Appendix A
For Chapter 2

A.1 Derivation of the Path Probability Density (2.4)

In this appendix, we define continuous time Markov processes by taking a continuous time limit in discrete time Markov processes. Then, we derive the path probability density (2.4) in this formulation. Finally, with the obtained path probability density, we show that the probability distribution $P(\boldsymbol{n}, t)$ satisfies the Master equation (2.2).

We consider a finite set Ω, on which a discrete time Markov process is defined. We set an initial time $t = 0$, and fix the initial condition to be $\boldsymbol{n}_0 \in \Omega$. Then after each time interval Δt, we determine if the state $\boldsymbol{n}$ jumps to another state $\boldsymbol{n}'$ or remains in the same state $\boldsymbol{n}$, according to the transition probability

$$\mathrm{Prob}_{\Delta t}(\boldsymbol{n}'|\boldsymbol{n}) = \delta_{\boldsymbol{n}',\boldsymbol{n}}\left[1 - \lambda(\boldsymbol{n})\Delta t\right] + w(\boldsymbol{n} \to \boldsymbol{n}')\Delta t, \tag{A.1}$$

where $w(\boldsymbol{n} \to \boldsymbol{n}')$ is the transition rate introduced in Sect. 2.2.1, and $\lambda(\boldsymbol{n})$ is the escape rate defined in (2.1). This means that the state remains in the same state with a probability $\lambda(\boldsymbol{n})\Delta t$, and on the other hand, when the system jumps, the place after the jump is determined by a probability that is proportional to $w(\boldsymbol{n} \to \boldsymbol{n}')$. After a certain time t, we have the trajectory of the state, ω, which can be specified by the total number of the jumps n, a collection of transition times $(t_i)_{i=1}^{n}$ ($t_i = k\Delta t$ with some integers k), and a sequence of states $(\boldsymbol{n}_i)_{i=0}^{n}$, where $\boldsymbol{n}_i = \boldsymbol{n}(t)$ for $t_i \le t \le t_{i+1}$ with $t_0 = 0$, $t_{n+1} = t$. See Fig. 2.1 for the schematic diagram. From the definition of the transition probability (A.1), the path probability $\mathrm{Prob}(\omega|\boldsymbol{n}_0)$ is calculated as

$$\mathrm{Prob}(\omega|\boldsymbol{n}_0) = \left[1 - \lambda(\boldsymbol{n}_0)\Delta t\right]^{\frac{t_1}{\Delta t}} \prod_{i=1}^{N} \left[w(\boldsymbol{n}_{i-1} \to \boldsymbol{n}_i)\Delta t \left[1 - \lambda(\boldsymbol{n}_i)\Delta t\right]^{\frac{t_{i+1}-t_i}{\Delta t}} \right] \tag{A.2}$$

Now we take $\Delta t \to 0$ limit in this formulation. The obtained model is a continuous time Markov process. The path probability density $P(\omega|\boldsymbol{n}_0)$ is obtained from (A.2) as

T. Nemoto, *Phenomenological Structure for the Large Deviation Principle in Time-Series Statistics*, Springer Theses,
DOI 10.1007/978-981-287-811-3

$$P(\omega|\boldsymbol{n}_0) \equiv \lim_{\Delta t \to 0} \frac{\mathrm{Prob}(\omega|\boldsymbol{n}_0)}{\Delta t^N} = e^{-\lambda(\boldsymbol{n}_0)t_1} \prod_{i=1}^{N} \left[w(\boldsymbol{n}_{i-1} \to \boldsymbol{n}_i) e^{-\lambda(\boldsymbol{n}_i)(t_{i+1}-t_i)} \right], \tag{A.3}$$

which corresponds to (2.3).

Lastly, we derive the master equation (2.2) from this formulation. By using the path probability density (2.3), the distribution function of $\boldsymbol{n}(t)$ is given as

$$\begin{aligned} \left\langle \delta_{\boldsymbol{n}(t),\boldsymbol{n}} \right\rangle = \sum_{N=0}^{\infty} \sum_{\boldsymbol{n}_0,\boldsymbol{n}_1,\ldots,\boldsymbol{n}_N} \int_0^t dt_N \int_0^{t_N} dt_{N-1} \cdots \int_0^{t_2} dt_1\, P_0(\boldsymbol{n}_0) \\ \delta_{\boldsymbol{n}(t),\boldsymbol{n}} e^{-\int_0^t d\tilde{t}\lambda(\boldsymbol{n}(\tilde{t}))} \prod_{i=1}^{N} \left[w(\boldsymbol{n}_{i-1} \to \boldsymbol{n}_i) \right]. \end{aligned} \tag{A.4}$$

Then, by using $\delta_{\boldsymbol{n}(t),\boldsymbol{n}}$, we rewrite this as follows:

$$\begin{aligned} \left\langle \delta_{\boldsymbol{n}(t),\boldsymbol{n}} \right\rangle = & \sum_{N=0}^{\infty} \sum_{\boldsymbol{n}_0,\boldsymbol{n}_1,\ldots,\boldsymbol{n}_{N-1}} \int_0^t dt_N \int_0^{t_N} dt_{N-1} \cdots \int_0^{t_2} dt_1 e^{-\int_0^t d\tilde{t}\lambda(\boldsymbol{n}(\tilde{t}))} P_0(\boldsymbol{n}_0) \\ & \times \prod_{i=1}^{N-1} \left[w(\boldsymbol{n}_{i-1} \to \boldsymbol{n}_i) \right] w(\boldsymbol{n}_{N-1} \to \boldsymbol{n}) \end{aligned} \tag{A.5}$$

Finally, we take the derivative of this expression with respect to t. By noticing that the derivative is taken at the end-time of the integral with respect to t_N and at the exponential of the time-integral of the escape rate, and by calculating a little bit after that, we obtain

$$\frac{\partial \left\langle \delta_{\boldsymbol{n}(t),\boldsymbol{n}} \right\rangle_{\boldsymbol{n}_0}}{\partial t} = \sum_{\boldsymbol{n}'} \left\langle \delta_{\boldsymbol{n}(t),\boldsymbol{n}'} \right\rangle_{\boldsymbol{n}_0} w(\boldsymbol{n}' \to \boldsymbol{n}) - \lambda(\boldsymbol{n}) \left\langle \delta_{\boldsymbol{n}(t),\boldsymbol{n}} \right\rangle_{\boldsymbol{n}_0}, \tag{A.6}$$

which is the Master equation (2.2).

A.2 Derivation of the Variational Principle (2.29) and (2.30) from Donsker–Varadhan Formula

Here, we derive the variational principle (2.29) and (2.30) from Donsker–Varadhan formula. This derivation was given by C. Maes in a private discussion in Nordita, Stockholm.

A.2.1 Donsker–Varadhan Functional

We first introduce Donsker–Varadhan formula [2]. For a given path ω, we define an empirical measure $\hat{\rho}(\boldsymbol{n};\omega)$ by

$$\hat{\rho}(\boldsymbol{n},\omega) \equiv \frac{1}{t}\int_0^t ds\delta_{\boldsymbol{n}(s),\boldsymbol{n}}. \tag{A.7}$$

In the limit $t \to \infty$, due to the law of large numbers, $\hat{\rho}(\boldsymbol{n},t)$ is equal to the stationary distribution function with probability 1. For the large but finite time t, $\hat{\rho}(\boldsymbol{n},t)$ takes almost the same form as the stationary distribution function, it certainly deviates from it. The probability distribution of this deviation is given as a large deviation principle. Donsker and Varadhan proved it for general Markov dynamics with a mathematically rigorous manner, and derived a formula determining the large deviation functional: For large t, the probability of $\hat{\rho}(\boldsymbol{n},t)$ satisfies a large deviation principle

$$\mathrm{Prob}[\hat{\rho}(\boldsymbol{n},t)=\rho(\boldsymbol{n})] \sim e^{-tI[\rho(\boldsymbol{n})]}. \tag{A.8}$$

with a large deviation functional

$$I[\rho] = -\min_{\tilde{\phi}>0}\sum_{\boldsymbol{n},\boldsymbol{n}'}\rho(\boldsymbol{n})\frac{\left[(w(\boldsymbol{n}\to\boldsymbol{n}')-\lambda(\boldsymbol{n})\delta_{\boldsymbol{n},\boldsymbol{n}'}\right]\tilde{\phi}(\boldsymbol{n}')}{\tilde{\phi}(\boldsymbol{n})}, \tag{A.9}$$

where $\tilde{\phi}(\boldsymbol{n})$ is a variational parameter (vector), where each component takes a positive value. By defining a potential $\tilde{V}(\boldsymbol{n})$ as $\tilde{V}(\boldsymbol{n}) = -2\log\tilde{\phi}(\boldsymbol{n})$, we rewrite it in terms of the modified transition rate $\tilde{w}_h^{\tilde{V}}$ and the corresponding escape rate $\tilde{\lambda}_h^{\tilde{V}}$ introduced in (2.28). The result is

$$I[\rho] = -\min_{\tilde{V}}\sum_{\boldsymbol{n}}\rho(\boldsymbol{n})\left[\tilde{\lambda}_0^{\tilde{V}}(\boldsymbol{n})-\lambda(\boldsymbol{n})\right], \tag{A.10}$$

Below, we show that the variational formula (2.29) and (2.30) are derived from this Donsker–Varadhan formula.

A.2.2 Correspondence Between Biased Ensemble and Modified System

The key idea to derive (2.29) and (2.30) is to construct a correspondence between the biased ensemble and the modified system. We denote the path probability density of $\tilde{w}_h^{\tilde{V}}$-system by $P_h^{\tilde{V}}(\omega|\boldsymbol{n}_0)$ with a given initial condition $\boldsymbol{n}_0$. The path probability density is given as

$$
\begin{aligned}
P_h^{\tilde{V}}(\omega|\boldsymbol{n}_0) &= e^{-\int_0^t d\tilde{t}\tilde{\lambda}_h^{\tilde{V}}(\boldsymbol{n}(\tilde{t}))} \prod_{i=1}^{N} \left[w(\boldsymbol{n}_{i-1} \to \boldsymbol{n}_i) \right] \prod_{i=1}^{N} \left[e^{h\alpha(\boldsymbol{n}_{i-1}\to\boldsymbol{n}_i)-(1/2)\tilde{V}(\boldsymbol{n}_i)+(1/2)\tilde{V}(\boldsymbol{n}_{i-1})} \right] \\
&= e^{htA(\omega)-(1/2)\tilde{V}(\boldsymbol{n}_N)+(1/2)\tilde{V}(\boldsymbol{n}_0)-\int_0^t d\tilde{t}\tilde{\lambda}_h^{\tilde{V}}(\boldsymbol{n}(\tilde{t}))} \prod_{i=1}^{N} \left[w(\boldsymbol{n}_{i-1} \to \boldsymbol{n}_i) \right] \\
&= e^{htA(\omega)-(1/2)\tilde{V}(\boldsymbol{n}_N)+(1/2)\tilde{V}(\boldsymbol{n}_0)-\int_0^t d\tilde{t}\left[\tilde{\lambda}_h^{\tilde{V}}(\boldsymbol{n}(\tilde{t}))-\lambda(\boldsymbol{n}(\tilde{t}))\right]} P(\omega|\boldsymbol{n}_0).
\end{aligned}
\tag{A.11}
$$

We then neglect $-(1/2)\tilde{V}(\boldsymbol{n}_N) + (1/2)\tilde{V}(\boldsymbol{n}_0)$, since these terms are not proportional to t. Finally, by expressing the time integral with the empirical measure $\hat{\rho}(\boldsymbol{n}, \omega)$, we arrive at

$$
e^{htA(\omega)} P(\omega|\boldsymbol{n}_0) \simeq P_h^{\tilde{V}}(\omega|\boldsymbol{n}_0) e^{t\sum_{\boldsymbol{n}} \hat{\rho}(\boldsymbol{n},\omega)\left[\tilde{\lambda}_h^{\tilde{V}}(\boldsymbol{n})-\lambda(\boldsymbol{n})\right]} \tag{A.12}
$$

for large t.

A.2.3 Derivation of (2.29) and (2.30)

We set $\tilde{V}$ to be 0 in (A.12). Then, we substitute it into the definition of the cumulant generating function (2.8). We consider the Donsker–Varadhan formula in $\tilde{w}_h^0$-system, for which we denote the Donsker–Varadhan functional in this system by $I_h^0[\rho]$. By using $I_h^0[\rho]$ in the obtained expression and evaluating it by using a saddle-point method, we obtain the cumulant generating function as

$$
G(h) = \max_{\tilde{\rho}} \left[\sum_{\boldsymbol{n}} \tilde{\rho}(\boldsymbol{n}) \left(\tilde{\lambda}_h^0(\boldsymbol{n}) - \lambda(\boldsymbol{n}) \right) - I_h^0[\tilde{\rho}] \right]. \tag{A.13}
$$

Here, we recall the expression of the Donsker–Varadhan formula (A.9). We denote by V_h^* the optimal $\tilde{V}$ that appeared in the variational principle in the Donsker–Varadhan formula:

$$
V_h^* = \operatorname*{argmin}_{\tilde{V}} \left[\sum_{\boldsymbol{n}} \rho(\boldsymbol{n}) \left(\tilde{\lambda}_h^{\tilde{V}}(\boldsymbol{n}) - \lambda_h^0(\boldsymbol{n}) \right) \right]. \tag{A.14}
$$

By combining this definition with (A.13), we thus obtain

$$
G(h) = \max_{\tilde{\rho}} \left[\sum_{\boldsymbol{n}} \tilde{\rho}(\boldsymbol{n}) \left[\tilde{\lambda}_h^{V_h^*}(\boldsymbol{n}) - \lambda(\boldsymbol{n}) \right] \right]. \tag{A.15}
$$

The obtained formula (A.15) is basically the same as (2.30). The only difference in it is the variational parameter $\tilde{\rho}$. As the final step, we replace this variational parameter $\tilde{\rho}$ by a potential. First, we consider the variational equation for the minimization problem (A.14), which determines $V_h^*(\boldsymbol{n})$. The equation is

$$\sum_{\boldsymbol{n},\boldsymbol{n}'} \tilde{\rho}(\boldsymbol{n}) \left[\delta\tilde{V}(\boldsymbol{n}) - \delta\tilde{V}(\boldsymbol{n}') \right] \tilde{w}_h^{V_h^*}(\boldsymbol{n} \to \boldsymbol{n}') = 0. \tag{A.16}$$

It is further rewritten as

$$\sum_{\boldsymbol{n}} \delta\tilde{V}(\boldsymbol{n}) \sum_{\boldsymbol{n}'} \left[\tilde{\rho}(\boldsymbol{n}) \tilde{w}_h^{\tilde{V}_h^*}(\boldsymbol{n} \to \boldsymbol{n}') - \tilde{\rho}(\boldsymbol{n}') \tilde{w}_h^{V_h^*}(\boldsymbol{n}' \to \boldsymbol{n}) \right] = 0. \tag{A.17}$$

Since $\delta\tilde{V}$ is arbitrary, we obtain

$$\sum_{\boldsymbol{n}'} \left[\tilde{\rho}(\boldsymbol{n}) \tilde{w}_h^{V_h^*}(\boldsymbol{n} \to \boldsymbol{n}') - \tilde{\rho}(\boldsymbol{n}') \tilde{w}_h^{V_h^*}(\boldsymbol{n}' \to \boldsymbol{n}) \right] = 0. \tag{A.18}$$

It means that $V_h^*(\boldsymbol{n})$ is determined in such a way that $\tilde{\rho}(\boldsymbol{n})$ is the stationary probability for the modified system with $\tilde{w}_h^{V_h^*}$. (For to the uniqueness and existence, see Proposition III.1 in Ref. [1].) Since the stationary probability ρ_h^V is determined uniquely for a given $\tilde{w}_h^V$, we find the one-to-one correspondence between V_h^* and $\tilde{\rho}$. Thus, we may rewrite (A.15) as

$$G(\sigma) = \max_V \left[\sum_x \mu_\sigma^V(x) \left(D_\sigma^V(x) - D(x) \right) \right]. \tag{A.19}$$

This is exactly the same as (2.30), which directly leads to (2.29).

A.3 Derivation of (2.36)

Here, we show that $\left\langle \delta_{\boldsymbol{n}(t),\boldsymbol{n}} e^{htA(\omega)} \right\rangle_{\boldsymbol{n}_0}$ satisfies the same equations as (2.34) and (2.35), which leads to (2.36). For the derivation of (2.34), by substituting t in $\left\langle \delta_{\boldsymbol{n}(t),\boldsymbol{n}} e^{htA(\omega)} \right\rangle_{\boldsymbol{n}_0}$ by 0, we can easily check $\left\langle \delta_{\boldsymbol{n}(t),\boldsymbol{n}} e^{htA(\omega)} \right\rangle_{\boldsymbol{n}_0} = \delta_{\boldsymbol{n},\boldsymbol{n}_0}$. On the other hand, for (2.35), we start with the path probability density (2.4). From the expression, $\left\langle \delta_{\boldsymbol{n}(t),\boldsymbol{n}} e^{htA(\omega)} \right\rangle_{\boldsymbol{n}_0}$ is written as

$$\begin{aligned} \left\langle \delta_{\boldsymbol{n}(t),\boldsymbol{n}} e^{htA(\omega)} \right\rangle_{\boldsymbol{n}_0} = & \sum_{N=0}^{\infty} \sum_{\boldsymbol{n}_1,\boldsymbol{n}_2,\ldots,\boldsymbol{n}_N} \int_0^t dt_N \int_0^{t_N} dt_{N-1} \cdots \int_0^{t_2} dt_1 \\ & \delta_{\boldsymbol{n}(t),\boldsymbol{n}} e^{htA(\omega)} e^{-\int_0^t d\tilde{t} \lambda(\boldsymbol{n}(\tilde{t}))} \prod_{i=1}^{N} \left[w(\boldsymbol{n}_{i-1} \to \boldsymbol{n}_i) \right]. \end{aligned} \tag{A.20}$$

Then, in this expression, we make $\delta_{\boldsymbol{n}(t),\boldsymbol{n}} e^{htA(\omega)}$ included inside the path-probability. The result is

$$\left\langle \delta_{\boldsymbol{n}(t),\boldsymbol{n}} e^{htA(\omega)} \right\rangle_{\boldsymbol{n}_0} = \sum_{N=0}^{\infty} \sum_{\boldsymbol{n}_1,\boldsymbol{n}_2,\ldots,\boldsymbol{n}_{N-1}} \int_0^t dt_N \int_0^{t_N} dt_{N-1} \cdots \int_0^{t_2} dt_1 e^{-\int_0^t d\tilde{t}\lambda(\boldsymbol{n}(\tilde{t}))} \\ \times \prod_{i=1}^{N-1} \left[w(\boldsymbol{n}_{i-1} \to \boldsymbol{n}_i) e^{h\alpha(\boldsymbol{n}_{i-1}\to\boldsymbol{n}_i)} \right] w(\boldsymbol{n}_{N-1} \to \boldsymbol{n}) e^{h\alpha(\boldsymbol{n}_{N-1}\to\boldsymbol{n})} \tag{A.21}$$

We differentiate this expression with respect to t. The differentiation is taken at the end-time of the integral with respect to t_N and at the exponential of the time-integral of the escape rate. After some calculations, we arrive at

$$\frac{\partial \left\langle \delta_{\boldsymbol{n}(t),\boldsymbol{n}} e^{htA(\omega)} \right\rangle_{\boldsymbol{n}_0}}{\partial t} \\ = \sum_{\boldsymbol{n}'} \left\langle \delta_{\boldsymbol{n}(t),\boldsymbol{n}'} e^{htA(\omega)} \right\rangle_{\boldsymbol{n}_0} w(\boldsymbol{n}' \to \boldsymbol{n}) e^{h\alpha(\boldsymbol{n}'\to\boldsymbol{n})} - \lambda(\boldsymbol{n}) \left\langle \delta_{\boldsymbol{n}(t),\boldsymbol{n}} e^{htA(\omega)} \right\rangle_{\boldsymbol{n}_0}, \tag{A.22}$$

which is the corresponding equation to (2.35).

A.4 Derivation of the Theoretical Basis (2.48) of the Rare-Event Sampling Method

We denote by $\phi(\boldsymbol{n}; h)$ the left-eigenvector corresponding to the largest eigenvalue of the matrix $L^h_{\boldsymbol{n}',\boldsymbol{n}}$ defined in (2.25). In this appendix, then, we derive

$$\phi(\boldsymbol{n}; l\delta h) \propto \prod_{k=0}^{l-1} \left\langle e^{\tau \delta h A(\omega)} \right\rangle_{\boldsymbol{n}}^{k\delta h} \tag{A.23}$$

for $l = 1, 2, 3, \ldots$ From this relation, it is easy to check (2.48) follows. Indeed, by combining this relation with the argument in Sect. 2.2.5, we obtain (2.48).

First, we define

$$w^h(\boldsymbol{n} \to \boldsymbol{n}') = w(\boldsymbol{n} \to \boldsymbol{n}') e^{h\alpha(\boldsymbol{n}\to\boldsymbol{n}')} \frac{\phi(\boldsymbol{n}'; l\delta h)}{\phi(\boldsymbol{n}; l\delta h)}. \tag{A.24}$$

Then, by using $w^h(\boldsymbol{n} \to \boldsymbol{n}')$, we define a matrix

$$L^{h,h'}_{\boldsymbol{n},\boldsymbol{n}'} \equiv w^{h'}(\boldsymbol{n}' \to \boldsymbol{n}) e^{h\alpha(\boldsymbol{n}'\to\boldsymbol{n})} - \lambda^{w^{h'}}(\boldsymbol{n}) \delta_{\boldsymbol{n},\boldsymbol{n}'}, \tag{A.25}$$

where $\lambda^{w^{h'}}(\boldsymbol{n}) \equiv \sum_{\boldsymbol{n}'} w^{h'}(\boldsymbol{n} \to \boldsymbol{n}')$. Let $K^{h,h'}$ and $\phi^{h,h'}$ be the largest eigenvalue and the corresponding left-eigenvector of (A.25). Then, we can prove the following *multiplicative property* for the eigenvector and *additive property* for the eigenvalue of the matrix (A.25):

$$\phi^{h+h',0} = \phi^{h,h'}\phi^{h',0}, \tag{A.26}$$

$$K^{h+h',0} = K^{h,h'} + K^{h',0}. \tag{A.27}$$

- Proof:
First, we write the eigenvalue equations for $\phi^{h+h',0}$, $\phi^{h',0}$, and $\phi^{h,h'}$. Those are

$$\sum_{\boldsymbol{n}'} w(\boldsymbol{n} \to \boldsymbol{n}')e^{(h+h')\alpha(\boldsymbol{n}\to\boldsymbol{n}')}\frac{\phi^{h+h',0}(\boldsymbol{n}')}{\phi^{h+h',0}(\boldsymbol{n})} - \lambda(\boldsymbol{n}) = K^{h+h',0}, \tag{A.28}$$

$$\sum_{\boldsymbol{n}'} w(\boldsymbol{n} \to \boldsymbol{n}')e^{h'\alpha(\boldsymbol{n}\to\boldsymbol{n}')}\frac{\phi^{h',0}(\boldsymbol{n}')}{\phi^{h',0}(\boldsymbol{n})} - \lambda(\boldsymbol{n}) = K^{h',0}, \tag{A.29}$$

$$\begin{aligned}&\sum_{\boldsymbol{n}'} w(\boldsymbol{n} \to \boldsymbol{n}')e^{(h+h')\alpha(\boldsymbol{n}\to\boldsymbol{n}')}\frac{\phi^{h,h'}(\boldsymbol{n}')\phi^{h',0}(\boldsymbol{n}')}{\phi^{h,h'}(\boldsymbol{n})\phi^{h',0}(\boldsymbol{n})} - \lambda^{u^{h'}}(\boldsymbol{n})\\ &= K^{h,h'}.\end{aligned} \tag{A.30}$$

We sum up (A.29) and (A.30). Since the first term of (A.29) is the same as the second one in (A.30), these terms cancel each other. The result is

$$\begin{aligned}&\sum_{\boldsymbol{n}'} w(\boldsymbol{n} \to \boldsymbol{n}')e^{(h+h')\alpha(\boldsymbol{n}\to\boldsymbol{n}')}\frac{\phi^{h,h'}(\boldsymbol{n}')\phi^{h',0}(\boldsymbol{n}')}{\phi^{h,h'}(\boldsymbol{n})\phi^{h',0}(\boldsymbol{n})} - \lambda(\boldsymbol{n})\\ &= K^{h,h'} + K^{h',0}.\end{aligned} \tag{A.31}$$

From the Perron-Frobenius theory for irreducible matrices [3], the eigenvector of $L^{h,0}_{\boldsymbol{n},\boldsymbol{n}'}$ that takes positive value is unique, and the corresponding eigenvalue should become the largest one $G^{h+h',0}$. Thus, by comparing (A.28) with (A.31), we obtain (A.26) and (A.27).

Next, we denote by $\langle\ \rangle^{w^h}$ the expected value in the stationary state of the system generated by the transition rate $w^h(\boldsymbol{n} \to \boldsymbol{n}')$. We will show that this expected value and the expected value $\langle\ \rangle^h$ defined in Sect. 2.3.4 are equal. Meanwhile, we show a formula for $\langle\ \rangle^{w^h}$ and finally we will show the equivalence. First, we have a relation

$$\phi^{h,h'}(\boldsymbol{n}) \propto \left\langle e^{thA(\omega)}\right\rangle_{\boldsymbol{n}}^{w^{h'}} \tag{A.32}$$

for large $t(> t_a)$. The derivation is the same as the one for (2.38), so we don't repeat it here. Then, by combining (A.32) with (A.28), we obtain

$$\phi^{h+h',0}(\boldsymbol{n}) \propto \left\langle e^{thA(\omega)}\right\rangle_{\boldsymbol{n}}^{w^{h'}} \phi^{h',0}(\boldsymbol{n}), \tag{A.33}$$

or equivalently,

$$\phi^{h+h',0}(\boldsymbol{n}) \propto \left\langle e^{thA(\omega)} \right\rangle_{\boldsymbol{n}}^{w^{h'}} \left\langle e^{th'A(\omega)} \right\rangle_{\boldsymbol{n}}^{0}, \tag{A.34}$$

where we used (A.32). The equivalence between $\langle\ \rangle^{w^h}$ and $\langle\ \rangle^h$ is easily checked. Indeed, by using (2.38) in the definition of $\langle\ \rangle^h$, the relation follows. We thus have

$$\phi^{h+h',0}(\boldsymbol{n}) \propto \left\langle e^{thA(\omega)} \right\rangle_{\boldsymbol{n}}^{h'} \left\langle e^{th'A(\omega)} \right\rangle_{\boldsymbol{n}}^{0}. \tag{A.35}$$

The generalisation to the case, where a given h is divided into l pieces, is straightforward. Therefore, we arrive at (A.23).

A.5 Application of Our Method to Obtain a L Dependence of $\tilde{G}(\tilde{h})$ in (2.72)

For investigating the singular behaviour of $G(h)$ in greater detail, a scaled biasing parameter $\tilde{h} \equiv hL$ has been used [4, 5]. The problem was in the numerical study of it because the population dynamics method does not exhibit good convergence of $\tilde{G}(\tilde{h})$ for relatively large L [5]. In this appendix, we show that our method can also be applied to obtain the reliable L dependence of $\tilde{G}(\tilde{h})$ even in this case.

A.5.1 (i) Sufficiently Large Truncating Number r for Severe L

We investigate r dependence of $\tilde{G}(\tilde{h})$ for several L. We will conclude that $r = 4$ is sufficiently large to obtain $\tilde{G}(\tilde{h}) \equiv G(L^{-1}\tilde{h})$ for $\tilde{h} < 0$. First, we define $\tilde{G}(\tilde{h}, r)$ as the obtained cumulant generating function with the truncating number r. Then, we show the numerical examples of $\tilde{G}(\tilde{h}, r)$ in Fig. A.1. In the figure, we plot $\tilde{G}(\tilde{h}, 1)$, $\tilde{G}(\tilde{h}, 2)$, $\tilde{G}(\tilde{h}, 3)$, and $\tilde{G}(\tilde{h}, 4)$ for $L = 10$ (i), 20 (ii), 30 (iii), and 60 (iv). In the same figure, we also plot the straight line (Ex) with the slope $4c^2(1-c)$, which is the expected value of the activity in the unmodified system ($\tilde{h} = 0$). We note that the straight line corresponds to $\tilde{G}(\tilde{h}, 0)$, because $r = 0$ means that there are no modifications. From the figure, we find that the differences between $\tilde{G}(\tilde{h}, 3)$ and $\tilde{G}(\tilde{h}, 4)$ are small even for larger L (say $L = 20$, 30, and 60).

We then quantitatively evaluate those small differences. For this, we introduce a difference function

$$\delta\tilde{G}(\tilde{h}, r) = \tilde{G}(\tilde{h}, r+1) - \tilde{G}(\tilde{h}, r). \tag{A.36}$$

Then, in Fig. A.2, we plot the logarithm of $\delta G(r, h)$ as a function of r for $\tilde{h} = -0.225, -0.45, -0.675$ and -0.9 with $L = 20$ (i), 30 (ii), and 60 (iii). We can see

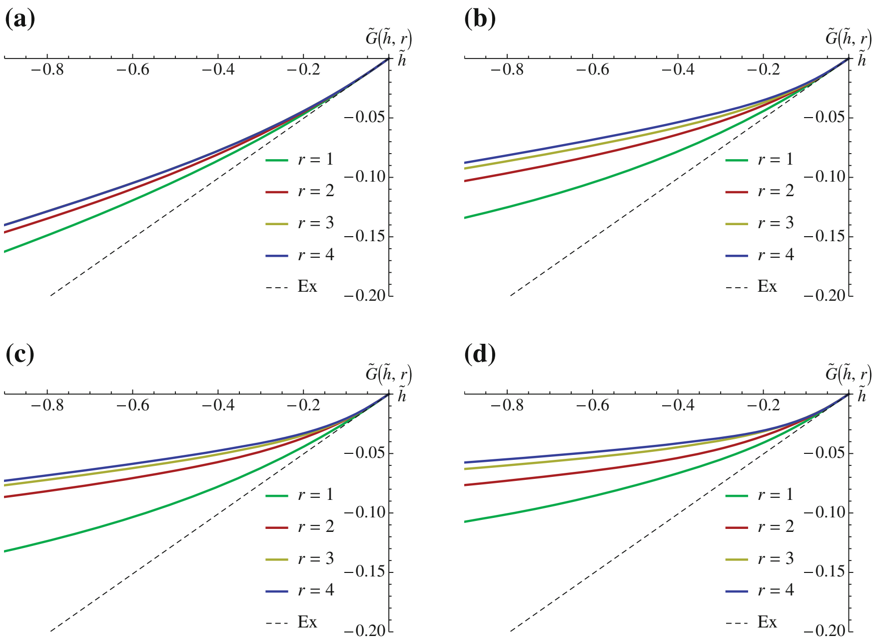

Fig. A.1 $\tilde{G}(\tilde{h}, r)$ for $r = 1, 2, 3, 4$ with $L = 10$ (**a**), 20 (**b**), 30 (**c**), and 60 (**d**), where we set $c = 0.3$. We also plot the *straight line* of slope $4c^2(1-c)$ (Ex) in each figure, which corresponds to $\tilde{G}(\tilde{h}, 0)$

the linear dependence of $\log[\delta G(r, h)]$ on r. This means exponentially fast decay of $\delta\tilde{G}$, which may indicate that larger r isn't needed to obtain the correct $\tilde{G}(\tilde{h})$.

Finally, we evaluate the error due to the truncation of $r = 4$. From Fig. A.2, we assume that the decaying of $\delta\tilde{G}(\tilde{h}, r)$ with r is well described by an exponential function:

$$\delta\tilde{G}_{\rm lin}(\tilde{h}, r) = e^{a(\tilde{h})r + b(\tilde{h})}, \tag{A.37}$$

where $a(\tilde{h})$ and $b(\tilde{h})$ are coefficients determined from the least squares fit of data points $\delta\tilde{G}(\tilde{h}, r)$. The examples of the least squares fit are shown in Fig. A.2. By using this difference function, we define an (exponential decaying) approximation function of $\tilde{G}(\tilde{h}, r)$ as

$$\tilde{G}_{\rm lin}(\tilde{h}, r) \equiv \tilde{G}(\tilde{h}, 4) + \sum_{s=4}^{r-1} \delta\tilde{G}_{\rm lin}(\tilde{h}, s) \tag{A.38}$$

for $r = 5, 6, \ldots$ Especially, here, we denote $\tilde{G}_{\rm lin}(\tilde{h}, L/2 - 1)$ by $\tilde{G}_{\rm lin}(\tilde{h})$:

$$\tilde{G}_{\rm lin}(\tilde{h}) \equiv \tilde{G}_{\rm lin}(\tilde{h}, L/2 - 1). \tag{A.39}$$

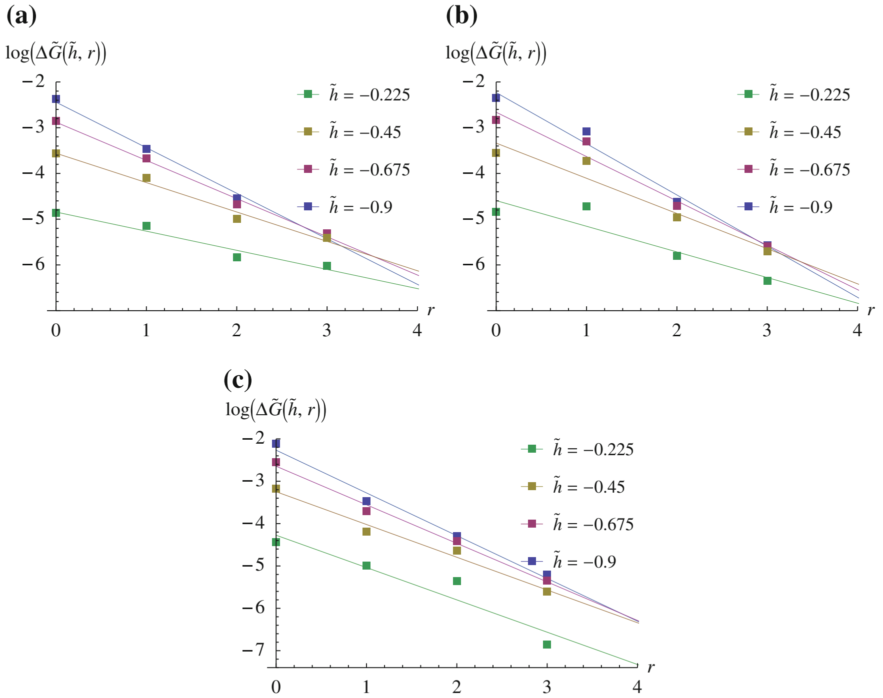

Fig. A.2 The logarithm of the difference function $\delta\tilde{G}(\tilde{h}, r)$ for $r = 0, 1, 2$, and 3 with $L = 20$ (**a**), 30 (**b**), and 60 (**c**). We set $c = 0.3$. We also plot *straight lines* obtained from a least squares fit of those data points

Since $\tilde{G}(\tilde{h}, r)$ with $r \simeq L/2$ is equal to $\tilde{G}(\tilde{h})$, we regard $\tilde{G}_{\mathrm{lin}}(\tilde{h})$ as an approximation function of $\tilde{G}(\tilde{h})$. We plot $\tilde{G}_{\mathrm{lin}}(\tilde{h})$ and $\tilde{G}(\tilde{h}, r)$ for $r = 1, 2, 3, 4$ in Fig. A.3. The figure shows that the differences between $\tilde{G}_{\mathrm{lin}}(\tilde{h})$ and $\tilde{G}(\tilde{h}, 4)$ are quite small even for large L. Therefore, we judge that $r = 4$ is sufficiently large to obtain $\tilde{G}(\tilde{h})$ even for those large L.

A.5.2 (ii) *L* Dependence of $\tilde{G}(\tilde{h})$ Obtained from Truncation

In the previous subsection, we judged that $r = 4$ was sufficiently large to obtain $\tilde{G}(\tilde{h})$. Here, by using this result, we show the L dependence of $\tilde{G}(\tilde{h})$. In Fig. A.4, we plot $\tilde{G}(\tilde{h})$ with $r = 4$ for various values of L . Even though it was conjectured that $\tilde{G}(\tilde{h})$ has a non-differentiable point in the limit $L \to \infty$ [4], our result does not show any clear sign of such a cusp. More and more larger system sizes are required for investigating the singular behaviour of $\tilde{G}(\tilde{h})$.

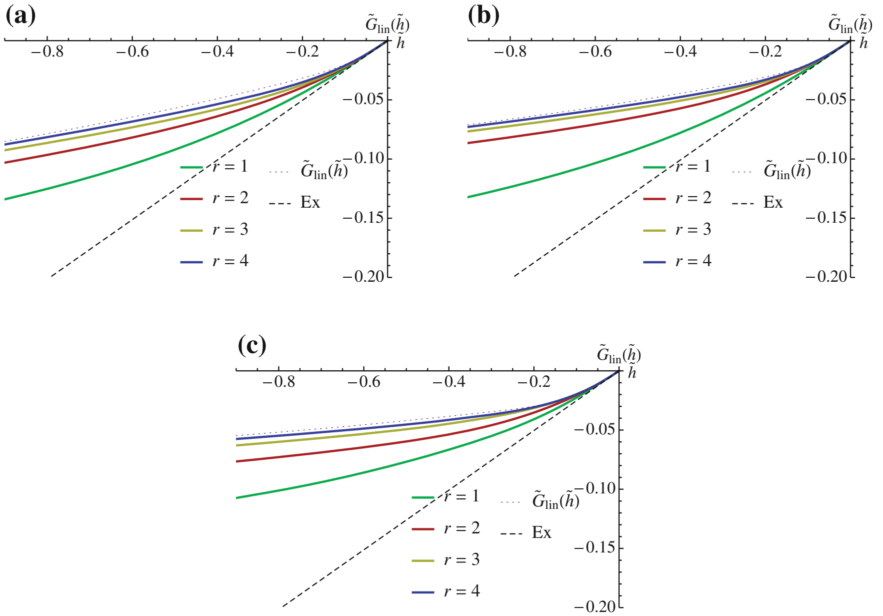

Fig. A.3 $\tilde{G}(\tilde{h}, r)$ and $\tilde{G}_{\rm lin}(\tilde{h})$ for $r = 1, 2, 3, 4$ with $L = 20$ (**a**), 30 (**b**), and 60 (**c**), where we set $c = 0.3$

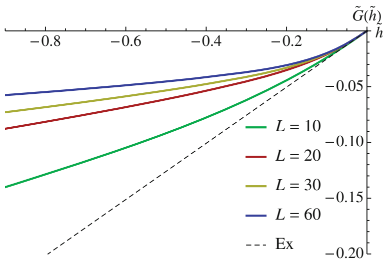

Fig. A.4 $\tilde{G}(\tilde{h})$ for various values of L with $r = 4$ fixed. We also plot the *straight line* (Ex) of slope $4c^2(1 - c)$ that corresponds to the expected value of the activity in the unbiased system ($h = 0$)

References

[1] C. Maes, K. Netočný, B. Wynants, J. Phys. A: Math. Theor. **45**, 455001 (2012)
[2] M.D. Donsker, S.R. Varadhan, Commun. Pure Appl. Math. **28**, 1 (1975)
[3] E. Seneta, *Non-Negative Matrices and Markov Chains*, 2nd edn. (Springer, New York, 2006)
[4] T. Bodineau, C. Toninelli, Commun. Math. Phys. **311**, 357 (2012)
[5] T. Bodineau, V. Lecomte, C. Toninelli, J. Stat. Phys. **147**, 1 (2012)

Appendix B
For Chapter 3

B.1 Derivation of Finite-Size Corrections to Modifying Free Energy $\Delta f_{s_c}^{(1)}(\rho)$ in (3.55) and (3.56)

In this appendix, we derive the finite-size correction $\Delta f_{s_c}^{(1)}(\rho)$ given in (3.55) and (3.56).

B.1.1 For the Region $\rho > \rho_c^L$

We first focus on the region $\rho > \rho_c^L$, which corresponds to (3.56). In this region, from the expression of the leading order of the free energy (3.50), we know that $\phi(n)$ doesn't satisfy any large deviation principle. We thus define $\tilde{\phi}(\rho) = \phi(\rho L)$ and assume the differentiability of it:

$$\tilde{\phi}(\rho \pm 1/L) = \tilde{\phi}(\rho) \pm \frac{\partial \tilde{\phi}}{\partial \rho} \frac{1}{L} + O(1/L^2). \tag{B.1}$$

By using this scaling property, we rewrite the left-hand side of (3.13). The result is

$$\tilde{\phi}(\rho)\left\{-\tilde{s}_c \rho \left[c + (1-2c)\rho\right] + c\right\} + \frac{\partial \tilde{\phi}(\rho)}{\partial \rho} \rho(c-\rho) + O(1/L^2) = 0, \tag{B.2}$$

where we defined $\tilde{s}_c \equiv sL$. By solving this differential equation, we obtain the modifying free energy with finite-size correction, which we denote by $-2\log\tilde{\phi}$, as follows:

T. Nemoto, *Phenomenological Structure for the Large Deviation Principle in Time-Series Statistics*, Springer Theses,
DOI 10.1007/978-981-287-811-3

$$
\begin{aligned}
&-2\log\tilde{\phi}(\rho)\\
&\quad=-2\left[\tilde{s}_{\mathrm{c}}\rho(2c-1)-\log\rho+(-\tilde{s}_{\mathrm{c}}2c(1-c)+1)\log|c-\rho|\right]+\text{const.}\\
&\quad=-2\left[\frac{\rho(2c-1)}{2c(1-c)}-\log\rho\right]+\text{const.},
\end{aligned}
\tag{B.3}
$$

where we used a relation $\tilde{s}_{\mathrm{c}}=1/(2c(1-c))+O(1/L)$ from the second to the third line. Since the leading term to the modifying free energy is constant as shown in (3.50), we arrive at

$$
\Delta f_{s_{\mathrm{c}}}^{(1)}(\rho)=-2\left[\frac{\rho(2c-1)}{2c(1-c)}-\log\rho\right]+\text{const.}
\tag{B.4}
$$

for $\rho>\rho_{\mathrm{c}}^{L}$, which corresponds to (3.56).

B.1.2 For the Region $\rho\le\rho_{\mathrm{c}}^{L}$

Next, we focus on the region $\rho\le\rho_{\mathrm{c}}^{L}$, which corresponds to (3.55). In this region, $\phi(n)$ satisfies a large deviation principle. We thus directly substitute the definition (3.54) with the aide of the result of the leading order give by (3.49) into the eigenvalue equation (3.43). In the calculation, we evaluate $\phi(n+1)/\phi(n)$ as

$$
\begin{aligned}
\frac{\phi(n+1)}{\phi(n)}&=e^{\partial f_{\mathrm{e}}/\partial\rho+1/(2L)\partial^2 f_{\mathrm{e}}/\partial\rho^2-1/(2L)\partial\Delta f_{s_{\mathrm{c}}}^{(1)}/\partial\rho}\\
&=\frac{(1-c)\rho}{c(1-\rho)}e^{(1/(2L\rho(1-\rho)))}e^{-1/(2L)\partial\Delta f_{s_{\mathrm{c}}}^{(1)}/\partial\rho}.
\end{aligned}
$$

By using this expression in (3.43), we arrive at a differential equation

$$
\begin{aligned}
\frac{\partial\Delta f_{s_{\mathrm{c}}}^{(1)}(\rho)}{\partial\rho}&=\frac{2}{\rho}-2\frac{\frac{c}{\rho}+[(1-c)\rho+c(1-\rho)]\left[\frac{1}{2\rho(1-\rho)}-\frac{1}{2c(1-c)}\right]}{-(1-c)\rho+c(1-\rho)}\\
&=-\frac{1}{\rho}+\frac{1}{1-\rho}-\frac{2}{c-\rho}-\frac{1-2c}{c(1-c)},
\end{aligned}
\tag{B.5}
$$

which leads to

$$
\Delta f_{s_{\mathrm{c}}}^{(1)}(\rho)=-\log\frac{\rho(1-\rho)}{(c-\rho)^2}-\frac{\rho(1-2c)}{c(1-c)}+\text{const.}
\tag{B.6}
$$

This corresponds to (3.55).

B.2 Derivation of $a^*(s)$ Given in (3.77)

Here, we derive the expression $a^*(s)$ given as (3.77) by solving (3.74).

In order to evaluate the variational function $\Psi(a)$ given in (3.75), we divide the domain of the summation in (3.75) into three parts:

(i) $n < n_\mathrm{c}$

$$\Psi_<(a) \equiv (1/L) \sum_{n<n_\mathrm{c}} P^s(n) \left[\tilde{r}(n) - r(n)\right], \tag{B.7}$$

(ii) $n > n_\mathrm{c} + 1$

$$\Psi_>(a) \equiv (1/L) \sum_{n>n_\mathrm{c}} P^s(n) \left[\tilde{r}(n) - r(n)\right], \tag{B.8}$$

(iii) $n \equiv n_\mathrm{c}, n_\mathrm{c} + 1$

$$\Psi_=(a) \equiv (1/L) \sum_{n=n_\mathrm{c},n_\mathrm{c}+1} P^s(n) \left[\tilde{r}(n) - r(n)\right]. \tag{B.9}$$

The dependence of a in (i) and (ii) is linear because $\tilde{r}(n)$ is an independent function of a. For these part, we define two constants $\Omega_<$ and $\Omega_>$ that doesn't depend on a:

$$\Omega_< = \frac{1}{2L} \left\langle \tilde{r}_\mathrm{i} \mathrm{e}^{-s} - r \right\rangle_\mathrm{i}, \tag{B.10}$$

$$\Omega_> = \frac{1}{2L} \left\langle \tilde{r}_\mathrm{a} \mathrm{e}^{-s} - r \right\rangle_\mathrm{a}, \tag{B.11}$$

where $\tilde{r}_{\mathrm{i,a}}(n)$ is defined as

$$\tilde{r}_{\mathrm{i,a}}(n) = nc\left(1 - \frac{n}{L}\right) \frac{\mathrm{e}^{-Lf_{\mathrm{i,a}}((n+1)/L)/2}}{\mathrm{e}^{-Lf_{\mathrm{i,a}}(n/L)/2}} + n(1-c)\left(\frac{n}{L} - \frac{1}{L}\right) \frac{\mathrm{e}^{-Lf_{\mathrm{i,a}}((n-1)/L)/2}}{\mathrm{e}^{-Lf_{\mathrm{i,a}}(n/L)/2}}. \tag{B.12}$$

By using these constants, $\Psi_<(a)$ and $\Psi_>(a)$ are written as

$$\Psi_<(a) = (1+a)\Omega_<, \tag{B.13}$$

$$\Psi_>(a) = (1-a)\Omega_>. \tag{B.14}$$

On the other hand, the dependence of a in (iii) is more complicated. We first write down $\Psi_=(a)$ as

$$\begin{aligned} \Psi_=(a) = &\frac{n_\mathrm{c}}{L} c(1 - \frac{n_\mathrm{c}}{L}) \frac{\phi(n_\mathrm{c}+1)}{\phi(n_\mathrm{c})} P^s(n_\mathrm{c}) \mathrm{e}^{-s} \\ &+ \frac{(n_c+1)}{L}(1-c) \frac{n_\mathrm{c}}{L} \frac{\phi(n_\mathrm{c})}{\phi(n_\mathrm{c}+1)} P^s(n_\mathrm{c}+1) \mathrm{e}^{-s} + \cdots, \end{aligned} \tag{B.15}$$

where $\cdots$ represents the terms that are proportional to a. Since this linear dependence of a is exponentially smaller than the one in $\Psi_<(a)$ and $\Psi_>(a)$, we omit this part hereafter. Next, with a relation between $P^s(n)$ and $\phi(n)$

$$P^s(n+1)\frac{\phi(n)}{\phi(n+1)} = P^s(n)\frac{\phi(n+1)P_{\rm eq}(n+1)}{\phi(n)P_{\rm eq}(n)}, \tag{B.16}$$

we find that the first term and the second term in the right hand side of (B.15) are equal. Also, by recalling a dependence of $P^s(n)$ and $\phi(n)$, we find that this term is proportional to $\sqrt(1-a^2)$. Thus, by defining the coefficient of it, $\Omega_=(a)$, as

$$\Omega_= \equiv 2\frac{n_{\rm c}}{L}c(1-\frac{n_{\rm c}}{L})P_{\rm i}(n_{\rm c})\frac{{\rm e}^{-Lf_{\rm a}((n_{\rm c}+1)/L)/2}}{{\rm e}^{-Lf_{\rm i}(n_{\rm c}/L)/2}}{\rm e}^{-s}, \tag{B.17}$$

we arrive at

$$\Psi_=(a) = \sqrt{1-a^2}\,\Omega_=. \tag{B.18}$$

The non-linear dependence in $\Psi_=(a)$ is important. Even though $\Psi_=(a)$ is exponentially small compared with the other parts $\Psi_>(a)$, $\Psi_<(a)$, due to the non-liner dependence, we need to consider this term. As seen in the main text, this smallness of $\Psi_=(a)$ is the origin of the exponentially small width of the coexistence region.

From (B.13), (B.14), and (B.18), we obtain

$$\Psi(a) = \Omega_< + \Omega_> + a(\Omega_< - \Omega_>) + \sqrt{1-a^2}\,\Omega_=. \tag{B.19}$$

By maximising $\Psi(a)$ with respect to a, we finally obtain the expression of $a^*(s)$ as

$$a^*(s) = \frac{A}{\sqrt{1+A^2}}, \tag{B.20}$$

with

$$A = (\Omega_< - \Omega_>)/\Omega_=. \tag{B.21}$$

This is (3.77).

B.3 Analytical Expressions of the Scaling Functions for $\partial G(s)/\partial s$ and $\partial^2 G(s)/\partial s^2$

Here, we derive the scaling function for $\partial G(s)/\partial s$ and $\partial^2 G(s)/\partial s^2$.

We first recall the relation between $\partial G(s)/\partial s$ and the expected values of $\rho = n/L$. As already introduced in Sect. 3.2.1, we have

$$
\begin{aligned}
-\frac{\partial G(s)}{\partial s} &= \sum_n \sum_{n'} P^s(n) w(n \to n') = \sum_n P^s(n)\lambda(n) \\
&= \rho(s)(Lc + c - 1) + \rho(s)^2 L(1-2c) + (1-2c)\chi(s).
\end{aligned}
\tag{B.22}
$$

Then, by substituting these $\rho(s)$ and $\chi(s)$ by the one in (3.70) and (3.71), changing the variables to x, and using (3.77) and (3.86), we rewrite it as

$$
\begin{aligned}
&-\frac{\partial G(s)}{\partial s}\Big|_{s=s_c+\kappa^{-1}x} \\
&= \frac{1}{2}\left[\langle\rho\rangle_i + \langle\rho\rangle_a\right](Lc + c - 1) + \frac{1-2c}{2} L\left[\langle\rho^2\rangle_i + \langle\rho^2\rangle_a\right] \\
&\quad + \frac{2x\left[\langle\rho_a\rangle - \langle\rho\rangle_i\right]^{-1}}{\sqrt{1 + 4x^2\left[\langle\rho_a\rangle - \langle\rho\rangle_i\right]^{-2}}}\left\{\frac{1}{2}\left[\langle\rho\rangle_i - \langle\rho\rangle_a\right](Lc + c - 1)\right. \\
&\qquad\qquad\qquad \left. + \frac{1-2c}{2} L\left[\langle\rho^2\rangle_i - \langle\rho^2\rangle_a\right].\right\}
\end{aligned}
$$

Then, by taking $L \to \infty$, we arrive at

$$
-\lim_{L\to\infty} \frac{1}{L}\frac{\partial G(s)}{\partial s}\Big|_{s=s_c+\kappa^{-1}x} = c^2(1-c)\left[1 - \frac{2xc^{-1}}{\sqrt{1+4x^2c^{-2}}}\right]. \tag{B.23}
$$

Since the expression of $\partial G(s)/\partial s$ is determined, we can obtain $\partial^2 G(s)/\partial s^2$ just by taking the derivative of it. By noticing that $\partial^2 G(s)/\partial s^2$ is not directly connected to the equilibrium distribution function $P^s(n)$ from the definition, this property is suggestive. We thus obtain

$$
\lim_{L\to\infty} \frac{1}{L\kappa}\frac{\partial^2 \psi(s)}{\partial s^2}\Big|_{s=s_c+\kappa^{-1}x} = 2c(1-c)\frac{1}{(1+4x^2c^{-2})^{3/2}}. \tag{B.24}
$$

B.4 Variational Principle to Determine the Ground State Energy in Quantum Systems

Here, we show that the variational principle (3.21) that gives the cumulant generating function, when the system satisfies detailed balance condition, is reduced to the one for determining the ground energy in quantum systems. The key is a symmetrisation of the matrix $L^s_{n',n}$ in the largest eigenvalue problem (3.20). Thanks to the detailed balance condition, such a symmetrisation is possible. For symmetric matrices, a variational principle for determining the largest (or lowest) eigenvalue problem is well known in quantum mechanics. We apply this variational principle to the system, and show that this formula and (3.21) is equivalent.

The detailed balance condition is given as (3.7). By using this condition, we rewrite $L^s_{n',n}$ as

$$\begin{aligned} L^s_{n',n} &= w(n \to n')e^{-s} - \delta_{n,n'}\lambda(n) \\ &= P_{\rm eq}(n')w(n' \to n)P_{\rm eq}(n)^{-1}e^{-s} - \delta_{n,n'}\lambda(n) \\ &= P_{\rm eq}(n')^{1/2}\left[P_{\rm eq}(n')^{1/2}w(n' \to n)P_{\rm eq}(n)^{-1/2}e^{-s} - \delta_{n,n'}\lambda(n)\right]P_{\rm eq}(n)^{-1/2} \\ &= P_{\rm eq}(n')^{1/2}\tilde{L}^s_{n',n}P_{\rm eq}(n)^{-1/2}, \end{aligned} \tag{B.25}$$

where we defined

$$\tilde{L}^s_{n',n} = P_{\rm eq}(n')^{1/2}w(n' \to n)P_{\rm eq}(n)^{-1/2}e^{-s} - \delta_{n,n'}\lambda(n). \tag{B.26}$$

Here, we note that the matrix $\tilde{L}^s_{n'n,n}$ is symmetric. Indeed, by using detailed balance condition, we have

$$\begin{aligned} \tilde{L}^s_{n',n} &= P_{\rm eq}(n')^{1/2}w(n' \to n)P_{\rm eq}(n)^{-1/2}e^{-s} - \delta_{n,n'}\lambda(n) \\ &= P_{\rm eq}(n')^{1/2}\left[P_{\rm eq}(n)w(n \to n')P_{\rm eq}(n')^{-1}\right]P_{\rm eq}(n)^{-1/2}e^{-s} - \delta_{n,n'}\lambda(n) \\ &= P_{\rm eq}(n)^{1/2}w(n \to n')P_{\rm eq}(n')^{-1/2}e^{-s} - \delta_{n',n}\lambda(n) \\ &= \tilde{L}^s_{n,n'}. \end{aligned} \tag{B.27}$$

Then, we recall a variational principle for symmetric matrices, which is well-known in quantum physics. By applying it to $\tilde{L}^s_{n',n}$, we have

$$E = \max_{\Phi^0>0} \frac{\sum_{n,n'} \Phi^0(n')\tilde{L}^s_{n',n}\Phi^0(n)}{\sum_n \Phi^0(n)^2}, \tag{B.28}$$

where E is the largest eigenvalue of $\tilde{L}^s_{n,n'}$. In this variational principle, by introducing a variational function $\Delta\widetilde{F}(n)$ by a relation

$$\Phi^0(n)^2 = P_{\rm eq}(n)e^{-\Delta\widetilde{F}(n)}, \tag{B.29}$$

we change the variational parameter from $\Phi^0(n)$ to $\Delta\widetilde{F}(n)$. The variational functional is also rewritten as follows

$$\frac{\sum_{n,n'} \Phi^0(n')\tilde{L}^s_{n',n}\Phi^0(n)}{\sum_n \Phi^0(n)^2} = \frac{\sum_{n,n'} \Phi^0(n')P_{\rm eq}(n')^{-1/2}L^s_{n',n}P_{\rm eq}(n)^{1/2}\Phi^0(n)}{\sum_n \Phi^0(n)^2}$$
$$= \frac{\sum_{n,n'} e^{-\Delta\widetilde{F}(n')/2}L^s_{n',n}P_{\rm eq}(n)e^{-\Delta\widetilde{F}(n)/2}}{\sum_n P_{\rm eq}(n)e^{-\Delta\widetilde{F}(n)}}$$
$$= \sum_n \frac{P_{\rm eq}(n)e^{-\Delta\widetilde{F}(n)}}{\sum_{n'} P_{\rm eq}(n')e^{-\Delta\widetilde{F}(n')}} \sum_{n'} e^{-\Delta\widetilde{F}(n')/2}L^s_{n',n}e^{\Delta\widetilde{F}(n)/2}. \tag{B.30}$$

Thus, by introducing $\widetilde{P}(n)$ as

$$\widetilde{P}(n) = \frac{P_{\rm eq}(n)e^{-\Delta\widetilde{F}(n)}}{\sum_{n'} P_{\rm eq}(n')e^{-\Delta\widetilde{F}(n')}}, \tag{B.31}$$

and by using the explicit expression of the transition rate $w(n \to n')$ in the variational functional, we arrive at

$$\frac{\sum_{n,n'} \Phi^0(n')\tilde{L}^s_{n',n}\Phi^0(n)}{\sum_n \Phi^0(n)^2} = \sum_n \tilde{P}(n)\left[\tilde{\lambda}(n) - \lambda(n)\right], \tag{B.32}$$

where $\tilde{\lambda}(n)$ is defined as (3.23). This functional is exactly the same expression as the variational functional in (3.21). Therefore, by combining it with (B.28) and noticing the fact that the largest eigenvalue of $L^s_{n',n}$ and $\tilde{L}^s_{n',n}$ are the same, we obtain (3.21).

Appendix C
For Chapter 4

C.1 Derivation of (4.64)

Here, we derive (4.64). We consider a joint distribution function of $y(0)$, $y(t)$, $W(t)$ and $Q(t)$ defined as

$$P(y_0, y, W, Q, t|p) = p(y_0)\left\langle \delta(y(t)-y)\delta(W(t)-W)\delta(Q(t)-Q)\right\rangle_{y_0}. \tag{C.1}$$

By using the Langevin equations (4.6), (4.15), and (4.16), we derive the Fokker-Planck equation for $P(y_0, y, W, Q, t|p)$ as

$$\frac{\partial P}{\partial t} = \mathcal{L}_{\mathrm{FP}}^{(y,W,Q)} \cdot P \tag{C.2}$$

with the Fokker-Planck operator $\mathcal{L}_{\mathrm{FP}}^{(y,W,Q)}$ given as

$$\begin{aligned}\mathcal{L}_{\mathrm{FP}}^{(y,W,Q)} \cdot \varphi = &-\frac{\partial}{\partial y}\left[\left(-\frac{1}{\gamma}\frac{\partial}{\partial y}U(y)+v\right)\varphi\right] - v\left(\frac{\partial}{\partial y}U(y)\right)\frac{\partial}{\partial W}\varphi \\ &-\left[\frac{1}{\gamma}\left(\frac{\partial U(y)}{\partial y}\right)^2 - \frac{T}{\gamma}\frac{\partial^2}{\partial y^2}U(y)\right]\frac{\partial}{\partial Q}\varphi + \frac{T}{\gamma}\frac{\partial^2}{\partial y^2}\varphi \\ &+\frac{T}{\gamma}\left(\frac{\partial U(y)}{\partial y}\right)^2\frac{\partial^2}{\partial Q^2}\varphi - \frac{2T}{\gamma}\frac{\partial^2}{\partial Q \partial y}\left[\frac{\partial U(y)}{\partial y}\varphi\right]. \end{aligned} \tag{C.3}$$

We multiply (C.3) by $e^{Wh_w+Qh_q}$ and integrate it with respect W and Q. Finally by noticing the definitions of the biased distribution function P_{h_w,h_q} and q_{h_w,h_q} given as (4.20) and (4.63), we obtain (4.64).

T. Nemoto, *Phenomenological Structure for the Large Deviation Principle in Time-Series Statistics*, Springer Theses,
DOI 10.1007/978-981-287-811-3

C.2 Derivation of (4.77)–(4.79) with the Cole-Hopf Transformation

Here, (4.77)–(4.79) are derived from the largest eigenvalue problems of $\mathcal{L}^{(y)}_{h_w,h_q}$ and $\mathcal{L}^{(y)\dagger}_{h_w,h_q}$. We note that Similar calculations were seen in Refs. [1, 2], for example.

We first consider the largest eigenvalue problem of $\mathcal{L}^{(y)\dagger}_{h_w,h_q}$,

$$\mathcal{L}^{(y)\dagger}_{h_w,h_q} \cdot \phi_0 = \nu_0 \phi_0. \tag{C.4}$$

We divide this equation by ϕ_0 and simplify the obtained expression. The result is

$$\begin{aligned} \nu_0 = (h_q + h_w) v \frac{\partial U}{\partial y} + \frac{T}{\gamma} \left[\frac{\partial}{\partial y} \left(\log \phi_0 - h_q U \right) \right]^2 \\ + \left(-\frac{1}{\gamma} \frac{\partial}{\partial y} U + v \right) \frac{\partial}{\partial y} \left(\log \phi_0 - h_q U \right) + \frac{T}{\gamma} \frac{\partial}{\partial y} \left[\frac{\partial}{\partial y} \left(\log \phi_0 - h_q U \right) \right]. \end{aligned} \tag{C.5}$$

We then introduce a potential function $V_0(y)$ as

$$V_0(y) = -2T \left(\log \phi_0(y) - h_q U(y) \right). \tag{C.6}$$

This means we change the function we consider, $\phi_0(y)$, to the new one $V_0(y)$. This changing (or transformation) is called the Cole-Hopf transformation. We substitute this potential function (C.6) into (C.5), and we combine it with (4.61) and (4.70). This leads to

$$\mathcal{M}_{h_w+h_q,v} \cdot \left(-\frac{\partial V_0}{\partial y} \right) = 2T G_{\text{scaled}}(h_w, h_q), \tag{C.7}$$

Next, the largest eigenvalue problem of $\mathcal{L}^{(y)}_{h_w,h_q}$,

$$\mathcal{L}^{(y)}_{h_w,h_q} \cdot \psi_0 = \mu_0 \psi_0 \tag{C.8}$$

is considered. By dividing this equation by ψ_0 and simplifying it, we obtain

$$\begin{aligned} \mu_0 = -\left(\frac{1}{T} + h_q + h_w \right) (-v) \frac{\partial U}{\partial y} + \frac{T}{\gamma} \left[\frac{\partial}{\partial y} \left(\log \psi_0 + \left(h_q + \frac{1}{T} \right) U \right) \right]^2 \\ + \left(-\frac{1}{\gamma} \frac{\partial U}{\partial y} - v \right) \frac{\partial}{\partial y} \left[\log \psi_0 + \left(h_q + \frac{1}{T} \right) U \right] \\ + \frac{T}{\gamma} \frac{\partial}{\partial y} \left[\frac{\partial}{\partial y} \left(\log \psi_0 + \left(h_q + \frac{1}{T} \right) U \right) \right]. \end{aligned} \tag{C.9}$$

We then define

$$\tilde{V}_0(y) = -2T\left[\log\psi_0(y) + \left(h_q + \frac{1}{T}\right)U(y)\right] \tag{C.10}$$

and substitute it into (C.9). Finally, by combining it with (4.70), we arrive at

$$\mathcal{M}_{-\beta-h_w-h_q,-v}\cdot\left(-\frac{\partial\tilde{V}_0}{\partial y}\right) = 2TG_{\mathrm{scaled}}(h_w, h_q). \tag{C.11}$$

We note that the sign of the velocity in the left-hand side of (C.11) is reversed, which represents a reversed protocol of moving the periodic potential.

From these results (C.6), (C.7), (C.10) and (C.11), we obtain (4.77)–(4.79). Here, we mention that the uniqueness of the solution of the non-linear eigenvalue problem (4.74) is guaranteed from the Perron-Frobenius theory, since (4.74) can be rewritten as the same form as (C.4), by following the (reversed) calculation, from (C.7) to (C.4).

C.3 Derivation of (4.83) and (4.84) with Boundary Layer Analysis

Here, we derive (4.83) and (4.84), by using boundary layer analysis. First, the outer solution of (4.74), $\tilde{w}^{\mathrm{o}}_{h,v}(Y)$, is considered, which satisfies

$$\left|\frac{\partial\tilde{w}^{\mathrm{o}}_{h,v}(Y)}{\partial Y}\right| \approx \left|\tilde{w}^{\mathrm{o}}_{h,v}(Y)\right|, \tag{C.12}$$

where $\approx$ means that the left-hand side and the right-hand side are the same order of magnitude with respect to L. We then solve (4.74) for $\tilde{w}_{h,v}(Y)$, which can be easily done, because this is just a quadratic equation of $\tilde{w}_{h,v}(Y)$. The result is

$$\tilde{w}_{h,v}(Y) = \gamma\left[-v + \frac{1}{\gamma}\frac{\partial U(YL)}{\partial y} \pm \left|v(1+2Th) - \frac{1}{\gamma}\frac{\partial U(YL)}{\partial y}\right|\sqrt{1+R(Y)}\right], \tag{C.13}$$

where $R(Y)$ is defined as

$$R(Y) \equiv \frac{-v^2 4Th(1+Th) + 2K_{h,v}/\gamma}{[v(1+2Th) - (1/\gamma)\partial U(YL)/\partial y]^2} + \frac{-2T/(\gamma^2 L)\partial\tilde{w}_{h,v}/\partial Y}{[v(1+2Th) - (1/\gamma)\partial U(YL)/\partial y]^2}. \tag{C.14}$$

Now, we set $\tilde{w}_{h,v}(Y) = \tilde{w}^{\mathrm{o}}_{h,v}(Y)$ in the right-hand side of this $R(Y)$. Then, according to the following argument, we find that $R(Y)$ is negligible: By using (C.12) and (4.7), we neglect the second term of $R(Y)$. We then assume that $K_{h,v} = O(1)$, where $K_{h,v}$ is equal to $2TG_{\text{scaled}}(h_w, h_q)|_{h_w+h_q=h}$.[1] By using this assumption and (4.7), we also neglect the first term of $R(Y)$. Therefore, we omit $R(Y)$ in $\tilde{w}_{h,v}$. The result leads to an expression of the outer solution $\tilde{w}^{\text{o}}_{h,v}(Y)$ as

$$\tilde{w}^{\text{o}}_{h,v}(Y) = \begin{cases} 2T\gamma hv \\ 2\gamma\left[-v(1+Th) + (1/\gamma)\partial U(YL)/\partial y\right]. \end{cases} \tag{C.15}$$

Finally, we connect these two expressions in $\tilde{w}^{\text{o}}_{h,v}(Y)$. For this purpose, we assume the following things: Firstly, the number of the connecting points is minimized. Secondly, the one of the connecting points is located as the one where the potential has a discontinuity. Especially, in the case of $U = U_{\text{harmo}}$ and $U = U_{\text{quart}}$, the derivatives of the potentials have discontinuities at $Y = \pm 1$.[2] By combining these two assumptions with the normalization condition (4.75), we uniquely determine the solution $\tilde{w}_{h,v}(Y)$ as (4.83) and (4.84).

References

[1] T. Nemoto, S. Sasa, Phys. Rev. E **83**, 030105(R) (2011)
[2] T. Nemoto, S. Sasa, Phys. Rev. E **84**, 061113 (2011)

[1] At least, we can check that this assumption is self-consistent, because we can see that $G_{\text{scaled}} = O(1)$ as shown in (4.37), which is confirmed numerically.

[2] This assumption is also confirmed numerically.

Curriculum Vitae

École Normale Supérieure de Lyon, 15 parvis René Descartes, 69007 Lyon, France
Kyoto University, Yoshidahonmachi, Sakyo Ward, Kyoto, 6068501, JAPAN
e-mail: takahiro.nemoto@ens-lyon.fr

Appointments

- JSPS Research Fellowship for Young Scientists (DC1), (April, 2012–March, 2015)
- Post-doctoral research fellow with Prof. Freddy BOUCHET, École Normale Supérieure de Lyon, (April, 2015 –).

Education

- Doctor of Science (Ph.D.) in Division of Physics and Astronomy, Graduate School of Science, Kyoto University, (2012–2015), supervisor: Shin-ichi Sasa
- Master of Arts (MA) in Group of Condensed Matter theory and Statistical Mechanics (C group), Graduate School of Arts and Sciences, The University of Tokyo, (2010–2012), supervisor: Shin-ichi Sasa
- Bachelor of Science in Department of Physics, Tokyo Institute of Technology (2007–2010).

Research Statement

A realisability of rare-event samplings in real experiments had not been explored yet. If such a sampling method existed, it certainly would lead to many applications in several fields. My work was the first positive result about it. I proposed a new rare-event sampling method to calculate large deviation functions of time-averaged quantities. The method was based on thermodynamics, so that it could potentially be used in real experiments.

Major Honors and Awards

- 2015 Young Scientist Award of the Physical Society of Japan.

T. Nemoto, *Phenomenological Structure for the Large Deviation Principle in Time-Series Statistics*, Springer Theses,
DOI 10.1007/978-981-287-811-3