*J. Appl. Prob.* **44,** 938–949 (2007)
*Printed in England*


# 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 $\Delta_\infty := \{\boldsymbol{x} \in [0,1]^{\mathbb{N}} : \sum_{i\geq 1} x_i = 1\}$ 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

$$\Delta_\infty = \left\{\boldsymbol{x} = (x_1, x_2, \ldots) \in [0,1]^{\mathbb{N}} : \sum_{k=1}^{\infty} x_k = 1\right\},$$

and let

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

The Poisson–Dirichlet distribution with parameter $\theta > 0$ is a probability measure $\Pi_\theta$ on $\nabla$. We use $\boldsymbol{P}(\theta) = (P_1(\theta), P_2(\theta), \ldots)$ 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

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 $\mathrm{Beta}(1, \theta)$, and set

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

Clearly $(X_1^\theta, X_2^\theta, \ldots)$ is in space $\Delta_\infty$. The law of $X_1^\theta, X_2^\theta, \ldots$ is called the one-parameter GEM distribution and is denoted by $\Pi_\theta^{\mathrm{GEM}}$. The descending order of $X_1^\theta, X_2^\theta, \ldots$ has distribution $\Pi_\theta$. The sequence $X_k^\theta$, $k = 1, 2, \ldots$, 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 $\mathrm{Dirichlet}(\theta, \nu)$ distribution, and it is the reversible distribution of the Fleming–Viot process with mutation operator (see [2])

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

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

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

The law of $X_1^{\theta,\alpha}, X_2^{\theta,\alpha}, \ldots$ is called the two-parameter GEM distribution and is denoted by $\Pi_{\alpha,\theta}^{\mathrm{GEM}}$. The law of the descending order statistic of $X_1^{\theta,\alpha}, X_2^{\theta,\alpha}, \ldots$ is called the two-parameter Poisson–Dirichlet distribution (henceforth denoted by $\Pi_{\alpha,\theta}$), which was studied thoroughly in [12] . The sequence $X_k^{\theta,\alpha}$, $k = 1, 2, \ldots$, 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 $X_k^\theta$ being replaced by $X_k^{\theta,\alpha}$. We call the law of $\Theta_{\theta,\alpha,\nu}$ a $\mathrm{Dirichlet}(\theta, \alpha, \nu)$ 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 $\mathrm{Dirichlet}(\theta, \alpha, \nu)$ distribution.

In this paper we firstly construct a class of reversible infinite-dimensional diffusion processes, the GEM processes, so that both $\Pi_\theta^{\mathrm{GEM}}$ and its two-parameter generalization $\Pi_{\alpha,\theta}^{\mathrm{GEM}}$ 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

https://doi.org/10.1239/jap/1197908815 Published online by Cambridge University Press

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 $\Pi_\theta^{\text{GEM}}$ and $\Pi_{\alpha,\theta}^{\text{GEM}}$ 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

$$\inf_i b_i \geq \tfrac{1}{2}. \tag{2.1}$$

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

$$\mathrm{d}X_i(t) = (a_i - (a_i + b_i)X_i(t))\,\mathrm{d}t + \sqrt{X_i(t)(1 - X_i(t))}\,\mathrm{d}B_i(t), \qquad X_i(0) \in [0, 1],$$

where $\{B_i(t) \colon i = 1, 2, \dots\}$ 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

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

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 $E = [0, 1)^{\mathbb{N}}$. The process

$$\boldsymbol{X}(t) = (X_1(t), X_2(t), \dots)$$

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

$$\Phi \colon E \to \bar{\Delta}_\infty, \qquad \boldsymbol{x} = (x_1, x_2, \dots) \to (\varphi_1(\boldsymbol{x}), \varphi_2(\boldsymbol{x}), \dots),$$

with

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

Clearly $\Phi$ is a bijection and the process $\boldsymbol{Y}(t) = \Phi(\boldsymbol{X}(t))$ 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 $C_{\text{cl}}^2(\bar{E})$ 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 $C_{\text{cl}}^2(E)$ will be the respective restrictions of $C(\bar{E})$ and $C_{\text{cl}}^2(\bar{E})$ on $E$. Then the generator of process $\boldsymbol{X}(t)$ is given by

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

and can be extended to $C_{\text{cl}}^2(\bar{E})$. The sets $B(E)$ and $B(\Delta_\infty)$ are bounded measurable functions on $E$ and $\Delta_\infty$, respectively.

https://doi.org/10.1239/jap/1197908815 Published online by Cambridge University Press

Let $\boldsymbol{a} = (a_1, a_2, \ldots)$ and $\boldsymbol{b} = (b_1, b_2, \ldots)$, and let

$$\mu_{\boldsymbol{a},\boldsymbol{b}} = \prod_{k=1}^{\infty} \pi_{a_k,b_k} \quad \text{and} \quad \Xi_{\boldsymbol{a},\boldsymbol{b}} = \mu_{\boldsymbol{a},\boldsymbol{b}} \circ \Phi^{-1}.$$

Then we have the following result.

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

*Proof.* The reversibility of $\boldsymbol{X}(t)$ 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 $\boldsymbol{X}(t)$, we have, for any $t > 0$,

$$\begin{aligned}\int_{\Delta_\infty} f(\boldsymbol{y}) E_{\boldsymbol{y}}[g(\boldsymbol{y}(t))] \Xi_{\boldsymbol{a},\boldsymbol{b}}(\mathrm{d}\boldsymbol{y}) &= \int_E f(\Phi(\boldsymbol{x})) E_{\boldsymbol{x}}[g(\Phi(\boldsymbol{x}(t)))] \mu_{\boldsymbol{a},\boldsymbol{b}}(\mathrm{d}\boldsymbol{x}) \\ &= \int_E g(\Phi(\boldsymbol{x})) E_{\boldsymbol{x}}[f(\Phi(\boldsymbol{x}(t)))] \mu_{\boldsymbol{a},\boldsymbol{b}}(\mathrm{d}\boldsymbol{x}) \\ &= \int_{\Delta_\infty} g(\boldsymbol{y}) E_{\boldsymbol{y}}[f(\boldsymbol{y}(t))] \Xi_{\boldsymbol{a},\boldsymbol{b}}(\mathrm{d}\boldsymbol{y}).\end{aligned}$$

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

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

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

Let

$$\bar{\Delta}_\infty := \left\{ \boldsymbol{x} \in [0,1]^{\mathbb{N}} \colon \sum_{i=1}^{\infty} x_i \le 1 \right\}$$

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

Now, for $\boldsymbol{x} \in \bar{\Delta}_\infty$ such that

$$\sum_{i=1}^{n} x_i < 1 \quad \text{for all finite } n,$$

let

$$L(\boldsymbol{x}) = \sum_{i,j=1}^{\infty} a_{ij}(\boldsymbol{x}) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{\infty} b_i(\boldsymbol{x}) \frac{\partial}{\partial x_i},$$

where

$$a_{ij}(\boldsymbol{x}) := x_i x_j \sum_{k=1}^{i\wedge j} \frac{(\delta_{ki}(1-\sum_{l=1}^{k-1} x_l) - x_k)(\delta_{kj}(1-\sum_{l=1}^{k-1} x_l) - x_k)}{x_k(1-\sum_{l=1}^{k} x_l)},$$

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

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

It follows, from direct calculation, that

$$\sum_{i,j=1}^{\infty} |a_{ij}(x)| \le 3, \qquad |b_i(x)| \le \sum_{k=1}^{i} (b_k x_k + a_k), \qquad x \in \bar{\Delta}_\infty. \tag{3.1}$$

Indeed, since $1 - \sum_{l=1}^{i-1} x_l \ge x_i$ and $\sum_{1\le i<j<\infty} x_i x_j \le \frac{1}{2}$, we obtain

$$\begin{aligned}
\sum_{i,j=1}^{\infty} |a_{ij}(x)| &= \sum_{i=1}^{\infty} a_{ii}(x) + 2 \sum_{1\le i<j<\infty} |a_{ij}(x)| \\
&\le \sum_{i=1}^{\infty} x_i^2 \bigg( \frac{1-\sum_{l=1}^{i} x_l}{x_i} + \sum_{k=1}^{i-1} \frac{x_k}{1-\sum_{l=1}^{k} x_l} \bigg) \\
&\quad + 2 \sum_{1\le i<j<\infty} x_i x_j \bigg( 1 + \sum_{k=1}^{i-1} \frac{x_k}{1-\sum_{l=1}^{k} x_l} \bigg) \\
&\le \sum_{i=1}^{\infty} x_i \bigg( 1 - \sum_{l=1}^{i} x_l + \sum_{k=1}^{i-1} x_k \bigg) \\
&\quad + 2 \sum_{i=1}^{\infty} x_i \sum_{j=i+1}^{\infty} x_j \bigg( 1 + \frac{\sum_{k=1}^{i-1} x_k}{\sum_{l=i+1}^{\infty} x_l} \bigg) \\
&\le 1 + 2 \\
&= 3.
\end{aligned}$$

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 $\Gamma(f,f) \in C_b(\bar{\Delta}_\infty)$ for any $f \in C_b^1(\bar{\Delta}_\infty)$.

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,

$$\pi_{a,b}(f^2 \log f^2) \le \frac{1}{\alpha_{a,b}} \int_0^1 x(1-x) f'(x)^2 \pi_{a,b}(\mathrm{d}x) + \pi_{a,b}(f^2) \log \pi_{a,b}(f^2), \tag{3.2}$$

holds. According to Lemma 2.7 of [13], we have $\alpha_{a,b} \geq (a \wedge b)/320$. Moreover, it is easy to see that, for $a, b > 0$, the operator

$$r(1-r)\frac{\mathrm{d}^2}{\mathrm{d}r^2} + (a - (a+b)r)\frac{\mathrm{d}}{\mathrm{d}r}$$

on $[0, 1]$ has a spectral gap $a + b$ with eigenfunction $h(r) := a - (a+b)r$. So, the Poincaré inequality,

$$\pi_{a,b}(f^2) \leq \frac{1}{a+b}\int_0^1 x(1-x)f'(x)^2 \pi_{a,b}(\mathrm{d}x) + \pi_{a,b}(f)^2, \tag{3.3}$$

holds.

Let $C_{\mathrm{cl}}^\infty([0,1]^{\mathbb{N}})$ denote the set of all bounded, $C^\infty$ cylindrical functions on $[0,1]^{\mathbb{N}}$, and

$$FC_b^\infty = \{f|_{\bar{\Delta}_\infty} : f \in C_{\mathrm{cl}}^\infty([0,1]^{\mathbb{N}})\}.$$

Now we have the following theorem.

**Theorem 3.1.** *For any $f, g \in FC_b^\infty$, we have*

$$\mathcal{E}(f,g) := \Xi_{\boldsymbol{a},\boldsymbol{b}}(\Gamma(f,g)) = -\Xi_{\boldsymbol{a},\boldsymbol{b}}(fLg). \tag{3.4}$$

*Consequently, $(\mathcal{E}, FC_b^\infty)$ is closable in $L^2(\bar{\Delta}_\infty; \Xi_{\boldsymbol{a},\boldsymbol{b}})$, and its closure is a conservative regular Dirichlet form which satisfies the Poincaré inequality*

$$\Xi_{\boldsymbol{a},\boldsymbol{b}}(f^2) \leq \frac{1}{\inf_{i\geq 1}(a_i+b_i)}\mathcal{E}(f,f), \qquad f \in D(\mathcal{E}),\ \Xi_{\boldsymbol{a},\boldsymbol{b}}(f) = 0.$$

*Moreover, if* $\inf\{a_i \wedge b_i : i \geq 1\} > 0$, *the log-Sobolev inequality*

$$\Xi_{\boldsymbol{a},\boldsymbol{b}}(f^2 \log f^2) \leq \frac{1}{\beta_{\boldsymbol{a},\boldsymbol{b}}}\mathcal{E}(f,f), \qquad f \in D(\mathcal{E}),\ \Xi_{\boldsymbol{a},\boldsymbol{b}}(f^2) = 1, \tag{3.5}$$

*holds for some* $\beta_{\boldsymbol{a},\boldsymbol{b}} \geq \inf\{(a_i \wedge b_i)/320 : i \geq 1\} > 0$.

*Proof.* For any $f, g \in FC_b^\infty$, there exists $n \geq 1$ such that

$$f(\boldsymbol{x}) = f(x_1,\dots,x_n), \qquad g(\boldsymbol{x}) = g(x_1,\dots,x_n), \\ \boldsymbol{x} = (x_1,\dots,x_n,\dots) \in [0,1]^{\mathbb{N}}. \tag{3.6}$$

Let

$$\varphi^{(n)}(\boldsymbol{x}) = (\varphi_1(\boldsymbol{x}),\dots,\varphi_n(\boldsymbol{x})),$$

which maps $[0,1]^n$ onto $\Delta_n := \{x \in [0,1]^n : \sum_{i=1}^n x_i \leq 1\}$. Define

$$L_n := \sum_{i=1}^n x_i(1-x_i)\frac{\partial}{\partial x_i^2} + \sum_{i=1}^n (a_i - (a_i+b_i)x_i)\frac{\partial}{\partial x_i},$$

and

$$\pi_{\boldsymbol{a},\boldsymbol{b}}^n = \prod_{i=1}^n \pi_{a_i,b_i} \quad \text{and} \quad \Xi^n = \pi_{\boldsymbol{a},\boldsymbol{b}}^n \circ \varphi^{(n)^{-1}}.$$

Then, regarding $\{\Xi^n := \pi^n_{\boldsymbol{a},\boldsymbol{b}} \circ \varphi^{(n)^{-1}} : n \geq 1\}$ as probability measures on $\bar{\Delta}_\infty$ and by letting $\Xi^n := \Xi^n(\mathrm{d}x_1, \ldots, \mathrm{d}x_n) \times \delta_0(\mathrm{d}x_{n+1}, \ldots)$, it converges weakly to $\Xi_{\boldsymbol{a},\boldsymbol{b}}$. Since $L_n$ is symmetric with respect to $\pi^n_{\boldsymbol{a},\boldsymbol{b}}$, we have

$$\begin{aligned}\int_{[0,1]^n} \sum_{i=1}^n x_i(1-x_i)\bigg(\frac{\partial}{\partial x_i} f \circ \varphi^{(n)}\bigg)\bigg(\frac{\partial}{\partial x_i} g \circ \varphi^{(n)}\bigg)\,\mathrm{d}\pi^n_{\boldsymbol{a},\boldsymbol{b}} \\ = -\int_{[0,1]^n} g \circ \varphi^{(n)} L_n f \circ \varphi^{(n)}\,\mathrm{d}\pi^n_{\boldsymbol{a},\boldsymbol{b}}.\end{aligned}$$

Noting that

$$\varphi_i(\boldsymbol{x}) = x_i \prod_{l=1}^{i-1}(1-x_l) \quad \text{and} \quad x_i = \frac{\varphi_i(\boldsymbol{x})}{1-\sum_{l=1}^{i-1}\varphi_l(\boldsymbol{x})}, \qquad i \geq 1,$$

we have

$$\frac{\mathrm{d} f \circ \varphi^{(n)}(\boldsymbol{x})}{\mathrm{d}x_i} = \sum_{j \geq i} \frac{(\delta_{ij} - x_i)\varphi_j(\boldsymbol{x})}{x_i(1-x_i)} \frac{\mathrm{d} f}{\mathrm{d}\varphi_j} \circ \varphi^{(n)}(\boldsymbol{x}).$$

Therefore,

$$\begin{aligned}\int_{[0,1]^n} \sum_{i=1}^n x_i(1-x_i)\bigg(\frac{\partial}{\partial x_i} f \circ \varphi^{(n)}\bigg)\bigg(\frac{\partial}{\partial x_i} g \circ \varphi^{(n)}\bigg)\,\mathrm{d}\pi^n_{\boldsymbol{a},\boldsymbol{b}} \\ = \int_{[0,1]^n} \Gamma(f,g) \circ \varphi^{(n)}\,\mathrm{d}\pi^n_{\boldsymbol{a},\boldsymbol{b}} \\ = \int_{\Delta_n} \Gamma(f,g)\,\mathrm{d}\Xi^n.\end{aligned} \tag{3.7}$$

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

$$\lim_{n\to\infty} \int_{\Delta_n} \Gamma(f,g)\,\mathrm{d}\Xi^n = \int_{\bar{\Delta}_\infty} \Gamma(f,g)\,\mathrm{d}\Xi_{\boldsymbol{a},\boldsymbol{b}}. \tag{3.8}$$

Similarly, by straightforward calculations we find that

$$L_n f \circ \varphi^{(n)}(\boldsymbol{x}) = (Lf) \circ \varphi^{(n)}(\boldsymbol{x}).$$

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

$$\lim_{n\to\infty} \int_{\Delta_n} g \circ \varphi^{(n)} L_n f \circ \varphi^{(n)}\,\mathrm{d}\pi^n_{\boldsymbol{a},\boldsymbol{b}} = \int_{\bar{\Delta}_\infty} gLf\,\mathrm{d}\Xi_{\boldsymbol{a},\boldsymbol{b}}.$$

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 $\bar{\Delta}_\infty$ under the usual metric

$$\rho(\boldsymbol{x}, \boldsymbol{y}) := \sum_{i=1}^\infty 2^{-i}|x_i - y_i|.$$

https://doi.org/10.1239/jap/1197908815 Published online by Cambridge University Press

Indeed, it is trivial that $D(\mathcal{E})\cap C_0([0,1]^{\mathbb{N}})\supset FC_b^\infty$, which is dense in $D(\mathcal{E})$ under $\mathcal{E}_1^{1/2}$ given by

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

Moreover, for any $F\in C(\bar{\Delta}_\infty)=C_0(\bar{\Delta}_\infty)$, by its uniform continuity owing to the compactness of the space,

$$\bar{\boldsymbol{\Delta}}_\infty\ni \boldsymbol{x}\mapsto F_n(\boldsymbol{x}):=F(x_1,\dots,x_n,0,0,\dots),\qquad n\ge 1,$$

is a sequence of continuous cylindric functions converging uniformly to $F$. Since a cylindric continuous function can be uniformly approximated by functions in $FC_b^\infty$ under the uniform norm, it follows that $FC_b^\infty$ is dense in $C_0(\bar{\Delta}_\infty)$ 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]),

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

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

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

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

$$\Xi^n({f^{(n)}}^2\log {f^{(n)}}^2)\le\frac{1}{\beta^n_{\boldsymbol{a},\boldsymbol{b}}}\int_{\Delta_n}\Gamma^{(n)}(f,f)\,\mathrm{d}\Xi^n+\Xi^n({f^{(n)}}^2)\log\Xi^n({f^{(n)}}^2).$$

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 $\mathcal{E}_1^{1/2}$.

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 $L^2(\Xi_{\boldsymbol{a},\boldsymbol{b}})$. 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 $\Xi_{\boldsymbol{a},\boldsymbol{b}}$ exponentially fast in entropy; more precisely (see, e.g. [1, Proposition 2.1]),

$$\Xi_{\boldsymbol{a},\boldsymbol{b}}(P_tf\log P_tf)\le\exp(-4\beta_{\boldsymbol{a},\boldsymbol{b}}t)\Xi_{\boldsymbol{a},\boldsymbol{b}}(f\log f),\qquad f\ge 0,\ \Xi_{\boldsymbol{a},\boldsymbol{b}}(f)=1.$$

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 $\Xi_{\boldsymbol{a},\boldsymbol{b}}$ 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.

https://doi.org/10.1239/jap/1197908815 Published online by Cambridge University Press

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 $S^{\mathbb{N}}$. More precisely, let $X_t := (X_1(t), \dots, X_n(t), \dots)$ be the Markov process on $\Delta_\infty$ associated to $(\mathcal{E}, D(\mathcal{E}))$, and let $\xi_t := (\xi_1(t), \dots, \xi_n(t), \dots)$ be a Markov process on $S^{\mathbb{N}}$, 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 $\mathrm{M}_1$ denote the set of all probability measures on $S$. Then the state space of this process is

$$\mathrm{M}_0 := \{\gamma \in \mathrm{M}_1 \colon \operatorname{supp}\gamma \text{ contains at most countably many points}\},$$

which is dense in $\mathrm{M}_1$ under the weak topology.

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

$$\psi \colon \Delta_\infty \times S^{\mathbb{N}} \to \mathrm{M}_0, \qquad \psi(\boldsymbol{x}, \xi) := \sum_{i=1}^{\infty} x_i \delta_{\xi_i}.$$

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

$$\mathcal{E}_{\mathrm{M}_0}(F, G) := \int_{S^{\mathbb{N}}} \mathcal{E}(F_\xi, G_\xi)\nu(\mathrm{d}\xi) + \int_{\Delta_\infty} \mathcal{E}_{S^{\mathbb{N}}}(F_{\boldsymbol{x}}, G_{\boldsymbol{x}})\pi_{a,b}(\mathrm{d}\boldsymbol{x}),$$

$$\begin{aligned} F, G \in D(\mathcal{E}_{\mathrm{M}_0}) &:= \{H \in L^2(\Pi) \colon H_{\boldsymbol{x}} := H \circ \psi(x, \cdot) \in D(\mathcal{E}_{S^{\mathbb{N}}}) \text{ for } \Xi_{\boldsymbol{a},\boldsymbol{b}}\text{-almost surely (a.s.) } \boldsymbol{x}, \\ &\qquad H_\xi := H \circ \psi(\cdot, \xi) \in D(\mathcal{E}) \text{ for } \nu\text{-a.s. } \xi, \text{ such that } \mathcal{E}_{\mathrm{M}_0}(H, H) < \infty\}. \end{aligned}$$

Since $\Pi$ has full mass on $\mathrm{M}_0$, to make the state space complete we may also consider the above defined form to be a symmetric form on $L^2(\mathrm{M}_1; \Pi)$ $(= L^2(\mathrm{M}_0; \Pi))$.

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

$$\nu(f^2 \log f^2) \le \frac{1}{\alpha}\mathcal{E}_{S^{\mathbb{N}}}(f, f) + \nu(f^2)\log\nu(f^2), \qquad f \in D(\mathcal{E}_{S^{\mathbb{N}}}),$$

*holds, then*

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

*Moreover, if* $D(\mathcal{E}_{\mathrm{M}_0}) \subset L^2(\mathrm{M}_1; \Pi)$ *is dense then* $(\mathcal{E}_{\mathrm{M}_0}, D(\mathcal{E}_{\mathrm{M}_0}))$ *is a conservative Dirichlet form on* $L^2(\mathrm{M}_0; \Pi)$*, so that the associated Markov semigroup* $P_t$ *satisfies*

$$\Pi(P_t F \log P_t F) \le \Pi(F \log F)\exp(-(\beta_{\boldsymbol{a},\boldsymbol{b}} \wedge \alpha)t), \qquad t \ge 0,\ F \ge 0,\ \Pi(F) = 1, \tag{4.2}$$

https://doi.org/10.1239/jap/1197908815 Published online by Cambridge University Press

*and* $(\mathcal{E}_{\mathrm{M}_0}, D(\mathcal{E}_{\mathrm{M}_0}))$ *is regular provided that the space* $(\mathcal{E}_{S^{\mathbb{N}}}, D(\mathcal{E}_{S^{\mathbb{N}}}))$ *is regular and* $S$ *is compact.*

*Proof.* Let

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

where

$$\tilde{\mathcal{E}}(\tilde{F}, \tilde{G}) := \int_{\Delta_\infty} \mathcal{E}_{S^{\mathbb{N}}}(\tilde{F}(\boldsymbol{x}, \cdot), \tilde{G}(\boldsymbol{x}, \cdot))\Xi_{\boldsymbol{a},\boldsymbol{b}}(\mathrm{d}\boldsymbol{x}) + \int_{S^{\mathbb{N}}} \mathcal{E}(\tilde{F}(\cdot, \xi), \tilde{G}(\cdot, \xi))\nu(\mathrm{d}\xi).$$

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

$$(\Xi_{\boldsymbol{a},\boldsymbol{b}} \times \nu)(\tilde{F}^2 \log \tilde{F}^2) \leq \frac{1}{\beta_{\boldsymbol{a},\boldsymbol{b}} \wedge \alpha}(\Xi_{\boldsymbol{a},\boldsymbol{b}} \times \nu)(\tilde{F}^2), \qquad \tilde{F} \in D(\tilde{\mathcal{E}}),\ (\Xi_{\boldsymbol{a},\boldsymbol{b}} \times \nu)(\tilde{F}^2) = 1. \tag{4.3}$$

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])

$$(\Xi_{\boldsymbol{a},\boldsymbol{b}} \times \nu)(\tilde{P}_t G \log \tilde{P}_t G) \leq (\Xi_{\boldsymbol{a},\boldsymbol{b}} \times \nu)(G \log G) \exp(-4(\beta_{\boldsymbol{a},\boldsymbol{b}} \wedge \alpha)t)$$

for all $t \geq 0$ and nonnegative function $G$ with $(\Xi_{\boldsymbol{a},\boldsymbol{b}} \times \nu)(G) = 1$. Since $F \in D(\mathcal{E}_{\mathrm{M}_0})$ if and only if $F \circ \psi \in D(\tilde{\mathcal{E}})$, and

$$\mathcal{E}_{\mathrm{M}_0}(F, F) = \tilde{\mathcal{E}}(F \circ \psi, F \circ \psi),$$

(4.1) follows from (4.3). By the same reasoning and noting that $(\tilde{\mathcal{E}}, D(\mathrm{E}))$ is a Dirichlet form, we conclude that $(\mathcal{E}_{\mathrm{M}_1}, D(\mathcal{E}_{\mathrm{M}_0}))$ is a Dirichlet form provided it is densely defined on $L^2(\mathrm{M}_1; \Pi)$. Finally, if $S$ is compact then so is $\mathrm{M}_1$ (under the weak topology). Thus, as explained in the proof of Theorem 3.1, for regular $(\mathcal{E}_{S^{\mathbb{N}}}, D(\mathcal{E}_{S^{\mathbb{N}}}))$, the set

$$\{f(\langle \cdot, g_1\rangle, \ldots, \langle \cdot, g_n\rangle) \colon n \geq 1,\ f \in C_b^1(\mathbb{R}^n),\ g_i \in C(S),\ 1 \leq i \leq n\} \subset C_0(\mathrm{M}_0) \cap D(\mathcal{E}_{\mathrm{M}_1})$$

is dense both in $C_0(\mathrm{M}_1)(= C(\mathrm{M}_1))$ under the uniform norm and in $D(\mathcal{E}_{\mathrm{M}_1})$ under the Sobolev norm.

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

$$\nu(f^2) \leq \frac{1}{\lambda}\mathcal{E}_{S^{\mathbb{N}}}(f, f) + \nu(f)^2, \qquad f \in D(\mathcal{E}_{S^{\mathbb{N}}}),$$

holds then

$$\Pi(F^2) \leq \frac{1}{\lambda \wedge \inf_{i \geq 1}(a_i + b_i)}\mathcal{E}_{\mathrm{M}_0}(F, F) + \Pi(F)^2, \qquad F \in D(\mathcal{E}_{\mathrm{M}_0}).$$

To see that the above theorem applies to a class of measure-valued processes on $S$, we present below a concrete condition on $\mathcal{E}_{S^{\mathbb{N}}}$ such that the assertions of Theorem 4.1 apply. In particular, it is the case if $\mathcal{E}_{S^{\mathbb{N}}}$ 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*

$$S_0 := \left\{ g \in C_0(S) \colon g^{(i)} \in D(\mathcal{E}_{S^{\mathbb{N}}}),\ \sup_{i \geq 1} \mathcal{E}_{S^{\mathbb{N}}}(g^{(i)}, g^{(i)}) < \infty \right\}$$

*is dense in* $C_0(S)$*. Then* $(\mathcal{E}_{\mathrm{M}_0}, D(\mathcal{E}_{\mathrm{M}_0}))$ *is a symmetric Dirichlet form.*

*Proof.* Under the assumption and the fact that $C^2_{\mathrm{cl}}(\Delta_\infty)$ is dense in $L^2(\mathrm{M}_0; \Pi)$, the set

$$S := \{ f(\langle \cdot, g_1 \rangle, \dots, \langle \cdot, g_n \rangle) \colon n \geq 1,\ f \in C_b^1(\mathbb{R}^n),\ g_i \in S_0,\ 1 \leq i \leq n \}$$

is dense in $L^2(\mathrm{M}_0; \Pi)$. Therefore, by Theorem 4.1 it suffices to show that $S \subset D(\mathcal{E}_{\mathrm{M}_0})$; that is, for $F := f(\langle \cdot, g_1 \rangle, \dots, \langle \cdot, g_n \rangle) \in S$, we have $F \circ \psi \in D(\tilde{\mathcal{E}})$. Let

$$F_m(\boldsymbol{x}) = F\left( \sum_{i=1}^m x_i g_1(\xi_i), \dots, \sum_{i=1}^m x_i g_n(\xi_i) \right), \qquad \boldsymbol{x} \in \Delta_\infty,\ m \geq 1.$$

Since, for fixed $\xi \in S^{\mathbb{N}}$,

$$\partial_{x_i} F \circ \psi(\cdot, \xi)(\boldsymbol{x}) = \sum_{k=1}^n \partial_k f g_k(\xi_i), \qquad i \geq 1,$$

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 $\xi \in S^{\mathbb{N}}$. Thus, $F \circ \psi(\cdot, \xi) \in D(\mathcal{E})$ for each $\xi \in S^{\mathbb{N}}$ and

$$\sup_{\xi} \mathcal{E}(F \circ \psi(\cdot, \xi),\ F \circ \psi(\cdot, \xi)) \leq C. \tag{4.4}$$

Conversely, since $g_k \in S_0,\ 1 \leq k \leq n$, noting that, for any $\boldsymbol{x} \in \Delta_\infty$,

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

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

$$\mathcal{E}_{S^{\mathbb{N}}}(F \circ \psi(\boldsymbol{x}, \cdot),\ F \circ \psi(\boldsymbol{x}, \cdot)) \leq C'$$

for some $C' > 0$ independent of $\boldsymbol{x}$. 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

[1] Bakry, D. (1997). On Sobolev and logarithmic Sobolev inequalities for Markov semigroups. In *New Trends in Stochastic Analysis*, eds K. D. Elworthy *et al.*, World Scientific, Singapore, pp. 43–75.
[2] Ethier, S. N. and Kurtz, T. G. (1981). The infinitely-many-neutral-alleles diffusion model. *Adv. Appl. Prob.* **13,** 429–452.
[3] Feng, S. and Hoppe, F. M. (1998). Large deviation principles for some random combinatorial structures in population genetics and Brownian motion. *Ann. Appl. Prob.* **8,** 975–994.
[4] Fukushima, M., Oshima, Y. and Takeda, M. (1994). *Dirichlet Forms and Symmetric Markov Processes*. De Gruyter, Berlin.
[5] Gross, L. (1976). Logarithmic Sobolev inequalities. *Amer. J. Math.* **97,** 1061–1083.
[6] Gross, L. (1993). *Logarithmic Sobolev Inequalities and Contractivity Properties of Semigroups* (Lecture Notes Math. **1563**). Springer, Berlin.
[7] Hoppe, F. M. (1984). Pólya-like urns and the Ewens sampling formula. *J. Math. Biol.* **20,** 91–94.
[8] Kingman, J. F. C. *et al.* (1975). Random discrete distribution. *J. R. Statist. Soc. Ser. B* **37,** 1–22.
[9] Ma, Z. M. and Röckner, M. (1992). *Introduction to the Theory of (Nonsymmetric) Dirichlet Forms.* Springer, Berlin.
[10] Pitman, J. (1995). Exchangeable and partially exchangeable random partitions. *Prob. Theory Relat. Fields* **102,** 145–158.
[11] Pitman, J. (1996). Random discrete distributions invariant under size-biased permutation. *Adv. Appl. Prob.* **28,** 525–539.
[12] Pitman, J. and Yor, M. (1997). The two-parameter Poisson–Dirichlet distribution derived from a stable subordinator. *Ann. Prob.* **25,** 855–900.
[13] Stannat, W. (2000). On the validity of the log-Sobolev inequality for symmetric Fleming–Viot operators. *Ann. Prob.* **28,** 667–684.