S}$ into the product $\boldsymbol{P}$. From the perspective of a single enzyme molecule, this enzyme kinetics can be modeled as n-step Markov chain $(\xi_l)_{l\geq 0}$, with finite state space S defined on some space $(\Omega, \mathcal{F}, P)$. When $n = 2$, this Markov chain only have two state $\boldsymbol{E}$ and $\boldsymbol{ES}$, we say that the state space $S = \{1, 2\}$.

Definition 2.1. Let $\mathbb{Z}$ be the set of integers, and a periodic function f which maps $\mathbb{Z}$ to S is called circuit function. If s is the smallest positive integer which satisfied $f(n+s) = f(n)$ for $\forall n \in \mathbb{Z}$, then we called it the period of f.

Definition 2.2. Two circuit functions f and g in S are called equivalent if there exists some $m \in \mathbb{Z}$ such that $g(n) = f(n+m)$ for $\forall n \in \mathbb{Z}$.

Definition 2.3. For a circuit function f in S with period s that satisfies $f(1) = i_1$, $f(2) = i_2$, $f(3) = i_3$. The equivalence class that f belongs is a cycle $c = (i_1, i_2, \ldots i_s)$.

Therefore, according to this definitions, $c_1 = (1, 2, 3), c_2 = (3, 1, 2)$ and $c_3 = (2, 3, 1)$ represent the same cycle.

For presentation purposes, if the order sequence $i_1, i_2, \ldots, i_s$ occurs in the cycle c continuously, we denote that $[i_1, i_2, \ldots, i_s] \in c$. Specially, if $[i_1] \in c$, the point i_1 occurrs in c, and $[i_1, i_2] \in c$ denotes the edge $i_1 i_2$ exists in the cycle c. For the cycle $c_1 = (1, 2)$, we use k_{12} to denote the number of cycle c_1.

Definition 2.4. Let $\mathcal{C}_n(\omega)$ be the class of cycles occurring along the sample path $(\xi_l)_{l\geq 0}$ until time n. Then we use $\mathcal{C}_\infty$ to represent the limit of $\mathcal{C}_n$ as $n \to \infty$. This convergence has been proofed in **?**.

Definition 2.5. Let $k_{c,n}$ represent the number of time that cycle c is formed by a Markov chain up to time n. Then the sample circulation J_n^c along cycle c by time t is defined as

$$J_n^c = \frac{1}{n} k_{c,n} \quad \forall c \in \mathcal{C}_\infty$$

and the circulation w^c along cycle c is a nonnegative real number defined as the following almost sure limit:

$$w_c = \lim_{n \to \infty} J_n^c \quad \forall c \in \mathcal{C}_\infty, \quad a.s.$$

which represents the number of times that cycle c ic formed per unit time. Let $J_n = (J_n^c)_{c \in C_\infty}$ and $w = (w_c)_{c \in C_\infty}$.

For the enzyme kinetics model, if the state space $S = \{1, 2, \dots, m\}$, then $\mathcal{C}_\infty$ has $2m + 2$ cycles, including n 1-state cycles, n two-state cycles and two n-state cycles

Let $|c|$ denotes the length of cycle c, and $E_m = \{\mu = (\mu_c)_{c \in \mathcal{C}_\infty} \in [0, 1]^r : \sum_{c \in \mathcal{C}_\infty} |c|\mu_c = 1\}$ for m-state Markov chains, then the circulation distribution $w = (w_c)_{c \in \mathcal{C}_\infty} \in E_m$.

Definition 2.6. For m-state Markov chains, we say that J_n^c satisfies a large deviation principle with rate n and good rate function $I : E_m \to [0, \infty]$ if:

$$\lim_{n \to \infty} \frac{1}{n} \log \mathbb{P}(J_n^c = \nu_c, c \in \mathcal{C}_\infty) = -I(\nu), \quad \forall \nu \tag{2}$$

where $\sum_{c \in \mathcal{C}_\infty} |c|\nu_c = 1$, and $\nu = (\nu_c)_{c \in \mathcal{C}_\infty} \in E_m$.

3 Large deviation of circulation for Monocyclic Markov chains

3.1 Large deviation of circulation for three state Markov chains

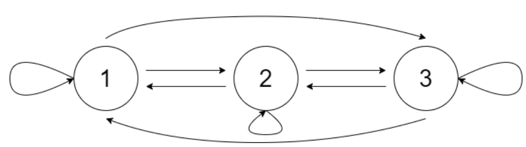

Figure 1: 3-state transition diagram

Theorem 3.1. *For the three-state Markov chains, let $S = \{1, 2, 3\}$, J_n satisfies a large deviation principle, and its rate function is*

$$I_3^c(\nu) = \sum_{i \in S} \left(\nu_i \log \frac{\nu_i}{w_i} + (\nu^i - \nu_i) \log \frac{\nu^i - \nu_i}{w^i - w_i} - \nu^i \log \frac{\nu^i}{w^i} \right) + \left(\sum_{c \in \mathcal{C}_\infty, c \neq (i)} \nu_c \log \frac{\nu_c}{w_c} \right) - \tilde{\nu} \log \frac{\tilde{\nu}}{\tilde{w}}$$

where $S = \{1, 2, 3\}$ is the state space for Markov chains

$$\mathcal{C}_\infty^3 = \{(1), (2), (3), (1, 2), (2, 3), (1, 3), (1, 2, 3), (1, 3, 2)\}$$

is the class of all cycles occurring. ν_c is the frequence of cycle c occurring, w_c is the cycle skipping rate on c.

Let $\nu^i = \sum_{c\supseteq[i]} \nu_c$, such as $\nu^1 = \nu_1 + \nu_{12} + \nu_{13} + \nu_{123} + \nu_{132}$. Let $\tilde{\nu} = \nu_{12} + \nu_{13} + \nu_{23} + \nu_{123} + \nu_{132} = \sum_{c\in\mathcal{C}_\infty, c\neq(i)} \nu_c$ represent the sum of all the elements of ν. And w^i, w_i is similarly defined for w.

Proof. For the first n-step path of Markov chains, after counting all the cycles, maybe it still remains some points which haven't form cycles, we call remains the derived chains. Refer to the Appendix, $\mathcal{E}(G^3(k))$ represents the amount of path with k cycle occuring, and $\Pi_{i,j\in S} p_{ij}^{\sum_{c\ni[i,j]} k_c}$ is the probability of the part formed all cycles. Obviously, the length of the derived chains is no more than two (the size of state space is three), then $\min_{\{i,j\}} p_{ij}^2$, $\max_{\{i,j\}} p_{ij}^2$ are the lower and upper bound of the probability of the derived chains occuring respectively. So

$$\mathbb{P}(J_n^c = \frac{k_c}{n}, c \in \mathcal{C}_\infty) \geq \mathcal{E}(G^3(k))\Pi_{i,j\in S} p_{ij}^{\sum_{c\ni[i,j]} k_c} \min_{\{i,j\}} p_{ij}^2$$

The length of the derived chains is no more than 2. And the steps in the derived chains is included in the n steps, so

$$\mathbb{P}(J_n^c = \frac{k_c}{n}, c \in \mathcal{C}_\infty) \leq 3\binom{n}{2}\mathcal{E}(G^3(k))\Pi_{i,j\in S} p_{ij}^{\sum_{c\ni[i,j]} k_c} \max_{\{i,j\}} p_{ij}^2$$

We know

$$\frac{1}{n}\log\binom{n}{2} \leq \frac{1}{n}\log\frac{n^2}{(2)^2} = O(\frac{\log n}{n})$$

so

$$\mathbb{P}(J_n^c = \frac{k_c}{n}, c \in \mathcal{C}_\infty) = \exp(O(\log n))\mathcal{E}(G(k))\Pi_{i,j\in S} p_{ij}^{\sum_{c\ni[i,j]} k_c}$$

We could neglect the influence of the derived chains.

Let

$$K_n = \left\{k = (k_c)_{c\in\mathcal{C}_\infty} : \sum_{c\in\mathcal{C}_\infty} k_c|c| \leq n\right\},$$

and this set includes all possible situations of each cycles occuring amount. It can easily observe that the size of this set $|K_n| \leq n^3$, $\frac{1}{n}K_n \in E$.

For $\forall k \in K_n$, let $\mu_n(k) = \frac{1}{n}k \in E$. Let us put

$$Q_n(a) = \max_{k\in K_n:\mu_n(k)\in B_a(\nu)} \mathcal{E}(G(k))\Pi_{i,j\in S} p_{ij}^{\sum_{c\ni[i,j]} k_c}$$

where $B_a(\nu)$ is the open neighborhood of ν with the total vaiation distance $d(\alpha,\beta) = \frac{1}{2}\sum_{s=1}^r |\alpha_s - \beta_s|$ and radius a. For enough large n, clearly

$$Q_n(a) \leq \mathbb{P}(J_n \in B_a(\nu), c \in \mathcal{C}_\infty) \leq |K_n|Q_n(a).$$

Stirling's formula gives $\frac{1}{n}\log\binom{k}{k'} = h(\frac{k}{n}) - h(\frac{k'}{n}) - h(\frac{k-k'}{n}) + O(\frac{\log n}{n})$ where $h(x) = x\log(x)$. We find that

$$
\begin{aligned}
&\frac{1}{n}\log\mathcal{E}(G^3(k))\Pi_{i,j}p_{ij}^{\sum_{c\ni[i,j]}k_c}\\
&=h(\nu_{12}+\nu_{13}+\nu_{23}+\nu_{123}+\nu_{132})+h(\nu_1+\nu_{12}+\nu_{13}+\nu_{123}+\nu_{132})\\
&+h(\nu_2+\nu_{12}+\nu_{123}+\nu_{132}+\nu_{23})+h(\nu_3+\nu_{13}+\nu_{123}+\nu_{132}+\nu_{23})\\
&-[h(\nu_1)+h(\nu_2)+h(\nu_3)+h(\nu_{12})+h(\nu_{13})+h(\nu_{23})+h(\nu_{123})+h(\nu_{132})]\\
&-\Bigg(h(\nu_{12}+\nu_{13}+\nu_{123}+\nu_{132})+h(\nu_{12}+\nu_{123}+\nu_{132}+\nu_{23})\\
&+h(\nu_{13}+\nu_{123}+\nu_{132}+\nu_{23})\Bigg)+\sum_{i,j,p_{ij}\neq 0}(\sum_{c\ni[i,j]}\nu_c)\log p_{ij}+O(\frac{\log n}{n})
\end{aligned}
$$

Merging all the same items, then

$$
\begin{aligned}
&\frac{1}{n}\log\mathcal{E}(G^3(k))\Pi_{i,j}p_{ij}^{\sum_{c\ni[i,j]}k_c}\\
&=-\Bigg(\nu_1\log(\frac{1}{p_1}\frac{\nu_1}{\nu^1})+\nu_2\log(\frac{1}{p_2}\frac{\nu_2}{\nu^2})+\nu_3\log(\frac{1}{p_3}\frac{\nu_3}{\nu^3})\\
&+\nu_{12}\log(\frac{1}{p_{12}p_{21}}\frac{\nu_{12}}{\nu_{12}+\nu_{23}+\nu_{13}+\nu_{123}+\nu_{132}}\frac{\nu^1-\nu_1}{\nu^1}\frac{\nu^2-\nu_2}{\nu^2})\\
&+\nu_{13}\log(\frac{1}{p_{13}p_{31}}\frac{\nu_{13}}{\nu_{12}+\nu_{23}+\nu_{13}+\nu_{123}+\nu_{132}}\frac{\nu^1-\nu_1}{\nu^1}\frac{\nu^3-\nu_3}{\nu^3})\\
&+\nu_{23}\log(\frac{1}{p_{23}p_{32}}\frac{\nu_{23}}{\nu_{12}+\nu_{23}+\nu_{13}+\nu_{123}+\nu_{132}}\frac{\nu^3-\nu_3}{\nu^3}\frac{\nu^2-\nu_2}{\nu^2})\\
&+\nu_{123}\log(\frac{1}{p_{12}p_{23}p_{31}}\frac{\nu_{123}}{\nu_{12}+\nu_{23}+\nu_{13}+\nu_{123}+\nu_{132}}\frac{\nu^1-\nu_1}{\nu^1}\frac{\nu^2-\nu_2}{\nu^2}\frac{\nu^3-\nu_3}{\nu^3})\\
&+\nu_{132}\log(\frac{1}{p_{13}p_{32}p_{21}}\frac{\nu_{132}}{\nu_{12}+\nu_{23}+\nu_{13}+\nu_{123}+\nu_{132}}\frac{\nu^1-\nu_1}{\nu^1}\frac{\nu^2-\nu_2}{\nu^2}\frac{\nu^3-\nu_3}{\nu^3})\Bigg)
\end{aligned}
$$

Refer to **?**, the calculation formula for circulation is:

$$
w_c=p_{i_1i_2}p_{i_2i_3}\cdots p_{i_{m-1}i_1}\frac{D(\{i_1,i_2,\cdots i_s\}^c)}{\sum_{j\in S}D(\{j\}^c)}
$$

With this formula, we know

$$
\begin{aligned}
w_{12}+w_{13}+w_{23}+w_{123}+w_{132}&=\frac{(1-p_{11})(1-p_{22})(1-p_{33})}{\sum_{i\in S}D(\{i\}^c)}\\
&=\frac{\Pi_{i,j}D(\{i,j\}^c)}{\sum_{i\in S}D(\{i\}^c)}
\end{aligned}
$$

Since the property of circulation, $p_{ij} = \frac{\sum_{c\ni[i,j]} w_c}{\sum_{c\ni(i)} w_c}$, So

$$\frac{1}{n}\log\mathcal{E}(G^3(k))\Pi_{i,j}p_{ij}^{\sum_{c\ni[i,j]}k_c}$$
$$= -(\sum_{c\in\mathcal{C}_\infty}\nu_c\log\frac{\nu_c}{w_c} + \sum_{i\in S}(\nu^i-\nu_i)\log\frac{\nu^i-\nu_i}{w^i-w_i} - \tilde{\nu}\log\frac{\tilde{\nu}}{\tilde{w}} - \sum_{i\in S}\nu^i\log(\frac{\nu^i}{w^i}))$$
$$= -\left[\sum_{i\in S}\left(\nu_i\log\frac{\nu_i}{w_i} + (\nu^i-\nu_i)\log\frac{\nu^i-\nu_i}{w^i-w_i} - \nu^i\log\frac{\nu^i}{w^i}\right) + \left(\sum_{c\in\mathcal{C}_\infty, c\neq(i)}\nu_c\log\frac{\nu_c}{w_c}\right) - \tilde{\nu}\log\frac{\tilde{\nu}}{\tilde{w}}\right]$$

Now we find that

$$\frac{1}{n}\mathbb{P}(J_n\in B_a(\nu)) = O(\frac{\log n}{n}) + \frac{1}{n}\log Q_n(a)$$
$$= O(\frac{\log n}{n}) - \min_{k\in K_n:\mu_n(k)\in B_a(\nu)} I_3^c(\mu_n(k))$$

And we know: (i) $\bigcup_{n\in\mathbb{N}}\{\mu_n(k): k\in K_n\}\bigcap E_3$ is dense in E_3. (ii) $\mu\to I_3^c(\mu)$ is continuous on E_3. It is analogous to the proof of Sanov's Theorem, then

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{P}(J_n\in B_a(\nu)) = -\inf_{k\in K_n:\mu_n(k)\in B_a(\nu)} I_3^c(\nu)$$

If the size of neighborhood $B_a(\nu)$ is enough small,

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{P}(J_n=\nu) = -I_3^c(\nu)$$

□

Corollary. *For 3-state Markov chains, $\bar{J}_n = J_n^{123} - J_n^{132}$ satisfies the net circulation large deviation.*

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{P}(\bar{J}_n=\bar{\nu}) = -\bar{I}_3^c(\bar{\nu})$$

and

$$\bar{I}_3^c(\bar{\nu}) = \min_{\nu\in\bar{E}(\bar{\nu})} I_3^c(\nu)$$

where net cycle frequence $\bar{\nu} = \nu_{123} - \nu132$, in that $(1,3,2)$ is the anti cycle of $(1,3,2)$, and $\bar{\nu}$ belongs the space $\bar{E}_3(\bar{\nu}) = \{\nu = (\nu_c)_{c\in\mathcal{C}_\infty}\in E_3|\nu_1+\nu_2+\nu_3+2(\nu_{12}+\nu_{13}+\nu_{23})+3(2\nu_{132}+\bar{\nu})=1\}$

Proof. Let $f(\nu) = \nu^{123} - \nu^{132}$, so $f: E_3\to[-\frac{1}{3},\frac{1}{3}]$ is a continuous function. Let

$$\bar{I}_3^c(\bar{\mu}) = \inf\{I_3^c(\nu): \nu\in E_3, \mu = f(\nu)\}$$

According to contraction principle, $\bar{I}_3^c(\bar{\mu})$ is also a good rate function on $f(E_3)$. □

Corollary. *For the two-state Markov chains, we have:*

$$I_2^c(\nu) = \nu_1\log(\frac{\nu_1}{\nu_1+\nu_{12}}/\frac{w_1}{w_1+w_{12}}) + \nu_2\log(\frac{\nu_2}{\nu_2+\nu_{12}}/\frac{w_2}{w_1+w_{12}})$$
$$+\nu_{12}\log(\frac{\nu_{12}}{\nu_1+\nu_{12}}/\frac{w_{12}}{w_1+w_{12}}) + \nu_{12}\log(\frac{\nu_{12}}{\nu_2+\nu_{12}}/\frac{w_{12}}{w_2+w_{12}})$$

Corollary. *I_3^c satisfies the fluctuation theorem, for $\nu,\mu\in E$, if $\nu_{123}=\mu_{132}$, $\nu_{132}=\mu_{123}$ and $\nu_c=\mu_c$ for*

$c \notin \{(1,2,3),(1,3,2)\}$ *the following formula holds:*

$$I_3^c(\nu) = I_3^c(\mu) - (\nu_{123} - \nu_{132}) \log \frac{\gamma^{123}}{\gamma^{132}} \tag{3}$$

and net circulation $\bar{I}_3^c$ *also satisfies the fluctuation theorem*

$$\bar{I}_3^c(\bar{\nu}) = \bar{I}_3^c(-\bar{\nu}) - \rho^{123}(\bar{\nu})$$

where $\gamma^c = p_{i_1 i_2} p_{i_2 i_3} \cdots p_{i_s i_1}$ *is the strength of cycle* $c = (i_1, i_2, \dots i_s)$*, and* $\rho^c = \frac{\gamma^c}{\gamma^{c^-}}$*,* c^- *is the inverse cycle of c.*

Proof. With the expression of I_3^c, we know

$$I_3^c(\nu) = I_3^c(\mu) - (\nu_{123} - \nu_{132}) \log \frac{w_{123}}{w_{132}}$$

and

$$w_{123} = \frac{p_{12} p_{23} p_{31}}{\sum_{j \in S} D(\{j\}^c)}$$

$$w_{132} = \frac{p_{13} p_{32} p_{21}}{\sum_{j \in S} D(\{j\}^c)}$$

3 can be obtained.

Then, we have

$$\begin{aligned} \bar{I}_3^c(\bar{\nu}) &= \min_{\nu_{123} - \nu_{132} = \bar{\nu}} I_3^c(\nu) \\ &= \min_{\nu_{123} - \nu_{132} = \bar{\nu}} I_3^c(\mu) - (\nu_{123} - \nu_{132}) \log \frac{\gamma^{123}}{\gamma^{132}} \\ &= \bar{I}_3^c(-\bar{\nu}) - \rho^{123} \log(\bar{\nu}) \end{aligned}$$

which gives the desired result. □

Corollary. I_3^c *is finite, continuous, positive and strictly convex on* E_3*, except along line segments* $\{\alpha\nu + (1-\alpha)\mu : \alpha \in [0,1]\}$*, between any* ν *and* μ *satisfying* $\nu_i/\nu^i = \mu_i/\mu^i, \forall i \in S$ *and* $\nu_c/(\tilde{\nu} - \sum_{i \in S} \nu_i) = \mu_c/(\tilde{\mu} - \sum_{i \in S} \nu_i), \forall c \in \mathcal{C}_\infty$*.* I_3^c *is also affine, i.e.* $I_3^c(\alpha\nu) = \alpha I_3^c(\nu)$*.*

Proof. Since $h(x) = x \log x$ is finite on the interval $[0,1]$, obviously, $I_3^c(\nu)$ is finite on E_3. Spliting the item $\sum_{c \in \mathcal{C}_\infty} \nu_c \log \frac{\nu_c}{w_c}$, because

$$\nu_i + (\nu^i - \nu_i) = \nu^i$$

$$\sum_{c \in \mathcal{C}_\infty, c \neq (i)} \nu_c = \tilde{\nu} - \sum_{i \in S} \nu_i$$

we employ Log sum inequality.

$$\begin{aligned}
I_3^c(\nu) &= \sum_{i\in S}\left(\nu_i \log\frac{\nu_i}{w_i} + (\nu^i-\nu_i)\log\frac{\nu^i-\nu_i}{w^i-w_i} - \nu^i\log\frac{\nu^i}{w^i}\right)\\
&+\left(\sum_{c\in\mathcal{C}_\infty, c\neq(i)} \nu_c\log\frac{\nu_c}{w_c}\right) - \tilde{\nu}\log\frac{\tilde{\nu}}{\tilde{w}}\\
&= \sum_{i\in S}\left(\nu_i \log\frac{\nu_i}{w_i} + (\nu^i-\nu_i)\log\frac{\nu^i-\nu_i}{w^i-w_i}\right) - \sum_{i\in S}\nu^i\log\frac{\nu^i}{w^i}\\
&+\left(\sum_{c\in\mathcal{C}_\infty, c\neq(i)} \nu_c\log\frac{\nu_c}{w_c}\right) - \tilde{\nu}\log\frac{\tilde{\nu}}{\tilde{w}}\\
&\geq \left(\sum_{i\in S}\nu^i\log\frac{\nu^i}{w^i} - \sum_{i\in S}\nu^i\log\frac{\nu^i}{w^i}\right)\\
&+\left(\tilde{\nu}\log\frac{\tilde{\nu}}{\tilde{w}} - \tilde{\nu}\log\frac{\tilde{\nu}}{\tilde{w}}\right)\\
&\geq 0
\end{aligned}$$

The convex can be proofed with the same way by Log sum inequality.

$$\begin{aligned}
I_3^c(\nu) &= \sum_{i\in S}\left(\nu_i \log\frac{\nu_i}{\nu^i} + (\nu^i-\nu_i)\log\frac{\nu^i-\nu_i}{\nu^i} - \nu^i\log\frac{w_i}{w^i} - \nu^i\log\frac{w^i-w_i}{w^i}\right)\\
&+\left(\sum_{c\in\mathcal{C}_\infty, c\neq(i)} \nu_c\log\frac{\nu_c}{\tilde{\nu}} - \tilde{\nu}\log(\frac{w_c}{\tilde{w}})\right)
\end{aligned}$$

Obviously, $I_3^c(\alpha\nu) = \alpha I_3^c(\nu)$, so I_3^c is also affine. The item involving w is linear for ν. For the remains, use log-sum inequality again, such as:

$$\lambda\nu_i\log\frac{\lambda\nu_i}{\lambda\nu^i} + (1-\lambda)\mu_i\log\frac{(1-\lambda)\mu_i}{(1-\lambda)\mu^i} \geq [\lambda\nu_i + (1-\lambda)\mu_i]\log\frac{\lambda\nu_i+(1-\lambda)\mu_i}{\lambda\nu^i+(1-\lambda)\mu^i}$$

□

3.2 Large deviation of circulation for finite state Markov chains

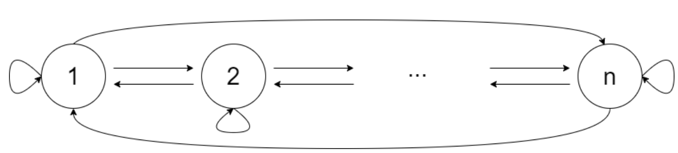

Figure 2: m-state transition diagram

Theorem 3.2. *For the m-state Markov chains, the good rate function is*

$$I_m^c(\nu) = [h(\nu_{12}+\nu_{1m}+\nu^+ +\nu^-) - h(\nu_{12}) - h(\nu_{1m}) - h(\nu^+) - h(\nu^-)] + \sum_{i\in S}[h(\nu^i) - h(\nu^i-\nu_i) - h(\nu_i)]$$

$$+ \max_{\nu_{ij}^+ + \nu_{ij}^- = \nu_{ij}, i\neq 1, j\neq m}\Big\{[h(\nu_{12}+\nu_{23}^+ +\nu^+) - h(\nu_{23}^+) - h(\nu_{12}+\nu^+)]$$

$$+ [h(\nu_{34}^+ + \nu_{23}^+ + \nu^+) - h(\mu_{34}^+) - h(\nu_{23}^+ + \nu^)] + \cdots +$$

$$+ [h(\nu_{m-1,m}^+ + \nu_{m-2,m-1}^+ + \nu^+) - h((\nu_{m-1,m}^+) - h(\nu_{m-2,m-1}^+ + \nu^+)]$$

$$+ [h(\nu_{1m} + \nu_{m-1,m}^- + \nu^-) - h(\nu_{m-1,m}^-) - h(\nu_{1m} + \nu^-)]$$

$$+ [h(\nu_{m-1,m}^- + \nu_{m-2,m-1}^- + \nu^-) - h(\nu_{m-2,m-1}^-) - h(\nu_{m-1,m}^- + \nu^-)]$$

$$+ \cdots + +h(\nu_{23}^- + \nu_{34}^- + \nu^-) - h(\nu_{23}^-) - h(\nu_{34}^- + \nu^-)\Big\} + \sum_{i,j}(\sum_{c\ni[i,j]}\nu_c)\log p_{ij}$$

Proof. According to contraction principle, the rate function exists, we know:

$$I_m^c(\nu) = \frac{1}{n}\log\mathcal{E}_1(G^m(k))\Pi_{i,j}p_{ij}^{\sum_{c\ni[i,j]}k_c}.$$

The accumulation part in $\mathcal{E}_1(G^m(k))$ has no more than n^{m-1} items, and $\frac{1}{n}\log n^{m-1} = O(\frac{\log n}{n})$. The other means for simplification has been mentioned in Theorem 1.1 many times. □

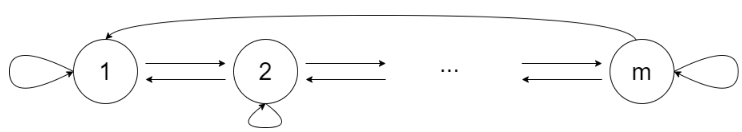

Figure 3: Simple m-state transition diagram

Theorem 3.3. *For the m-state Markov chains, and the non-zero terms only have* $\{p_{i,i+1}, i = 1,\dots m-1\}\bigcup\{p_{i,i-1}, i = 2,\dots m-1\}\bigcup\{p_{m,1}\}$, $(w_c)_{c\in\mathcal{C}_\infty}$ *satisfies a large deviation principle, and its rate function is*

$$I_{m'}^c(\nu) = \sum_{i\in S}\sum_{c\ni[i]}\nu_c\log\left(\frac{\nu_c}{\nu^i}/\frac{w_c}{w^i}\right)$$

Proof. The proof is similar in spirit to that of Theorem 1.1, the main difference lies in the simplification of rate function. Let

$$K_n = \left\{k = (k_c)_{c\in\mathcal{C}_\infty} : \sum_{c\in\mathcal{C}_\infty}k_c|c| \leq n\right\},$$

and

$$Q_n(a) = \max_{k\in K_n:\mu_n(k)\in B_a(\nu)}\mathcal{E}(G^{m'}(k))\Pi_{i,j\in S}p_{ij}^{\sum_{c\ni[i,j]}k_c}.$$

We only need to simply the following formula.

$$\frac{1}{n}\log\mathcal{E}(G^{m'}(k))\Pi_{i,j\in S}p_{ij}^{\sum_{c\ni[i,j]}k_c}$$

$$= -\Bigg(\sum_{i\in S}\nu_i\log(\frac{1}{p_i}\frac{\nu_i}{\nu^i}) + \nu_{12}\log\frac{1}{p_{12}p_{21}}\frac{\nu_{12}}{\nu_{12}+\nu_{23}+\nu_{12\ldots m}}\frac{\nu^1-\nu_1}{\nu^1}\frac{\nu^2-\nu_2}{\nu^2}$$

$$+\nu_{m-1,m}\log\frac{1}{p_{m-1,m}p_{n,m-1}}\frac{\nu_{m-1,m}}{\nu_{m-1,m}+\nu_{m-2,m-1}+\nu_{12\ldots m}}\frac{\nu^{m-1}-\nu_{m-1}}{\nu^{m-1}}\frac{\nu^m-\nu_m}{\nu^m}$$

$$+\sum_{i=2}^{m-2}\nu_{i,i+1}\log(\frac{1}{p_{i,i+1}p_{i+1,i}}\frac{\nu_{i,i+1}(\nu_{i,i+1}+\nu_{1,2,\ldots m})}{(\nu_{i,i+1}+\nu_{i+1,i+2}+\nu_{1,2,\ldots m})(\nu_{i-1,i}+\nu_{i,i+1}+\nu_{1,2,\ldots m})}\frac{\nu^i-\nu_i}{\nu^i}\frac{\nu^{i+1}-\nu_{i+1}}{\nu^{i+1}})$$

$$+\nu_{1,2,\ldots m}\log(\frac{1}{p_{12}p_{23}\cdots p_{m-1,s}p_{s,1}}\nu_{1,2,\ldots m}\frac{\Pi_{i=2}^{m-2}(\nu_{i,i+1}+\nu_{1,2,\ldots m})}{\Pi_{i=2}^{m-1}(\nu^i-\nu_i)}\Pi_{i=1}^{m}\frac{\nu^i-\nu_i}{\nu^i})\Bigg)$$

$$= -\Bigg(\sum_{i\in S}\nu_i\log(\frac{1}{p_i}\frac{\nu_i}{\nu^i}) + \nu_{1,2}\log(\frac{1}{p_{12}p_{21}}\frac{\nu_{12}(\nu^1-\nu_1)}{\nu^1\nu^2}) + \nu_{m-1,m}\log\frac{1}{p_{m-1,m}p_{n,m-1}}\frac{\nu_{m-1,m}(\nu^{m-1}-\nu_{m-1})}{\nu^{m-1}\nu^m}$$

$$+\sum_{i=2}^{m-2}\nu_{i,i+1}\log(\frac{1}{p_{i,i+1}p_{i+1,i}}\frac{\nu_{i,i+1}(\nu_{i,i+1}+\nu_{1,2,\ldots m})}{\nu^i\nu^{i+1}})$$

$$+\nu_{1,2,\ldots m}\log(\frac{1}{p_{12}p_{23}\cdots p_{m-1,m}p_{m,1}}\nu_{1,2,\ldots m}\frac{\Pi_{i=1}^{m-1}(\nu_{i,i+1}+\nu_{1,2,\ldots m})}{\Pi_{i=1}^{m}\nu^i}\Bigg)$$

Because $p_{ij} = \frac{\sum_{c\ni[i,j]}w_c}{\sum_{c\ni(i)}w_c}$, and w substitute for p_{ij} in th formula above, the rate function $I_n^{c'}(\nu)$ would be obtained. The rest proof is same as Theorem 1.1. $\square$

Corollary. *$I^c_{m'}(\nu)$ is finite, continuous, positive and strictly convex on $E_{m'}$, except along line segments $\{\alpha\nu+(1-\alpha)\mu : \alpha\in[0,1]\}$, between any ν and μ satisfying $\nu_c/\nu^i = \mu_c/\mu^i, \forall c\ni i$. $I^c_{m'}(\nu)$ is also affine, i.e. $I^c_{m'}(\alpha\nu) = \alpha I^c_{m'}(\nu)$. convex. The proof can be obtained by completely imitating it of $I^c_{m'}(\nu)$.*

4 conclusion

5 Appendix

Consider the first n-step of the above Markov chains (ξ_l), we assume that $\xi_n = \xi_1$. Let $k = (k_c)_{c\in\mathcal{C}_\infty}$ denote the number of each cycle in class $\mathcal{C}_\infty$ occurring. If marking each occurence of state pair (s,t) in $\xi_1,\ldots,\xi_n$ by drawing an arrow from s to t, we can obtain an oriented graph. Hence if we know the number of all state pair (s,t), the oriented graph is unique. Because we have known the amount of cycle $k = (k_c)_{c\in\mathcal{C}_\infty}$, the amount of each state pair can be denoted easily by k, let $G^m(k)$ denote the corresponding oriented graph. And we employ $\mathcal{E}(G^m(k))$ to denote the number of Euler circuits on $G^m(k)$ and which the cycle amount is k. By the way, $\mathcal{E}_i(G^m(k))$ denotes the number of Euler circuits which start from states i, i.e. $\xi_1 = i$.

Theorem 5.1. *For the graph $G^m(k)$ induced by the Markov chains $(\xi_l)_{l\geq 0}$, if the size of state space is s,* *then we have the following formula:*

$$\mathcal{E}_1(G^m(k)) = \binom{k_{12}+k_{1m}+k^++k^-}{k_{12},k_{1m},k^+,k^-}\binom{k_1+k_{12}+k_{1m}+k^++k^-}{k_{12}+k_{1m}+k^++k^-}\left[\Pi_{i\in S,i\neq 1}\binom{\sum_{c\ni[i]}k_c-1}{\sum_{c\ni[i]}k_c-k_i-1}\right]$$
$$\Bigg[\sum_{k_{23}^++k_{23}^-=k_{23}}\sum_{k_{34}^++k_{34}^-=k_{34}}\cdots\sum_{k_{m-1,m}^++k_{m-1,m}^-=k_{m-1,m}}\binom{k_{23}^++k_{12}+k^+-1}{k_{23}^+}\binom{k_{34}^++k_{23}^++k^+-1}{k_{34}^+}$$
$$\cdots\binom{k_{m-1,m}^++k_{m-2,m-1}^++k^+-1}{k_{m-1,m}^+}\binom{k_{1,m}+k_{m-1,m}^-+k^--1}{k_{m-1,m}^-}\binom{k_{m-1,m}^-+k_{m-2,m-1}^-+k^--1}{k_{m-2,m-1}^-}$$
$$\cdots\binom{k_{23}^-+k_{34}^-+k^--1}{k_{23}^-}\Bigg]$$

where k^+ *and* k^- *are the number of cycle* $\{1,2,\dots,s\}$, $\{s,m-1,\dots,1\}$ *respectively. And* $k_{s,t}^+$ *represents the number of cycles of which the state* s *occurring fistly among cycles* $\{s,t\}$*, so* $k_{s,t}^-$ *is number of remains.*

Proof. We use following three steps to count all the Euler circuits. 1. This Markov chains starts from state 1, hence we pick the cycles which includes state 1, i.e. $\{1\}, \{1,2\}, \{1,m\}, \{1,2,\dots,m\}, \{1,\dots,m-1,m,2\}$. And line them up, the number of permutations is:

$$\binom{k_1+k_{12}+k_{1m}+k^++k^-}{k_1,k_{12},k_{1m},k^+,k^-}$$

2. Next we need to insert the other 2-state cycles into it based on above permutation. For the cycle $(s-1,s)$, $s\in\{3,4,\dots,m\}$, it firstly be inserted into the cycle $(1,2,\dots,m)$ and $(s-2,s-1)$ which start from state $s-2$ if $s\neq 3$. The cycle $(s-1,s)$ that has been inserted all states from state $s-1$, next we insert the remains that starts from state s, the alternative cycles have $(1,m-1,m-2,\dots,2)$ and $(s,s+1)$ which start from state $s+1$ if $s\neq m$. 4 illustrates the way to insert 2-state cycles.
If we change the order of inserting, the cycle would not occurr at the inserting point, such as 5

$$\sum_{k_{23}^++k_{23}^-=k_{23}}\sum_{k_{34}^++k_{34}^-=k_{34}}\cdots\sum_{k_{m-1,m}^++k_{m-1,m}^-=k_{m-1,m}}\binom{k_{23}^++k_{12}+k^+-1}{k_{23}^+}\binom{k_{34}^++k_{23}^++k^+-1}{k_{34}^+}$$
$$\cdots\binom{k_{m-1,m}^++k_{m-2,m-1}^++k^+-1}{k_{m-1,m}^+}\binom{k_{1,m}+k_{m-1,m}^-+k^--1}{k_{m-1,m}^-}\binom{k_{m-1,m}^-+k_{m-2,m-1}^-+k^--1}{k_{m-2,m-1}^-}$$
$$\cdots\binom{k_{23}^-+k_{34}^-+k^--1}{k_{23}^-}$$

3. Consider the one of the above permutation, we insert the other 1-state cycles into it, i.e. (2), (3), ..., (s). The number of permutations is:

$$\Pi_{i\in S,i\neq 1}\binom{\sum_{c\ni[i]}k_c-1}{\sum_{c\ni[i]}k_c-k_i-1}$$

□

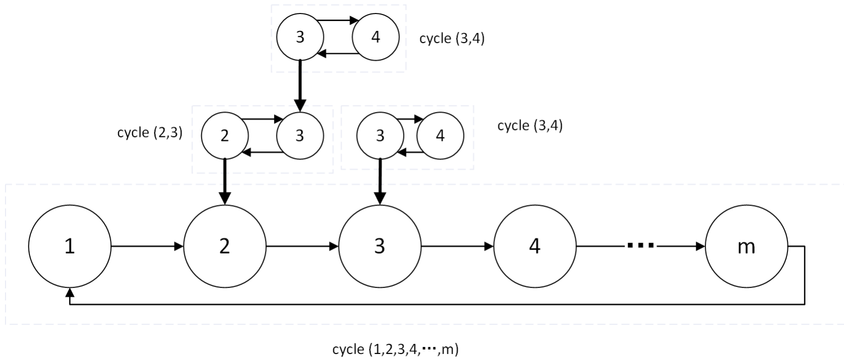

Figure 4: Correct Insertion

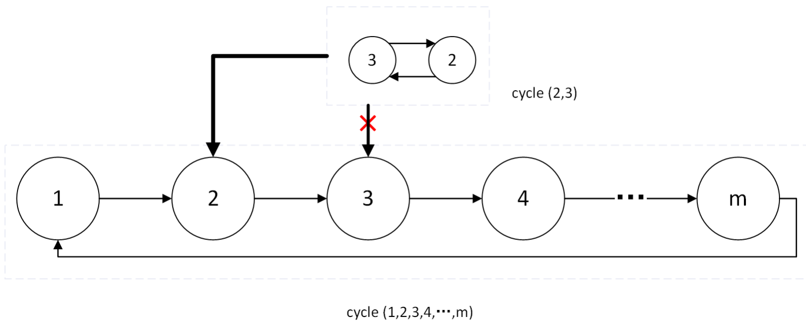

Figure 5: Wrong Insertion

Corollary. *If $p_{1m} = 0$, the number of Euler circuits is*

$$\mathcal{E}(G^{m'}(k)) = \exp(O \log n)\binom{k_{12} + k_{1,2\ldots,m}}{k_{12}}\Pi_{i=2}^{m-1}\binom{k_{1,2\ldots,m} + k_{i-1,i} + k_{i,i+1}}{k_{i,i+1}}\Pi_{i\in S}\binom{\sum_{c\ni[i]} k_c}{\sum_{c\ni[i]} k_c - k_i} \tag{4}$$

Proof. Because $p_{1m} = 0$, $k^- = 0$, we do not need to consider insert 2-state cycle into $(1, m)$ and $(1, m, m-1, \ldots, 2)$. Compairing to Theorem 2.1, only the second step is different, it is:

$$\Pi_{i=2}^{m-1}\binom{k_{1,2\ldots,m} + k_{i-1,i} + k_{i,i+1}}{k_{i,i+1}}$$

In addition, the path starting from other states which is not state 1 should be considered, so

$$\mathcal{E}(G^{m'}(k)) \leq n\binom{k_{12} + k_{1,2\ldots,m}}{k_{12}}\Pi_{i=2}^{m-1}\binom{k_{1,2\ldots,m} + k_{i-1,i} + k_{i,i+1}}{k_{i,i+1}}\Pi_{i\in S}\binom{\sum_{c\ni[i]} k_c}{\sum_{c\ni[i]} k_c - k_i}$$

We can get 4. □

Corollary. *Consider the graph $G(k)$ induced by the Markov chains $(\xi_l)_{l\geq 0}$. If $s = 2$, we have:*

$$\mathcal{E}(G^2(k)) = \exp(O \log n)\binom{k_1 + k_{12}}{k_1}\binom{k_2 + k_{12}}{k_2}$$

If $s = 3$*, we have:*

$$\begin{aligned}\mathcal{E}(G^3(k)) = &\exp(O(\log(n)))\binom{k_{12}+k_{13}+k_{123}+k_{132}}{k_{12},\ k_{13},\ k_{123},\ k_{132}}\\ &\binom{k_1+k_{12}+k_{13}+k_{123}+k_{132}}{k_1}\binom{k_{12}+k_{13}+k_{123}+k_{132}+k_{23}}{k_{23}}\\ &\binom{k_{12}+k_{123}+k_{132}+k_{23}+k_2}{k_2}\binom{k_{13}+k_{123}+k_{132}+k_{23}+k_3}{k_3}\end{aligned}$$