

ASYMPTOTIC BEHAVIOR OF THE MORAN PARTICLE SYSTEM

SHUI FENG,* *McMaster University*

JIE XIONG,** *University of Macau and University of Tennessee*

Abstract

The asymptotic behavior is studied for an interacting particle system that involves independent motion and random sampling. For a fixed sampling rate, the empirical process of the particle system converges to the Fleming–Viot process when the number of particles approaches ∞ . If the sampling rate approaches 0 as the number of particles becomes large, the corresponding empirical process will converge to the deterministic flow of the motion. In the main results of this paper, we study the corresponding central limit theorems and large deviations. Both the Gaussian limits and the large deviations depend on the sampling scales explicitly.

Keywords: Exponential tightness; Moran particle process; Fleming–Viot process; genealogy; large deviations; particle representation

2010 Mathematics Subject Classification: Primary 60F10

Secondary 92D10

1. Introduction

The Moran model introduced in [16] is a continuous-time birth–death Markov chain used to describe the evolution of biological populations of fixed size. The focus is on the allele frequency in the population and the evolution usually involves three basic forces: the mutation, the natural selection, and the random genetic drift (also known as random sampling or sampling replacement).

Let S be a Polish space. For each $n \geq 1$, denote by $C_b(S^n)$ and $C_b^2(S^n)$ the space of bounded continuous functions and the space of functions with bounded second-order continuous derivatives on S^n , respectively. Let A be the infinitesimal generator of a strong Markov process. Assume that the domain $D(A)$ of the A is dense in $C_b(S)$. For any $\epsilon > 0$, the Moran particle process considered in this paper is a Markov process $(Y_1(t), \dots, Y_n(t))$ with generator

$$L_\epsilon g(y_1, \dots, y_n) = \sum_{i=1}^n A_i g(y_1, \dots, y_n) + \frac{\epsilon}{2} \sum_{i,j=1}^n [\Phi_{ij} g(y_1, \dots, y_n) - g(y_1, \dots, y_n)],$$

where g is in $D(A)$, A_i is A acting on the i th coordinate, and $\Phi_{ij} f$ is obtained from f by replacing y_j with y_i . The number n is the population size. Each $Y_i(t)$ is the type of

Received 26 July 2011; revision received 31 August 2012.

* Postal address: Department of Mathematics and Statistics, McMaster University, Hamilton Hall #211, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada. Email address: shuifeng@univmail.cis.mcmaster.ca

Research supported by the Natural Science and Engineering Research Council of Canada.

** Postal address: Faculty of Science and Technology, University of Macau, Av. Padre Tomás Pereira, Taipa, Macau, China.

Research partially supported by NSF DMS-0906907.

individual i . The A_i s correspond to mutations while the random sampling Φ_{ij} corresponds to the replacement of the type of individual j with the type of individual i . The constant ϵ is called the sampling rate. The natural selection is not included for notational convenience.

Let $M_1(S)$ denote the space of all probability measures on S equipped with the weak topology. Then the empirical distribution (or allele frequency) of the process $(Y_1(t), \dots, Y_n(t))$ is an $M_1(S)$ -valued process. It is well known (cf. Chapter 2 of [3]) that, for large n , if the empirical initial distribution of the Moran particle system converges to μ in $M_1(S)$, then the empirical process of $(Y_1(t), \dots, Y_n(t))$ converges (in the sense of finite-dimensional distributions) to the probability-valued process introduced by Fleming and Viot in [12], which is known as the Fleming–Viot process.

Define $\mathcal{D} = \{F : F(\mu) = f(\langle \mu, \phi \rangle), f \in C_b^\infty(R), \phi \in D(A), \mu \in M_1(S)\}$. The generator of the Fleming–Viot process with mutation operator A and sampling rate ϵ has the form

$$\mathcal{L}_\epsilon F(\mu) = \int_S \left(A \frac{\delta F(\mu)}{\delta \mu(x)} \right) \mu(dx) + \frac{\epsilon}{2} \int_S \int_S \left(\frac{\delta^2 F(\mu)}{\delta \mu(x) \delta \mu(y)} \right) Q(\mu; dx, dy), \tag{1.1}$$

where

$$\begin{aligned} \frac{\delta F(\mu)}{\delta \mu(x)} &= \lim_{\gamma \rightarrow 0^+} \frac{F(\mu + \gamma \delta_x) - F(\mu)}{\gamma}, \\ \frac{\delta^2 F(\mu)}{\delta \mu(x) \delta \mu(y)} &= \lim_{\gamma_1 \rightarrow 0^+, \gamma_2 \rightarrow 0^+} \frac{F(\mu + \gamma_1 \delta_x + \gamma_2 \delta_y) - F(\mu)}{\gamma_1 \gamma_2}, \\ Q(\mu; dx, dy) &= \mu(dx) \delta_x(dy) - \mu(dx) \mu(dy), \end{aligned}$$

and δ_x stands for the Dirac measure at $x \in S$. The domain of \mathcal{L}_ϵ is \mathcal{D} . The space S is called the type space and the last term in (1.1) corresponds to the random sampling. It follows from direct calculation that, for $F(\mu) = f(\langle \mu, \phi \rangle)$,

$$\frac{\delta F(\mu)}{\delta \mu(x)} = f'(\langle \mu, \phi \rangle) \phi(x), \quad \frac{\delta^2 F(\mu)}{\delta \mu(x) \delta \mu(y)} = f''(\langle \mu, \phi \rangle) \phi(x) \phi(y).$$

The class of the Fleming–Viot processes is one of the two basic classes of measure-valued diffusions. The Fleming–Viot processes presented here are only a small portion of the class. More general models include forces such as natural selection, recombination, and interaction. Interested readers should consult [10] for a comprehensive survey on the subject.

The term Fleming–Viot process is also used in the literature for some finite particle systems. In [1], [2], and [13], finite particle systems are studied where each particle moves independently of the others until one particle hits a certain boundary or a catalyst or, more generally, is killed and a new birth occurs at a location uniformly chosen among the remaining particles. This sampling mechanism clearly resembles that of the Moran particle system or the Fleming–Viot process. These particle systems are closely related to the Moran particle system introduced above. A common new feature of the models in [1], [2], and [13] is that the sampling mechanism strongly depends on the motion or mutation. This strong interaction creates a lot of new challenges that do not present in our model.

Let us return to the Fleming–Viot process in (1.1). When the sampling rate ϵ goes to 0, the Fleming–Viot process converges to the solution of the forward equation associated with A . It is natural to consider limiting theorems, such as the central limit theorem and the Freidlin–Wentzell-type large deviations associated with this limiting procedure. These are useful in providing information on the stability of the mutation process under small sampling perturbations. The establishment of large deviations turns out to be a very challenging problem

due to the degenerate nature of the sampling mechanism. There have been many efforts in dealing with the problem. To the best of our knowledge, it is still an open problem except in the case when the mutation generator has the form

$$Af(s) = c \int_S (f(u) - f(s))\nu(du), \tag{1.2}$$

where $c \geq 0$ and ν is in $M_1(S)$ (cf. [4], [5], [11], and [18]).

Motivated by this problem, we consider the particle system $(Y_1(t), \dots, Y_n(t))$ when the sampling rate $\epsilon = \beta(n)$ such that $\lim_{n \rightarrow \infty} \beta(n) = 0$. The scale of $\beta(n)$ is divided into the following three cases.

Case A: $\lim_{n \rightarrow \infty} n\beta(n) = 0$.

Case B: $\lim_{n \rightarrow \infty} n\beta(n) = c \in (0, +\infty)$.

Case C: $\lim_{n \rightarrow \infty} n\beta(n) = +\infty$.

As expected, for large n , the empirical process of $(Y_1(t), \dots, Y_n(t))$ converges to the flow generated by A in all cases. But our main result shows that a critical phenomenon appears in both the central limit theorem and the large deviations depending on the scale of $\beta(n)$. Central limit theorems hold in all scales, but the Gaussian processes have different covariances. Large deviation principles with speed $1/n$ are considered in all three cases. The large deviation behavior in case A is shown to be equivalent to the independent particle system. In case C the large deviation behavior is trivial in a sense (see Section 3) at the speed $1/n$. In case B, we obtain the exponential tightness and the upper bound. The large deviation lower bound for this case is still open.

We conclude the introduction with an outline of this article. Central limit theorems are established in Section 2. The martingale methods are used in the proof. Three main factors contribute to the large deviations, namely, the noise in the initial distribution, the noise in the independent motion, and the random sampling. We call the first two averaging factors. In case A, the averaging factors are stronger than the sampling factor. In case B, all three factors act at the same scale. In case C, the averaging factors are weaker than the sampling factor. In Section 3 we first establish large deviation principles in case A. The proof is based on a detailed analysis of the genealogical structure of the particle system. The large deviation results in case B are obtained by first establishing the exponential tightness and then establishing the local estimates. A comparison is made between the Moran system in case A and the Fleming–Viot process in Section 4, revealing that the Freidlin–Wentzell-type large deviation for the Fleming–Viot process corresponds to the large deviations of the Moran particle model in case B.

2. Central limit theorem

Recall that in the Moran particle system $(Y_1(t), \dots, Y_n(t))$ each particle moves independently of all other particles according to A until an exponential clock with parameter $\binom{n}{2}\beta(n)$ rings. When the clock rings, an unordered pair $(Y_i(t), Y_j(t))$ of particles is chosen at random and the particle at position $Y_j(t)$ jumps to the position $Y_i(t)$. Before the clock rings again, each particle moves independently of all other particles according to A .

The empirical process of the particle system is defined as

$$\eta_n(t) = \frac{1}{n} \sum_{i=1}^n \delta_{Y_i(t)}.$$

We assume that $(Y_1(0), \dots, Y_n(0))$ are independent with identical distribution μ in $M_1(S)$. Thus,

$$\eta_n(0) \Rightarrow \mu.$$

For fixed ϵ , it was shown in [8] that $\eta_n(t)$ has the same distribution as the empirical process of a particle approximation of the Fleming–Viot process. Thus, for large n , $\eta_n(t)$ converges to the Fleming–Viot process starting at μ .

From now on, we assume that $\epsilon = \beta(n)$ and $\lim_{n \rightarrow \infty} \beta(n) = 0$.

Let $D([0, 1], M_1(S))$ be the space of all functions from $[0, 1]$ to $M_1(S)$ that are right continuous and have left limits (left continuous at 1) furnished with the Skorokhod topology. Let T_t be the semigroup associated with generator A . Then it follows (cf. [3]) from the law of large numbers that

$$\lim_{n \rightarrow \infty} \eta_n(t) = \mu T_t = \mu(t).$$

Let

$$\alpha(n) = \min\{\sqrt{n}, \beta^{-1}(n)\}$$

and

$$\xi^n(t) \equiv \alpha(n)(\eta_n(t) - \mu(t)).$$

The objective of this section is to establish a functional central limit theorem for the sequence of processes $\{\xi^n(\cdot)\}$. Based on the work of Mitoma [15], we need to find a nuclear Fréchet space E so that the process $\xi^n(\cdot)$ can be viewed as an element of the dual space of E . In this section we choose $E = \mathcal{S}$, the space of Schwartz functions on \mathbb{R}^d , even though the result can be relaxed. Since our main focus here is on the critical behavior of the particle system, the loss due to the restriction on space \mathcal{S} is minor.

Let \mathcal{S}' denote the dual space of \mathcal{S} . For $\kappa = 0, 1, 2, \dots$, define

$$\|f\|_\kappa^2 = \sum_{0 \leq |k| \leq \kappa} \int_{\mathbb{R}^d} (1 + |x|^2)^\kappa |D^k f(x)|^2 dx.$$

Let \mathcal{S}_κ be the completion of \mathcal{S} with respect to $\|\cdot\|_\kappa$. Let $\mathcal{S}_{-\kappa}$ be the dual space of \mathcal{S}_κ .

The process $\xi^n(\cdot)$ now belongs to the space $D([0, 1], \mathcal{S}')$ in which the central limit theorem will hold. To obtain the tightness in $D([0, 1], \mathcal{S}')$, we first consider the martingale problem of $\eta_n(t)$ and assume the following.

Assumption I. *The space $\mathcal{S} \subset D(A)$ is a core of A and there exist two indices, $\kappa' > \kappa$ and a constant K , such that, for any $f \in \mathcal{S}$, we have $Af \in \mathcal{S}$, $\|Af\|_{0,\infty} \leq K\|f\|_\kappa$, and $\|Af\|_\kappa \leq K\|f\|_{\kappa'}$, where $\|\cdot\|_{0,\infty}$ denotes the supremum norm.*

Assumption II. *The martingale problem associated with A is well posed.*

Examples that satisfy these assumptions include the Brownian motion, diffusions with bounded continuous coefficient, the random walk, and other Markov chains with bounded finite range jumps.

For each f in \mathcal{S} ,

$$N_t^n(f) \equiv \langle \eta_n(t), f \rangle - \langle \eta_n(0), f \rangle - \int_0^t \langle \eta_n(s), Af \rangle ds$$

is a martingale with quadratic variation

$$\begin{aligned} \langle N^n(f) \rangle_t &= \frac{1}{n} \int_0^t \langle \eta_n(s), A(f^2) - 2fA(f) \rangle ds \\ &\quad + 2\beta(n) \int_0^t (\langle \eta_n(s), f^2 \rangle - \langle \eta_n(s), f \rangle^2) ds. \end{aligned} \tag{2.1}$$

Lemma 2.1. *Under Assumptions I and II, the sequence $\{\xi^n(\cdot)\}$ is tight in $D([0, 1], \mathcal{S}')$.*

Proof. For each f in \mathcal{S} , set

$$\tilde{N}_t^n(f) \equiv \langle \xi^n(t), f \rangle - \langle \xi^n(0), f \rangle - \int_0^t \langle \xi^n(s), Af \rangle ds.$$

It follows from (2.1) that $\tilde{N}_t^n(f)$ is a martingale with quadratic variation process

$$\begin{aligned} \langle \tilde{N}^n(f) \rangle_t &= \frac{\alpha(n)^2}{n} \int_0^t \langle \eta_n(s), A(f^2) - 2fAf \rangle ds \\ &\quad + 2\alpha(n)^2\beta(n) \int_0^t (\langle \eta_n(s), f^2 \rangle - \langle \eta_n(s), f \rangle^2) ds. \end{aligned}$$

By convolution we obtain

$$\langle \xi^n(t), f \rangle = \langle \xi^n(0), T_t f \rangle + \int_0^t d\langle \tilde{N}^n(T_{t-s} f) \rangle_s,$$

where $\{T_t\}$ is the semigroup on $C_b(S)$ of the Feller Markov process with generator A . Then

$$\begin{aligned} \mathbb{E}(\langle \xi^n(t), f \rangle^2) &\leq K_1 \|T_t f\|_{0,\infty}^2 + K_1 \mathbb{E} \int_0^t \frac{\alpha(n)^2}{n} \langle \eta_n(s), A(f^2) - 2fAf \rangle ds \\ &\quad + c_1 \mathbb{E} \int_0^t 2\alpha(n)^2\beta(n) (\langle \eta_n(s), f^2 \rangle - \langle \eta_n(s), f \rangle^2) ds \end{aligned}$$

for $t \leq T$. Without loss of generality, we assume that $\kappa \geq 1$. Note that

$$\begin{aligned} \|T_t f\|_{0,\infty} &\leq \|f\|_{0,\infty} \leq K \|f\|_1 \leq K \|f\|_\kappa, \\ \|fAf\|_{0,\infty} &\leq \|f\|_{0,\infty} \|Af\|_{0,\infty} \leq K \|f\|_1 \|f\|_\kappa \leq K \|f\|_\kappa^2, \end{aligned}$$

and

$$\begin{aligned} \|A(f^2)\|_{0,\infty}^2 &\leq K \|f^2\|_\kappa^2 \\ &= K \sum_{0 \leq |k| \leq \kappa} \int_{\mathbb{R}^d} (1 + |x|^2)^\kappa |D^k f^2(x)|^2 dx \\ &= K \sum_{0 \leq |k| \leq \kappa} \int_{\mathbb{R}^d} (1 + |x|^2)^\kappa \left| \sum_{0 \leq \ell \leq k} \binom{k}{\ell} D^\ell f(x) D^{k-\ell} f(x) \right|^2 dx \\ &\leq K \|f\|_{0,\infty} \|f\|_\kappa \\ &\leq K \|f\|_\kappa^2. \end{aligned}$$

These estimates imply that

$$\mathbb{E}(\langle \xi^n(t), f \rangle^2) \leq K \|f\|_{\kappa}^2.$$

Then, for $u \leq \epsilon$,

$$\mathbb{E}(\langle \xi^n(t+u) - \xi^n(t), f \rangle^2 \mid \mathcal{F}_t) \leq \mathbb{E}(\zeta_f^{1,n}(\epsilon) + \zeta_f^{2,n}(\epsilon)),$$

where

$$\zeta_f^{1,n}(\epsilon) = \int_t^{t+\epsilon} \langle \xi^n(s), Af \rangle^2 ds$$

and

$$\begin{aligned} \zeta_f^{2,n}(\epsilon) &= \frac{\alpha(n)^2}{n} \int_t^{t+\epsilon} \langle \eta_n(s), A(f^2) - 2fAf \rangle ds \\ &\quad + 2\alpha(n)^2\beta(n) \int_t^{t+\epsilon} (\langle \eta_n(s), f^2 \rangle - \langle \eta_n(s), f \rangle^2) ds. \end{aligned}$$

Note that

$$\mathbb{E}\zeta_f^{1,n}(\epsilon) = \mathbb{E} \int_t^{t+\epsilon} \langle \xi^n(s), Af \rangle^2 ds \leq K\epsilon \|Af\|_{\kappa}^2 \leq K\epsilon \|f\|_{\kappa'}^2.$$

Similarly to the above, we have

$$\mathbb{E}\zeta_f^{2,n}(\epsilon) \leq K\epsilon \|f\|_{\kappa}^2.$$

We now have

$$\limsup_{\epsilon \rightarrow 0} \sup_n \mathbb{E}(\zeta_f^{1,n}(\epsilon) + \zeta_f^{2,n}(\epsilon)) = 0.$$

By Remark 8.7 of [9, p. 138], we obtain the tightness of $\langle \xi^n, f \rangle$ in $D([0, T], \mathbb{R})$ for every f .

Applying Doob’s inequality, we have

$$\mathbb{E} \sup_{t \leq T} \langle \xi^n(s) - \xi^n(0), f \rangle^2 \leq 8\mathbb{E} \int_0^T (\zeta_f^{1,n}(s) + T\zeta_f^{2,n}(s)) ds \leq K \|f\|_{\kappa'}^2.$$

Let $\kappa'' > \kappa'$ be such that the embedding from $\mathfrak{H}_{\kappa''}$ to $\mathfrak{H}_{\kappa'}$ is Hilbert–Schmidt. Let $\kappa_1 > \kappa''$ be such that the embedding from \mathfrak{H}_{κ_1} to $\mathfrak{H}_{\kappa''}$ is compact. Then $\{\xi^n(\cdot)\}$ satisfies the compact containment condition on $D([0, T], \mathfrak{H}_{-\kappa_1})$. It now follows from Theorem 4.1 of [15] that $\{\xi^n(\cdot)\}$ is tight in $D([0, 1], \mathfrak{H}')$. This completes the proof.

Let $\xi(\cdot)$ be a limit point. We now characterize $\xi(\cdot)$. To this end, we first prove the continuity of $\xi(\cdot)$ by adapting the proof of Theorem 8.1.1 of [14].

Lemma 2.2. *Under Assumptions I and II, the limit point $\xi(\cdot)$ belongs to the space $C([0, T], \mathfrak{H}')$ with probability 1.*

Proof. Let g be a nonnegative continuous function on \mathbb{R} , vanishing in a neighborhood of 0 and ∞ . Note that

$$\sum_{s \leq t} g(\Delta \langle \xi^n(s), f \rangle) = \int_0^t \int_U g\left(\frac{\alpha(n)}{n} (f(Y_i(s-)) - f(Y_j(s-)))\right) N^n(ds, d(i, j)),$$

where N^n is a Poisson random measure on $\mathbb{R}_+ \times U$ with intensity measure given by $\beta(n)$, multiplying the counting measure on

$$U = \{(i, j) : 1 \leq i \neq j \leq n\}.$$

Thus,

$$\begin{aligned} \mathbb{E} \sum_{s \leq t} g(\Delta \langle \xi^n(s), f \rangle) &= \beta(n) \sum_{1 \leq i \neq j \leq n} \mathbb{E} \int_0^t g\left(\frac{\alpha(n)}{n}(f(Y_i(s)) - f(Y_j(s)))\right) ds \\ &\leq \beta(n) \sum_{1 \leq i \neq j \leq n} \mathbb{E} \int_0^t \mathbf{1}\left\{\frac{\alpha(n)}{n}(f(Y_i(s)) - f(Y_j(s))) > b\right\} ds \\ &\leq \beta(n) \sum_{1 \leq i \neq j \leq n} b^{-p} \mathbb{E} \int_0^t \left|\frac{\alpha(n)}{n}(f(Y_i(s)) - f(Y_j(s)))\right|^p ds \\ &\leq Kn^{1-p/2} \\ &\rightarrow 0 \end{aligned}$$

by taking $p > 2$, where $b > 0$ such that $g(x) = 0$ when $|x| \leq b$. By Fatou’s lemma we obtain

$$\mathbb{E} \sum_{s \leq t} g(\Delta \langle \xi_s, f \rangle) = 0.$$

Let

$$\tilde{N}_t(f) \equiv \langle \xi_t, f \rangle - \langle \xi_0, f \rangle - \int_0^t \langle \xi_s, Af \rangle ds.$$

From above, we see that $\{\langle \xi^n(s), Af \rangle : s \in [0, T], n \geq 1\}$ is uniformly integrable. It is easy to show that $\{\langle \tilde{N}^n(f) \rangle_t : t \in [0, T], n \geq 1\}$ is uniformly integrable, and $(\tilde{N}^n(f), \langle \tilde{N}^n(f) \rangle)$ converges to $(\tilde{N}(f), \langle \tilde{N}(f) \rangle)$. Thus, $\tilde{N}_t(f)$ is a continuous martingale with quadratic variation process

$$\langle \tilde{N}(f) \rangle_t = \begin{cases} \int_0^t \langle \mu_s, A(f^2) - 2fAf \rangle ds & \text{if } \lim_{n \rightarrow \infty} n\beta(n) = 0, \\ 2 \int_0^t (\langle \mu_s, f^2 \rangle - \langle \mu_s, f \rangle^2) ds & \text{if } \lim_{n \rightarrow \infty} n\beta(n) = \infty, \end{cases}$$

and, if

$$\alpha(n) = \sqrt{n} \quad \text{and} \quad \lim_{n \rightarrow \infty} n\beta(n) = c,$$

then

$$\langle \tilde{N}(f) \rangle_t = \int_0^t (\langle \mu_s, A(f^2) - 2fAf \rangle + 2c(\langle \mu_s, f^2 \rangle - \langle \mu_s, f \rangle^2)) ds.$$

Theorem 2.1. Assume that $\beta(n)$ satisfies case A, B, or C. Then, under Assumptions I and II, the sequence $\{\xi^n(\cdot)\}$ converges to $\xi(\cdot)$ in distribution.

Proof. We only need to prove that ξ is the unique solution to the linear evolution equation

$$\langle \xi(t), f \rangle = \langle \xi(0), f \rangle + \int_0^t \langle \xi(s), Af \rangle ds + \tilde{N}_t(f) \quad \text{for all } f \in \mathcal{F}. \tag{2.2}$$

Let z_t be the difference between two solutions of (2.2). Then

$$\langle z_t, f \rangle = \int_0^t \langle z_s, Af \rangle ds.$$

This implies that $z_t = 0$.

3. Large deviations

In this section we study the large deviations for the empirical process of the Moran particle system in case A and case B with speed $1/n$. The initial distribution of $(Y_1(0), \dots, Y_n(0))$ is assumed to be independent and identically distributed (i.i.d.) with common distribution μ . A large deviation principle is established first in case A. For case B, we are able to establish the exponential tightness and the upper bound. We omit case C due to the following observation.

For appropriate function f ,

$$\exp \left\{ \langle \eta_n(t), f \rangle - \langle \eta_n(0), f \rangle - \int_0^t \mathcal{A}_n f(\eta_n(s)) ds \right\}$$

is an exponential martingale for

$$\mathcal{A}_n f(\eta_n) = n \langle \eta_n, e^{-f/n} A e^{f/n} \rangle + \frac{n^2 \beta(n)}{2} [\langle \eta_n, e^{f/n} \rangle \langle \eta_n, e^{-f/n} \rangle - 1].$$

In case C, it follows from a direct calculation that $\beta(n) \mathcal{A}_n f / \beta(n)$ has a limit for a large class of operators A when n tends to ∞ . This indicates that the large deviations in case C occur at the scale of $\beta(n)$, and aside from the speed differences it will be similar to case B. Hence, the local large deviations with speed $1/n$ in case C will produce a trivial rate function which is 0 in the effective domain of the large deviation rate function with speed $\beta(n)$.

3.1. Case A

For convenience, we will assume that the path of the motion process generated by A is continuous. Let $C([0, 1], S)$ denote the space of continuous functions from $[0, 1]$ to S equipped with the uniform convergence topology.

First, we consider the Moran particle system with sampling rate 0. This is simply an independent particle system. Let $X(t)$ be the Feller process generated by A with initial distribution μ . For each $n \geq 1$ and $1 \leq i \leq n$, let $X_1(t), \dots, X_n(t)$ be independent copies of $X(t)$. Define

$$\zeta_n(t) = \frac{1}{n} \sum_{i=1}^n \delta_{X_i(t)}.$$

Clearly, $\zeta_n(\cdot)$ is in the space $C([0, T], M_1(S))$. For each $s \geq 0$ and x in S , let $\mathbb{P}_{s,x}$ denote the law of the process generated by A starting from x at time s . For any $t \geq s \geq 0$ and f in $C_b(S)$, set

$$U_{s,t} f(x) = \int f(X(t)) d\mathbb{P}_{s,x},$$

and, for any $s \leq t_1 < t_2$ and any $\nu(t_1), \nu(t_2)$ in $M_1(S)$,

$$L(\nu(t_1), \nu(t_2)) = \sup_{f \in C_b(S)} \{ \langle \nu_{t_1}, f \rangle - \langle \mu, \log U_{t_1,t_2} e^f \rangle \}.$$

For any ν, μ in $M_1(S)$, let

$$H(\nu \mid \mu) = \sup\{\langle \nu, g \rangle - \log\langle \mu, e^g \rangle : g \in C_b(S)\}$$

denote the relative entropy of ν with respect to μ .

Theorem 3.1. *Denote the law of $\zeta_n(\cdot)$ by \mathcal{Q}_n . Then the family $\{\mathcal{Q}_n : n = 1, \dots\}$ satisfies a large deviation principle on the space $C([0, 1], M_1(S))$ with rate function*

$$I_\mu(\nu(\cdot)) = H(\nu(0) \mid \mu) + \sup\left\{ \sum_{i=1}^m L(\nu(t_{i-1}), \nu(t_i)) : 0 = t_0 \leq t_1 \leq \dots \leq t_m = 1, m = 1, \dots \right\}. \tag{3.1}$$

Proof. For any $m \geq 1$ and $0 = t_0 < t_1 < \dots < t_m = 1$, we denote by $\pi^{(n)}(t_0, t_1, \dots, t_m)$ the joint distribution of $(\zeta_n(t_0), \dots, \zeta_n(t_m))$ on $M_1(S)^{\otimes(m+1)}$. It follows from Theorem 3.5 of [6] that, conditional on $\zeta_n(0)$ near a probability $\nu(0)$, the family $\{\pi^n(t_0, t_1, \dots, t_m) : n = 1, \dots\}$ satisfies a large deviation principle on the space $M_1(S)^{\otimes(m+1)}$ with speed $1/n$ and rate function

$$\tilde{L}(\nu(0), \nu(t_1), \dots, \nu(t_m)) = \sup_{g_0, \dots, g_m \in C_b(S)} \left[\sum_{i=0}^m \langle \nu(t_i), g_i \rangle - H(g_0, \dots, g_m) \right],$$

where

$$H(g_0, \dots, g_m) = \left\langle \nu(0), \log \mathbb{E}^{\mathbb{P}^{0, X(0)}} \left(\exp \left\{ \sum_{i=0}^m g_i(X(t_i)) \right\} \right) \right\rangle.$$

By the Markov property of $X(t)$, $H(g_0, \dots, g_m)$ can be written as

$$\langle \nu(0), \log \mathbb{E}^{\mathbb{P}^{0, X(0)}} (\exp\{g_0(X(t_0)) + h(g_1, \dots, g_m)(X(t_0))\}) \rangle,$$

where

$$h(g_1, \dots, g_m)(X(t_0)) = \log \mathbb{E}^{\mathbb{P}^{t_0, X(t_0)}} \left(\exp \left\{ \sum_{i=1}^m g_i(X(t_i)) \right\} \right).$$

Since $h(g_1, \dots, g_m) \in C_b(S)$, we conclude that

$$\begin{aligned} \tilde{L}(\nu(0), \nu(t_1), \dots, \nu(t_m)) &= \sup_{g_0, \dots, g_m \in C_b(S)} \left[\sum_{i=1}^m \langle \nu(t_i), g_i \rangle - \langle \nu(t_0), h(g_1, \dots, g_m) \rangle \right] \\ &= \sup_{g_1, \dots, g_m \in C_b(S)} \sum_{i=1}^m [\langle \nu(t_i), g_i \rangle - \langle \nu(t_0), h(g_1, \dots, g_m) \rangle]. \end{aligned}$$

It follows by induction that

$$\tilde{L}(\nu(0), \nu(t_1), \dots, \nu(t_m)) = \sum_{i=1}^m L(\nu(t_{i-1}), \nu(t_i)).$$

Since the initial deviations are controlled by the relative entropy, it follows that the family $\{\pi^n(t_0, t_1, \dots, t_m) : n = 1, \dots\}$ satisfies a large deviation principle on the space $M_1(S)^{\otimes(m+1)}$ with speed $1/n$ and rate function

$$H(\nu(0) \mid \mu) + \tilde{L}(\nu(0), \nu(t_1), \dots, \nu(t_m)).$$

The rest of the proof follows by the same argument as used in the proof of Lemma 4.7 of [6], namely a combination of the contraction principle and the projective limit technique. The path continuity of $X(t)$ guarantees that the projection map

$$\pi : M_1(C([0, 1], S)) \rightarrow C([0, 1], M_1(S))$$

is continuous. Thus, the contraction principle applies.

Remark. Several such examples can be found in [6] where the rate function in (3.1) has an explicit expression.

We now turn to the large deviations for the interacting particle system $(Y_1(t), \dots, Y_n(t))$ under the assumption that

$$\lim_{n \rightarrow \infty} n\beta(n) = 0.$$

Let $\sigma_k, k = 1, \dots, \binom{n}{2}$, be i.i.d. exponential random variables with parameter $\beta(n)$. Set

$$\tau = \inf \left\{ \sigma_k : k = 1, \dots, \binom{n}{2} \right\}.$$

Then τ is exponential with parameter $\binom{n}{2}\beta(n)$, and $\binom{n}{2}\beta(n)\tau$ is thus exponential with parameter 1. It follows from the Markov property that the waiting time for the first jump of the process $(Y_1(t), \dots, Y_n(t))$ is distributed as τ . Let τ_1, τ_2, \dots be i.i.d. copies of τ . The waiting time for the k th jump is thus given by $\sum_{i=1}^k \tau_i$. Since each jump results in the removal of one existing motion, the n particles after the first jump will no longer be independent. Particles that share a segment of path in the past are grouped together. Let $k_n(t)$ denote the number of groups that have more than one member at time t . For each $j = 1, \dots, k_n(t)$, let n_j denote the corresponding number of members. Set $r_n(t)$ to be the number of jumps of the process $(Y_1(t), \dots, Y_n(t))$ during the time period $[0, t]$. Then, clearly, $k_n(t) \leq r_n(t)$ and

$$\sum_{j=1}^{k_n(t)} n_j = r_n(t) + k_n(t).$$

For any $\kappa(n)$ converging to ∞ as n becomes large, it follows from the relation between the gamma distribution and the summation of i.i.d. exponential random variables that

$$\begin{aligned} \mathbb{P} \left\{ \sum_{k=1}^{\kappa(n)} \tau_i \leq t \right\} &= \mathbb{P} \left\{ \sum_{k=1}^{\kappa(n)} \binom{n}{2} \beta(n) \tau_i \leq \binom{n}{2} \beta(n) t \right\} \\ &= \int_0^{\binom{n}{2} \beta(n) t} \frac{1}{\Gamma(\kappa(n))} x^{\kappa(n)-1} e^{-x} dx \\ &\leq \frac{(\binom{n}{2} \beta(n) t)^{\kappa(n)}}{\Gamma(\kappa(n))}. \end{aligned} \tag{3.2}$$

Lemma 3.1. For some $0 < a < 1$, set

$$\gamma(n) = \beta(n) \vee (n \log n)^{-1} \quad \text{and} \quad \kappa(n) = \frac{n}{(-\log n \gamma(n))^a}.$$

Then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P} \{ r_n(t) \geq \kappa(n) \} = -\infty.$$

Proof. It is clear that $\kappa(n) = n/(-\log n\gamma(n))^a$ converges to ∞ as n becomes large. It follows from (3.2) that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\{r_n(t) \geq \kappa(n)\} \\ & \leq \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\left\{\sum_{k=1}^{\kappa(n)} \tau_i \leq t\right\} \\ & \leq \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(\frac{\binom{n}{2} \beta(n)t}{\kappa(n)}\right)^{\kappa(n)} \\ & \leq \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(\frac{\binom{n}{2} \gamma(n)t}{\kappa(n)}\right)^{\kappa(n)} \\ & = \lim_{n \rightarrow \infty} \frac{1}{(-\log n\gamma(n))^a} \left\{ \log \left[\frac{t}{2} (-\log n\gamma(n))^a \right] - (-\log(n-1)\gamma(n)) \right\} \\ & = -\infty. \end{aligned} \tag{3.3}$$

Consider a dense subset $\{f_i : i = 1, \dots\}$ of $C_b(S)$, and let $\{g_i : g_i = f_i/(|f_i| \vee 1)\}$. For any μ and ν in $M_1(S)$, define

$$\rho(\mu, \nu) = \sum_{i=1}^{\infty} \frac{|\langle \mu - \nu, g_i \rangle|}{2^i}. \tag{3.4}$$

Then ρ is a metric on $M_1(S)$ and generates the weak topology.

For any $\mu(\cdot)$ and $\nu(\cdot)$ in $D([0, T], M_1(S))$, set

$$d(\mu(\cdot), \nu(\cdot)) = \sup_{0 \leq t \leq 1} \rho(\mu(t), \nu(t))$$

and

$$d_s(\mu(\cdot), \nu(\cdot)) = \inf_{\phi \in \Phi} \left\{ d(\mu(\cdot), \nu(\phi(\cdot))) + \sup_{s \neq t, s, t \in [0, 1]} \left| \log \frac{\phi(s) - \phi(t)}{s - t} \right| \right\},$$

where Φ is the set of all strictly increasing continuous maps from $[0, 1]$ to $[0, 1]$. Then the Skorokhod topology on $D([0, 1], M_1(S))$ is generated by the metric d_s . Clearly,

$$d_s(\mu(\cdot), \nu(\cdot)) \leq d(\mu(\cdot), \nu(\cdot)).$$

Theorem 3.2. *For each $\delta > 0$, we can construct the processes $\eta_n(t)$ and $\zeta_n(t)$ on the same probability space such that*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \mathbb{P}\{d_s(\eta_n(\cdot), \zeta_n(\cdot)) \geq \delta\} = -\infty. \tag{3.5}$$

Using the terminology in [7], the process $\eta_n(\cdot)$ is exponentially equivalent to the process $\zeta_n(\cdot)$.

Proof. For each $n \geq 1$, let $(X_1(t), \dots, X_n(t))$ be n independent processes generated by A with respective starting points $X_i(0)$, $i = 1, \dots, n$, that are i.i.d. with common distribution. Construct the process $(Y_1(t), \dots, Y_n(t))$ as follows. Before the first ringing of the exponential clock, $(Y_1(t), \dots, Y_n(t))$ is the same as $(X_1(t), \dots, X_n(t))$. The waiting times of the clock are independent of $(X_1(t), \dots, X_n(t))$. Now both $\eta_n(t)$ and $\zeta_n(t)$ live in the same space.

For each fixed $\delta > 0$, choose M large enough so that $2^{-M} < \delta/2$. It then follows from the definitions of the metrics d_s and d that

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\{d_s(\eta_n(\cdot), \zeta_n(\cdot)) \geq \delta\} \\ & \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\{d(\eta_n(\cdot), \zeta_n(\cdot)) \geq \delta\} \\ & \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\left\{ \max_{1 \leq i \leq M} \left\{ \sup_{0 \leq t \leq 1} |\langle \eta_n(t) - \zeta_n(t), g_i \rangle| \right\} \geq \frac{\delta}{2M} \right\} \\ & \leq \max_{1 \leq i \leq M} \left\{ \limsup_{n \rightarrow \infty} \frac{1}{n} \mathbb{P}\left\{ \sup_{0 \leq t \leq 1} |\langle \eta_n(t) - \zeta_n(t), g_i \rangle| \geq \frac{\delta}{2M} \right\} \right\}. \end{aligned} \tag{3.6}$$

Applying (3.3), we obtain, for each i between 1 and M ,

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\left\{ \sup_{0 \leq t \leq 1} |\langle \eta_n(t) - \zeta_n(t), g_i \rangle| \geq \frac{\delta}{2M} \right\} \\ & \leq \max \left\{ \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\left\{ \sup_{0 \leq t \leq 1} |\langle \eta_n(t) - \zeta_n(t), g_i \rangle| \geq \frac{\delta}{2M}, r_n(1) \leq \kappa(n) \right\}, \right. \\ & \quad \left. \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\{r_n(1) > \kappa(n)\} \right\} \\ & = \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\left\{ \sup_{0 \leq t \leq 1} |\langle \eta_n(t) - \zeta_n(t), g_i \rangle| \geq \frac{\delta}{2M}, r_n(1) \leq \kappa(n) \right\}. \end{aligned} \tag{3.7}$$

On the set $\{r_n(1) \leq \kappa(n)\}$,

$$\sup_{0 \leq t \leq 1} |\langle \eta_n(t) - \zeta_n(t), g_i \rangle| \leq \frac{2(r_n(1) + k_n(1))}{n},$$

which combined with (3.7) implies (3.5) and the theorem.

Following Theorem 4.2.13 of [7], we obtain the following result.

Theorem 3.3. *Denote the law of $\eta_n(t)$ by \mathcal{P}_n . Then the family $\{\mathcal{P}_n : n = 1, \dots\}$ satisfies a large deviation principle on the space $D([0, 1], M_1(S))$ with speed $1/n$ and rate function $I_\mu(\mu(\cdot))$ defined in (3.1).*

Remark. It follows from the proof of Theorem 3.2 that the large deviation result still holds on $D([0, 1], M_1(S))$ when equipped with the uniform convergence topology.

3.2. Case B

In this subsection, the space S is assumed to be compact. This restriction guarantees the compact containment condition. To generalize to noncompact state spaces, one would need the existence of a Lyapunov function.

Let

$$D_e(A) = \{g \in C(S) : g, e^g \in D(A)\}.$$

We apply the following assumption to the motion generator A .

Assumption III. $D_e(A)$ is dense in $C(S)$ and, for any g in $D_e(A)$, there exists a positive constant C depending only on g such that, for all $\lambda \geq 0$,

$$\langle v, e^{-\lambda g} A e^{\lambda g} \rangle + \frac{1}{2}(\langle v, e^{\lambda g} \rangle \langle v, e^{-\lambda g} \rangle - 1) \leq C(e^{2\lambda|g|} - 1) \tag{3.8}$$

holds uniformly in v , where $|g| = \max\{|g(x)| : x \in S\}$.

Many generators satisfy this assumption. In particular, this includes the generator defined in (1.2), and the Laplacian operator.

Without loss of generality, we choose $\beta(n) = 1/n$. All the results below hold under Assumptions I, II, and III.

For each $n \geq 1$ and g in $D_e(A)$, set

$$Z(\mu(\cdot); t, n, g) = \exp \left\{ n \left[\langle \mu(t), g \rangle - \langle \mu(0), g \rangle - \int_0^t \langle \mu(s), e^{-s} A e^s \rangle ds - \frac{1}{2} \int_0^t (\langle \mu(s), e^s \rangle \langle \mu(s), e^{-s} \rangle - 1) ds \right] \right\}.$$

Then, by direct calculation, we have the following result.

Lemma 3.2. *The process $Z(\eta_n(\cdot); t, n, g)$ is a \mathcal{P}_n -martingale.*

Lemma 3.3. *For each g in $D_e(A)$, and any $\gamma > 0$, there is a compact subset $K(g, \gamma)$ of $D([0, 1], R)$ such that, for all n and any $\eta_n(0)$,*

$$\mathbb{P}_{\eta_n(0)} \{ \langle \eta_n(\cdot), g \rangle \in K(g, \gamma)^c \} \leq -e^{n\gamma}. \tag{3.9}$$

Proof. For any $\delta > 0$, set

$$W'_{\langle \mu(\cdot), g \rangle}(\delta) = \inf_{\{t_j\}} \max_j \sup_{t_{j-1} \leq t, s < t_j} | \langle \mu(t), g \rangle - \langle \mu(s), g \rangle |,$$

where the infimum is taken over the finite points $\{t_j\}$ satisfying

$$0 = t_0 < t_1 < \dots < t_m = 1, \quad t_j - t_{j-1} > \delta, \quad j = 1, \dots, m.$$

It follows from the Markov property that, for any $\varrho > 0$ and $\delta \in (0, \frac{1}{4}]$,

$$\begin{aligned} & \sup_{\eta_n(0)} \mathbb{P}_{\eta_n(0)} \{ W'_{\langle \eta_n(\cdot), g \rangle}(\delta) > \varrho \} \\ & \leq \sum_{j=0}^{\lceil 1/2\delta \rceil - 1} \sup_{\eta_n(0)} \mathbb{P}_{\eta_n(0)} \left\{ \sup_{t, s \in [2j\delta, 2(j+1)\delta]} | \langle \eta_n(t), g \rangle - \langle \eta_n(s), g \rangle | > \varrho \right\} \\ & \quad + \sup_{\eta_n(0)} \mathbb{P}_{\eta_n(0)} \left\{ \sup_{t, s \in [2\delta, 1]} | \langle \eta_n(t), g \rangle - \langle \eta_n(s), g \rangle | > \varrho \right\} \\ & \leq \sum_{j=0}^{\lceil 1/2\delta \rceil - 1} \sup_{\eta_n(0)} \mathbb{P}_{\eta_n(0)} \left\{ \sup_{t \in [2j\delta, 2(j+1)\delta]} | \langle \eta_n(t), g \rangle - \langle \eta_n(2j\delta), g \rangle | > \frac{\varrho}{2} \right\} \\ & \quad + \sup_{\eta_n(0)} \mathbb{P}_{\eta_n(0)} \left\{ \sup_{t \in [2\delta, 1]} \left| \langle \eta_n(t), g \rangle - \left\langle \eta_n \left(2\delta \left[\frac{1}{2\delta} \right] \right), g \right\rangle \right| > \frac{\varrho}{2} \right\} \end{aligned}$$

$$\begin{aligned} &\leq \frac{1}{2\delta} \sup_{\eta_n(0)} \mathbb{P}_{\eta_n(0)} \left\{ \sup_{t \in [0, 2\delta)} |\langle \eta_n(t), g \rangle - \langle \eta_n(0), g \rangle| > \frac{\varrho}{2} \right\} \\ &\leq \frac{1}{\delta} \sup_{\eta_n(0)} \mathbb{P}_{\eta_n(0)} \left\{ \sup_{t \in [0, 2\delta)} (\langle \eta_n(t), g \rangle - \langle \eta_n(0), g \rangle) > \frac{\varrho}{2} \right\}, \end{aligned}$$

where the control of the term

$$\sup_{\eta_n(0)} \mathbb{P}_{\eta_n(0)} \left\{ \sup_{t \in [2\delta[1/2\delta], 1]} \left| \langle \eta_n(t), g \rangle - \left\langle \eta_n \left(2\delta \left[\frac{1}{2\delta} \right] \right), g \right\rangle \right| > \frac{\varrho}{2} \right\}$$

is justified by the fact that $1 - 2\delta[1/2\delta] < 2\delta$.

This combined with assumption (3.8) and Lemma 3.2 implies that

$$\begin{aligned} &\mathbb{P}_{\eta_n(0)} \left\{ \sup_{t \in [0, 2\delta)} (\langle \eta_n(t), g \rangle - \langle \eta_n(0), g \rangle) > \frac{\varrho}{2} \right\} \\ &\leq \mathbb{P}_{\eta_n(0)} \left\{ \sup_{t \in [0, 2\delta)} Z(\eta_n(\cdot); t, n, \lambda g) > \beta \left(\frac{\varrho}{2} \right) - 2\delta C \frac{e^{2|g|\lambda} - 1}{\lambda} \right\} \\ &\leq \exp \left\{ -n \left[\frac{\beta\varrho}{2} - 2C\delta(e^{2|g|\beta} - 1) \right] \right\}. \end{aligned} \tag{3.10}$$

For $\varrho \geq 4C\delta|g|$, by minimizing the right-hand side of (3.10) with respect to $\lambda > 0$, we obtain

$$\begin{aligned} &\mathbb{P}_{\eta_n(0)} \left\{ \sup_{t \in [0, 2\delta)} (\langle \eta_n(t), g \rangle - \langle \eta_n(0), g \rangle) > \frac{\varrho}{2} \right\} \\ &\leq \exp \left\{ -n \frac{1}{2|g|} \left[\log \frac{\varrho}{4\delta C|g|} + (4\delta C|g| - \varrho) \right] \right\}. \end{aligned}$$

If we choose $\varrho_k = 1/k$ and $\delta_k = e^{-2|g|(k+1)}/4|g|Ck$, then we have

$$\sup_{\eta_n(0)} \mathbb{P}_{\eta_n(0)} \{ W'_{\langle \eta_n(\cdot), g \rangle}(\delta_k) > \varrho_k \} \leq \exp\{-nk\}.$$

For any $\gamma > 0$, set

$$\tilde{K}(g, \gamma) = \bigcap_{k > \gamma+1} \left\{ x(\cdot) \in D([0, 1], R) : \sup_{t \in [0, 1]} |x(t)| \leq 1, W'_{x(\cdot)}(\delta_k) \leq \varrho_k \right\}.$$

Clearly, the closure $K(g, \gamma)$ of $\tilde{K}(g, \gamma)$ is compact in $D([0, 1], R)$ and

$$\begin{aligned} \mathbb{P}_{\eta_n(0)} \{ \langle \eta_n(\cdot), g \rangle \in K(g, \gamma)^c \} &\leq \mathbb{P}_{\eta_n(0)} \{ \langle \eta_n(\cdot), g \rangle \in \tilde{K}(g, \gamma)^c \} \\ &\leq \sum_{k > \gamma+1} e^{-kn} \\ &\leq \frac{e^{-n}}{1 - e^{-n}} e^{-n\gamma}, \end{aligned}$$

which implies (3.9).

Lemma 3.4. *The family $\{\mathcal{P}_n : n = 1, 2, \dots\}$ is exponentially tight, i.e. for any $M > 0$, there exists a compact subset \mathcal{K}_M of $D([0, 1], M_1(S))$ such that*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathcal{P}_n\{\mathcal{K}_M^c\} \leq -M.$$

Proof. Let $\{g_i : i = 1, \dots, m\}$ be defined as in (3.4). For each $m \geq 1$, let $t_j = 1$ or -1 , $j = 1, \dots, m$. Set

$$h_m^{t_1, \dots, t_m}(s) = \sum_{j=1}^m \frac{t_j}{2^j} g_j(s).$$

For each $M > 0$, let

$$\tilde{\mathcal{K}}_M = \bigcap_{m=1}^{\infty} \bigcap_{l=mM+1}^{mM+2^m} \bigcap_{t_1, \dots, t_m} \{\mu(\cdot) \in D([0, 1], M_1(S)) : \langle \mu(\cdot), h_m^{t_1, \dots, t_m} \rangle \in K(h_m^{t_1, \dots, t_m}, l)\}.$$

Next we show that $\tilde{\mathcal{K}}_M$ is relatively compact. Since the space $M_1(S)$ is compact, it suffices to verify that

$$\lim_{\delta \rightarrow 0} \sup_{\mu(\cdot) \in \tilde{\mathcal{K}}_M} W'_{\mu(\cdot)}(\delta) = 0,$$

where

$$W'_{\mu(\cdot)}(\delta) = \inf_{\{t_i\}} \max_i \sup_{t_{i-1} \leq t, s < t_i} \rho(\mu(t), \mu(s)).$$

For any $\epsilon > 0$, choose m large enough so that $1/2^m \leq \epsilon/2$. Then we have

$$\begin{aligned} W'_{\mu(\cdot)}(\delta) &\leq \inf_{\{t_i\}} \max_i \sup_{t_{i-1} \leq t, s < t_i} \sum_{j=1}^m \frac{1}{2^j} |\langle \mu(t) - \mu(s), g_j \rangle| + \frac{\epsilon}{2} \\ &\leq \sum_{t_1, \dots, t_m} W'_{\langle \mu(\cdot), h_m^{t_1, \dots, t_m} \rangle}(\delta) + \frac{\epsilon}{2}. \end{aligned}$$

Based on the construction, we have, for each t_1, \dots, t_m ,

$$\lim_{\delta \rightarrow 0} \sup_{\mu(\cdot) \in \tilde{\mathcal{K}}_M} W'_{\langle \mu(\cdot), h_m^{t_1, \dots, t_m} \rangle}(\delta) = 0,$$

which implies that there exists $\delta_m > 0$ such that

$$\sup_{\mu(\cdot) \in \tilde{\mathcal{K}}_M} \max\{W'_{\langle \mu(\cdot), h_m^{t_1, \dots, t_m} \rangle}(\delta_m) : t_1, \dots, t_m\} \leq \frac{\epsilon}{2^{m+1}}$$

and

$$\sup_{\mu(\cdot) \in \tilde{\mathcal{K}}_M} W'_{\mu(\cdot)}(\delta_m) \leq \epsilon.$$

Thus, $\tilde{\mathcal{K}}_M$ is relatively compact in $D([0, 1], M_1(S))$.

Let \mathcal{K}_M be the closure of $\tilde{\mathcal{K}}_M$. Then \mathcal{K}_M is compact and

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\{\mathcal{K}_M^c\} \leq \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}\{\tilde{\mathcal{K}}_M^c\} \leq \lim_{n \rightarrow \infty} \frac{1}{n} \log \sum_{m=1}^{\infty} \sum_{l=mM+1}^{mM+2^m} e^{-nl} \leq e^{-M}.$$

Recall that $Y_1(0), \dots, Y_n(0)$ are independent with common distribution μ . For any ν in $M_1(S)$, g in $D_e(A)$, and $\mu(\cdot)$ in $D([0, 1], M_1(S))$, define

$$\begin{aligned}
 I_\mu(\nu, g) &= \langle \nu, g \rangle - \log \langle \mu, e^g \rangle, \\
 L(\mu(\cdot), g) &= \langle \mu(1), g \rangle - \langle \mu(0), g \rangle - \int_0^1 \langle \mu(s), e^{-g} A e^g \rangle ds \\
 &\quad - \frac{1}{2} \int_0^1 (\langle \mu(s), e^g \rangle \langle \mu(s), e^{-g} \rangle - 1) ds,
 \end{aligned}$$

and

$$\begin{aligned}
 H(\mu, \mu(\cdot), g) &= I_\mu(\mu(0), g) + L(\mu(\cdot), g), \\
 I_\mu(\mu(\cdot)) &= \sup\{H(\mu, \mu(\cdot), g) : g \in D_e(A)\}.
 \end{aligned}$$

Lemma 3.5. For any $\mu(\cdot)$ in $D([0, 1], M_1(S))$ and $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathcal{P}\{B(\mu(\cdot), \delta)\} \leq -I_\mu(\mu(\cdot)),$$

where $B(\mu(\cdot), \delta) = \{d_s(\nu(\cdot), \mu(\cdot)) < \delta\}$.

Proof. It follows by direct calculation that, for any g in $D_e(A)$,

$$\mathbb{E}(\exp\{n(I_\mu(\eta_n(0), g))\}) = 1.$$

This combined with Lemma 3.2 implies that

$$\mathbb{E}(\exp\{nH(\mu, \eta_n(\cdot), g)\}) = 1. \tag{3.11}$$

Since the functional $H(\mu, \nu(\cdot), g)$ is continuous in $\nu(\cdot)$, for any $\epsilon > 0$ and any $\mu(\cdot)$ in $D([0, 1], M_1(S))$, there exists $\delta > 0$ such that

$$|H(\mu, \nu(\cdot), g) - H(\mu, \mu(\cdot), g)| < \epsilon \quad \text{on } B(\mu(\cdot), \delta).$$

This combined with Chebyshev’s inequality and (3.11) implies that

$$\begin{aligned}
 \mathcal{P}_n\{B(\mu(\cdot), \delta)\} &\leq \frac{\mathbb{E}(\exp\{nH(\mu, \eta_n(\cdot), g)\})}{\inf_{\{\nu(\cdot) \in D([0, 1], M_1(S))\}} \exp\{nH(\mu, \nu(\cdot), g)\}} \\
 &\leq \exp\{-n(H(\mu), \mu(\cdot), g) - \epsilon)\}.
 \end{aligned}$$

The lemma follows from the fact that g is arbitrary.

Theorem 3.4. For any closed subset F of $D([0, 1], M_1(S))$,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathcal{P}_n\{F\} \leq - \inf_{\mu(\cdot) \in D([0, 1], M_1(S))} I_\mu(\mu(\cdot)). \tag{3.12}$$

Proof. The local estimate obtained in Lemma 3.5 guarantees that (3.12) holds for compact subset F . This combined with exponential tightness implies (3.12).

4. Comparison with the Fleming–Viot process

If the sampling rate $\beta(n) \equiv \gamma$ is a constant in the Moran particle system then the empirical process $\eta_n(t)$ is known to converge to the Fleming–Viot process with sampling rate γ when n goes to ∞ . Thus, the results considered in Section 3 are related to the asymptotic behavior of the Fleming–Viot process when the sampling rate approaches 0. In case A, the large deviation principle in Theorem 3.3 has speed $1/n$, while $\beta(n) = o(1/n)$. It is thus reasonable to expect a large deviation principle for the Fleming–Viot process with speed ϵ when the sampling rate is of the form $\epsilon b(\epsilon)$ with $b(\epsilon)$ converging to 0 as ϵ goes to 0. Next we show that the large deviation principle thus obtained has a trivial rate function which is essentially different from the particle model.

In the remainder of this section, S is assumed to be a compact set. Let the sampling rate be $\epsilon b(\epsilon)$, and let the corresponding Fleming–Viot process starting at μ be denoted by $\mu^\epsilon(t)$. The law of $\mu^\epsilon(t)$ is denoted by \mathcal{P}_μ^ϵ . The next theorem establishes the exponential tightness of the family $\{\mathcal{P}_\mu^\epsilon : \epsilon > 0\}$.

Theorem 4.1. *For any $M > 0$, there exists a compact subset \mathcal{K} of $C([0, 1], M_1(S))$ such that*

$$\limsup_{\epsilon \rightarrow 0} \epsilon b(\epsilon) \log \mathcal{P}_\mu^\epsilon(\mathcal{K}^c) \leq -M. \tag{4.1}$$

Proof. Since $\sup_{\epsilon > 0, t \in [0, T]} |\langle \mu^\epsilon(t), g \rangle|$ is uniformly bounded for each g in $\mathcal{D}(A)$, by [4] Schied’s criterion for exponential tightness, to prove the exponential tightness of the family $\{\langle \mu^\epsilon(\cdot), g \rangle\}$, it suffices to verify that there exist constants $\alpha > 0$ and $\beta > 0$ such that

$$\sup_{[s, t] \subset [0, T]} \mathbb{E} \left(\exp \left\{ \frac{\alpha}{\epsilon \sqrt{t-s}} |\langle \mu^\epsilon(t), g \rangle - \langle \mu^\epsilon(s), g \rangle| \right\} \right) \leq \beta^{1/\epsilon b(\epsilon)}. \tag{4.2}$$

For any constant θ , we have

$$\begin{aligned} & \mathbb{E}(\exp\{\theta|\langle \mu^\epsilon(t), g \rangle - \langle \mu^\epsilon(s), g \rangle|\}) \\ &= \mathbb{E}(\exp\{\theta(\langle \mu^\epsilon(t), g \rangle - \langle \mu^\epsilon(s), g \rangle)\} : \langle \mu^\epsilon(t), g \rangle - \langle \mu^\epsilon(s), g \rangle \geq 0) \\ & \quad + \mathbb{E}(\exp\{-\theta(\langle \mu^\epsilon(t), g \rangle - \langle \mu^\epsilon(s), g \rangle)\} : \langle \mu^\epsilon(t), g \rangle - \langle \mu^\epsilon(s), g \rangle < 0) \\ & \leq [\mathbb{E}(\exp\{2\theta(\langle \mu^\epsilon(t), g \rangle - \langle \mu^\epsilon(s), g \rangle)\})]^{1/2} \\ & \quad + [\mathbb{E}(\exp\{-2\theta(\langle \mu^\epsilon(t), g \rangle - \langle \mu^\epsilon(s), g \rangle)\})]^{1/2}. \end{aligned}$$

Define

$$\begin{aligned} M^\theta(s, t; g) &= \exp \left\{ \theta \left[\langle \mu^\epsilon(t), g \rangle - \langle \mu^\epsilon(s), g \rangle - \int_s^t \langle \mu^\epsilon(u), g \rangle du \right] \right. \\ & \quad \left. - \frac{\epsilon b(\epsilon) \theta^2}{2} \int_s^t \text{var}_{\mu^\epsilon(u)}(g) du \right\}. \end{aligned}$$

Then $M^\theta(s, t; g)$ is a \mathcal{P}_μ^ϵ local martingale, and thus a supermartingale. By direct calculation,

$$\begin{aligned} \mathbb{E}(\exp\{\theta|\langle \mu^\epsilon(t), g \rangle - \langle \mu^\epsilon(s), g \rangle|\}) &\leq \mathbb{E}(M^\theta(s, t; g)) \exp \left\{ (t-s) \left(c_1 \theta + c_2 \frac{\epsilon b(\epsilon) \theta^2}{2} \right) \right\} \\ &\leq \exp \left\{ (t-s) \left(c_1 |\theta| + c_2 \frac{\epsilon b(\epsilon) \theta^2}{2} \right) \right\}, \end{aligned}$$

where $c_1 = \sup_x |Ag(x)|$ and $c_2 = 2 \sup_x g^2(x)$. Choosing $\theta = \alpha/\epsilon b(\epsilon)\sqrt{t-s}$ and $\theta = -\alpha/\epsilon b(\epsilon)\sqrt{t-s}$, we respectively obtain

$$\begin{aligned} \mathbb{E}\left(\exp\left\{\frac{\alpha}{\epsilon b(\epsilon)\sqrt{t-s}}(\langle \mu^\epsilon(t), g \rangle - \langle \mu^\epsilon(s), g \rangle)\right\}\right) &\leq C^{1/\epsilon b(\epsilon)}, \\ \mathbb{E}\left(\exp\left\{-\frac{\alpha}{\epsilon b(\epsilon)\sqrt{t-s}}(\langle \mu^\epsilon(t), g \rangle - \langle \mu^\epsilon(s), g \rangle)\right\}\right) &\leq C^{1/\epsilon b(\epsilon)}, \end{aligned}$$

where $C = \exp\{c_1\alpha\sqrt{T} + \alpha^2c_2/2\}$. Now, by choosing $\beta = 2\sqrt{C}$, we obtain (4.2).

Finally, for any $M > 0$, and a countable dense subset $\{g_n\}$ of $\mathcal{D}(A)$, there exist compact sets K_{g_n} in $C([0, T], S)$ such that

$$\mathbb{P}\{\langle \xi^\epsilon, g_n \rangle \in K_{g_n}^c\} \leq \exp\left\{-M \frac{n}{\epsilon b(\epsilon)}\right\}.$$

Now set $\mathcal{K} = \{\mu(\cdot) \in C([0, T], M_1(S)) : \langle \mu(\cdot), g_n \rangle \in K_{g_n}\}$. Then it is clear that \mathcal{K} is compact in $C([0, T], M_1(S))$ and

$$\limsup_{\epsilon \rightarrow 0} \epsilon b(\epsilon) \log \mathcal{P}_\mu^\epsilon\{\mu^\epsilon(\cdot) \in \mathcal{K}^c\} \leq -M.$$

By a result in [17], every subsequence of \mathcal{P}_μ^ϵ has a subsequence that satisfies a large deviation principle with certain rate function. Focusing on a particular subsequence that has rate function $J(x)$, it follows that, for every $\mu(\cdot)$ in $C([0, T], M_1(E))$,

$$\begin{aligned} \lim_{\delta \rightarrow 0} \liminf_{\epsilon \rightarrow 0} \epsilon b(\epsilon) \log \mathcal{P}_\mu^\epsilon\{d(\mu(\cdot), \nu(\cdot)) < \delta\} &= \lim_{\delta \rightarrow 0} \limsup_{\epsilon \rightarrow 0} \epsilon b(\epsilon) \log \mathcal{P}_\mu^\epsilon\{d(\mu(\cdot), \nu(\cdot)) \leq \delta\} \\ &= -J(\mu(\cdot)), \end{aligned}$$

which implies that, for every $\mu(\cdot)$ satisfying $J(\mu(\cdot)) > 0$,

$$\lim_{\delta \rightarrow 0} \liminf_{\epsilon \rightarrow 0} \epsilon \log \mathcal{P}_\mu^\epsilon\{d(\mu(\cdot), \nu(\cdot)) < \delta\} = \lim_{\delta \rightarrow 0} \limsup_{\epsilon \rightarrow 0} \epsilon \log \mathcal{P}_\mu^\epsilon\{d(\mu(\cdot), \nu(\cdot)) \leq \delta\} = -\infty.$$

It follows from (4.1) that

$$\limsup_{\epsilon \rightarrow 0} \epsilon \log \mathcal{P}_\mu^\epsilon\{\mu^\epsilon(\cdot) \in \mathcal{K}^c\} \leq -\infty.$$

It is known (cf. [18]) that a full large deviation principle holds for the Fleming–Viot process with zero motion ($A = 0$). The calculation above shows that in this case large deviations hold for \mathcal{P}_μ^ϵ with speed ϵ and a rate function taking values 0 or ∞ .

Acknowledgement

We would like to thank an anonymous referee for very helpful comments and suggestions.

References

[1] ASSELAH, A., FERRARI, P. A. AND GROISMAN, P. (2011). Quasistationary distributions and Fleming–Viot processes in finite spaces. *J. Appl. Prob.* **48**, 322–332.
 [2] BURDZY, K., HOLYST, R. AND MARCH, P. (2000). A Fleming–Viot particle representation of the Dirichlet Laplacian. *Commun. Math. Phys.* **214**, 679–703.

- [3] DAWSON, D. A. (1993). Measure-valued Markov processes. In *Ecole d'Été de Probabilités de Saint Flour XXI* (Lecture Notes Math. **1541**), Springer, Berlin, pp. 1–260.
- [4] DAWSON, D. A. AND FENG, S. (1998). Large deviations for the Fleming-Viot process with neutral mutation and selection. *Stoch. Process. Appl.* **77**, 207–232.
- [5] DAWSON, D. A. AND FENG, S. (2001). Large deviations for the Fleming-Viot process with neutral mutation and selection. II. *Stoch. Process. Appl.* **92**, 131–162.
- [6] DAWSON, D. A. AND GÄRTNER, J. (1987). Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics* **20**, 247–308.
- [7] DEMBO, A. AND ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications* (Appl. Math. **38**), 2nd edn. Springer, New York.
- [8] DONNELLY, P. AND KURTZ, T. G. (1996). A countable representation of the Fleming-Viot measure-valued diffusion. *Ann. Prob.* **24**, 698–742.
- [9] ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes*. John Wiley, New York.
- [10] ETHIER, S. N. AND KURTZ, T. G. (1993). Fleming–Viot processes in population genetics. *SIAM J. Control Optimization* **31**, 345–386.
- [11] FENG, S. AND XIONG, J. (2002). Large deviations and quasi-potential of a Fleming-Viot process. *Electron. Commun. Prob.* **7**, 13–25.
- [12] FLEMING, W. H. AND VIOT, M. (1979). Some measure-valued Markov processes in population genetics theory. *Indiana Univ. Math. J.* **28**, 817–843.
- [13] GRIGORESCU, I. (2007). Large deviations for a catalytic Fleming-Viot branching system. *Commun. Pure Appl. Math.* **60**, 1056–1080.
- [14] KALLIANPUR, G. AND XIONG, J. (1995). *Stochastic Differential Equations on Infinite-Dimensional Spaces* (IMS Lecture Notes Monogr. Ser. **26**). Institute of Mathematical Statistics, Hayward, CA.
- [15] MITOMA, I. (1983). Tightness of probabilities on $C([0, 1]; \mathcal{S}')$ and $D([0, 1]; \mathcal{S}')$. *Ann. Prob.* **11**, 989–999.
- [16] MORAN, P. A. P. (1958). Random processes in genetics. *Proc. Camb. Phil. Soc.* **54**, 60–71.
- [17] PUHALSKII, A. (1991). On functional principle of large deviations. In *New Trends in Probability and Statistics*, Vol. 1, eds V. Sazonov and T. Shervashidze, VSP, Utrecht, pp. 198–218.
- [18] XIANG, K.-N. AND ZHANG, T.-S. (2005). Small time asymptotics for Fleming-Viot processes. *Infin. Dimens. Anal. Quantum Prob. Relat. Top.* **8**, 605–630.