Elect. Comm. in Probab. **7** (2002) 13–25

ELECTRONIC COMMUNICATIONS in PROBABILITY

LARGE DEVIATIONS AND QUASI-POTENTIAL OF A FLEMING-VIOT PROCESS

SHUI FENG[1]
Department of Mathematics and Statistics, McMaster University, Hamilton, ONT, Canada L8S 4K1
email: `shuifeng@mcmail.cis.mcmaster.ca`

JIE XIONG[2]
Department of Mathematics, University of Tennessee, Knoxville, TN 37996-1300
email: `jxiong@math.utk.edu`

submitted December 13, 2001 *Final version accepted* January 7, 2002

AMS 2000 Subject classification: Primary 60F10; secondary 92D10
Fleming-Viot process, large deviations, quasi-potential

Abstract

The large deviation principle is established for the Fleming-Viot process with neutral mutation when the process starts from a point on the boundary. Since the diffusion coefficient is degenerate on the boundary, the boundary behavior of the process is investigated in detail. This leads to the explicit identification of the rate function, the quasi-potential, and the structure of the effective domain of the rate function.

1 Introduction

Let $E = [0,1]$, and $M_1(E)$ be the set of all probability measures on E equipped with the weak topology. $C(E)$ is the set of all continuous functions on E. The set $C_b^2(R)$ contains all functions on the real line R that possess bounded continuous derivatives upto second order. For any $\theta > 0$ and ν_0 in $M_1(E)$, define

$$Af(x) = \frac{\theta}{2}\int (f(y) - f(x))\nu_0(dy), \quad f \in C(E).$$

Then A is clearly the generator of a Markov jump process on E. Let

$$\mathcal{D} = \{F(\mu) = f(\langle \mu, \varphi\rangle) : f \in C_b^2(R), \varphi \in C(E), \mu \in M_1(E)\}.$$

For any $\epsilon > 0$, consider the following operator on set $\mathcal{D}$:

$$\mathcal{L}^\epsilon F(\mu) = f'(\langle \mu, \varphi\rangle)\langle \mu, A\varphi\rangle + \frac{\epsilon}{2}\int\int f''(\langle \mu, \varphi\rangle)\varphi(x)\varphi(y)Q(\mu; dx, dy), \tag{1.1}$$

[1]RESEARCH SUPPORTED BY THE NATURAL SCIENCE AND ENGINEERING RESEARCH COUNCIL OF CANADA
[2]RESEARCH SUPPORTED PARTIALLY BY NSA

where $Q(\mu; dx, dy) = \mu(dx)\delta_x(dy) - \mu(dx)\mu(dy)$ and δ_x stands for the Dirac measure at $x \in E$. The measure-valued process generated by $\mathcal{L}^\epsilon$ is called the Fleming-Viot process with neutral mutation. It models the evolution of the distribution of the genotypes in a population under the influence of mutation and replacement sampling. The set E is called the type space or the space of alleles, the operator A describes the generation independent mutation process with mutation rate θ and mutation measure ν_0, and the last term describes the continuous sampling with sampling rate ϵ.

When A is replaced by the generator of any other Feller process, the resulting process is still called a Fleming-Viot process. In fact in the original Fleming-Viot process, A is the Laplacian operator. In the sequel, we will use the term Fleming-Viot (henceforth, FV) process solely for the neutral mutation case.

Let $C([0,+\infty), M_1(E))$ be the set of all continuous maps from $[0,+\infty)$ to $M_1(E)$. For any $\mu(\cdot)$ in $C([0,+\infty), M_1(E))$, any fixed $n \geq 2$, and $0 = x_0 < x_1, \cdots < x_n = 1$, let $X_k(t) = \mu(t)([x_{k-1}, x_k)), p_k = \nu_0([x_{k-1}, x_k))$ for $1 \leq k \leq n-1$ and $X_n(t) = \mu(t)([x_{n-1}, x_n]), p_n = \nu_0([x_{n-1}, x_n])$. It is well-known (cf. Ethier and Kurtz [4]) that under the law of the FV process, $X(t) = (X_1(t), \cdots, X_{n-1}(t))$ solves the following degenerate stochastic differential equation:

$$dX_k(t) = \frac{\theta}{2}(p_k - X_k(t))dt + \sqrt{\epsilon}\sum_{j=1}^{n-1} \sigma_{kj}(X(t))dB_j(t), 1 \leq k \leq n-1,$$

where for any $1 \leq i, j \leq n-1$, $B_j(t)$ are independent Brownian motions, and

$$\sum_{k=1}^{n-1} \sigma_{ik}(x)\sigma_{kj}(x) = x_i(\delta_{ij} - x_j).$$

Because of this partition property, it becomes natural to study the FV process through its finite dimensional analogue.

The first objective of this article is to establish the large deviation principle at the path level for the FV process when the sampling rate ϵ approaches 0. This problem has been studied in [1] and [2]. The only unresolved case is the lower bound when the mutation measure ν_0 is not absolutely continuous with respect to the initial point μ. In the finite dimensional case, this corresponds to the case when the process starts from a point on the boundary of the simplex $\{(x_1, \cdots, x_{n-1}) : \sum_{k=1}^{n-1} x_k = 1, x_k \geq 0, k = 1, \cdots, k\}$. The difficulty in resolving the issue comes from the degeneracy and non-Lipschitz property of $\sigma(x)$. Our proof of this result reveals that the derivatives of all possible large deviation paths at time zero have to be the same as the derivative of the solution of the limiting dynamic. This is very different from the non-degenerate case. An explicit expression of the rate function is also obtained. All definitions and terminology of large deviations follow those in [3].

Our second objective is to identify the quasi-potential of the process. Consider the following infinite dimensional dynamic:

$$\frac{d\mu(t)}{dt} = \frac{\theta}{2}(\nu_0 - \mu(t)), \mu(0) = \nu. \tag{1.2}$$

The solution of the dynamic is attracted to ν_0 as t goes to infinity. The FV process can be viewed as a random perturbation of this dynamic. It is thus natural to study the transition from ν_0 to another point ν in $M_1(E)$ under the perturbation. The quasi-potential $V(\nu_0; \nu)$, if exists, is the minimal energy needed for the transition from ν_0 to ν. It is known that the FV process has a unique reversible measure Π_ϵ which satisfies a full large deviation principle ([2]). Roughly speaking for suitable subset A of $M_1(E)$, we have

$$\Pi_\epsilon\{A\} = \exp\{-\frac{1}{\epsilon}\inf_{\nu\in A} I(\nu)(1+\circ(1))\},$$

where $I(\nu) = \theta H(\nu_0|\nu)$ and $H(\nu_0|\nu)$ is the relative entropy of ν_0 with respect to ν. If we set $F(\nu) = H(\nu_0|\nu)$, then $I(\nu) = \theta(F(\nu) - F(\nu_0))$. In physical terms, $F(\nu)$ is the free energy functional and θ plays the role of reciprocal temperature. Our second result shows that the quasi-potential equals to $I(\nu)$ and the minimal energy is attained by a time-reversed path of (1.2) connecting ν_0 with ν.
Large deviation lower bounds are obtained in Section 2, and a detailed description of the effective domain is given in Section 3. Result on quasi-potential is in Section 4.

2 Large Deviation at Path Level

For any fixed $T > 0$, let $C([0,T], M_1(E))$ be the space of all continuous $M_1(E)$-valued functions on $[0,T]$ equipped with the uniform convergence topology. In this section, we prove the large deviation lower bound for FV process at the path level, i.e. on space $C([0,T], M_1(E))$, under the assumption that the mutation measure ν_0 is not absolutely continuous with respect to the starting point μ of the process. Because of the partition property, we obtain the result through the study of the finite dimensional case.

2.1 Finite Dimension

For any fixed $n \geq 1$, let

$$\mathcal{S}_n = \left\{ x = (x_1, \cdots, x_{n-1}) : \ x_i \geq 0, \ \sum_{i=1}^{n-1} x_i \leq 1 \right\}.$$

For $1 \leq k \leq n-1$, consider

$$dX_k^\epsilon(t) = c(p_k - X_k^\epsilon(t))dt + \sqrt{\epsilon}\sum_{j=1}^{n-1}\sigma_{kj}(X^\epsilon(t))dB_j(t), \ X^\epsilon(0) = x,$$

where $c = \frac{\theta}{2}$, and $(\sigma_{kj}(x))$ is such that

$$\sigma(x)\sigma(x)' = (x_k(\delta_{kj} - x_j))_{1\leq k,j\leq n-1}.$$

Denote $p_n = 1 - \sum_{j=1}^{n-1} p_j$ and $x_n = 1 - \sum_{j=1}^{n-1} x_j$. Let

$$\mathcal{L}_p = \{x \in \mathcal{S}_n : \ 0 < x_k + p_k < 2, \ 1 \leq k \leq n; \ \text{there exists } 1 \leq i \leq n, \ \text{such that } x_i = 0\}.$$

In the remaining of this subsection, we assume that x is in $\mathcal{L}_p$.
Let

$$H_x^{\alpha,\beta} = \left\{ \phi \in C([\alpha,\beta], \mathcal{S}_n) : \ \phi(t) = x + \int_\alpha^t \dot{\phi}(s)ds \right\}.$$

Define

$$I_x^{\alpha,\beta}(\phi) = \begin{cases} \frac{1}{2}\int_\alpha^\beta \sum_{i=1}^n \frac{(\dot\phi_i(t)-c(p_i-\phi_i(t)))^2}{\phi_i(t)}dt, & \phi \in H_x^{\alpha,\beta} \\ \infty, & \phi \notin H_x^{\alpha,\beta} \end{cases}$$

where $\phi_n(t) = 1 - \sum_{j=1}^{n-1}\phi_j(t)$. Denote $I_x^{0,T}$ and $H_x^{0,T}$ by I_x and H_x respectively.

Lemma 2.1 *Suppose that $I_x(\phi) < \infty$.*
i) If $\phi_i(0) = 0$, then $\phi_i(t) \neq 0$, $t \in (0,T]$.
ii) If $\phi_i(0) = 1$, then $\phi_i(t) \neq 1$, $t \in (0,T]$.

Proof: i) Suppose that there exists $t_0 > 0$ such that $\phi_i(t_0) = 0$. If $\phi_i(t) = 0$ for all $t \in [0,t_0]$, then

$$\frac{(\dot\phi_i(t) - c(p_i - \phi_i(t)))^2}{\phi_i(t)} = \infty,$$

and hence $I_x(\phi) = \infty$. This contradicts the assumption. Therefore, we may choose $t_0 < t_0'$ such that $\phi_i(t_0) = \phi_i(t_0') = 0$ and $\phi_i(t) \neq 0$ for all $t \in (t_0, t_0')$.
Let $t_0 < t_1 < t_2 < t_0'$. We have

$$\begin{aligned} \infty &> I_x(\phi) \geq \int_{t_1}^{t_2} \frac{(\dot\phi_i(t) - c(p_i - \phi_i(t)))^2}{\phi_i(t)}dt \\ &\geq -2c\int_{t_1}^{t_2} \frac{\dot\phi_i(t)(p_i - \phi_i(t))}{\phi_i(t)}dt \\ &= -2c\int_{\phi_i(t_1)}^{\phi_i(t_2)} \frac{p_i - x}{x}dx \\ &\to \infty \end{aligned}$$

as t_2 increases to t_0'.
ii) For $j \neq i$, we have $\phi_j(0) = 0$ and hence, for all t in $(0,T]$, $\phi_j(t) \neq 0$. Therefore, for every t in $(0,T]$, $\phi_i(t) \neq 1$. □

Lemma 2.2 *If $I_x(\phi) < \infty$ and $\phi_i(0) = 0$, then there is $t_0 > 0$ such that for $0 \leq t \leq t_0$*

$$\phi_i(t) \leq 2(1+c)t. \tag{2.1}$$

Proof: Note that

$$\begin{aligned} \phi_i(t) &= \int_0^t (\dot\phi_i(s) - c(p_i - \phi_i(s)))ds + cp_i t - c\int_0^t \phi_i(s)ds \\ &\leq \left(\int_0^t \frac{(\dot\phi_i(s) - c(p_i - \phi_i(s)))^2}{\phi_i(s)}ds\right)^{1/2}\left(\int_0^t \phi_i(s)ds\right)^{1/2} + ct. \end{aligned}$$

We can choose $t_0 > 0$ such that for all t in $[0,t_0]$

$$\phi_i(t) \leq \sqrt{t}, \;\; \phi_i(t) \leq \frac{1}{2}\sqrt{\int_0^t \phi_i(s)ds} + ct.$$

By induction, it is easy to show that

$$\phi_i(t) \le \begin{cases} a_{n+1}t^{q_{n+1}} + ct, & \text{if } t \le (a_n/(2+c))^{1/(1-q_n)} \\ 2(1+c)t, & \text{if } t > (a_n/(2+c))^{1/(1-q_n)} \end{cases}$$

where $a_1 = \frac{2+c}{2(1+c)}$, $q_1 = \frac{1}{2}$, $a_{n+1} = \sqrt{\frac{(1+c)a_n}{2(2+c)(1+q_n)}}$ and $q_{n+1} = \frac{1+q_n}{2}$. Since $q_n = 1 - 2^{-n} \to 1$, we get that $a_n \le \sqrt{a_{n-1}/2}$, and $a_n \le \frac{1}{2}$. Therefore,(2.1) holds since $(a_n/(2+c))^{1/(1-q_n)} \to 0$.
□

The next lemma is the key in deriving the large deviation lower bound.

Lemma 2.3 *Suppose $I_x(\phi) < \infty$. If $\phi_i(0) = 0$, then $\dot\phi_i(0) = cp_i$; If $\phi_i(0) = 1$, then $\dot\phi_i(0) = -c(1-p_i)$.*

Proof: First we assume $\phi_i(0) = 0$. Note that

$$\begin{aligned}\left(\phi_i(t) - \int_0^t c(p_i - \phi_i(s))ds\right)^2 &= \left(\int_0^t (\dot\phi_i(s) - c(p_i - \phi_i(s)))ds\right)^2 \\ &\le \int_0^t \frac{(\dot\phi_i(s) - c(p_i - \phi_i(s)))^2}{\phi_i(s)} ds \int_0^t \phi_i(s)ds \\ &\le \int_0^t \frac{(\dot\phi_i(s) - c(p_i - \phi_i(s)))^2}{\phi_i(s)} ds(1+c)t^2.\end{aligned}$$

Then

$$\lim_{t\to 0} \frac{1}{t}\left(\phi_i(t) - \int_0^t c(p_i - \phi_i(s))ds\right) = 0.$$

The conclusion follows easily.
If $\phi_i(0) = 1$, then $\phi_j(0) = 0$ for all $j \ne i$ and hence, $\dot\phi_j(0) = cp_j$. Therefore

$$\dot\phi_i(0) = -\sum_{j\ne i} \dot\phi_j(0) = -c(1-p_i).$$

□

Lemma 2.4 *Let $\delta > 0$ and $\phi \in C([0,T], \mathcal{S}_n)$. If $X^\epsilon(0) = \phi(0)$, then*

$$\lim_{N\to\infty} \limsup_{\epsilon\to 0} \epsilon \log P\left(\sup_{0\le t\le 1/N} |X^\epsilon(t) - \phi(t)| > \delta\right) = -\infty.$$

Proof: Take N large so that $\delta > \frac{4c}{N}$ and $\sup_{0\le t\le 1/N} |\phi(t) - \phi(0)| < \frac{\delta}{4}$. Note that

$$\begin{aligned}P\left(\sup_{0\le t\le 1/N} |X^\epsilon(t) - \phi(t)| > \delta\right) &\le P\left(\sup_{0\le t\le 1/N} \sum_{i=1}^{n-1}\left(\sum_{j=1}^{n-1}\int_0^t \sqrt{N}\sigma_{ij}(X^\epsilon(s))dB_s^j\right)^2 \ge \frac{N\delta^2}{4\epsilon}\right) \\ &\le e^{-\lambda\frac{N\delta^2}{4\epsilon}} E\exp\left(\lambda\sum_{i=1}^{n-1}\sup_{t\le 1}(W_t^i)^2\right) \\ &\equiv c_1 e^{-\lambda\frac{N\delta^2}{4\epsilon}}\end{aligned}$$

where W^i are Brownian motions such that

$$\sum_{j=1}^{n-1}\int_0^t \sqrt{N}\sigma_{ij}(X^\epsilon(s))dB_s^j = W^i_{\tau_t^i}, \qquad \tau_t^i = \int_0^t NX^{\epsilon,i}(s)(1-X^{\epsilon,i}(s))ds,$$

λ is a constant determined by Fernique's theorem (cf. Kuo [5]) such that c_1 is finite. Therefore

$$\limsup_{\epsilon\to 0}\epsilon\log P\left(\sup_{0\le t\le 1/N}|X^\epsilon(t)-\phi(t)|>\delta\right)\le -\lambda N\delta^2. \tag{2.2}$$

The conclusion then follows by taking $N\to\infty$. □

We are now ready to prove the lower bound.

Theorem 2.5 *For $\delta_0>0$ and any ϕ satisfying $I_x(\phi)<\infty$, we have*

$$\liminf_{\epsilon\to 0}\epsilon\log P\{\sup_{0\le t\le T}|X_t^\epsilon-\phi(t)|<2\delta_0\}\ge -I_x(\phi) \tag{2.3}$$

Proof: Let $\psi_i(t)=x_ie^{-ct}+p_i(1-e^{-ct})$. Then it is clear that $X^\epsilon(t)$ converges to $\psi(t)$ for all t as ϵ goes to zero. For any $N\ge 1$, define

$$\varphi(t)=\begin{cases}\phi(t), & 1/N\le t\le T\\ \psi(1/2N)+2N(\phi(1/N)-\psi(1/2N))(t-1/2N), & 1/2N\le t\le 1/N\\ \psi(t), & 0\le t\le 1/2N\end{cases}$$

Choosing N large enough such that $\varphi(t)\in\mathcal{S}_n^\circ$, the interior of $\mathcal{S}_n$, for $t\in[1/2N,1/N]$, and

$$\{\sup_{0\le t\le T}|X^\epsilon(t)-\varphi(t)|<\delta_0\}\subset\{\sup_{0\le t\le T}|X^\epsilon(t)-\phi(t)|<2\delta_0\}.$$

For any $\delta<\delta_0$, we have

$$\begin{aligned}
&2\max\left(P\{\sup_{0\le t\le T}|X^\epsilon(t)-\varphi(t)|<\delta_0\},P\{\sup_{0\le t\le 1/2N}|X^\epsilon(t)-\varphi(t)|\ge\delta_0\}\right)\\
&\qquad\ge P\{\sup_{0\le t\le T}|X^\epsilon(t)-\varphi(t)|<\delta_0\}+P\{\sup_{0\le t\le 1/2N}|X^\epsilon(t)-\varphi(t)|\ge\delta_0\}\\
&\qquad\ge P\{\sup_{1/2N\le t\le T}|X^\epsilon(t)-\varphi(t)|<\delta\}\\
&\qquad\ge P\{|X^\epsilon(1/2N)-\psi(1/2N)|<\delta/2\}\\
&\qquad\quad\times\inf_{|y-\varphi(1/2N)|\le\delta/2}P_y\{\sup_{0\le t\le T-1/2N}|X^\epsilon(t)-\varphi(t+1/2N)|<\delta\}
\end{aligned}$$

where δ is small enough such that $|y-\varphi(1/2N)|\le\delta/2$ implies $y\in\mathcal{S}_n^\circ$. Since $X^\epsilon(1/2N)$ converges to $\psi(1/2N)$ as ϵ goes to zero, by Lemma 2.4, we get

$$\begin{aligned}
&\liminf_{\epsilon\to 0}\epsilon\log P\{\sup_{0\le t\le T}|X^\epsilon(t)-\phi(t)|<2\delta_0\}\\
&\ge\liminf_{\epsilon\to 0}\epsilon\log P\{\sup_{0\le t\le T}|X^\epsilon(t)-\varphi(t)|<\delta_0\}\\
&\ge -I^{1/2N,T}_{\psi(1/2N)}(\varphi)\ge -I^{1/2N,1/N}_{\psi(1/2N)}(\varphi)-I_x(\phi)
\end{aligned}\tag{2.4}$$

where the next to last inequality follows from Theorem 3.3 in [1].
By direct calculation, we have

$$I^{1/2N,1/N}_{\psi(1/2N)}(\varphi) = \frac{1}{2}\sum_{i=1}^{n}\int_{1/2N}^{1/N}\frac{(\dot{\varphi}_i(t)-c(p_i-\varphi_i(t))^2}{\varphi_i(t)}dt.$$

If $x_i \neq 0$, then $\varphi_i(0) \neq 0$ and it is clear that

$$\int_{1/2N}^{1/N}\frac{(\dot{\varphi}_i(t)-c(p_i-\varphi_i(t))^2}{\varphi_i(t)}dt \to 0$$

as $N \to \infty$. If $x_i = 0$, then $\psi_i(0) = \phi_i(0) = 0$. By lemma 2.3, $\dot{\phi}_i(0) = cp_i$. It is clear that $\dot{\psi}_i(0) = cp_i$. Hence

$$\lim_{N\to\infty} 2N\psi_i(1/(2N)) = \lim_{N\to\infty} N\varphi_i(1/N) = cp_i.$$

Then

$$\begin{aligned}
&\int_{1/2N}^{1/N}\frac{(\dot{\varphi}_i(t)-c(p_i-\varphi_i(t))^2}{\varphi_i(t)}dt\\
=\;&\int_{1/2N}^{1/N}\frac{(\dot{\varphi}_i-cp_i)^2}{\varphi_i(t)}dt+\int_{1/2N}^{1/N}(2c\dot{\varphi}_i+c^2\varphi_i(t)-2p_ic^2)dt\\
=\;&\frac{(\dot{\varphi}_i-cp_i)^2}{\dot{\varphi}_i}\log\frac{\dot{\varphi}_i+2N\psi_i(1/2N)}{2N\psi_i(1/2N)}+\int_{1/2N}^{1/N}(2c\dot{\varphi}_i+c^2\varphi_i(t)-2p_ic^2)dt\\
\to\;&0
\end{aligned}$$

as $N \to \infty$. □

2.2 Infinite Dimension

Let $C^{1,0}([0,T]\times E)$ denote the set of all continuous functions on $[0,T]\times E$ with continuous first order derivative in time t. For any $\mu \in M_1(E)$, let $C_\mu([0,T], M_1(E))$ be the set of all $M_1(E)$-valued continuous functions on $[0,T]$ starting at μ.
In this subsection, we will assume that ν_0 is not absolutely continuous with respect to μ. Without loss of generality, we also assume that the support of ν_0 is E.
For any $\mu(\cdot)$ in $C_\mu([0,T], M_1(E))$, define

$$S_\mu(\mu(\cdot)) = \sup_{g\in C^{1,0}([0,T]\times E)} J^\mu_{0,T}(g) \tag{2.5}$$

where

$$\begin{aligned}
J^\mu_{0,T}(g) &= \langle\mu(T), g(T)\rangle - \langle\mu(0), g(0)\rangle\\
&- \int_0^T\langle\mu(s), (\frac{\partial}{\partial s}+A)g(s)\rangle\, ds - \frac{1}{2}\int_0^T\int\int g(s,x)\,g(s,y)\,Q(\mu(s);dx,dy)\,ds.
\end{aligned}$$

Recall that for any μ in $M_1(E)$, we have $Q(\mu;dx,dy) = \mu(dx)\delta_x(dy) - \mu(dx)\mu(dy)$.
Let $B^{1,0}([0,T]\times E)$ be the set of all bounded measurable functions on $[0,T]\times E$ with continuous first order derivative in time. Then by approximation, we have

$$S_\mu(\mu(\cdot)) = \sup_{g\in B^{1,0}([0,T])\times E)} J^\mu_{0,T}(g) \tag{2.6}$$

Definition 2.1 *Let $C^\infty(E)$ denote the set of all continuous functions on E possessing continuous derivatives of all order. An element $\mu(\cdot)$ in $C([0,T],M_1(E))$ is said to be absolutely continuous as a distribution-valued function if there exist $M>0$ and an absolutely continuous function $h_M:[0,T]\to R$ such that for all $t,s\in[0,T]$*

$$\sup_f |\langle \mu(t),f\rangle - \langle \mu(s),f\rangle| \le |h_M(t)-h_M(s)|,$$

where the supremum is taken over all f in $C^\infty(E)$ with absolute value bounded by M.

Let

$$\mathcal{H}_\mu = \{\mu(\cdot)\in C([0,T],M_1(E)) : S_\mu(\mu(\cdot))<\infty\}$$

be the effective domain of $S_\mu(\cdot)$. Then any $\mu(\cdot)$ in $\mathcal{H}_\mu$ is absolutely continuous, and by theorem 5.1 of [2]

$$S_\mu(\mu(\cdot)) = \int_0^T ||\dot{\mu}(s) - A^*(\mu(s))||^2_{\mu(s)}\, ds, \tag{2.7}$$

where A^* is the formal adjoint of A defined through the equality $\langle A^*(\mu),f\rangle = \langle \mu, Af\rangle$, and for any linear functional ϑ on space $C^\infty(E)$

$$||\vartheta||^2_\mu = \sup_{f\in C^\infty(E)} [\langle \vartheta, f\rangle - \frac{1}{2}\int_E\int_E f(x)\,f(y)\,Q(\mu;dx,dy)].$$

Theorem 2.6 *For any μ in $M_1(E)$, $\mu(\cdot)$ in $\mathcal{H}_\mu$, and any open neighborhood U of $\mu(\cdot)$ in space $C_\mu([0,T],M_1(E))$, we have*

$$\liminf_{\epsilon\to 0} \epsilon \log P_\mu\{U\} \ge -S_\mu(\mu(\cdot)), \tag{2.8}$$

where P_μ is the law of the FV process starting at μ.

Proof: The result is clear when $\mu(\cdot)$ is not in the effective domain $\mathcal{H}_\mu$. For $\mu(\cdot)$ in $\mathcal{H}_\mu$, both $\mu(t,[a,b))$ and $\mu(t,[a,b])$ are absolutely continuous in t for $a,b\in E$.
Let $\{f_n\in C(E) : n\ge 1\}$ be a countable dense subset of $C(E)$. Define a metric d on $C_\mu([0,T],M_1(E))$ as follows:

$$d(\nu(\cdot),\mu(\cdot)) = \sup_{t\in[0,T]} \sum_{n=1}^\infty \frac{1}{2^n}(|\langle \nu(t),f_n\rangle - \langle \mu(t),f_n\rangle| \wedge 1). \tag{2.9}$$

Then d generates the uniform convergence topology on $C_\mu([0,T],M_1(E))$. Clearly for any $\delta>0$, there exists a $k\ge 1$ such that

$$\{\mu(\cdot)\in C_\mu([0,T],M_1(E)) : \sup_{t\in[0,T]}\{|\langle \nu(t),f_n\rangle - \langle \mu(t),f_n\rangle|\} \le \delta/2,$$
$$n=1,\cdots,k\} \subset \{d(\nu(\cdot),\mu(\cdot))<\delta\}.$$

Let

$$\mathcal{P} = \{A_1,\cdots,A_n : n\ge 1, A_1,\cdots,A_n \text{ is a partition of } E\},$$

and element of $\mathcal{P}$ is denoted by $\imath,\jmath$, etc. For any ν in $M_1(E)$,$\mu(\cdot)$ in $C([0,T],M_1(E))$, and $\imath=\{A_1,\cdots A_n\}$ and $\jmath=\{[x_0,x_1),\cdots,[x_m,x_{m+1}]\}$ in $\mathcal{P}$, set

$$\pi_\imath(\nu) = (\nu(A_1),\nu(A_2),\cdots,\nu(A_n)),$$

and

$$\pi_\jmath(\mu(\cdot)) = (\mu(\cdot)([0,x_1)), \cdots, \mu(\cdot)([x_m,1])).$$

Next we choose $\jmath$ with $x_0 = 0,\ x_{m+1} = 1$ such that

$$\max_{0\leq i\leq m, x,y\in[x_i,x_{i+1}]}\{|f_n(y) - f_n(x)| : 1 \leq n \leq k\} \leq \frac{\delta}{6},$$

and let

$$U_\jmath(\mu(\cdot), \frac{\delta}{6\Gamma}) = \Big\{\mu(\cdot) \in C_\mu([0,T], M_1(E)) :$$
$$\sup_{t\in[0,T],0\leq i\leq m} |\nu(t)([x_i,x_{i+1})) - \mu(t)([x_i,x_{i+1}))| \leq \frac{\delta}{6\Gamma}\Big\},$$

where $\Gamma = \sup_{x\in E,\ 1\leq n\leq k}|f_n(x)|$. Then we have

$$U_\jmath(\mu(\cdot), \frac{\delta}{6\Gamma}) \subset \{d(\nu(\cdot),\mu(\cdot)) < \delta\}. \tag{2.10}$$

The partition property of the FV process implies that $P_\mu \circ \pi_\jmath^{-1}$ is the law of a FV process with finite type space. This combined with Theorem 2.5, implies

$$\begin{aligned}\liminf_{\epsilon\to 0} \epsilon \log P_\mu\{U\} &\geq \liminf_{\delta\to 0}\liminf_{\epsilon\to 0} \epsilon \log P_\mu\{d(\nu(\cdot),\mu(\cdot)) < \delta\}\\ &\geq \liminf_{\delta\to 0}\liminf_{\epsilon\to 0} \epsilon \log P_\mu\{U_\jmath(\mu(\cdot), \frac{\delta}{6\Gamma})\}\\ &\geq -I_{\pi_\jmath(\mu)}(\pi_\jmath(\mu(\cdot))) \geq -S_\mu(\mu(\cdot)),\end{aligned} \tag{2.11}$$

where the last inequality follows from Theorem 3.4 in [1]. □

3 Effective Domain

In this subsection, we will analyze the structure of the effective domain. We start with another representation of the rate function. For any $\mu(\cdot) \in C_\mu([0,T], M_1(E))$, let

$$L^2_{\mu(\cdot)}([0,T]\times E) = \{f : \|f\|^2_{L^2} = \int_0^T \int_E f^2(t,x)\mu(t,dx)dt < \infty\}.$$

Let $\mathcal{H}^0_\mu$ be the collection of all absolutely continuous $\mu(\cdot)$ in $C_\mu([0,T], M_1(E))$ such that $\dot\mu(t) - A^*(\mu(t))$ is absolutely continuous with respect to $\mu(t)$ as Schwartz distribution with derivative $h(t) = \frac{d(\dot\mu(t)-A^*(\mu(t)))}{d\mu(t)}$ in space $L^2_{\mu(\cdot)}([0,T]\times E)$ satisfying

$$\left\langle \mu(t), \frac{d(\dot\mu(t) - A^*(\mu(t)))}{d\mu(t)}\right\rangle = 0, \qquad t\in[0,T].$$

Define

$$F_\mu(\mu(\cdot)) = \begin{cases} \frac{1}{2}\left\|\frac{d(\dot\mu(t)-A^*(\mu(t)))}{d\mu(t)}\right\|^2_{L^2} & \text{if } \mu(\cdot)\in\mathcal{H}^0_\mu\\ \infty & \text{otherwise.}\end{cases}$$

Theorem 3.1

$$S_\mu(\mu(\cdot)) = F_\mu(\mu(\cdot)).$$

Proof: First we assume that $\mu(\cdot) \in \mathcal{H}^0_\mu$. Using integration by parts, we have

$$\begin{aligned} J^\mu_{0,T}(g) &= \int_0^T \langle \dot\mu(r), g(r)\rangle\, dr - \int_0^T \left\langle \mu(r), Ag(r) + \frac{1}{2}g^2(r)\right\rangle dr + \frac{1}{2}\int_0^T \langle \mu(r), g(r)\rangle^2\, dr \\ &= \int_0^T \left(\left\langle \mu(r), h(r)g(r) - \frac{1}{2}g^2(r)\right\rangle + \frac{1}{2}\langle \mu(r), g(r)\rangle^2\right) dr \\ &= \frac{1}{2}\int_0^T \langle \mu(r), h^2(r)\rangle\, dr - \frac{1}{2}\int_0^T \left(\langle \mu(r), (g(r)-h(r))^2\rangle - \langle \mu(r), g(r)-h(r)\rangle^2\right) dr \\ &\le F_\mu(\mu(\cdot)) \end{aligned}$$

and equality hold for $g = h$.
On the other hand, suppose that $S_\mu(\mu(\cdot)) < \infty$. As in [1], define

$$\ell^\mu_{s,t}(g) = \langle \mu(t), g(t)\rangle - \langle \mu(s), g(s)\rangle - \int_s^t \langle \mu(r), \dot g(r) + Ag(r)\rangle\, dr$$

so that

$$J^\mu_{s,t}(g) = \ell^\mu_{s,t}(g) - \frac{1}{2}\int_s^t \left(\langle \mu(r), g^2(r)\rangle - \langle \mu(r), g(r)\rangle^2\right) dr.$$

Similar to the Appendix in [1], we have

$$\ell^\mu_{s,t}(g) = \int_s^t \langle \mu(r), h(r)g(r)\rangle\, dr$$

for some $h \in L^2_{\mu(\cdot)}([0,T]\times E)$. It follows that $\mu(\cdot)$ is absolutely continuous and $\dot\mu(t) - A^*(\mu(t)) = h(t)\mu(t)$. Note that

$$\langle \mu(t), \phi\rangle - \langle \mu(s), \phi\rangle = \int_s^t \langle \mu(r), A\phi + h(r)\phi\rangle\, dr.$$

Taking $\phi = 1$, we get $0 = \int_s^t \langle \mu(r), h(r)\rangle\, dr$, and hence $\langle \mu(r), h(r)\rangle = 0$, i.e., $\mu(\cdot) \in \mathcal{H}^0_\mu$. □

Let ϕ be any measurable function on E taking values between 0 and 1.

Theorem 3.2 *Assume that $\langle \mu, \phi\rangle = 0$. If $\mu(\cdot)$ is in the effective domain $\mathcal{H}^0_\mu$ of S_μ, then*

$$\frac{d}{dt}\langle \mu(t), \phi\rangle\Big|_{t=0} = \langle \mu, A\phi\rangle.$$

Proof: Let $a_t = \langle \mu(t), \phi\rangle$. Note that

$$\langle \mu(t), A\phi\rangle = \int\int (\phi(y) - \phi(x))\nu_0(dy)\mu(t,dx) \le 2.$$

Then

$$\begin{aligned} a_t &= \int_0^t \langle \mu(s), A\phi \rangle\, ds + \int_0^t \langle \mu(s), h(s)\phi \rangle\, ds \\ &\leq 2t + \sqrt{\int_0^t \langle \mu(s), h^2(s) \rangle\, ds} \sqrt{\int_0^t a_s ds}. \end{aligned}$$

Similar to Lemma 2.2, we have $a_t \leq ct$. So

$$\left| a_t - \int_0^t \langle \mu(s), A\phi \rangle\, ds \right|^2 \leq \int_0^t \langle \mu(s), h^2(s) \rangle\, ds \frac{c}{2} t^2.$$

The conclusion of the theorem follows from the same arguments as in the proof of Lemma 2.3. □

4 Quasi-Potential

For any ν in $M_1(E)$, the quasi-potential $V(\nu_0;\nu)$ is defined as

$$V(\nu_0;\nu) = \inf\{S_{\nu_0}^{T_1,T_2}(\mu(\cdot)) : \mu(T_1) = \nu_0, \mu(T_2) = \nu, -\infty \leq T_1 < T_2 \leq \infty\}.$$

It is the least amount of energy needed for a transition from ν_0 to ν under large deviations. In this section the quasi-potential of the FV process is identified explicitly.

Theorem 4.1 *The quasi-potential of the FV process is given by*

$$V(\nu_0;\nu) = \theta H(\nu_0|\nu). \tag{4.1}$$

Proof: First consider the case of $E = \{1, \cdots, n\}$ for some $n \geq 2$. For any $1 \leq k \leq n$, let $p_k = \nu_0(k)$, $x_k = \nu(k)$. Then $H(\nu_0|\nu) = \sum_{k=1}^n p_k \log \frac{p_k}{x_k}$ is smooth in ν. For any $\mu(\cdot)$ in $C([T_1,T_2], M_1(E))$ satisfying $\mu(T_1) = \nu_0, \mu(T_2) = \nu, S_{\nu_0}(\mu(\cdot)) < \infty$, define $x_k(t) = \mu(t,k)$ and $g(t,k) = -\theta \frac{p_k}{x_k(t)}$. Then it is clear that g is in $C^{1,0}([T_1,T_2] \times E)$. By direct calculation, we have

$$\begin{aligned} &\int_{T_1}^{T_2} \langle \mu(s), (Ag(s) \rangle\, ds + \frac{1}{2} \int_{T_1}^{T_2} \int\int g(s,x)\, g(s,y)\, Q(\mu(s); dx, dy)\, ds = \\ &\quad -\frac{\theta^2}{2} \int_{T_1}^{T_2} [\sum_{k=1}^n (\frac{p_k^2}{x_k(s)} - p_k) - \sum_{k=1}^n (x_k(s)(\frac{p_k}{x_k(s)})^2) + (\sum_{k=1}^n p_k)^2] ds = 0. \end{aligned}$$

Thus it follows from (2.5) that

$$\begin{aligned} S_{\nu_0}^{T_1,T_2}(\mu(\cdot)) \geq -\theta \sum_{k=1}^n [x_k(T_2) \frac{p_k}{x_k(T_2)} - p_k] \\ +\theta \int_{T_1}^{T_2} \sum_{k=1}^n x_k(s)(-\frac{p_k}{x_k^2(s)}) \dot{x}_k(s) d s = \theta H(\nu_0|\nu), \end{aligned}$$

which implies that $V(\nu_0;\nu) \geq \theta H(\nu_0|\nu)$.
Next we consider the general case. Note that for any partition $\imath$ of E,

$$S^{T_1,T_2}_{\pi_\imath(\nu_0),\pi_\imath(\nu)}(\pi_\imath(\mu(\cdot))) \geq V(\pi_\imath(\nu_0),\pi_\imath(\nu)),$$

and hence

$$\sup_\imath\{S^{T_1,T_2}_{\pi_\imath(\nu_0),\pi_\imath(\nu)}(\pi_\imath(\mu(\cdot)))\} \geq \sup_\imath\{V(\pi_\imath(\nu_0),\pi_\imath(\nu))\}.$$

Therefore

$$\begin{aligned}
V(\nu_0,\nu) &= \inf\{S^{T_1,T_2}_{\nu_0,\nu}(\mu(\cdot)) : \mu(T_1)=\nu_0, \mu(T_2)=\nu, -\infty \leq T_1 < T_2 \leq \infty\} \\
&\geq \inf\left\{\sup_\imath\{S^{T_1,T_2}_{\pi_\imath(\nu_0),\pi_\imath(\nu)}(\pi_\imath(\mu(\cdot)))\} : \mu(T_1)=\nu_0, \mu(T_2)=\nu, -\infty \leq T_1 < T_2 \leq \infty\right\} \\
&\geq \sup_\imath\{V(\pi_\imath(\nu_0),\pi_\imath(\nu))\} \\
&= \sup_\imath\{\theta H(\pi_\imath(\nu_0)|\pi_\imath(\nu)) \\
&= \theta H(\nu_0|\nu).
\end{aligned}$$

Finally we derive the upper bound. Let $\mu(\cdot)$ be the solution of the equation on $[T_1,T_2]$.

$$\frac{d\mu(t)}{dt} = \frac{\theta}{2}(\mu(t)-\nu_0), \mu(T_2)=\nu. \tag{4.2}$$

Then

$$\mu(t) = e^{-.5\theta(T_2-t)}\nu + (1-e^{-.5\theta(T_2-t)})\nu_0.$$

It is not hard to see that $\mu(T_1) \to \nu_0$ as T_1 approaches $-\infty$. From the expression (2.7), we get

$$S^{T_1,T_2}_{\mu(T_1),\nu}(\mu(\cdot)) = \int_{T_1}^{T_2} \sup_{f\in B^{1,0}([0,T]\times E)} \{\theta\langle\mu(s)-\nu_0,f\rangle - \frac{1}{2}\left(\langle\mu(s),f^2\rangle - \langle\mu(s),f\rangle^2\right)\}ds. \tag{4.3}$$

For f in $B^{1,0}([0,T]\times E)$, let

$$M(f)(s) = \theta\langle\mu(s)-\nu_0,f\rangle - \frac{1}{2}\left(\langle\mu(s),f^2\rangle - \langle\mu(s),f\rangle^2\right).$$

Define $g(s,x) = -\theta\frac{d\nu_0}{d\mu(s)}(x)$. It is clear that g is in $B^{1,0}([0,T]\times E)$, and

$$M(f)(s) = M(g+f-g)(s) = M(g)(s) - \frac{1}{2}[\langle\mu(s),(f-g)^2\rangle - \langle\mu(s),f-g\rangle^2] \leq M(g)(s).$$

Hence the sup in (4.3) is attained at g, and we have

$$S^{T_1,T_2}_{\mu(T_1),\nu}(\mu(\cdot)) = \int_{T_1}^{T_2}\{\theta\langle\mu(s)-\nu_0,g\rangle - \frac{1}{2}\left(\langle\mu(s),g^2\rangle - \langle\mu(s),g\rangle^2\right)\}ds = \int_{T_1}^{T_2}\langle\dot\mu(s),g\rangle ds. \tag{4.4}$$

By integration by parts, we get

$$\begin{aligned}
\int_{T_1}^{T_2}\langle\dot\mu(s),g\rangle ds &= \langle\mu(T_2),g(T_2)\rangle - \langle\mu(T_1),g(T_1)\rangle - \int_{T_1}^{T_2}\langle\mu(s),\dot g\rangle ds \\
&= -\int_{T_1}^{T_2}\langle\mu(s),\dot g\rangle ds = \theta\int_{T_1}^{T_2}(H(\nu_0|\mu(s)))'ds = \theta[H(\nu_0|\nu) - H(\nu_0|\mu(T_1))].
\end{aligned}$$

The upper bound then follows by letting T_1 go to $-\infty$. $\square$

Acknowledgements. We would like to thank the referee and the associate editor for helpful suggestions and comments.

References

[1] D. Dawson and S. Feng (1998). Large deviations for the Fleming-Viot process with neutral mutation and selection. *Stochastic Process. Appl.* **77**, 207–232.

[2] D. Dawson and S. Feng (2001). Large deviations for the Fleming-Viot process with neutral mutation and selection, II. *Stochastic Process. Appl.* **92**, 131–162.

[3] A. Dembo and O. Zeitouni. (1998). *Large deviations techniques and applications.* Second edition. Applications of Mathematics, **38**. Springer-Verlag, New York.

[4] Ethier, S.N. and Kurtz, T.G.(1994). Convergence to Fleming-Viot processes in the weak atomic topology. *Stoch. Proc. Appl.* , **54**, 1–27.

[5] H. H. Kuo (1975). Gaussian Measures in Banach Spaces, *Lecture Notes in Mathematics* **463**, Springer-Verlag.