

# Circulation theory of enzyme kinetics

Yuhao Jiang

Date: July 10, 2021

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



where  $E$  is an enzyme turning the substrate  $S$  into the product  $P$ . 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  $E$  and  $ES$ , we say that the state space  $S = \{1, 2\}$ .

**Definition 0.1.** 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 0.2.** 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 0.3.** 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  $c = (i_1, i_2, \dots, i_s)$

According to above definitions,  $c_1 = (1, 2, 3), c_2 = (3, 1, 2)$  and  $c_3 = (2, 3, 1)$  represent the same cycle. So a cycle is also a equivalence class on the space of all circuit functions under the equivalence relation.

For presentation purposes, if the order sequence  $i_1, i_2, \dots, i_s$  occurs in the cycle  $c$  continuously, we denote that  $[i_1, i_2, \dots, i_s] \in c$ . Specially, if  $[i_1] \in c$ , the point  $i_1$  occurs 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 0.4.** 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  $n \rightarrow \infty$ . This convergence has been proofed in [].

**Definition 0.5.** Let  $k_{c,n}$  denote 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:

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

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

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  $E = \{\mu = (\mu_c)_{c \in \mathcal{C}_\infty} \in [0, 1]^r : \sum_{c \in \mathcal{C}_\infty} |c| \mu_c = 1\}$ , then the circulation distribution  $w = (w_c)_{c \in \mathcal{C}_\infty} \in E$ .

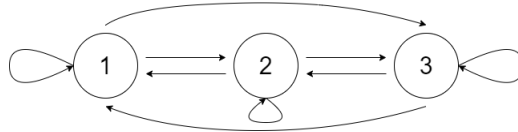
**Definition 0.6.** We say that  $J_n^c$  satisfies a large deviation principle with rate  $n$  and good rate function  $I : E \rightarrow [0, \infty]$  if:

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

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

## 1 Large deviation of circulation for finite Markov chains

### 1.1 Large deviation of circulation for three state Markov chains



**Figure 1:** 3-state transition diagram

**Theorem 1.1.** For the three-state Markov chains, let  $S = \{1, 2, 3\}$ ,  $J_n$  satisfies a large deviation principle, and its rate function is

$$I_3^c(\nu) = \sum_{c \in \mathcal{C}_\infty} \nu_c \log \frac{\nu_c}{w_c} + \sum_{i \in S} (\nu^i - \nu_i) \log \frac{\nu^i - \nu_i}{w^i - w_i} \quad (3)$$

$$- (\tilde{\nu} - \sum_{i \in S} \nu_i) \log \left( \frac{\tilde{\nu} - \sum_{i \in S} \nu_i}{\tilde{w} - \sum_{i \in S} w_i} \right) \quad (4)$$

$$- \sum_{i \in S} \nu^i \log \left( \frac{\nu^i}{w^i} \right) \quad (5)$$

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.  $\nu_c$  is the frequency of cycle  $c$  occurring,  $w_c$  is the cycle skipping rate on  $c$ .

Let  $\nu^i = \sum_{c \ni i} \nu_c$ , such as  $\nu^1 = \nu_1 + \nu_{12} + \nu_{13} + \nu_{123} + \nu_{132}$ . Let  $\tilde{\nu} = \nu_1 + \nu_2 + \nu_3 + \nu_{12} + \nu_{13} + \nu_{23} + \nu_{123} + \nu_{132}$  represent the sum of all the elements of  $\nu$ . And  $w^i$ ,  $w_i$  is similarly defined for  $w$ .

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

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

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

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

We know

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

so

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

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 occurring 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

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

where  $B_a(\nu)$  is the open neighborhood of  $\nu$  with the total variation 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  $\frac{1}{n} \log \binom{k}{k'} = h\left(\frac{k}{n}\right) - h\left(\frac{k'}{n}\right) - h\left(\frac{k-k'}{n}\right) + O\left(\frac{\log n}{n}\right)$  where  $h(x) = x \log(x)$ . We find that

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

Merging all the same items, then

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

Refer to [Jiang], the calculation formula for circulation is:

$$w_c = p_{i_1 i_2} p_{i_2 i_3} \cdots p_{i_{m-1} i_1} \frac{D(\{i_1, i_2, \dots, i_s\}^c)}{\sum_{j \in S} D(\{j\}^c)}$$

With this formula, we know

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

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

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

Now we find that

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

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

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

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

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

□

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

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

and

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

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

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

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

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

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

$$\begin{aligned} I_2^c(\nu) &= \nu_1 \log\left(\frac{\nu_1}{\nu_1 + \nu_{12}} / \frac{w_1}{w_1 + w_{12}}\right) + \nu_2 \log\left(\frac{\nu_2}{\nu_2 + \nu_{12}} / \frac{w_2}{w_1 + w_{12}}\right) \\ & \quad + \nu_{12} \log\left(\frac{\nu_{12}}{\nu_1 + \nu_{12}} / \frac{w_{12}}{w_1 + w_{12}}\right) + \nu_{12} \log\left(\frac{\nu_{12}}{\nu_2 + \nu_{12}} / \frac{w_{12}}{w_2 + w_{12}}\right) \end{aligned}$$

**Corollary.**  $I_3^c$  satisfies 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:

$$I_3^c(\nu) = I_3^c(\mu) - (\nu_{123} - \nu_{132}) \log \frac{\gamma^{123}}{\gamma^{132}} \quad (6)$$

$\gamma^c = p_{i_1 i_2} p_{i_2 i_3} \cdots p_{i_s i_1}$  is the strength of cycle  $c = (i_1, i_2, \dots, i_s)$ .

**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

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

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

6 can be obtained. □

**Corollary.**  $I_3^c$  is finite, continuous, positive and strictly convex.

**Proof.** Since  $h(x) = x \log x$  is finite on the interval  $[0, 1]$ , obviously,  $I_3^c(\nu)$  is finite on  $E$ . Splitting item  $\sum_{c \in \mathcal{C}_\infty} \nu_c \log \frac{\nu_c}{w_c}$ , and employ Log sum inequality.

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

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

## 1.2 Large deviation of circulation for finite state Markov chains

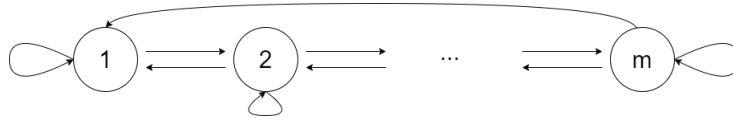
**Theorem 1.2.** *For the  $m$ -state Markov chains, if the rate function exists, it is*

$$\begin{aligned}
I_m^c(\nu) = & [h(\nu_{12} + \nu_{1m} + \nu^+ + \nu^-) - h(\nu_{12}) - h(\nu_{1m}) + h(\nu^+) + h(\nu^-)] + \sum_{i \in S} [h(\nu^i) - h(\nu^i - \nu_i)] \\
& + \max_{\substack{\nu_{ij}^+ + \nu_{ij}^- = \nu_{ij}, i \neq 1, j \neq m}} \left\{ [h(\nu_{12} + \nu_{23}^+ + \nu^+) - h(\nu_{23}^+) - h(\nu_{12} + \nu^+)] \right. \\
& + [h(\nu_{34}^+ + \nu_{23}^+ + \nu^+) - h(\nu_{34}^+) - h(\nu_{23}^+ + \nu^+)] + \cdots + \\
& + [h(\nu_{m-1,m}^+ + \nu_{m-2,m-1}^+ + \nu^+) - h(\nu_{m-1,m}^+) - h(\nu_{m-2,m-1}^+ + \nu^+)] \\
& + [h(\nu_{1m} + \nu_{m-1,m}^- + \nu^-) - h(\nu_{m-1,m}^-) - h(\nu_{1m} + \nu^-)] \\
& + [h(\nu_{m-1,m}^- + \nu_{m-2,m-1}^- + \nu^-) - h(\nu_{m-2,m-1}^-) - h(\nu_{m-1,m}^- + \nu^-)] \\
& \left. + \cdots + [h(\nu_{23}^- + \nu_{34}^- + \nu^-) - h(\nu_{23}^-) - h(\nu_{34}^- + \nu^-)] \right\} + \sum_{i,j} \left( \sum_{c \ni [i,j]} \nu_c \right) \log p_{ij}
\end{aligned}$$

**Proof.** If the rate function exists, we know:

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

The accumulation part in  $\mathcal{E}(G^m(k))$  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.  $\square$



**Figure 2:**  $n$ -state transition diagram

**Theorem 1.3.** *For the  $m$ -state Markov chains, and the non-zero terms only have  $\{p_{i,i+1}, i = 1, \dots, m-1\} \cup \{p_{i,i-1}, i = 2, \dots, m-1\} \cup \{p_{m,1}\}$ ,  $(w_c)_{c \in \mathcal{C}_\infty}$  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

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

We only need to simply the following formula.

$$\begin{aligned}
& \frac{1}{n} \log \mathcal{E}(G^{m'}(k)) \prod_{i,j \in S} p_{ij}^{\sum_{c \ni [i,j]} k_c} \\
&= - \left( \sum_{i \in S} \nu_i \log \left( \frac{1}{p_i} \frac{\nu_i}{\nu^i} \right) + \nu_{12} \log \frac{1}{p_{12} p_{21}} \frac{\nu_{12}}{\nu_{12} + \nu_{23} + \nu_{12 \dots m}} \frac{\nu^1 - \nu_1}{\nu^1} \frac{\nu^2 - \nu_2}{\nu^2} \right. \\
&+ \nu_{m-1,m} \log \frac{1}{p_{m-1,m} p_{n,m-1}} \frac{\nu_{m-1,m}}{\nu_{m-1,m} + \nu_{m-2,m-1} + \nu_{12 \dots m}} \frac{\nu^{m-1} - \nu_{m-1}}{\nu^{m-1}} \frac{\nu^m - \nu_m}{\nu^m} \\
&+ \sum_{i=2}^{m-2} \nu_{i,i+1} \log \left( \frac{1}{p_{i,i+1} p_{i+1,i}} \frac{\nu_{i,i+1} (\nu_{i,i+1} + \nu_{1,2 \dots m})}{(\nu_{i,i+1} + \nu_{i+1,i+2} + \nu_{1,2 \dots m}) (\nu_{i-1,i} + \nu_{i,i+1} + \nu_{1,2 \dots m})} \frac{\nu^i - \nu_i}{\nu^i} \frac{\nu^{i+1} - \nu_{i+1}}{\nu^{i+1}} \right) \\
&+ \nu_{1,2 \dots m} \log \left( \frac{1}{p_{12} p_{23} \dots p_{m-1,s} p_{s,1}} \nu_{1,2 \dots m} \frac{\prod_{i=2}^{m-2} (\nu_{i,i+1} + \nu_{1,2 \dots m})}{\prod_{i=2}^{m-1} (\nu^i - \nu_i)} \prod_{i=1}^m \frac{\nu^i - \nu_i}{\nu^i} \right) \\
&= - \left( \sum_{i \in S} \nu_i \log \left( \frac{1}{p_i} \frac{\nu_i}{\nu^i} \right) + \nu_{1,2} \log \left( \frac{1}{p_{12} p_{21}} \frac{\nu_{12} (\nu^1 - \nu_1)}{\nu^1 \nu^2} \right) + \nu_{m-1,m} \log \frac{1}{p_{m-1,m} p_{n,m-1}} \frac{\nu_{m-1,m} (\nu^{m-1} - \nu_{m-1})}{\nu^{m-1} \nu^m} \right. \\
&+ \sum_{i=2}^{m-2} \nu_{i,i+1} \log \left( \frac{1}{p_{i,i+1} p_{i+1,i}} \frac{\nu_{i,i+1} (\nu_{i,i+1} + \nu_{1,2 \dots m})}{\nu^i \nu^{i+1}} \right) \\
&+ \nu_{1,2 \dots m} \log \left( \frac{1}{p_{12} p_{23} \dots p_{m-1,m} p_{m,1}} \nu_{1,2 \dots m} \frac{\prod_{i=1}^{m-1} (\nu_{i,i+1} + \nu_{1,2 \dots m})}{\prod_{i=1}^m \nu^i} \right) \\
\end{aligned}$$

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

**Corollary.**  $I'_{n'}(\nu)$  is finite, continuous, positive and strictly convex

**Proof.** The finiteness and continuity is obvious, and we also can employ Log-sum inequality to proof its positivity convexity.  $\square$

## 2 Appendix

Consider the first  $n$ -step of the above Markov chains  $(\xi_l)$ , we assume that  $\xi_n = \xi_1$ , and the number of cycle  $c$  occurring is  $k_c$ . If marking each occurrence of state pair  $(s, t)$  in  $\xi_1, \dots, \xi_n$  by drawing an arrow from  $s$  to  $t$ , we can obtain an oriented graph  $G^m(k)$ . And we employ  $\mathcal{E}(G^m(k))$  to denote the number of Euler circuits on  $G^m(k)$ , 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 2.1.** 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:

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

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 firstly 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 optional cycles have  $(1, s, m-1, \dots, 2)$  and  $(s, s+1)$  which start from state  $s+1$  if  $s \neq m$ . ?? illustrates the way to insert 2-state cycles.

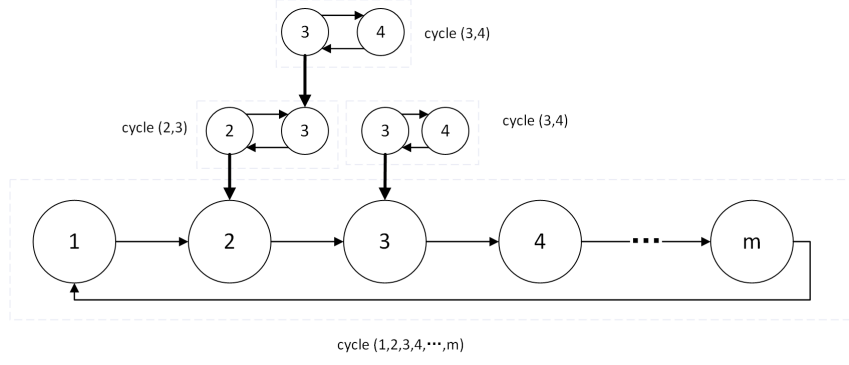
If we change the order of inserting, the cycle would not occur at the inserting point, such as ??

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

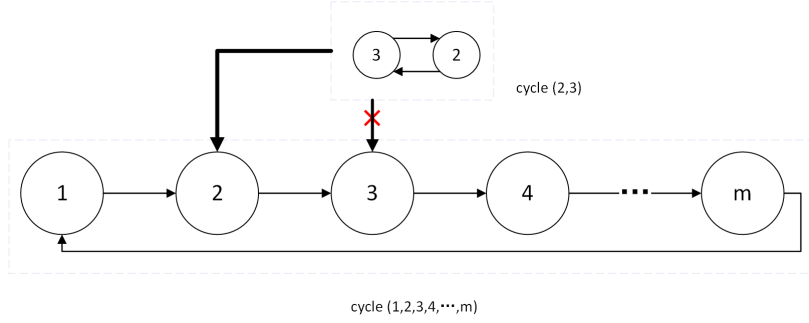
3. Consider the one of the above permutation, we insert the other 1-state cycles into it, i.e.  $(2), (3), \dots, (s)$ . The number of permutations is:

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

□



**Figure 3: Right Insertion**



**Figure 4: Wrong Insertion**

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

$$\mathcal{E}(G^{m'}(k)) = \exp(O \log n) \binom{k_{12} + k_{1,2,\dots,m}}{k_{12}} \prod_{i=2}^{m-1} \binom{k_{1,2,\dots,m} + k_{i-1,i} + k_{i,i+1}}{k_{i,i+1}} \prod_{i \in S} \binom{\sum_{i \ni c} k_c}{\sum_{i \ni c} k_c - k_i} \quad (7)$$

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

$$\prod_{i=2}^{m-1} \binom{k_{1,2,\dots,m} + k_{i-1,i} + k_{i,i+1}}{k_{i,i+1}}$$

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

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

We can get 7. □

**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:

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