Circulation theory of enzyme kinetics

Yuhao Jiang

Date: July 24, 2021

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 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 $c = (i_1, i_2, \dots i_s)$.

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 $i_1, i_2, \dots, i_s$ 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 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 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 $n \to \infty$. This convergence has been proofed in **?**.

Definition 0.5. 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 $S = \{1, 2, \ldots, m\}$, 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 0.6. 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:

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

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

1 Large deviation of circulation for finite Markov chains

1.1 Large deviation of circulation for finite state Markov chains

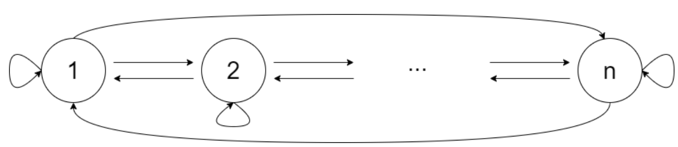

Figure 1: m-state transition diagram

Theorem 1.1. *For the m-state Markov chains, the good rate function is*

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

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

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

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. $\square$