A CLASS OF INFINITE-DIMENSIONAL DIFFUSION PROCESSES WITH CONNECTION TO POPULATION GENETICS

SHUI FENG,* McMaster University
FENG-YU WANG,** Beijing Normal University and Swansea University

Abstract
Starting from a sequence of independent Wright–Fisher diffusion processes on [0, 1], we construct a class of reversible infinite-dimensional diffusion processes on \( \Delta_\infty := \{ x \in [0, 1]^\mathbb{N} : \sum_{k \geq 1} x_k = 1 \} \) with GEM distribution as the reversible measure. Log-Sobolev inequalities are established for these diffusions, which lead to the exponential convergence of the corresponding reversible measures in the entropy. Extensions are made to a class of measure-valued processes over an abstract space \( S \). This provides a reasonable alternative to the Fleming–Viot process, which does not satisfy the log-Sobolev inequality when \( S \) is infinite as observed by Stannat (2000).

Keywords: Poisson–Dirichlet distribution; GEM distribution; Fleming–Viot process; log-Sobolev inequality

2000 Mathematics Subject Classification: Primary 60F10 Secondary 92D10

1. Introduction
Population genetics is concerned with the distribution and evolution of gene frequencies in a large population at a particular locus. The infinitely-many-neutral-alleles model describes the evolution of the gene frequencies under generation independent mutation and resampling. In statistical equilibrium the distribution of gene frequencies is the well-known Poisson–Dirichlet distribution introduced by Kingman [8]. When a sample of size \( n \) genes is selected from a Poisson–Dirichlet population, the distribution of the corresponding allelic partition is given explicitly by the Ewens sampling formula. This provides an important tool in testing neutrality of a population.

Let
\[ \Delta_\infty = \left\{ x = (x_1, x_2, \ldots) \in [0, 1]^\mathbb{N} : \sum_{k = 1}^{\infty} x_k = 1 \right\} , \]
and let
\[ \nabla = \left\{ x = (x_1, x_2, \ldots) \in [0, 1]^\mathbb{N} : x_1 \geq x_2 \geq \cdots \geq 0, \sum_{k = 1}^{\infty} x_k = 1 \right\} . \]

The Poisson–Dirichlet distribution with parameter \( \theta > 0 \) is a probability measure \( \Pi_\theta \) on \( \nabla \). We use \( P(\theta) = (P_1(\theta), P_2(\theta), \ldots) \) to denote the \( \nabla \)-valued random variable with distribution \( \Pi_\theta \).

Received 30 August 2007.
* Postal address: Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada L8S 4K1. Email address: shuifeng@mcmaster.ca
** Postal address: Department of Mathematics, Swansea University, Singleton Park, Swansea SA2 8PP, UK. Email address: f.y.wang@swansea.ac.uk

938
The component $P_k(\theta)$ represents the proportion of the $k$th most frequent alleles. If $u$ denotes the individual mutation rate and $N$ denotes the effective population size then the parameter $\theta = 4Nu$ denotes the population mutation rate. An alternative way of describing the distribution is through the following size-biased sampling. Let $U_k$, $k = 1, 2, \ldots$, be a sequence of independent, identically distributed (i.i.d.) random variables with common distribution Beta$(1, \theta)$, and set

$$X_1^0 = U_1, \quad X_n^0 = (1 - U_1) \cdots (1 - U_{n-1}) U_n, \quad n \geq 2.$$  

Clearly $(X_1^0, X_2^0, \ldots)$ is in space $\Delta_\infty$. The law of $X_1^0, X_2^0, \ldots$ is called the one-parameter GEM distribution and is denoted by $\Pi^\text{GEM}_\theta$. The descending order of $X_1^0, X_2^0, \ldots$ has distribution $\Pi_\theta$. The sequence $X_k^0$, $k = 1, 2, \ldots$, has the same distribution as the size-biased permutation of $\Pi_\theta$.

Let $\xi_k$, $k = 1, 2, \ldots$, be a sequence of i.i.d. random variables with common diffusive distribution $\nu$ on $[0, 1]$, i.e. $\nu(x) = 0$ for every $x$ in $[0, 1]$. Set

$$\Theta_{\theta, \nu} = \sum_{k=1}^{\infty} P_k(\theta) \delta_{\xi_k}.$$  

It is known that the law of $\Theta_{\theta, \nu}$ is a Dirichlet$(\theta, \nu)$ distribution, and it is the reversible distribution of the Fleming–Viot process with mutation operator (see [2])

$$Af(x) = \frac{\theta}{2} \int_0^1 (f(y) - f(x))\nu(dx). \quad (1.1)$$  

For $0 \leq \alpha < 1$ and $\theta > -\alpha$, let $\{V_k : k = 1, 2, \ldots\}$ be a sequence of independent random variables such that $V_k$ is a Beta$(1 - \alpha, \theta + k\alpha)$ random variable for each $k$. Set

$$X_1^{0,\alpha} = V_1, \quad X_n^{0,\alpha} = (1 - V_1) \cdots (1 - V_{n-1}) V_n, \quad n \geq 1. \quad (1.2)$$  

The law of $X_1^{0,\alpha}, X_2^{0,\alpha}, \ldots$ is called the two-parameter GEM distribution and is denoted by $\Pi^\text{GEM}_{\alpha,\theta}$. The law of the descending order statistic of $X_1^{0,\alpha}, X_2^{0,\alpha}, \ldots$ is called the two-parameter Poisson–Dirichlet distribution (henceforth denoted by $\Pi_{\alpha,\theta}$), which was studied thoroughly in [12]. The sequence $X_k^{0,\alpha}$, $k = 1, 2, \ldots$, has the same distribution as the size-biased permutation of $\Pi_{\alpha,\theta}$. In [11] it was shown that the two-parameter Poisson–Dirichlet distribution is the most general distribution whose size-biased permutation has the same distribution as the GEM representation (1.2). A two-parameter ‘Ewens sampling formula’ was obtained in [10]. Let $\Theta_{\theta,\alpha,\nu}$ be defined similarly to $\Theta_{\theta,\nu}$ with $X_k^0$ being replaced by $X_k^{0,\alpha}$. We call the law of $\Theta_{\theta,\alpha,\nu}$ a Dirichlet$(\theta, \alpha, \nu)$ distribution.

The Poisson–Dirichlet distribution and its two-parameter generalization have many similar structures including the urn construction in [3] and [7], GEM representation, sampling formula, etc. However, we have not seen a stochastic dynamic model similar to the infinitely-many-neutral-alleles model and the Fleming–Viot process developed for the two-parameter Poisson–Dirichlet distribution and the Dirichlet$(\theta, \alpha, \nu)$ distribution.

In this paper we firstly construct a class of reversible infinite-dimensional diffusion processes, the GEM processes, so that both $\Pi^\text{GEM}_\theta$ and its two-parameter generalization $\Pi^\text{GEM}_{\alpha,\theta}$ appear as the reversible measures for appropriate parameters.

In [13] the log-Sobolev inequality is studied for the Fleming–Viot process with the motion given by (1.1). It turns out that the log-Sobolev inequality holds only when the type space is finite. In the second result of this paper we first construct a measure-valued process that...
has the Dirichlet$(\theta, \nu)$ distribution as reversible measure. Then we establish the log-Sobolev inequality for this process.

The rest of the paper is organized as follows. The GEM processes associated with $\Gamma_{\pi}^{\text{GEM}}$ and $\Gamma_{\nu}^{\text{GEM}}$ are introduced in Section 2. Section 3 includes the proof of uniqueness and the log-Sobolev inequality of the GEM process. Finally, in Section 4 the measure-valued process is introduced and the corresponding log-Sobolev inequality is established.

2. GEM processes

For any $i \geq 1$, let $a_i$ and $b_i$ denote two strictly positive numbers. We assume that

$$\inf_i b_i \geq \frac{1}{2}. \quad (2.1)$$

Let $X_i(t)$ denote the unique strong solution of the stochastic differential equation

$$dX_i(t) = (a_i - (a_i + b_i)X_i(t)) \, dt + \sqrt{X_i(t)(1 - X_i(t))} \, dB_i(t), \quad X_i(0) \in [0, 1],$$

where $\{B_i(t) : i = 1, 2, \ldots\}$ are independent one-dimensional Brownian motions. It is known that the process $X_i(t)$ is reversible with reversible measure $\pi_{a_i,b_i} = \text{Beta}(2a_i, 2b_i)$. By direct calculation, the scale function of $X_i(\cdot)$ is given by

$$s_i(x) = \left(\frac{1}{4}\right)^{a_i+b_i} \int_{1/2}^{x} \frac{dy}{y^{2a_i}(1-y)^{2b_i}}.$$

By (2.1) we have $\lim_{x \to 1} s_i(x) = \infty$ for all $i$. Thus, starting from the interior of $[0, 1]$, the process $X_i(t)$ will not hit the boundary 1 with probability 1. Let $E = [0, 1]^N$. The process

$$X(t) = (X_1(t), X_2(t), \ldots)$$

is then an $E$-valued Markov process. Consider the map

$$\Phi : E \to \hat{\Delta}_\infty, \quad x = (x_1, x_2, \ldots) \rightarrow (\varphi_1(x), \varphi_2(x), \ldots),$$

with

$$\varphi_1(x) = x_1, \quad \varphi_n(x) = x_n(1 - x_1) \cdots (1 - x_{n-1}), \quad n \geq 2.$$

Clearly $\Phi$ is a bijection and the process $Y(t) = \Phi(X(t))$ is thus a Markov process. Let $\bar{E} := [0, 1]^\mathbb{N}$ denote the closure of $E$, let $C(\bar{E})$ denote the set of all continuous functions on $\bar{E}$, and let $C^2_{cl}(\bar{E})$ denote the set of cylindrical functions in $C(\bar{E})$ that have second-order continuous derivatives and depend only on a finite number of coordinates. The sets $C(E)$ and $C^2_{cl}(E)$ will be the respective restrictions of $C(\bar{E})$ and $C^2_{cl}(\bar{E})$ on $E$. Then the generator of process $X(t)$ is given by

$$Lf(x) = \sum_{k=1}^{\infty} \left\{ x_k(1 - x_k) \frac{\partial^2 f}{\partial x_k^2} + (a_k - (a_k + b_k)x_k) \frac{\partial f}{\partial x_k} \right\}, \quad f \in C^2_{cl}(E),$$

and can be extended to $C^2_{cl}(\bar{E})$. The sets $B(E)$ and $B(\Delta_{\infty})$ are bounded measurable functions on $E$ and $\Delta_{\infty}$, respectively.
Let $a = (a_1, a_2, \ldots)$ and $b = (b_1, b_2, \ldots)$, and let
\[ \mu_{a, b} = \prod_{k=1}^{\infty} \pi_{a_k, b_k} \quad \text{and} \quad \Xi_{a, b} = \mu_{a, b} \circ \Phi^{-1}. \]

Then we have the following result.

**Theorem 2.1.** The processes $X(t)$ and $Y(t)$ are reversible with respective reversible measures $\mu_{a, b}$ and $\Xi_{a, b}$.

**Proof.** The reversibility of $X(t)$ follows from the reversibility of each $X_i(t)$. Now, for any two $f$ and $g$ in $B(\Delta_\infty)$, the two functions $f \circ \Phi$ and $g \circ \Phi$ are in $B(E)$. From the reversibility of $X(t)$, we have, for any $t > 0$,
\[
\int_{\Delta_\infty} f(y) E_y[g(y(t))]|\Xi_{a, b}(dy) = \int_E f(\Phi(x)) E_x[g(\Phi(x(t)))]|\mu_{a, b}(dx) \\
= \int_E g(\Phi(x)) E_x[f(\Phi(x(t)))]|\mu_{a, b}(dx) \\
= \int_{\Delta_\infty} g(y) E_y[f(y(t))]|\Xi_{a, b}(dy).
\]

Hence, $Y(t)$ is reversible with reversible measure $\Xi_{a, b}$.

**Remark.** The one-parameter GEM distribution, $\Pi_1^\text{GEM}$, corresponds to $a_i = \frac{1}{2}$ and $b_i = \theta/2$, and the two-parameter GEM distribution, $\Pi_{a, \theta}^\text{GEM}$, corresponds to $a_i = (1 - \alpha)/2$ and $b_i = (\theta + i\alpha)/2$.

### 3. Uniqueness and Poincaré/log-Sobolev inequalities

Let
\[ \tilde{\Delta}_\infty := \left\{ x \in [0, 1]^N : \sum_{i=1}^{\infty} x_i \leq 1 \right\} \]

be the closure of space $\Delta_\infty$ in $\mathbb{R}^N$ under the topology induced by cylindrically continuous functions. The probability $\Xi_{a, b}$ can be extended to the space $\tilde{\Delta}_\infty$. For simplicity, the same notation is used to denote this extended probability measure.

Now, for $x \in \tilde{\Delta}_\infty$ such that
\[ \sum_{i=1}^{n} x_i < 1 \quad \text{for all finite } n, \]

let
\[ L(x) = \sum_{i,j=1}^{\infty} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{\infty} b_i(x) \frac{\partial}{\partial x_i}. \]
where

\[ a_{ij}(x) := x_i x_j \sum_{k=1}^i \frac{\delta_{ki}(1 - \sum_{l=1}^{k-1} x_l) - x_k}{x_k(1 - \sum_{l=1}^k x_l)} (\delta_{kj}(1 - \sum_{l=1}^{k-1} x_l) - x_k) \]

\[ b_i(x) := x_i \sum_{k=1}^i \frac{\delta_{ik}(1 - \sum_{l=1}^{k-1} x_l) - x_k}{x_k(1 - \sum_{l=1}^k x_l)} (a_k(1 - \sum_{l=1}^{k-1} x_l) - (a_k + b_k)x_k) \]

Here and in the sequel, we set \( \sum_{i=1}^0 = 1 \) and \( \prod_{i=1}^0 = 1 \) by convention. By treating 0 as 1, the definition of \( L(x) \) can be extended to all points in \( \Delta\infty \). Through direct calculation we can see that \( L(x) \) is the generator of the GEM process.

It follows, from direct calculation, that

\[ \sum_{i,j=1}^{\infty} |a_{ij}(x)| \leq 3, \quad |b_i(x)| \leq \sum_{k=1}^i (b_kx_k + a_k), \quad x \in \Delta\infty. \]  

(3.1)

Indeed, since \( 1 - \sum_{i,j=1}^{i,j-1} x_i \geq x_i \) and \( \sum_{1 \leq i < j < \infty} x_i x_j \leq \frac{1}{2} \), we obtain

\[ \sum_{i,j=1}^{\infty} |a_{ij}(x)| = \sum_{i=1}^{\infty} a_{ij}(x) + 2 \sum_{1 \leq i < j < \infty} |a_{ij}(x)| \]

\[ \leq \sum_{i=1}^{\infty} x_i^2 \left( \frac{1 - \sum_{l=1}^{i-1} x_l}{x_i} + \sum_{k=1}^{i-1} \frac{x_k}{1 - \sum_{l=1}^k x_l} \right) \]

\[ + 2 \sum_{1 \leq i < j < \infty} x_i x_j \left( 1 + \sum_{k=1}^{i-1} \frac{x_k}{1 - \sum_{l=1}^k x_l} \right) \]

\[ \leq \sum_{i=1}^{\infty} x_i \left( 1 - \sum_{l=1}^i x_l + x_k \right) \]

\[ + 2 \sum_{i=1}^{\infty} x_i \sum_{j=i+1}^{\infty} x_j \left( 1 + \sum_{k=i+1}^{\infty} \frac{x_k}{1 - \sum_{l=i+1}^k x_l} \right) \]

\[ \leq 1 + 2 \]

\[ = 3. \]

Thus, the first inequality in (3.1) holds. Similarly, the second inequality also holds.

Let

\[ \Gamma(f, g)(x) = \sum_{i,j=1}^{\infty} a_{ij}(x) \frac{\partial f(x)}{\partial x_i} \frac{\partial g(x)}{\partial x_j}. \]

Then \( \Gamma(f, g) \in C_b(\Delta\infty) \) for any \( f \in C_b^1(\Delta\infty) \).

For each \( a > 0 \) and \( b > 0 \), let \( \alpha_{a,b} \) be the largest constant such that, for \( f \in C_b^1([0, 1]) \), the log-Sobolev inequality,

\[ \pi_{a,b}(f^2 \log f^2) \leq \frac{1}{\alpha_{a,b}} \int_0^1 x(1-x) f'(x)^2 \pi_{a,b}(dx) + \pi_{a,b}(f^2) \log \pi_{a,b}(f^2), \]  

(3.2)
holds. According to Lemma 2.7 of [13], we have \( \alpha_{a,b} \geq \frac{(a \wedge b)}{320} \). Moreover, it is easy to see that, for \( a, b > 0 \), the operator
\[
\frac{r(1-r)}{dr^2} + (a - (a + b)r) \frac{d}{dr}
\]
on \([0, 1]\) has a spectral gap \( a + b \) with eigenfunction \( h(r) := a - (a + b)r \). So, the Poincaré inequality,
\[
\pi_{a,b}(f^2) \leq \frac{1}{a + b} \int_0^1 x(1-x)f'(x)^2 \pi_{a,b}(dx) + \pi_{a,b}(f)^2,
\]
holds.

Let \( C^\infty_c([0, 1]^N) \) denote the set of all bounded, \( C^\infty_c \) cylindrical functions on \([0, 1]^N\), and
\[
FC^\infty_b = \{ f |_{\Delta_n} : f \in C^\infty_c([0, 1]^N) \}.
\]

Now we have the following theorem.

**Theorem 3.1.** For any \( f, g \in FC^\infty_b \), we have
\[
\mathcal{E}(f, g) := \mathbb{E}_{a,b}(\Gamma(f, g)) = -\mathbb{E}_{a,b}(f L g).
\]

Consequently, \( (\mathcal{E}, \mathcal{F}) \) is closable in \( L^2(\Delta_n; \mathbb{E}_{a,b}) \), and its closure is a conservative regular Dirichlet form which satisfies the Poincaré inequality
\[
\pi_{a,b}(f^2) \leq \frac{1}{\inf_{i \geq 1}(a_i + b_i)} \mathcal{E}(f, f), \quad f \in D(\mathcal{E}), \quad \mathbb{E}_{a,b}(f) = 0.
\]

Moreover, if \( \inf\{a_i \wedge b_i : i \geq 1\} > 0 \), the log-Sobolev inequality
\[
\mathbb{E}_{a,b}(f^2 \log f^2) \leq \frac{1}{\beta_{a,b}} \mathcal{E}(f, f), \quad f \in D(\mathcal{E}), \quad \mathbb{E}_{a,b}(f^2) = 1,
\]
holds for some \( \beta_{a,b} \geq \inf\{(a_i \wedge b_i)/320 : i \geq 1\} > 0 \).

**Proof.** For any \( f, g \in FC^\infty_b \), there exists \( n \geq 1 \) such that
\[
f(x) = f(x_1, \ldots, x_n), \quad g(x) = g(x_1, \ldots, x_n), \quad x = (x_1, \ldots, x_n, \ldots) \in [0, 1]^N.
\]

Let
\[
\phi^{(n)}(x) = (\phi_1(x), \ldots, \phi_n(x)),
\]
which maps \([0, 1]^n\) onto \( \Delta_n := \{ x \in [0, 1]^n : \sum_{i=1}^n x_i \leq 1 \} \). Define
\[
L_n := \sum_{i=1}^n x_i(1-x_i) \frac{\partial}{\partial x_i} + \sum_{i=1}^n (a_i - (a_i + b_i)x_i) \frac{\partial}{\partial x_i},
\]
and
\[
\pi_{a,b}^n = \prod_{i=1}^n \pi_{a_i,b_i} \quad \text{and} \quad \mathbb{E}^n = \pi_{a,b}^n \circ \phi^{(n)-1}.
\]
Then, regarding \( \Xi_n := \pi_n \circ \varphi^{(n)}_a \circ \varphi^{(n)}_b : n \geq 1 \) as probability measures on \( \tilde{\Delta}_\infty \) and by letting \( \Xi_n := \Xi_n(d\mathbf{x}_1, \ldots, d\mathbf{x}_n) \times \delta_0(d\mathbf{x}_{n+1}, \ldots) \), it converges weakly to \( \Xi_{a,b} \). Since \( L_n \) is symmetric with respect to \( \pi_n \), we have

\[
\int_{[0,1]^n} \sum_{i=1}^n x_i (1 - x_i) \left( \frac{\partial}{\partial x_i} f \circ \varphi^{(n)} \right) \left( \frac{\partial}{\partial x_i} g \circ \varphi^{(n)} \right) d\pi_{a,b}^n
= - \int_{[0,1]^n} g \circ \varphi^{(n)} L_n f \circ \varphi^{(n)} d\pi_{a,b}^n.
\]

Noting that
\[
\varphi_i(x) = x_i \prod_{j=1}^{i-1} (1 - x_j) \quad \text{and} \quad x_i = \frac{\varphi_i(x)}{1 - \sum_{j=1}^{i-1} \varphi_j(x)}, \quad i \geq 1,
\]
we have
\[
\frac{df \circ \varphi^{(n)}(x)}{dx_i} = \sum_{j \geq i} \frac{(x_j - x_i) \varphi_j(x)}{x_i (1 - x_i)} \frac{df}{d\varphi_j} \circ \varphi^{(n)}(x).
\]

Therefore,

\[
\int_{[0,1]^n} \sum_{i=1}^n x_i (1 - x_i) \left( \frac{\partial}{\partial x_i} f \circ \varphi^{(n)} \right) \left( \frac{\partial}{\partial x_i} g \circ \varphi^{(n)} \right) d\pi_{a,b}^n
= \int_{[0,1]^n} \Gamma(f, g) \circ \varphi^{(n)} d\pi_{a,b}^n
= \int_{\tilde{\Delta}_n} \Gamma(f, g) d\Xi^n.
\]  \hspace{1cm} (3.7)

By (3.1) and (3.6), we have \( \Gamma(f, g) \in C_b(\tilde{\Delta}_\infty) \), so that the weak convergence of \( \Xi^n \) to \( \Xi_{a,b} \) implies that

\[
\lim_{n \to \infty} \int_{\Delta_n} \Gamma(f, g) d\Xi^n = \int_{\tilde{\Delta}_\infty} \Gamma(f, g) d\Xi_{a,b}.
\]  \hspace{1cm} (3.8)

Similarly, by straightforward calculations we find that

\[
L_n f \circ \varphi^{(n)}(x) = (Lf) \circ \varphi^{(n)}(x).
\]

Moreover, (3.1) and (3.6) imply that \( gLf \in C_b(\tilde{\Delta}_\infty) \). Thus, we arrive at

\[
\lim_{n \to \infty} \int_{\Delta_n} g \circ \varphi^{(n)} L_n f \circ \varphi^{(n)} d\pi_{a,b}^n = \int_{\tilde{\Delta}_\infty} gLf d\Xi_{a,b}.
\]

Therefore, (3.4) follows by combining this with (3.7) and (3.8). This implies the closability of \( (\mathcal{E}, FC_{b}^\infty) \), while the regularity of its closure follows from the compactness of \( \tilde{\Delta}_\infty \) under the usual metric

\[
\rho(x, y) := \sum_{i=1}^{\infty} 2^{-i} |x_i - y_i|.
\]
Indeed, it is trivial that $D(\mathcal{E}) \cap C_0([0,1]^n) \supset FC_{\mathcal{E}}^\infty$, which is dense in $D(\mathcal{E})$ under $\mathcal{E}_{1/2}^{1/2}$ given by

$$\mathcal{E}_1(f, f) = \mathcal{E}(f, f) + \|f\|_2^2.$$ 

Moreover, for any $F \in C(\tilde{\Delta}_\infty) = C(\Delta_\infty)$, by its uniform continuity owing to the compactness of the space,

$$\tilde{\Delta}_\infty \ni x \mapsto F_n(x) := F(x_1, \ldots, x_n, 0, 0, \ldots), \quad n \geq 1,$$

is a sequence of continuous cylindric functions converging uniformly to $F$. Since a cylindric continuous function can be uniformly approximated by functions in $FC_{\mathcal{E}}^\infty$ under the uniform norm, it follows that $FC_{\mathcal{E}}^\infty$ is dense in $C(\Delta_\infty)$ under the uniform norm. That is, the Dirichlet form $(\mathcal{E}, D(\mathcal{E}))$ is regular.

Next, the desired Poincaré and log-Sobolev inequalities can be deduced from (3.3) and (3.2), respectively. For simplicity, we only prove the latter. By the additivity property of the log-Sobolev inequality (see [6]),

$$\mu^n(h^2 \log h^2) \leq \frac{1}{\beta^n_{a, b}} \int_{[0,1]^n} \sum_{i=1}^n x_i(1 - x_i) \left( \frac{\partial h}{\partial x_i} \right)^2 \, d\pi^n_{a, b} + \mu^n(h^2) \log \pi^n_{a, b}(h^2)$$

holds for all $h \in C^1_b([0,1]^n)$, where

$$\beta^n_{a, b} = \inf \{\alpha_{a, b} : i = 1, \ldots, n\} \quad \text{and} \quad f^{(n)}(x) = f(x_1, \ldots, x_n, 0, \ldots).$$

Combining this with (3.7) for any $f \in D$, the domain of $L$, we have

$$\mathbb{E}^n(f^{(n)} \log f^{(n)}) \leq \frac{1}{\beta^n_{a, b}} \int_{\Delta_n} \Gamma^{(n)}(f, f) \, d\mathbb{E}^n + \mathbb{E}^n(f^{(n)}^2) \log \mathbb{E}^n(f^{(n)}^2).$$

Therefore, as explained above, for $f \in D$, (3.5) follows immediately by letting $n$ tend to $\infty$. Hence, the proof is completed since $D(\mathcal{E})$ is the closure of $D$ under $\mathcal{E}^{1/2}$.

We remark that since $(\mathcal{E}, D(\mathcal{E}))$ is regular, according to [4] and [9], $(L, D)$ generates a Hunt process whose semigroup $P_t$ is unique in $L^2(\mathbb{E}_{a, b})$. Thus, the GEM process constructed in Section 2 is the unique Feller process generated by $L$. Moreover, it is well known that the log-Sobolev inequality, (3.5), implies that $P_t$ converges to $\mathbb{E}_{a, b}$ exponentially fast in entropy; more precisely (see, e.g. [1, Proposition 2.1]),

$$\mathbb{E}_{a, b}(P_t f \log P_t f) \leq \exp(-4\beta_{a,b}^n) \mathbb{E}_{a, b}(f \log f), \quad f \geq 0, \quad \mathbb{E}_{a, b}(f) = 1.$$ 

Moreover, owing to [5], the log-Sobolev inequality is also equivalent to the hypercontractivity of $P_t$.

Thus, according to Theorem 3.1, we have constructed a diffusion process which converges to its reversible distribution $\mathbb{E}_{a, b}$ in entropy exponentially fast.

4. Measure-valued process

It was shown in [13] that the log-Sobolev inequality fails to hold for the Fleming–Viot process with parent independent mutation when there are an infinite number of types. In this section we will construct a class of measure-valued processes for which the log-Sobolev inequality holds even when the number of types is infinity.
Let us first consider a measure-valued process on a Polish space $S$ induced by the above constructed process and a proper Markov process on $S^N$. More precisely, let $X_t := (X_1(t), \ldots, X_n(t), \ldots)$ be the Markov process on $\Delta_\infty$ associated to $(\mathcal{E}, D(\mathcal{E}))$, and let $\xi_t := (\xi_1(t), \ldots, \xi_n(t), \ldots)$ be a Markov process on $S^N$, independent of $X_t$. We consider the measure-valued process

$$\eta_t := \sum_{i=1}^{\infty} X_i(t) \delta_{\xi_i(t)},$$

where $X_i$ can be viewed as the proportion of the $i$th family in the population, and $\xi_i$ can be viewed as its type or label. Then the above process describes the evolution of all (countably many) families on the space $S$. Let $M_1$ denote the set of all probability measures on $S$. Then the state space of this process is

$$M_0 := \{ \gamma \in M_1 : \text{supp} \gamma \text{ contains at most countably many points} \},$$

which is dense in $M_1$ under the weak topology.

Owing to Theorem 3.1, if $\xi_t$ converges to its unique invariant probability measure $\nu$ on $S^N$ then $\eta_t$ converges to $\Pi := (\Xi_{a,b} \times \nu) \circ \psi^{-1}$ for

$$\psi : \Delta_\infty \times S^N \rightarrow M_0, \quad \psi(x, \xi) := \sum_{i=1}^{\infty} x_i \delta_{\xi_i}.$$  

Unfortunately the process $\eta_t$ is in general non-Markovian. So we like to modify the construction using Dirichlet forms.

Let $\nu$ denote a probability measure on $S^N$ and $(\mathcal{E}_{\mathcal{M}_0}, D(\mathcal{E}_{\mathcal{M}_0}))$ denote a conservative symmetric Dirichlet form on $L^2(\nu)$. We then construct the corresponding quadratic form on $L^2(M_0; \Pi)$ as follows:

$$E_{\mathcal{M}_0}(F, G) := \int_{S^N} \mathcal{E}(F, G) \nu(d\xi) + \int_{\Delta_\infty} \mathcal{E}_{\mathcal{M}_0}(F, G) \nu(d\xi),$$

$$F, G \in D(\mathcal{E}_{\mathcal{M}_0})$$

$$:= \{ H \in L^2(\Pi) : H_x := H \circ \psi(x, \cdot) \in D(\mathcal{E}_{\mathcal{M}_0}) \text{ for } \Xi_{a,b} \text{-almost surely (a.s.) } x, \}$$

$$H_\xi := H \circ \psi(\cdot, \xi) \in D(\mathcal{E}) \text{ for } \nu \text{-a.s. } \xi, \text{ such that } \mathcal{E}_{\mathcal{M}_0}(H, H) < \infty \}.$$  

Since $\Pi$ has full mass on $M_0$, to make the state space complete we may also consider the above defined form to be a symmetric form on $L^2(M_0; \Pi) := L^2(M_0; \Pi)$.  

**Theorem 4.1.** Assume that there exists $\alpha > 0$ such that

$$\nu(f^2 \log f^2) \leq \frac{1}{\alpha} \mathcal{E}_{\mathcal{M}_0}(f, f) + \nu(f^2 \log f^2), \quad f \in D(\mathcal{E}_{\mathcal{M}_0}),$$

holds, then

$$\Pi(F^2 \log F^2) \leq \frac{1}{\alpha \wedge \beta_{a,b}} \mathcal{E}_{\mathcal{M}_0}(F, F) + \Pi(F^2 \log F^2), \quad F \in D(\mathcal{E}_{\mathcal{M}_0}). \quad (4.1)$$

Moreover, if $D(\mathcal{E}_{\mathcal{M}_0}) \subset L^2(M_0; \Pi)$ is dense then $(\mathcal{E}_{\mathcal{M}_0}, D(\mathcal{E}_{\mathcal{M}_0}))$ is a conservative Dirichlet form on $L^2(M_0; \Pi)$, so that the associated Markov semigroup $P_t$ satisfies

$$\Pi(P_t F \log P_t F) \leq \Pi(F \log F) \exp(-\beta_{a,b} \wedge \alpha t), \quad t \geq 0, \quad F \geq 0, \quad \Pi(F) = 1, \quad (4.2)$$
and \((\mathcal{E}_{\xi_M}, D(\mathcal{E}_{\xi_M}))\) is regular provided that the space \((\mathcal{E}_{S^N}, D(\mathcal{E}_{S^N}))\) is regular and \(S\) is compact.

**Proof.** Let

\[
D(\tilde{\mathcal{E}}) = \{ \tilde{F} \in L^2(\Xi_{a,b} \times \nu): \tilde{F}(x, \cdot) \in D(\mathcal{E}_{S^N}) \text{ for } \Xi_{a,b}\text{-a.s. } x, \tilde{F}(\cdot, \xi) \in D(\mathcal{E}) \text{ for } \nu\text{-a.s. } \xi, \text{ such that } \tilde{E}(\tilde{F}, \tilde{F}) < \infty \},
\]

where

\[
\tilde{E}(\tilde{F}, \tilde{G}) := \int_{\Delta_N} \mathcal{E}_{S^N}(\tilde{F}(x, \cdot)), \tilde{G}(x, \cdot)) d\nu(x) + \int_{S^N} \mathcal{E}(\tilde{F}(\cdot, \xi)), \tilde{G}(\cdot, \xi)) d\nu(\xi).
\]

Then \((\tilde{\mathcal{E}}, D(\tilde{\mathcal{E}}))\) is a symmetric Dirichlet form on \(L^2(\Delta_N \times S^N; \Xi_{a,b} \times \nu)\) and (see, e.g. [6, Theorem 2.3])

\[
(\Xi_{a,b} \times \nu)(\tilde{F}^2 \log \tilde{F}^2) \leq \frac{1}{\alpha a_{a,b} \vee \alpha}(\Xi_{a,b} \times \nu)(\tilde{F}^2), \quad \tilde{F} \in D(\tilde{\mathcal{E}}), (\Xi_{a,b} \times \nu)(\tilde{F}^2) = 1. \tag{4.3}
\]

Let \(\tilde{P}_t\) denote the Markov semigroup associated to \((\tilde{\mathcal{E}}, D(\tilde{\mathcal{E}}))\). Then (4.2) follows from the fact that \(\eta_t = \psi(X(t), \xi(t))\), and (4.3) implies that (see [1, Proposition 2.1])

\[
(\Xi_{a,b} \times \nu)(\tilde{P}_t G \log \tilde{P}_t G) \leq (\Xi_{a,b} \times \nu)\left( G \log G \right) \exp(-4(\alpha a_{a,b} \vee \alpha)t)
\]

for all \(t \geq 0\) and nonnegative function \(G\) with \((\Xi_{a,b} \times \nu)(G) = 1\). Since \(F \in D(\mathcal{E}_{\xi_M})\) if and only if \(F \circ \psi \in D(\tilde{\mathcal{E}})\), and

\[
\mathcal{E}_{\xi_M}(F, F) = \tilde{E}(F \circ \psi, F \circ \psi),
\]

(4.1) follows from (4.3). By the same reasoning and noting that \((\tilde{\mathcal{E}}, D(\tilde{\mathcal{E}}))\) is a Dirichlet form, we conclude that \((\mathcal{E}_{\xi_M}, D(\mathcal{E}_{\xi_M}))\) is a Dirichlet form provided it is densely defined on \(L^2(M_1; \Pi)\). Finally, if \(S\) is compact then so is \(M_1\) (under the weak topology). Thus, as explained in the proof of Theorem 3.1, for regular \((\mathcal{E}_{S^N}, D(\mathcal{E}_{S^N}))\), the set

\[
\{ f(\cdot, g_1), \ldots, f(\cdot, g_n): n \geq 1, f \in C^1_b(\mathbb{R}^n), g_i \in C(S), 1 \leq i \leq n \} \subset C_0(M_0) \cap D(\mathcal{E}_{\xi_M})
\]

is dense both in \(C_0(M_1)(= C(M_1))\) under the uniform norm and in \(D(\mathcal{E}_{\xi_M})\) under the Sobolev norm.

**Remark.** Obviously, we have a similar assertion for the Poincaré inequality: if there exists \(\lambda > 0\) such that

\[
\nu(f^2) \leq \frac{1}{\lambda} \mathcal{E}_{S^N}(f, f) + \nu(f)^2, \quad f \in D(\mathcal{E}_{S^N}),
\]

holds then

\[
\Pi(F^2) \leq \frac{1}{\lambda \wedge \inf_{i \geq 1}(a_i + b_i)} \mathcal{E}_{\xi_M}(F, F) + \Pi(F)^2, \quad F \in D(\mathcal{E}_{\xi_M}).
\]

To see that the above theorem applies to a class of measure-valued processes on \(S\), we present below a concrete condition on \(\mathcal{E}_{S^N}\) such that the assertions of Theorem 4.1 apply. In particular, it is the case if \(\mathcal{E}_{S^N}\) is the Dirichlet form of a particle system without interactions.
**Proposition 4.1.** Let \( \nu_i \) be the \( i \)th marginal distribution of \( \nu \) and, for a function \( g \) on \( S \), let \( g^{(i)}(\xi) := g(\xi_i), \ i \geq 1 \). Assume that

\[
S_0 := \left\{ g \in C_0(S) : g^{(i)} \in D(\mathcal{E}^{\mathbb{S}^n}), \ \sup_{i \geq 1} \mathcal{E}^{\mathbb{S}^n}(g^{(i)}, g^{(i)}) < \infty \right\}
\]

is dense in \( C_0(S) \). Then \((\mathcal{E}_{M_0}, D(\mathcal{E}_{M_0}))\) is a symmetric Dirichlet form.

**Proof.** Under the assumption and the fact that \( C^2_{cl}(\Delta^{\infty}) \) is dense in \( L^2(M_0; \Pi) \), the set

\[
S := \left\{ f(\langle \cdot, g_1 \rangle, \ldots, \langle \cdot, g_n \rangle) : n \geq 1, \ f \in C^1_b(\mathbb{R}^n), \ g_i \in S_0, \ 1 \leq i \leq n \right\}
\]

is dense in \( L^2(M_0; \Pi) \). Therefore, by Theorem 4.1 it suffices to show that \( S \subset D(\mathcal{E}) \); that is, for \( F := f(\langle \cdot, g_1 \rangle, \ldots, \langle \cdot, g_n \rangle) \in S \), we have \( F \circ \psi \in D(\tilde{\mathcal{E}}) \). Let

\[
F_m(x) = F\left( \sum_{i=1}^m x_i g_1(\xi_i), \ldots, \sum_{i=1}^m x_i g_n(\xi_i) \right), \quad x \in \Delta^{\infty}, \ m \geq 1.
\]

Since, for fixed \( \xi \in \mathbb{S}^n \),

\[
\partial_{\xi_i} F(\cdot, \xi)(x) = \sum_{k=1}^n \partial_k f g_k(\xi_i), \quad i \geq 1,
\]

is uniformly bounded, we have \( F_m \in D(\mathcal{E}) \) and (3.1) yields

\[
\mathcal{E}(F_m, F_m) \leq C
\]

for some constant \( C > 0 \) and all \( m \geq 1 \) and \( \xi \in \mathbb{S}^n \). Thus, \( F \circ \psi(\cdot, \xi) \in D(\mathcal{E}) \) for each \( \xi \in \mathbb{S}^n \) and

\[
\sup_{\xi} \mathcal{E}(F \circ \psi(\cdot, \xi), F \circ \psi(\cdot, \xi)) \leq C.
\]

Conversely, since \( g_k \in S_0 \), \( 1 \leq k \leq n \), noting that, for any \( x \in \Delta^{\infty} \),

\[
|F \circ \psi(x, \xi) - F \circ \psi(x, \xi')|^2 \leq \left( \sum_{k=1}^n \|\partial_k f\|_{\infty} \right)^2 \sum_{i=1}^\infty x_i |g_k(\xi_i) - g_k(\xi'_i)|^2.
\]

we conclude, in the spirit of Proposition I-4.10 of [9], that \( F \circ \psi(x, \cdot) \in D(\mathcal{E}^{\mathbb{S}^n}) \) and

\[
\mathcal{E}^{\mathbb{S}^n}(F \circ \psi(x, \cdot), F \circ \psi(x, \cdot)) \leq C'
\]

for some \( C' > 0 \) independent of \( x \). Combining this with (4.4) we obtain \( F \circ \psi \in D(\tilde{\mathcal{E}}) \).

**Acknowledgements**

The research of the first author was supported by NSERC of Canada. The research of the second author was supported by NNSFC(10121101), RFDP(20040027009), and the 973-Project of P. R. China.
References


