The Annals of Probability
1995, Vol. 23, No. 1, 1–36

AN ANALYTIC APPROACH TO FLEMING–VIOT PROCESSES WITH INTERACTIVE SELECTION

BY LUDGER OVERBECK, MICHAEL RÖCKNER AND BYRON SCHMULAND

Universität Bonn, Universität Bonn and University of Alberta

We study a class of (nonsymmetric) Dirichlet forms $(\mathscr{E}, D(\mathscr{E}))$ having a space of measures as state space E and derive some general results about them. We show that under certain conditions they "generate" diffusion processes $\mathbf{M}$. In particular, if $\mathbf{M}$ is ergodic and $(\mathscr{E}, D(\mathscr{E}))$ is symmetric w.r.t. quasi-every starting point, the large deviations of the empirical distribution of $\mathbf{M}$ are governed by $\mathscr{E}$. We apply all of this to construct Fleming–Viot processes with interactive selection and prove some results on their behavior. Among other things, we show some support properties for these processes using capacitary methods.

1. Introduction. The aims of this paper are the following: The first is the construction and analysis of a certain class of measure-valued diffusions using techniques from the theory of Dirichlet forms. The second is a thorough study of the underlying class of Dirichlet forms that serve as a case study for Dirichlet forms on infinite dimensional "nonflat manifolds" and also as a basis of an infinite dimensional calculus for measure-valued processes.

The theory of Dirichlet forms became available for the measure-valued diffusions studied in this paper due to recent developments in [1, 20] that extend the classical theory of Fukushima [14, 15], Silverstein [27], Carrillo-Menendez [5], Le Jan [18] and Boulean and Hirsch [4]. The classical theory deals with regular Dirichlet forms on locally compact state spaces, whereas the extension covers arbitrary state spaces and the regularity condition is dropped. Regularity is replaced by another analytical condition called *quasi-regularity*, which is the appropriate condition because it was shown [1, 20] to be necessary and sufficient for the existence of an associated nice Markov process. All this in turn has been extended to semi-Dirichlet forms [19]. (See Section 2 for the terminology.)

The class of semi-Dirichlet forms analyzed in this paper is of the type (closure of)

$$\mathscr{E}(u,v) := \int_E \big[\langle \nabla u(\mu), \nabla v(\mu)\rangle_\mu + \langle b(\mu), \nabla u(\mu)\rangle_\mu v(\mu) + \alpha u(\mu) v(\mu)\big] m(d\mu), \tag{1.1}$$

where u, v are finitely based smooth functions on $L^2(E; m)$. Here E denotes the set of all probability measures on a Polish space S, m is a probability

Received July 1993.
AMS 1991 *subject classifications*. Primary 60G57; secondary 31C25, 60K35, 92O15, 60J45.
Key words and phrases. Fleming-Viot processes, Dirichlet forms, exceptional sets.

measure on the Borel σ-algebra on E (generated by the weak topology), α is a sufficiently large constant and for bounded measurable functions f, g on S and $\mu \in E$, we set

$$\langle f, g\rangle_\mu = \int fg\,d\mu - \int f\,d\mu \int g\,d\mu.$$

In (1.1), $b\colon S \times E \to \mathbb{R}$ is a measurable map such that $\sup_\mu \langle b(\cdot, \mu), b(\cdot, \mu)\rangle_\mu < \infty$ and the gradient ∇u of a finitely based smooth function $u(\mu) = \varphi(\int f_1\,d\mu, \dots, \int f_k\,d\mu)$, $\mu \in E$, $k \in \mathbb{N}$, $\varphi \in C_b^\infty(\mathbb{R}^k)$, $f_1, \dots, f_k \in C_b(S)$, is defined by

$$\nabla u(\mu) := \left(\frac{d}{ds} u(\mu + s\varepsilon_x) \Big|_{s=0} \right)_{x \in S}$$

(with ε_x equal to the Dirac measure at x); hence $\nabla u(\mu) \in L^2(S;\mu)$. Geometrically one can think of E as an infinite dimensional manifold that is the base space of a (tangent) bundle with (Hilbert) fiber $(L^2(S;\mu), \langle\ ,\ \rangle_\mu)$.

The main results of Sections 2 and 3 of this paper state that provided the form (1.1) is closable, its closure is indeed a semi-Dirichlet form. It is local and quasi-regular; hence, it gives rise to an associated diffusion process with state space E (cf. Theorems 2.4, and 3.5).

The class of measure-valued diffusions that we obtain, in particular, contains a new type of Fleming–Viot process. A Fleming–Viot process is a (probability) measure-valued process that describes the genetic evolution of a population. It is known that the Fleming–Viot process with neutral mutation has a unique reversible probability measure m^{FV} (cf. [13]). As an application of our general results, we adopt this measure as a starting point by taking m in (1.1) to be equal to $\varphi^2 \cdot m^{\mathrm{FV}}$ for some "weakly differentiable" (not necessarily bounded or continuous functions) $\varphi\colon E \to \mathbb{R}$. If L^{FV} denotes the generator of the Fleming–Viot process with neutral mutation, the new diffusion process associated with $\mathscr{E}$ in (1.1) will have a generator on $L^2(E;m)$ given by

$$Lu(\mu) = L^{\mathrm{FV}}u(\mu) + \langle \varphi^{-1}\nabla\varphi(\mu) - b(\mu), \nabla u(\mu)\rangle_\mu - \alpha u(\mu)$$

for bounded smooth cylinder functions $u\colon E \to \mathbb{R}$. Therefore, we call these new processes Fleming–Viot processes with neutral mutation and *interactive selection*; they are perturbations of the classical Fleming–Viots. One of the first to consider Fleming–Viot processes with interactive selection was Shiga [28]. He constructed generalized Fleming–Viot processes with bounded selection, whereas we allow unbounded selection, but we have to assume neutral mutation. For another recent work on generalized Fleming–Viot processes we refer to [7], and for excellent recent survey papers on this subject to [6] and [13]. The details on all this are contained in Section 5.

Sections 6 and 7 are devoted to a more detailed analysis of the processes related to the form $\mathscr{E}$ in (1.1). There, we study support properties of these measure-valued diffusions using the capacity of the associated Dirichlet form. This is done both in the general case (Section 6) and in the special case of our generalized Fleming–Viot processes (Section 7). For example, in the general case, we show that if m-a.e. μ is purely atomic, then X_t is purely atomic for all t and that the sample paths of (X_t) are continuous in the variation norm. This holds in particular in the Fleming–Viot case, where, in addition, we prove necessary and sufficient conditions (in terms of the mutation measure ν_0 and mutation intensity θ) that certain sets of measures in E are not hit by (X_t). The ideas in these sections are closely related to those in [25].

There are also a number of results on these measure-valued diffusions that one obtains immediately from the general theory, once one knows that $\mathscr{E}$ in (1.1) is closable and that its closure is a quasi-regular semi-Dirichlet form. For example, one can write the martingale problems solved by the diffusions [cf. Theorem 3.5(iv) and Section 5]. One can also give conditions for the existence of a dual process [cf. Remark 2.6(iii), Theorem 3.5(iii) and Section 5.3]. Finally, if $\mathscr{E}$ in (1.1) is symmetric, general results about the large deviations of the empirical distribution of the associated diffusions (cf. [8, 20]) apply (cf. Section 4).

2. A class of Dirichlet forms having a space of measures as state space. If $\mathscr{E}$ denotes a bilinear form with domain D on a Hilbert space H with scalar product $(\cdot,\cdot)$, we define for $\alpha > 0$,

$$\mathscr{E}_\alpha := \mathscr{E} + \alpha(\ ,\ ). \tag{2.1}$$

We set $\mathscr{E}(u) := \mathscr{E}(u, u)$, $u \in D$. The *symmetric* part of $\mathscr{E}$ is given by

$$\tilde{\mathscr{E}}(u,v) := \tfrac{1}{2}(\mathscr{E}(u,v) + \mathscr{E}(v,u)), \qquad u, v \in D, \tag{2.2}$$

and its *antisymmetric* part by

$$\check{\mathscr{E}}(u,v) := \tfrac{1}{2}(\mathscr{E}(u,v) - \mathscr{E}(v,u)), \qquad u, v \in D. \tag{2.3}$$

Let $E := \mathscr{M}_1(S)$ be the space of probability measures on a Polish space S with its Borel σ-algebra $\mathscr{B}(S)$. We equip E with the weak topology and its Borel σ-algebra $\mathscr{B}(E)$. Note that E is then also a Polish space (cf. [3]). If f, g are bounded $\mathscr{B}(S)$-measurable functions on S and $\mu \in E$, we define

$$\mu(f) := \int f\,d\mu, \tag{2.4}$$

$$\langle f, g\rangle_\mu := \int fg\,d\mu - \int f\,d\mu \int g\,d\mu = \mathrm{cov}_\mu(f,g) \tag{2.5}$$

and

$$\|f\|_\mu = \sqrt{\langle f, f\rangle_\mu}\,, \tag{2.6}$$

The set $\mathscr{F}C_b^\infty$ of *finitely based smooth functions* on E is defined by

$$u \in \mathscr{F}C_b^\infty \quad \Leftrightarrow \quad \begin{array}{ll} u(\mu) = \varphi(\mu(f_1),\dots,\mu(f_k)), & k \in \mathbb{N}, \\ f_i \in C_b(S), \quad 1 \le i \le k, & \varphi \in C_0^\infty(\mathbb{R}^k). \end{array} \tag{2.7}$$

Let m be a finite (positive) measure on $(E, \mathscr{B}(E))$. We suppose that supp$[m]$ $= E$. Finally, let

$$b(\cdot,\cdot) : S \times E \to \mathbb{R} \tag{2.8}$$

be a measurable function such that

$$\sup_{\mu \in E} \|b(\mu)\|_\mu < \infty, \tag{2.9}$$

where

$$b(\mu)(x) := b(x, \mu). \tag{2.10}$$

Now we can define the bilinear form, which is the central object of this paper.

DEFINITION 2.1. For $u, v \in \mathscr{F}C_b^\infty$ let

$$\begin{aligned} \mathscr{E}^b(u,v) := \int (\langle \nabla u(\mu), \nabla v(\mu) \rangle_\mu \\ + \langle b(\mu), \nabla u(\mu) \rangle_\mu v(\mu)) m(d\mu), \end{aligned} \tag{2.11}$$

where

$$\nabla u(\mu) := (\nabla_x u(\mu))_{x \in S} = \left(\frac{\partial u}{\partial \varepsilon_x}(\mu) \right)_{x \in S} \tag{2.12}$$

and

$$\nabla_x u(\mu) := \frac{\partial u}{\partial \varepsilon_x}(\mu) := \frac{d}{ds} u(\mu + s\varepsilon_x)\bigg|_{s=0} = \sum_{i=1}^k \frac{\partial \varphi}{\partial y_i}(\mu(f_1), \ldots, \mu(f_k)) f_i(x),$$

$$u(\mu) = \varphi(\mu(f_1), \ldots, \mu(f_k)), \qquad \varepsilon_x := \text{Dirac measure at } x.$$

Here we consider the natural extension of $u \in \mathscr{F}C_b^\infty$ to all finite positive measures on $(S, \mathscr{B}(S))$. Set $\mathscr{E}(u, v) := \mathscr{E}^0(u, v)$ (i.e., the form with $b \equiv 0$).

REMARK 2.2. (i) Definition 2.1 is motivated by the following geometrical setup. One should consider E as "the manifold" and $L^2(S; \mu)$ as the tangent space at $\mu \in E$. The "Riemannian structure" is given by $\langle\ ,\ \rangle_\mu$, $\mu \in E$. In this spirit ∇u is then "the gradient vector field" of the smooth function u on E.

(ii) In succeeding text, for two functions $b_1, b_2 \colon S \times E \to \mathbb{R}$, we denote the function $\mu \mapsto \langle b_1(\mu), b_2(\mu) \rangle_\mu$ (if this is defined) by $\langle b_1, b_2 \rangle$.

We now briefly recall the terminology of semi-Dirichlet forms. Let $(E, \mathscr{B}, m)$ be a measure space. Let $\mathscr{E}$ be a positive definite bilinear form with domain $D(\mathscr{E})$ on the (real) Hilbert space $L^2(E; m)$ [with the natural inner product $(\ ,\ )$].

DEFINITION 2.3. $(\mathscr{E}, D(\mathscr{E}))$ is called a *coercive closed form* if $D(\mathscr{E})$ is dense in $L^2(E; m)$ and the following conditions (i) and (ii) hold.

(i) $(\tilde{\mathscr{E}}, D(\mathscr{E}))$ is positive definite and closed on $L^2(E; m)$.

(ii) (*Weak sector condition*). There exists a constant $K > 0$ such that $|\mathscr{E}_1(u, v)| \le K \mathscr{E}_1(u, u)^{1/2} \mathscr{E}_1(v, v)^{1/2}$ for all $u, v \in D(\mathscr{E})$.

$(\mathscr{E}, D(\mathscr{E}))$ is called a *semi-Dirichlet form* on $L^2(E; m)$ if in addition:

(iii) (*Semi-Dirichlet property*). For every $u \in D(\mathscr{E})$, $u^+ \wedge 1 \in D(\mathscr{E})$ and $\mathscr{E}(u + u^+ \wedge 1, u - u^+ \wedge 1) \ge 0$.

(iv) If also the dual form $\check{\mathscr{E}}(u, v) := \mathscr{E}(v, u)$ satisfies (iii), then $(\mathscr{E}, D(\mathscr{E}))$ is called a *Dirichlet form* (cf. [20], I.4).

In succeeding text, $(\mathscr{E}, D(\mathscr{E}))$ is always equipped with the norm $\tilde{\mathscr{E}}_1^{1/2}$.

We recall that a positive definite bilinear form $\mathscr{E}$ with domain $D \subset L^2(E; m)$ is called *closable* if every $\tilde{\mathscr{E}}^{1/2}$-Cauchy sequence in D that converges to 0 in $L^2(E; m)$ also converges to 0 with respect to $\tilde{\mathscr{E}}^{1/2}$. We call the smallest closed extension $(\mathscr{E}, \overline{D})$ of a closable form $(\mathscr{E}, D)$ the *closure* of $(\mathscr{E}, D)$ (cf. [19], I.3.2).

Now let us return to the forms $\mathscr{E}$ and $\mathscr{E}^b$ introduced in Definition 2.1. For later use we define Γ: $\mathscr{F}C_b^\infty \times \mathscr{F}C_b^\infty \to L^1(E; m)$ and β: $\mathscr{F}C_b^\infty \times \mathscr{F}C_b^\infty \to L^1(E; m)$ by

$$\Gamma(u, v) = \langle \nabla u, \nabla v \rangle, \qquad u, v \in \mathscr{F}C_b^\infty, \tag{2.13}$$

$$\beta(u, v) = \langle b, \nabla u \rangle v, \qquad u, v \in \mathscr{F}C_b^\infty. \tag{2.14}$$

Hence

$$\mathscr{E}^b(u, v) = \int (\Gamma(u, v) + \beta(u, v))\, dm, \qquad u, v \in \mathscr{F}C_b^\infty.$$

We now formulate the main result of this section.

THEOREM 2.4. (i) *If b satisfies (2.9) and m is such that $(\mathscr{E}, \mathscr{F}C_b^\infty)$ is closable on $L^2(E; m)$, then there exists $\alpha > 0$ such that $(\mathscr{E}_\alpha^b, \mathscr{F}C_b^\infty)$ is a densely defined positive definite closable form and its closure $(\mathscr{E}_\alpha^b, D(\mathscr{E}_\alpha^b))$ is a semi-Dirichlet form on $L^2(E; m)$.*

(ii) *If additionally,*

$$\int (\langle b, \nabla v \rangle + \alpha v)\, dm \geq 0 \quad \text{for } v \in \mathscr{F}C_b^\infty, v \geq 0, \tag{2.15}$$

then the closure $(\mathscr{E}_\alpha^b, D(\mathscr{E}_\alpha^b))$ of $(\mathscr{E}_\alpha^b, \mathscr{F}C_b^\infty)$ is a Dirichlet form on $L^2(E; m)$.

For the proof of Theorem 2.4 we need the following simple lemma.

LEMMA 2.5. *Under the assumptions of Theorem 2.4(i) we have:*

(i) *Let $c_0 := \sup_{\mu \in E} \|b(\mu)\|_\mu$. Then for all $u, v \in \mathscr{F}C_b^\infty$,*

$$|\beta(u, v)| \leq \left(\tfrac{1}{2}\Gamma(u, u) + 2c_0^2 u^2\right)^{1/2} \left(\tfrac{1}{2}\Gamma(v, v) + 2c_0^2 v^2\right)^{1/2}.$$

(ii) *There exist $c, \alpha \in]0, \infty[$ such that*

$$\frac{1}{c}\mathscr{E}_1(u, u) \leq \mathscr{E}_\alpha^b(u, u) \leq c\mathscr{E}_1(u, u) \quad \text{for all } u \in \mathscr{F}C_b^\infty.$$

PROOF. (i) We have $u, v \in \mathscr{F}C_b^\infty$ and $\mu \in E$ that

$$\begin{aligned} |\beta(u, v)(\mu)| &= |\langle b(\mu), \nabla u(\mu) \rangle_\mu v(\mu)| \\ &\leq \|b(\mu)\|_\mu \|\nabla u(\mu)\|_\mu |v(\mu)| \\ &\leq \left(\tfrac{1}{2}\|\nabla u(\mu)\|_\mu^2 + 2c_0^2 u(\mu)^2\right)^{1/2} \left(\tfrac{1}{2}\|\nabla v(\mu)\|_\mu^2 + 2c_0^2 v(\mu)^2\right)^{1/2}. \end{aligned}$$

(ii) An immediate consequence of (i). □

PROOF OF THEOREM 2.4.

1. Because $\mathscr{F}C_b^\infty$ separates the points of E and contains the constant function 1, a monotone class argument implies that $\mathscr{F}C_b^\infty$ is dense in $L^2(E; m)$.
2. Choose $\alpha > 0$ as in Lemma 2.5(ii). The positivity of $\mathscr{E}_\alpha^b$ is already contained in Lemma 2.5(ii). Lemma 2.5(i) yields that there exists $K > 0$ such that
$$|\mathscr{E}_\alpha^b(u, v)| \le K \mathscr{E}_\alpha^b(u, u)^{1/2} \mathscr{E}_\alpha^b(v, v)^{1/2} \quad \text{for all } u, v \in \mathscr{F}C_b^\infty.$$
Hence $\mathscr{E}_\alpha^b$ is coercive. Because we assume the closability of $(\mathscr{E}, \mathscr{F}C_b^\infty)$, the closability of $(\mathscr{E}_\alpha^b, \mathscr{F}C_b^\infty)$ follows by Lemma 2.5(ii).
3. *Semi-Dirichlet property*. Let $(\varphi_\varepsilon)_{\varepsilon > 0}$ be as in [20], II.2.7, a smooth approximation of the unit contraction, that is $\varphi_\varepsilon \colon \mathbb{R} \to [-\varepsilon, 1 + \epsilon]$ such that $\varphi_\varepsilon(t) = t$ for $t \in [0, 1]$, $0 \le \varphi_\varepsilon(t) + \varphi_\varepsilon(s) \le t - s$ for all $s, t \in \mathbb{R}$, $t \ge s$, $\varphi_\varepsilon(t) = 1 + \varepsilon$ for $t \in [1 + 2\varepsilon, \infty[$ and $\varphi_\varepsilon(t) = -\varepsilon$ for $t \in]-\infty, -2\varepsilon]$. By (the proof of) [20], I.4.7, it suffices to show that
$$\liminf_{\varepsilon \downarrow 0} \mathscr{E}_\alpha^b(\varphi_\varepsilon \circ u, u - \varphi_\varepsilon \circ u) \ge 0.$$
However, by the chain rule,
$$\begin{aligned} &\mathscr{E}_\alpha^b(\varphi_\varepsilon \circ u, u - \varphi_\varepsilon \circ u) \\ &\quad = \int \{\varphi_\varepsilon' \circ u(1 - \varphi_\varepsilon' \circ u)(\nabla u, \nabla u) + \langle b, \nabla u\rangle \varphi_\varepsilon' \circ u(u - \varphi_\varepsilon \circ u) \\ &\qquad\qquad + \alpha \varphi_\varepsilon \circ u(u - \varphi_\varepsilon \circ u)\}\, dm. \end{aligned}$$
The first summand is positive because $0 \le \varphi_\varepsilon' \le 1$; the second summand converges to 0 because
$$\varphi_\varepsilon' \circ u(u - \varphi_\varepsilon \circ u) \le 1_{[-2\varepsilon, 1+2\varepsilon]} \circ u(u - \varphi_\varepsilon \circ u) \xrightarrow[\varepsilon \downarrow 0]{} 0;$$
the third summand is positive because
$$\varphi_\varepsilon \circ u(u - \varphi_\varepsilon \circ u) \ge 0.$$
4. *Dirichlet property*. Assume now that (2.15) holds. To prove the Dirichlet property, by [20], I.4.7, it is sufficient to show that
$$\liminf_{\varepsilon \downarrow 0} \mathscr{E}_\alpha^b(u - \varphi_\varepsilon \circ u, \varphi_\varepsilon \circ u) \ge 0.$$
By the product rule for ∇, we have that
$$\begin{aligned} &\int \{\langle b, \nabla(u - \varphi_\varepsilon \circ u)\rangle \varphi_\varepsilon \circ u + \alpha(u - \varphi_\varepsilon \circ u)\varphi_\varepsilon \circ u\}\, dm \\ &\quad = \int \{\langle b, \nabla((\varphi_\varepsilon \circ u)(u - \varphi_\varepsilon \circ u))\rangle + \alpha(u - \varphi_\varepsilon \circ u)\varphi_\varepsilon \circ u\}\, dm \\ &\qquad - \int \varphi_\varepsilon' \circ u(u - \varphi_\varepsilon \circ u)\langle b, \nabla u\rangle\, dm. \end{aligned}$$

The first term is nonnegative by the assumption in (ii) and the positivity of $(u - \varphi_\varepsilon \circ u)\varphi_\varepsilon \circ u$; the second term converges to 0 as in part 3 of this proof. □

REMARK 2.6. (i) Consider the situation of Theorem 2.4. Clearly Γ extends as a continuous bilinear map from $D(\mathscr{E}_\alpha^b) \times D(\mathscr{E}_\alpha^b)$ to $L^1(E; m)$ and so does β by Lemma 2.5(i). Hence

$$\mathscr{E}_\alpha^b(u, v) = \int (\Gamma(u, v) + \beta(u, v) + \alpha uv)\, dm, \qquad u, v \in D(\mathscr{E}_\alpha^b). \tag{2.16}$$

Furthermore, Lemma 2.5(ii) shows that $D(\mathscr{E}_b^\alpha) = D(\mathscr{E})$. Hence (e.g., by [20], I.4.15),

$$uv \in D(\mathscr{E}_\alpha^b) \tag{2.17}$$

for all bounded $u, v \in D(\mathscr{E}_\alpha^b) = D(\mathscr{E})$. Furthermore, because any element in $D(\mathscr{E}_b^\alpha)$ can be approximated by elements in $\mathscr{F}C_b^\infty$ with respect to $(\tilde{\mathscr{E}}_\alpha^b)^{1/2}$, for bounded $u, v, w \in D(\mathscr{E}_\alpha^b)$ we have

$$\begin{aligned} \Gamma(wu, v) &= w\Gamma(u, v) + u\Gamma(w, v), \\ \Gamma(u, wv) &= w\Gamma(u, v) + v\Gamma(u, w); \end{aligned} \tag{2.18}$$

$$\begin{aligned} \beta(wu, v) &= w\beta(u, v) + u\beta(w, v), \\ \beta(u, wv) &= w\beta(u, v). \end{aligned} \tag{2.19}$$

Note also that for all $u, v \in D(\mathscr{E}_\alpha^b)$, setting $\Gamma(u) := \Gamma(u, u)$, we have that

$$|\Gamma(u, v)| \le \Gamma(u)^{1/2}\Gamma(v)^{1/2}$$

and hence $v_n \xrightarrow[n\to\infty]{} v$ in $D(\mathscr{E}_\alpha^b)$ implies that

$$\Gamma(u, v_n) \xrightarrow[n\to\infty]{} \Gamma(u, v) \quad \text{in } L^2(E; m),$$

provided $\Gamma(u)$ is bounded. Furthermore, for a Lipschitz function φ: $\mathbb{R} \to \mathbb{R}$ and $u, v \in D(\mathscr{E}_\alpha^b)$,

$$\Gamma(\varphi(u), v) = \varphi'(u)\Gamma(u, v),$$

because this trivially holds for $u, v \in \mathscr{F}C_b^\infty$.

(ii) If in Theorem 2.4 we replace $\mathscr{E}_\alpha^b(u, v)$ for $u, v \in \mathscr{F}C_b^\infty$ by

$$\mathscr{E}_\alpha^b(u, v) + \int \langle \nabla v, d\rangle u\, dm, \tag{2.20}$$

where d: $S \times E \to \mathbb{R}$ satisfies (2.9) and

$$\int (\langle d, \nabla v\rangle + \alpha v)\, dm \ge 0 \quad \text{for all } v \in \mathscr{F}C_b^\infty, v \ge 0,$$

then the assertions of Theorem 2.4 remain true. The proof is entirely analogous.

(iii) Condition (2.15) obviously holds, for example, if $b = \nabla w$ for $w \in D(L)$ with

$$Lw \le \alpha. \tag{2.21}$$

Here $(L, D(L))$ denotes the generator of $(\mathscr{E}, D(\mathscr{E}))$.

There is an obvious relation between the *generators* (cf. [20], I.2.9 and I.2.16) of the (semi-) Dirichlet forms $\mathscr{E}$ and $\mathscr{E}_\alpha^b$ in Theorem 2.4.

PROPOSITION 2.7. *Let $(L, D(L))$ be the generator of $(\mathscr{E}, D(\mathscr{E}))$ and $(L_\alpha^b, D(L_\alpha^b))$ be the generator of $(\mathscr{E}_\alpha^b, D(\mathscr{E}_\alpha^b))$. Then $D(L) \subset D(L_\alpha^b)$ and*

$$L_\alpha^b u = Lu - \beta(u, 1) - \alpha u \quad \textit{for all } u \in D(L).$$

PROOF. We have to show that for $u \in D(L)$,

$$\mathscr{E}_\alpha^b(u, v) = (-Lu, v) + \int (\beta(u, 1)v + \alpha uv)\, dm \qquad \text{for all } v \in D(\mathscr{E}_\alpha^b) \tag{2.22}$$

(cf., e.g., [20], I.2.16). If $v \in \mathscr{F}C_b^\infty$, then (2.22) is evident by the definition of $\mathscr{E}_\alpha^b$. Let $v_n \in \mathscr{F}C_b^\infty$, $n \in \mathbb{N}$, such that $\mathscr{E}_\alpha^b(v_n - v, v_n - v) \to 0$. [The existence of the sequence $(v_n)_{n \in \mathbb{N}}$ follows because $D(\mathscr{E}_\alpha^b)$ is the closure of $\mathscr{F}C_b^\infty$ with respect to $(\tilde{\mathscr{E}}_\alpha^b)^{1/2}$.] Then

$$\begin{aligned} \mathscr{E}_\alpha^b(u, v) &= \lim_{n \to \infty} \mathscr{E}_\alpha^b(u, v_n) \\ &= \lim_{n \to \infty} \int \left(-v_n Lu + \beta(u, v_n) + \alpha u v_n \right) dm \\ &= \int \left(-vLu + \beta(u, v) + \alpha uv \right) dm \end{aligned}$$

because $(\tilde{\mathscr{E}}_\alpha^b)$-convergence implies L^2-convergence [and by definition $Lu \in L^2(E; m)$]. □

3. Associated diffusion processes. We recall the following notions from [20]. For the first three definitions in this section, we only need to assume that E is a topological space, $\mathscr{B}(E)$ is its Borel σ-algebra and m is a σ-finite positive measure on $\mathscr{B}(E)$. For simplicity, however, and because it is sufficient for the purposes of this paper (i.e., for Theorem 3.5), we assume that E is a Borel subset of a Polish space. For the same reason we shall also consider diffusions in succeeding text. For the general case involving special standard processes, we refer to [20], [2] and [19]. We fix a semi-Dirichlet form $(\mathscr{E}, D(\mathscr{E}))$ on $L^2(E; m)$ with generator $(L, D(L))$. Let $T_t := e^{tL}$, $t > 0$, be the strongly continuous contraction semigroup generated by L (cf. [20], I, Sections 1 and 2).

DEFINITION 3.1. (i) A sequence $(F_n)_{n \in \mathbb{N}}$ of closed sets in E is called an *$\mathscr{E}$-nest* if $\bigcup_{n \geq 1} D(\mathscr{E})_{F_n}$ is $\tilde{\mathscr{E}}_1^{1/2}$-dense in $D(\mathscr{E})$, where for $n \in \mathbb{N}$,

$$D(\mathscr{E})_{F_n} := \{u \in D(\mathscr{E}) \mid u = 0 \ m\text{-a.e. on } E \setminus F_n\}.$$

(ii) $N \subset E$ is called *$\mathscr{E}$-exceptional* if $N \subset E \setminus \bigcup F_n$ for some $\mathscr{E}$-nest $(F_n)_{n \in \mathbb{N}}$. A property of points in E is said to hold *$\mathscr{E}$-quasi-everywhere* (abbreviated $\mathscr{E}$-q.e.) if it holds outside an $\mathscr{E}$-exceptional set.

(iii) An $\mathscr{E}$-q.e. defined function $f: E \to \mathbb{R}$ is called *$\mathscr{E}$-quasi-continuous* if $f \in C(\{F_n\})$ for some $\mathscr{E}$-nest $(F_n)_{n \in \mathbb{N}}$, where

$$C(\{F_n\}) := \{f: A \to \mathbb{N} \mid \bigcup F_n \subset A \subset E \text{ and } f|_{F_n} \text{ is continuous for all } n \in \mathbb{N}\}$$

(and $f|_{F_n}$ denotes the restriction of f to F_n).

(iv) $(\mathscr{E}, D(\mathscr{E}))$ is said to have the *local property* if $\mathscr{E}(u, v) = 0$ for all $u, v \in D(\mathscr{E})$ with $\operatorname{supp}[|u| \cdot m] \cap \operatorname{supp}[|v| \cdot m] = \varnothing$.

By [20], III.2.3, every $\mathscr{E}$-exceptional set has m-measure zero.

DEFINITION 3.2. $(\mathscr{E}, D(\mathscr{E}))$ is called *quasi-regular* if:

(i) There exists an $\mathscr{E}$-nest consisting of compact sets.

(ii) There exists an $\tilde{\mathscr{E}}_1^{1/2}$-dense subset of $D(\mathscr{E})$ whose elements have $\mathscr{E}$-quasi-continuous m-versions.

(iii) There exist $u_n \in D(\mathscr{E})$, $n \in \mathbb{N}$, having $\mathscr{E}$-quasi-continuous m-versions $\tilde{u}_n$, $n \in \mathbb{N}$, such that $\{\tilde{u}_n \mid n \in \mathbb{N}\}$ separates the points of $E \setminus N$ for some $\mathscr{E}$-exceptional set N.

Note that if $(\mathscr{E}, D(\mathscr{E}))$ is quasi-regular, every $u \in D(\mathscr{E})$ has an $\mathscr{E}$-quasi-continuous m-version $\tilde{u}$ (cf., e.g., [20], IV.3.3(ii)).

DEFINITION 3.3. (i) A normal strong Markov process $\mathbf{M} := (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \geq 0}, (X_t)_{t \geq 0}, (P_z)_{z \in E})$ with state space E is called a *diffusion process* if its sample paths are continuous (up to its lifetime ζ).

(ii) A diffusion process $\mathbf{M} := (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \geq 0}, (X_t)_{t \geq 0}, (P_z)_{z \in E})$ with state space E is called *properly associated* with $(\mathscr{E}, D(\mathscr{E}))$ if $z \mapsto E_z[\hat{u}(X_t)]$ is an $\mathscr{E}$-quasi-continuous version of $T_t u$ for all $t > 0$ and all m-versions $\hat{u}$ of any $u \in L^2(E; m)$.

Note that in Definition 3.3(ii), $\mathbf{M}$ is (essentially) unique (cf. [20], IV.6.4). We need the following general result, which can be deduced from [20] and [19] (see also [1, 2]).

THEOREM 3.4. *The following are equivalent*:

(i) *$(\mathscr{E}, D(\mathscr{E}))$ is quasi-regular and has the local property.*

(ii) *There exists a diffusion process* $\mathbf{M}$ *properly associated with $(\mathscr{E}, D(\mathscr{E}))$.*

In this case, if in addition $1 \in D(\mathscr{E})$ and $\mathscr{E}(1, u) = 0$ for all $u \in D(\mathscr{E})$, then $\mathbf{M}$ *is* conservative, *that is, $P_z[\zeta = \infty] = 1$ for $\mathscr{E}$-q.e. $z \in E$.*

PROOF. Except for the last part, this follows from [19], Theorems 3.8 and 3.9, and [20], V.1.5 (see [19], Remark 3.10). The last part follows because in this case (cf., e.g., [20], I.2.16]) $1 \in D(L)$ and $e^{tL}1 = 1$, $t > 0$. Hence $\mathbf{M}$ is conservative (cf., e.g., [21], Lemma 9). □

Now we will apply these results to the Dirichlet forms of Section 2. So let $E = \mathscr{M}_1(S)$ and m, b, $\mathscr{E}$ and $\mathscr{E}^b$ be as in Section 2.

THEOREM 3.5. *Suppose that b satisfies* (2.9) *and that* $(\mathscr{E}, \mathscr{F}C_b^\infty)$ *is closable on* $L^2(E; m)$. *Let* $\alpha > 0$ *and* $(\mathscr{E}_\alpha^b, D(\mathscr{E}_\alpha^b))$ *be as in Theorem* 2.4(i). *Then*:

(i) $(\mathscr{E}_\alpha^b, D(\mathscr{E}_\alpha^b))$ *is quasi-regular and has the local property.*

(ii) *There exists a diffusion process* $\mathbf{M}_\alpha^b := (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \geq 0}, (X_t)_{t \geq 0}, (P_\mu)_{\mu \in E})$ *properly associated with* $(\mathscr{E}_\alpha^b, D(\mathscr{E}_\alpha^b))$. *If* $\alpha = 0$, $\mathbf{M}_0^b$ *is conservative.*

(iii) *If, in addition, condition* (2.15) *holds, there exist diffusion processes* $\tilde{\mathbf{M}}_\alpha^b$ $(\hat{\mathbf{M}}_\alpha^b)$ *associated with* $(\tilde{\mathscr{E}}_\alpha^b, D(\mathscr{E}_\alpha^b))$ [*resp.* $(\hat{\mathscr{E}}_\alpha^b, D(\mathscr{E}_\alpha^b))$], *where,* $\hat{\mathscr{E}}_\alpha^b(u, v) := \mathscr{E}_\alpha^b(v, u)$, $u, v \in D(\mathscr{E}_\alpha^b)$. $\mathbf{M}_\alpha^b$ *is in duality with* $\mathbf{M}_\alpha^b$ *w.r.t.* m.

(iv) *The diffusion process* $\mathbf{M}_\alpha^b$ *in* (ii) *solves the martingale problem for* $(L_\alpha^b, D(L_\alpha^b))$ *in the sense that for every* $u \in D(L_\alpha^b)$,

$$\tilde{u}(X_t) - \tilde{u}(X_0) - \int_0^t L_\alpha^b u(X_s)\, ds, \qquad t \geq 0, \tag{3.1}$$

is an $(\mathscr{F}_t)_{t \geq 0}$*-martingale under* P_μ *for* $\mathscr{E}$*-q.e.* $\mu \in E$. *Here* $\tilde{u}$ *denotes an* $\mathscr{E}$*-quasi-continuous* m*-version of* u.

REMARK 3.6. (i) Theorem 3.5 immediately extends to the more general semi-Dirichlet form described in Remark 2.6.

(ii) As we shall see, Theorem 3.5(iv) follows easily from Theorem 3.5(i) and (ii). It is clear that $\mathbf{M}_\alpha^b$ is the unique solution of this martingale problem. However, this information is useless because, in general, we do not know $D(L_\alpha^b)$ explicitly. The interesting question is whether $\mathbf{M}_\alpha^b$ is the unique solution with $\mathscr{F}C_b^\infty$ replacing $D(L_\alpha^b)$. This is, in general, a very difficult problem. A solution is only known in very special cases (cf. following text and also [22, 23]) for the solution in the "flat case").

(iii) If $\alpha = 0$ and $b = 0$, then m is the invariant measure of $\mathbf{M}_0^0$.

PROOF OF THEOREM 3.5. (i) We first note that to show quasi-regularity, because of Lemma 2.5(ii) it suffices to consider the case $b \equiv 0$, that is, to show that $(\mathscr{E}, D(\mathscr{E}))$ is quasi-regular. Because $\mathscr{F}C_b^\infty$ is $\tilde{\mathscr{E}}_1^{1/2}$-dense in $D(\mathscr{E})$ and separates the points of E, properties Definition 3.2(ii) and (iii) are clear. Let $\{\mu_j \mid j \in \mathbb{N}\}$ be a dense subset of E and let ρ be a bounded, complete metric on E compatible with its topology. Condition 3.2(i) follows from [22], Theorem 3.4 if we can find $f_{ij} \in \mathscr{F}C_b^\infty$, $i, j \in \mathbb{N}$, such that

$$\sup_{i,j} \|\nabla f_{ij}\|_\mu \in L^2(E; m) \tag{3.2}$$

and

$$\sup_i f_{ij}(\mu) = \rho(\mu, \mu_j) \quad \text{for all } j \in \mathbb{N}. \tag{3.3}$$

Let us now show that this is possible. Recall (see [9]) that if (S, d) is a complete, separable metric space, then $E = \mathscr{M}_1(S)$ equipped with the weak topology can be metrized by

$$\rho(\mu, \nu) = \sup\left\{\int f d(\mu - \nu) \mid \|f\|_{BL} \le 1\right\}. \tag{3.4}$$

where

$$\|f\|_{BL} := \sup_{x \neq y} |f(x) - f(y)|/d(x, y) + \sup_{x} |f(x)|, \tag{3.5}$$

and that (E, ρ) is a complete, separable metric space with ρ bounded. For every $j, k \in \mathbb{N}$, select a sequence $(g_i^{jk})_{i \in \mathbb{N}}$ with $\|g_i^{jk}\|_{BL} \le 1$ so that

$$\rho(\mu_j, \mu_k) = \sup_i \int g_i^{jk} \, d(\mu_j - \mu_k). \tag{3.6}$$

Now rewrite the whole collection $(g_i^{jk})_{i,j,k \in \mathbb{N}}$ with a single index $(g_i)_{i \in \mathbb{N}}$ and for μ, ν in E, define

$$\tilde{\rho}(\mu, \nu) := \sup_i \int g_i \, d(\mu - \nu). \tag{3.7}$$

Then clearly $\tilde{\rho}(\mu, \nu) \le \rho(\mu, \nu)$ on all of E, whereas for the members of the dense set $\{\mu_j \mid j \in \mathbb{N}\}$ we have $\tilde{\rho}(\mu_j, \mu_k) = \rho(\mu_j, \mu_k)$. We claim that, in fact, $\rho = \tilde{\rho}$. To see this, let μ, ν in E and choose μ_j, μ_k so that $\rho(\mu, \mu_j) \le \delta$ and $\rho(\nu, \mu_k) \le \delta$. Therefore, also $\tilde{\rho}(\mu, \mu_j) \le \delta$, $\tilde{\rho}(\nu, \mu_k) \le \delta$ and thus $\tilde{\rho}(\mu_j, \mu_k) \le \tilde{\rho}(\mu, \nu) + 2\delta$, but $\tilde{\rho}(\mu_j, \mu_k) = \rho(\mu_j, \mu_k)$ and so

$$\begin{aligned} \rho(\mu, \nu) &\le \rho(\mu, \mu_j) + \rho(\mu_j, \mu_k) + \rho(\mu_k, \nu) \\ &\le \delta + \tilde{\rho}(\mu, \nu) + 2\delta + \delta \\ &= \tilde{\rho}(\mu, \nu) + 4\delta. \end{aligned}$$

Letting $\delta \to 0$ shows that $\tilde{\rho} = \rho$. Finally, setting $f_{ij}(\mu) := \int g_i \, d(\mu - \mu_j)$ gives us a collection satisfying (3.2) and (3.3), which concludes the proof of the quasi-regularity of $(\mathscr{E}_\alpha^b, D(\mathscr{E}_\alpha^b))$.

Now we show that $(\mathscr{E}_\alpha^b, D(\mathscr{E}_\alpha^b))$ has the local property. Let $u, v \in D(\mathscr{E}_\alpha^b)$ with $\operatorname{supp}[|u| \cdot m] \cap \operatorname{supp}[|v| \cdot m] = \varnothing$. We may assume that u, v are bounded. Because $(\mathscr{E}_\alpha^b, D(\mathscr{E}_\alpha^b))$ is quasi-regular, by [19], Remark 3.10, and [20], V.1.7, there exists $\chi \in D(\mathscr{E}_\alpha^b)$ such that $0 \le \chi \le 1_{E \setminus \operatorname{supp}[|u| \cdot m]}$ and $\chi > 0$ on $E \setminus \operatorname{supp}[|u| \cdot m]$. Hence, by (2.18),

$$0 = \Gamma(\chi u, v) = \chi \Gamma(u, v) + u \Gamma(\chi, v),$$

which implies that $\operatorname{supp}[|\Gamma(u, v)| \cdot m] \subset \operatorname{supp}[|u| \cdot m]$. Similarly, $\operatorname{supp}[|\Gamma(u, v)| \cdot m] \subset \operatorname{supp}[|v| \cdot m]$; hence, $\Gamma(u, v) = 0$. By the same argument using (2.19) we obtain that $\beta(u, v) = 0$. Hence by (2.16), $\mathscr{E}_\alpha^b(u, v) = 0$.

(ii) This follows from (i) and Theorem 3.4.

(iii) Because (2.15) implies that $(\mathscr{E}_\alpha^b, D(\mathscr{E}_\alpha^b))$ is a Dirichlet form, so are $(\tilde{\mathscr{E}}_\alpha^b, D(\mathscr{E}_\alpha^b))$ and $(\hat{\mathscr{E}}_\alpha^b, D(\mathscr{E}_\alpha^b))$ (cf. [20], I.4.6). They obviously also have the local property. The quasi-regularities of $(\mathscr{E}_\alpha^b, D(\mathscr{E}_\alpha^b))$, $(\tilde{\mathscr{E}}_\alpha^b, D(\mathscr{E}_\alpha^b))$ and $(\hat{\mathscr{E}}_\alpha^b, D(\mathscr{E}_\alpha^b))$ are equivalent and consequently the assertion follows by (i) and Theorem 3.4.

(iv) Let $u \in D(L)$. First note that $\tilde{u}(X_t) - \tilde{u}(X_0)$ does not depend on the $\mathscr{E}$-quasi-continuous m-version of u that we take. Second, also $\int_0^t L_\alpha^b u(X_s)\,ds$, $t \geq 0$, does not depend on the m-version of $L_\alpha^b u$ [$\in L^2(E;\, m)$] we choose, because for all $f \in L^2(E;\, m)$ and any m-version $\hat{f}$ of f we have that

$$\int\int_0^T |\hat{f}|(X_s)\,ds\,dP_\mu \leq e^t \int_0^\infty e^{-s} E_\mu\big[|\hat{f}|(X_s)\big]\,ds$$

and by Definition 3.3(ii) and [20], IV.2.8, the right-hand side is an $\mathscr{E}$-quasi-continuous version of

$$e^t \int_0^\infty e^{-s} T_s f\,ds.$$

Hence if $f = 0$, then $\int_0^t |\hat{f}|(X_s)\,ds = 0$ for all $t \geq 0$ P_μ-a.s. for $\mathscr{E}$-q.e. $\mu \in E$ (cf. [19], IV.3.3(iii)). The rest of the proof is now standard by the Markov property (and Fubini's theorem). □

4. A result on large deviations for Fleming–Viot processes. A central question in the theory of large deviations for an ergodic Markov process $X = (X_t)_{t \geq 0}$ is the deviation of the empirical distribution $L = (L_t)_{t \geq 0}$ from its ergodic behavior (cf. [8]). If E is the state space of X, the process L has state space $\mathscr{M}_1(E)$ and is defined by

$$L_t(A) := \frac{1}{t}\int_0^t 1_A(X_s)\,ds \quad \text{for all } A \in \mathscr{B}(E). \tag{4.1}$$

Let the space $\mathscr{M}_1(E)$ be equipped with the τ-topology, which is generated by open sets of the form

$$U(M;\, \varepsilon, F) := \left\{L \in \mathscr{M}_1(E)\ \Big|\ \left|\int F\,dL - \int F\,dM\right| < \varepsilon\right\},$$

where $\varepsilon > 0$, $M \in \mathscr{M}(E)$ and F bounded, $\mathscr{B}(E)$-measurable.

We now return to the situation of Section 2. So let $E = \mathscr{M}_1(S)$ and m be as in Section 1. We only consider our basic bilinear form $(\mathscr{E}, \mathscr{F}C_b^\infty)$ on $L^2(E;\, m)$ (i.e., $b \equiv 0$; cf. Definition 2.1).

THEOREM 4.1. *Suppose that $(\mathscr{E}, \mathscr{F}C_b^\infty)$ is closable on $L^2(E;\, m)$ with closure $(\mathscr{E}, D(\mathscr{E}))$. Let $\mathbf{M} := (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \geq 0}, (X_t)_{t \geq 0}, (P_\mu)_{\mu \in E})$ be the diffusion process associated with $(\mathscr{E}, D(\mathscr{E}))$. Assume that the following condition holds*:

(4.2) *If $u \in D(\mathscr{E})$ satisfies $\mathscr{E}(u, u) = 0$, then u is constant m-a.e.*

Let U be a τ-open and K a τ-compact subset of $\mathscr{M}_1(E)$. Then for $\mathscr{E}$-q.e. $\mu \in E$ we have that

$$\liminf_{t \to \infty} \frac{1}{t} \log P_\mu[L_t \in U] \geq -\inf\{\mathscr{E}(\psi, \psi) | \psi \in D(\mathscr{E}),\ \psi^2 m \in U\}$$

and

$$\inf\left\{\sup_{\mu\in E\setminus N}\ \limsup_{t\to\infty}\frac{1}{t}\log P_\mu[L_t\in K]\ \middle|\ N\subset E,\ N\ \mathscr{E}\text{-exceptional}\right\}$$
$$\le -\inf\{\mathscr{E}(\psi,\psi)\mid \psi\in D(\mathscr{E}),\ \psi^2 m\in K\}.$$

REMARK 4.2. It is well-known that (4.2) is equivalent with the ergodicity of P_m, that is, all shift invariant events have P_m-measure 0 or 1 (cf. [16] or [21], Lemma 12).

PROOF OF THEOREM 4.1. The assertion follows a special case from [21], Theorems 1 and 2. □

REMARK 4.3. (i) (4.2) or equivalently the ergodicity of P_m holds if $m = m^{\mathrm{FV}}$, where m^{FV} is the reversible measure of the Fleming–Viot process with neutral mutation and without selection (cf. Section 5.1 and [13], Chapters 5, and 8). This is more generally true if $m = \varphi^2 m^{\mathrm{FV}}$, where $\varphi \in L^2(E;\ m^{\mathrm{FV}})$, $\varphi > 0$ m-a.e., because trivially,

$$\int\langle\nabla u,\nabla u\rangle\varphi^2\,dm = 0 \quad\Rightarrow\quad \langle\nabla u,\nabla u\rangle = 0\ m\text{-a.e.}$$

$$\Rightarrow\quad \int\langle\nabla u,\nabla u\rangle\,dm = 0 \quad\Rightarrow\quad u = \text{constant}\ m\text{-a.e.}$$

(ii) Theorems 1 and 2 in [21] are valid for m-symmetric right processes $\mathbf{M}$ that are ergodic and properly associated with symmetric quasi-regular Dirichlet forms $(\mathscr{E}, D(\mathscr{E}))$ on $L^2(E; m)$, where $m \in \mathscr{M}(E)$ and the state space E is a topological Hausdorff space on which Baire and Borel σ-algebras coincide. It extends results of [17] and [8]. In [17] it is assumed that E is locally compact and that the Dirichlet form is regular, which is not the case for $E = \mathscr{M}_1(S)$ [resp. $(\mathscr{E}, D(\mathscr{E}))$] in our situation. In the special case $m = m^{\mathrm{FV}}$ considered in Section 5.1, the lower bound in Theorem 4.1 also follows from [8], Theorem 5.3.10. The assumption made there that "$\varepsilon_\mu p_t$ is not singular to m for some $t > 0$" is (unlike in the examples in [21]) fulfilled in this case. This can be easily seen from the specific form of m^{FV} and $(p_t)_{t>0}$ (cf. Section 5.1 and [11], Theorem 1.1). Here $(p_t)_{t>0}$ denotes the transition semigroup of $\mathbf{M}$.

(iii) We emphasize that by exactly the same arguments Theorem 4.1 remains true in the case $b \not\equiv 0$ [b satisfying (2.9)] if (2.15) holds and $(\mathscr{E}, D(\mathscr{E}))$ and $\mathbf{M}$ are replaced by $(\tilde{\mathscr{E}}_\alpha^b, D(\mathscr{E}_\alpha^b))$ and $\mathbf{M}_\alpha^b$, respectively (cf. Theorems 2.4 and 3.5).

5. Examples. The motivation of the analysis of the forms defined in Sections 2 and 3 is the construction and investigation of Fleming–Viot processes with an *interactive selection*. Let us therefore first consider the basic model without selection. We use [13] and [6] as the main references for Fleming–Viot processes.

5.1. *Fleming–Viot process with neutral mutation and without selection.* The neutral mutation is described by a probability measure ν_0 on S and a mutation intensity $\theta > 0$. The effect of a mutation on an allele $x \in S$ is always distributed according to ν_0; in other words, the mutation operator has the form

$$Af(x) := \frac{\theta}{2}\int_S (f(\xi) - f(x))\nu_0(d\xi).$$

The Fleming–Viot process is the unique solution of the martingale problem for $(L^{\mathrm{FV}}, \mathscr{F}C_b^\infty)$, where L^{FV} is defined on $\mathscr{F}C_b^\infty$ by

$$\begin{aligned} L^{\mathrm{FV}}u(\mu) &:= \frac{1}{2}\int_S\int_S \mu(dx)(\varepsilon_x(dy) - \mu(dy))\nabla_x(\nabla_y u)(\mu) \\ &\quad + \frac{\theta}{2}\int_S \mu(dx)\int_S (\nabla_y u(\mu) - \nabla_x u(\mu))\nu_0(dy) \end{aligned} \tag{5.1}$$

[cf. [13], (3.13), (3.14) and (8.1)].

Then there is a unique reversible stationary measure m^{FV} associated with L^{FV}, that is, $\forall\, u, v \in \mathscr{F}C_b^\infty$,

$$\int uL^{\mathrm{FV}}v\, dm^{\mathrm{FV}} = \int vL^{\mathrm{FV}}u\, dm^{\mathrm{FV}}; \tag{5.2}$$

hence (because $L^{\mathrm{FV}}1 = 0$),

$$\int L^{\mathrm{FV}}u\, dm^{\mathrm{FV}} = 0. \tag{5.3}$$

The measure m^{FV} is the distribution of the random measure $\sum_{i=0}^\infty \rho_i \varepsilon_{\xi_i}$, where $(\rho_1, \rho_2, \dots) \in\,]0,1[^{\mathbb{N}}$ has the Poisson–Dirichlet distribution with parameter θ and $(\xi_1, \xi_2, \dots) \in S^{\mathbb{N}}$ is $\nu_0^{\otimes\mathbb{N}}$-distributed (cf. [13], Theorem 8.1). In particular, m^{FV} is a probability measure and its (closed) support equals $\mathscr{M}_1(\operatorname{supp}\nu_0)$. Hence, if $\operatorname{supp}\nu_0 \neq S$, we replace S by the support of ν_0. It is easy to see that the square field operator Γ^{FV} of L^{FV} is just $\frac{1}{2}\langle\nabla\cdot, \nabla\cdot\rangle$; that is, for all $u, v \in \mathscr{F}C_b^\infty$, $\mu \in E$,

$$\begin{aligned} L^{\mathrm{FV}}(uv)(\mu) &= v(\mu)L^{\mathrm{FV}}u(\mu) \\ &\quad + u(\mu)L^{\mathrm{FV}}v(\mu) + \langle\nabla u(\mu), \nabla v(\mu)\rangle_\mu. \end{aligned} \tag{5.4}$$

This implies (cf., e.g., [20], I.3.3) that

$$\mathscr{E}(u,v) := \tfrac{1}{2}\int\langle\nabla u, \nabla v\rangle\, dm^{\mathrm{FV}} = -\int L^{\mathrm{FV}}u\cdot v\, dm^{\mathrm{FV}} \tag{5.5}$$

is closable, where the second equality follows by (5.3) and (5.4). Hence, $(\mathscr{E}, \mathscr{F}C_b^\infty)$ can be taken as a starting point for our Theorem 2.4 (and also Theorem 3.5; cf. Section 5.2.3). We denote its closure by $(\mathscr{E}, D(\mathscr{E}))$. Note that by Remark 2.6, $\Gamma^{\mathrm{FV}}(\cdot,\cdot) = \langle\nabla\cdot, \nabla\cdot\rangle$ extends to all of $D(\mathscr{E})$.

REMARKS. (i) So far the reversibility of the Fleming–Viot process w.r.t. Π_{θ,ν_0} was only proved iff S is locally compact [13, 28]. If S is only Polish we consider its Stone–Čech compactification $\bar{S}$. Every function $f \in C_b(S)$ has an extension $\bar{f} \in C(\bar{S})$. The measure $\hat{\nu}_0(\cdot) := \nu_0(\cdot \cap S)$ is a measure on $(\bar{S}, \mathscr{B}(\bar{S}))$. We consider the Fleming–Viot process on $\mathscr{M}_1(\bar{S})$ with reversible measure $\Pi_{\theta,\hat{\nu}_0}$ and generator $\bar{L}$. Then we have

$$\int \bar{F}\bar{L}\bar{G}\, d\Pi_{\theta,\hat{\nu}_0} = \int \bar{G}\bar{L}\bar{F}\, d\Pi_{\theta,\hat{\nu}_0}$$

and

$$\int \bar{L}\bar{F}\, d\Pi_{\theta,\hat{\nu}_0} = 0, \qquad \forall\, \bar{F}, \bar{G} \in \overline{\mathscr{F}C_0^\infty}.$$

This implies for $F, G \in \mathscr{F}C_b^\infty$ that

$$\int FLG\, d\Pi_{\theta,\nu_0} = \int GLF\, d\Pi_{\theta,\nu_0}.$$

(ii) According to Theorem 3.1 in [13], the martingale problem with $\mathscr{F}C_b^\infty$ and the martingale problem with polynomials are equivalent.

5.2. *Perturbation by densities.* Let $\mathscr{E}$ and m^{FV} be as in Section 5.1. We consider $(\mathscr{E}^\varphi, \mathscr{F}C_b^\infty)$ on $L^2(E; \varphi^2 m^{\mathrm{FV}})$, where $\varphi \in L^2(E; m^{\mathrm{FV}})$ and

$$\mathscr{E}^\varphi(u, v) := \tfrac{1}{2}\int_E \langle \nabla u, \nabla v\rangle \varphi^2\, dm^{\mathrm{FV}}, \qquad u, v \in \mathscr{F}C_b^\infty. \tag{5.6}$$

5.2.1. *Closability.*

1. If $\varphi^2 \geq \varepsilon > 0$ m-a.e., then $(\mathscr{E}^\varphi, \mathscr{F}C_b^\infty)$ is closable on $L^2(E; \varphi^2 m^{\mathrm{FV}})$: Let $(u_n)_{n\in\mathbb{N}}$ in $\mathscr{F}C_b^\infty$ be an early $\mathscr{E}^\varphi$-Cauchy sequence such that $u_n \xrightarrow[n\to\infty]{} 0$ in $L^2(E; \varphi^2 m^{\mathrm{FV}})$. The strict positivity of φ^2 implies that $(u_n)_{n\in\mathbb{N}}$ is also an $\mathscr{E}$-Cauchy sequence and converges to 0 in $L^2(E; m^{\mathrm{FV}})$; hence, the closability of $(\mathscr{E}, \mathscr{F}C_b^\infty)$ yields $\int \langle \nabla u_n, \nabla u_n\rangle\, dm^{\mathrm{FV}} \to 0$ as $n \to \infty$, and also $\|\nabla u_{n_k}\|\varphi^2 \to 0$ m-a.e. as $k \to \infty$ for a subsequence $(u_{n_k})_{k\in\mathbb{N}}$ of $(u_n)_{n\in\mathbb{N}}$. Therefore,
$$\begin{aligned}\mathscr{E}^\varphi(u_n, u_n) &= \tfrac{1}{2}\int \lim_{k\to\infty} \langle \nabla(u_n - u_{n_k}), \nabla(u_n - u_{n_k})\rangle \varphi^2\, dm^{\mathrm{FV}} \\ &\leq \liminf_{k\to\infty} \tfrac{1}{2}\int \langle \nabla(u_n - u_{n_k}), \nabla(u_n - u_{n_k})\rangle \varphi^2\, dm^{\mathrm{FV}}.\end{aligned}$$
The last integral can be made arbitrarily small by choosing n large enough.
2. If $\varphi \in D(\mathscr{E})$ and $\varphi > 0$ m-a.e., then the form $(\mathscr{E}^\varphi, \mathscr{F}C_b^\infty)$ is also closable on $L^2(E; \varphi^2 m^{\mathrm{FV}})$. This can be deduced from a result in [10] by using the stochastic analysis of Dirichlet processes (cf. also [29]). We sketch the argument of [10] as follows: Let $\mathbf{M} = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, (X_t)_{t\geq 0}, (P_\mu)_{\mu\in E})$ be the Markov process associated with $(\mathscr{E}, D(\mathscr{E}))$. Up to the stopping time $\tau_\delta := \inf\{t > 0 \mid \tilde{\varphi}(X_t) \leq \delta\}$, where $\tilde{\varphi}$ is an $\mathscr{E}$-quasi-continuous version of φ, $\ln \tilde{\varphi}(X_t) = \ln(\tilde{\varphi} \vee \delta)(X_t)$. Clearly, $\ln(\varphi \vee \delta) \in D(\mathscr{E})$. For $\tau := \lim_{\delta\downarrow 0} \tau_\delta$

we define on $\{t < \tau\}$,

$$M_t^{[\ln \varphi]} := M_t^{[\ln(\varphi \vee \delta)]}.$$

Here for $\psi \in D(\mathscr{E})$ we let $M_t^{[\psi]}$ denote the continuous martingale additive functional of the Fukushima decomposition of $(\tilde{\psi}(X_t) - \psi(X_0))_{t>0}$ (cf. [15], Section 5.1 and [20], VI.2.5). We now consider the process $\mathbf{M}^\varphi$, which is the Girsanov transform of $\mathbf{M}$ with the multiplicative functional

$$L_t^{[\varphi]} = \exp\left(M_t^{[\ln \varphi]} - \tfrac{1}{2}\langle M^{[\ln \varphi]}\rangle_t\right) I_{\{t<\tau\}}$$

($\langle \cdot \rangle$ denotes the quadratic variation of a martingale). Then $\mathbf{M}^\varphi$ is associated with a Dirichlet form on $L^2(E; \varphi^2 m^{\mathrm{FV}})$ that extends $(\mathscr{E}^\varphi, \mathscr{F}C_b^\infty)$. Hence, $(\mathscr{E}^\varphi, \mathscr{F}C_b^\infty)$ is closable. [Note that the closure of $(\mathscr{E}^\varphi, \mathscr{F}C_b^\infty)$ may not necessarily be equal to the Dirichlet form of the process $\mathbf{M}^\varphi$.] Hence we can also take the forms $(\mathscr{E}^\varphi, \mathscr{F}C_b^\infty)$ on $L^2(E; \varphi^2 m^{\mathrm{FV}})$ as a starting point for our Theorem 2.4 (and also Theorem 3.5, cf. Section 5.2.3) and we can consider their perturbation in Section 5.3. Let us first consider their generator.

5.2.2. *Generator of* $(\mathscr{E}^\varphi, D(\mathscr{E}^\varphi))$. Let $\varphi \in D(\mathscr{E})$ with $\varphi > 0$ m^{FV}-a.e. Set $\varphi_n := \varphi \wedge n$, $n \in \mathbb{N}$. Then $\varphi_n \in D(\mathscr{E})$ and each φ_n is bounded. Let $u, v \in \mathscr{F}C_b^\infty$. Because $\varphi_n^2 \in D(\mathscr{E})$, we have by (2.18),

$$\begin{aligned}
\mathscr{E}^{\varphi_n}(u, v) &= \tfrac{1}{2}\int \Gamma^{\mathrm{FV}}(u, v)\varphi_n^2 \, dm^{\mathrm{FV}} \\
&= \tfrac{1}{2}\int \Gamma^{\mathrm{FV}}(u, v\varphi_n^2)\, dm^{\mathrm{FV}} - \tfrac{1}{2}\int \Gamma^{\mathrm{FV}}(u, \varphi_n^2) v \, dm^{\mathrm{FV}} \\
&= -\int L^{\mathrm{FV}} u \cdot v \varphi_n^2 \, dm^{\mathrm{FV}} - \int \Gamma^{\mathrm{FV}}(u, \varphi_n)\varphi_n v \, dm^{\mathrm{FV}}.
\end{aligned}$$

Clearly, $\mathscr{E}^{\varphi_n}(u, v) \to \mathscr{E}^\varphi(u, v)$ as $n \to \infty$ and

$$-\int L^{\mathrm{FV}} u \cdot v\varphi_n^2 \, dm^{\mathrm{FV}} \xrightarrow[n\to\infty]{} -\int L^{\mathrm{FV}} u \cdot v \varphi^2 \, dm,$$

because $L^{\mathrm{FV}} u \cdot v$ is bounded. Because $\Gamma^{\mathrm{FV}}(u, u) = \frac{1}{2}\langle \nabla u, \nabla u\rangle$ is bounded, the last part of Remark 2.6(ii) implies that

$$\int \Gamma^{\mathrm{FV}}(u, \varphi_n)\varphi_n v \, dm^{\mathrm{FV}} \xrightarrow[n\to\infty]{} \int \Gamma^{\mathrm{FV}}(u, \varphi)\varphi v \, dm^{\mathrm{FV}}.$$

Consequently, for all $u, v \in \mathscr{F}C_b^\infty$,

$$\mathscr{E}^\varphi(u, v) = -\int L^{\mathrm{FV}} u \cdot v\varphi^2 \, dm^{\mathrm{FV}} - \int \varphi^{-1}\Gamma^{\mathrm{FV}}(u, \varphi) v \varphi^2 \, dm^{\mathrm{FV}}.$$

Therefore, we obtain for the generator $(L^\varphi, D(L^\varphi))$ of $(\mathscr{E}^\varphi, D(\mathscr{E}^\varphi))$,

$$(5.7) \quad \mathscr{F}C_b^\infty \subset D(L^\varphi) \quad \text{and} \quad L^\varphi u = L^{\mathrm{FV}} u + \varphi^{-1}\Gamma^{\mathrm{FV}}(\varphi, u), \qquad u \in \mathscr{F}C_b^\infty.$$

5.2.3. *The process associated with* $(\mathscr{E}^\varphi, D(\mathscr{E}^\varphi))$. Let $\varphi \in D(\mathscr{E})$ with $\varphi > 0$ m^{FV}-a.e. According to Theorem 3.5 there exists a diffusion process $\mathbf{M}^\varphi = (\Omega,$

$\mathcal{F}$, $(\mathcal{F}_t)_{t\geq 0}$, $(X_t)_{t\geq 0}$, $(P_\mu)_{\mu\in E})$ that is properly associated with $(\mathcal{E}^\varphi, D(\mathcal{E}^\varphi))$ and that solves the martingale problem for $(L^\varphi, \mathcal{F}C_b^\infty)$ [cf. Theorem 3.5(iv)] and can be viewed as a Fleming–Viot process with interactive selection (cf. Section 5.3 for the definition of a Fleming–Viot process with selection). By Remark 3.6(iii), $\varphi^2 m$ is the invariant measure of $\mathbf{M}^\varphi$. We recall that Theorem 4.1 about large deviations also applies to $(\mathcal{E}^\varphi, D(\mathcal{E}^\varphi))$ and $\mathbf{M}^\varphi$ [cf. Remark 4.3(i)].

5.3. *Nonsymmetric perturbation.* So far we have only considered forms that we obtained as a kind of symmetric perturbation of the form associated with L^{FV}; more precisely, we only changed m^{FV} into $\varphi^2 m^{\mathrm{FV}}$. Our starting point in this section is a form $(\mathcal{E}^\varphi, D(\mathcal{E}^\varphi))$ of this type, that is, in the notation of Section 2, $m := \varphi^2 m^{\mathrm{FV}}$ and $(\mathcal{E}, D(\mathcal{E})) := (\mathcal{E}^\varphi, D(\mathcal{E}^\varphi))$ (which includes the case $\varphi \equiv 1$). Then we obtain for every b satisfying (2.9) by Theorem 2.4 that $(\mathcal{E}_\alpha^{\varphi,b}, D(\mathcal{E}_\alpha^{\varphi,b}))$ is a quasi-regular semi-Dirichlet form for α sufficiently big. According to Theorem 3.5 there exists an associated diffusion $\mathbf{M}_\alpha^b$ which by Proposition 2.7 solves the martingale problem for $(L^{\mathrm{FV}} + (1/\varphi)\Gamma^{\mathrm{FV}}(\varphi, \cdot) - \beta(\cdot, 1) - \alpha, \mathcal{F}C_b^\infty)$. Hence $\mathbf{M}_\alpha^b$ is a Fleming–Viot process with "interactive selection" in the following sense:

1. A Fleming–Viot process with selection, as defined in [6], 10.1.1 and [13], Chapter 3, is a solution of the martingale problem for $(L^s, \mathcal{F}C_b^\infty)$ with

$$L^s u(\mu) = L^{\mathrm{FV}} u(\mu) + \langle \nabla u(\mu), V(\mu)\rangle_\mu, \qquad u \in \mathcal{F}C_b^\infty, \tag{5.8}$$

   where $V(\mu, x) = \int_S v(y, x)\mu(dy)$. In [6], 10.1.1 (resp. [13], Chapter 3) the function $v(y, x)$ is called the *fitness function* (resp. *selection intensity*) and is assumed to be bounded. It describes the relative fitness of allele y compared with allele x. It is therefore reasonable to assume $v(y, x) = v(x, y)$.
2. We call a solution of the martingale problem $(L^s, \mathcal{F}C_b^\infty)$ with L^s defined in (5.8) and with an arbitrary measurable function $V(\mu, x)$ on $E \times S$ a Fleming–Viot process with *interactive selection*. The function $V(\mu, x)$ now quantifies the fitness of allele x compared with the fitness of all possible alleles in the support of μ. Hence our process $\mathbf{M}_\alpha^b$ (up to the killing α) has the fitness function

$$V(\mu, x) = \frac{\nabla_x \varphi(\mu)}{\varphi(\mu)} + b(\mu, x). \tag{5.9}$$

   If

$$\sup_\mu \sup_x \left(\frac{\nabla_x \varphi(\mu)}{\varphi(\mu)} + b(\mu, x)\right) < \infty,$$

this process can be constructed by Dawson's Girsanov transformation [6], 10.1.1, which gives also that the corresponding martingale problem is wellposed (hence unique). This implies that our process $\mathbf{M}_\alpha^b$ is (up to the killing α) equivalent with Dawson's in this case. Finally, we emphasize that for the *existence* of $\mathbf{M}_\alpha^b$ we only need to assume that $\varphi \in D(\mathscr{E})$, that $\varphi > 0$ m^{FV}-a.e. and that $\sup_\mu \|b(\mu)\|_\mu < \infty$. However, the price we have to pay is that, in contrast to [6], we have to assume neutral mutation (cf. Section 5.1).

REMARKS. (i) Note that in item 1, because v is symmetric, we have that

$$V(\mu, x) = \tfrac{1}{2}\nabla_x \psi(\mu) \quad \text{where } \psi(\mu) := \int\int v(x, y)\mu(dx)\mu(dy).$$

Hence, if $\varphi^2 := e^{\psi} \in D(\mathscr{E})$, then

$$V(\mu, x) = \varphi^{-1}\nabla_x \varphi(\mu)$$

and $(\mathscr{E}^{\varphi}, D(\mathscr{E}^{\varphi}))$ yields a Fleming–Viot process with selection and invariant measure $\varphi^2 m^{\mathrm{FV}}$, which is moreover the reversing measure.

(ii) We know that if b additionally satisfies

$$\int(\langle b, \nabla v\rangle + \alpha v)\varphi^2\, dm^{\mathrm{FV}} \geq 0 \quad \text{for all } v \in \mathscr{F}C_b^\infty,\, v \geq 0,$$

then $(\mathscr{E}_\alpha^b, D(\mathscr{E}_\alpha^b))$ is a Dirichlet form [cf. Remark 2.6(ii)]. Hence if $b = \nabla w$ for some function $w \in D \subset L^{\varphi}$ satisfying $L^{\varphi} w \leq \alpha$ (i.e., $L^{\mathrm{FV}} w \leq \alpha$ if $\varphi \equiv 1$, which is, e.g., fulfilled for a $w \in \mathscr{F}C_b^\infty$ if α is sufficiently big), then $\mathbf{M}_\alpha^b$ has a dual process $\mathbf{M}_\alpha^b$ [cf. Remark 2.6(iii) and Theorem 3.5(iii)].

6. Support properties in the general case. Consider the form $\mathscr{E}^b$ given in Definition 2.1, and assume that the conditions of Theorem 2.4(i) hold so that $(\mathscr{E}_\alpha^b, D(\mathscr{E}_\alpha^b))$ is a semi-Dirichlet form. Then by Theorem 3.5, $(\mathscr{E}_\alpha^b, D(\mathscr{E}_\alpha^b))$ is quasi-regular and there exists an E-valued diffusion process $\mathbf{M} := (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, (X_t)_{t \geq 0}, (P_\mu)_{\mu \in E})$ properly associated with $(\mathscr{E}_\alpha^b, D(\mathscr{E}_\alpha^b))$. An important consequence of this association is the following lemma. This is [20], Chapter IV, Proposition 5.30, but modified to exploit the fact that $(X_t)_{t\geq 0}$ has continuous sample paths.

LEMMA 6.1. *A set $B \in \mathscr{B}(E)$ is $\mathscr{E}_\alpha^b$-exceptional, if and only if*

$$P_\mu(\omega \mid \exists\, 0 < t < \zeta(\omega) \text{ so that } X_t(\omega) \in B) = 0 \quad \text{for } \mathscr{E}_\varphi^b \text{ q.e. } \mu.$$

This says that the process $(X_t)_{t\geq 0}$ will only hit $\mathscr{E}_\alpha^b$-nonexceptional sets. Therefore, we can study the same path properties of $(X_t)_{t\geq 0}$ by determining which sets are $\mathscr{E}_\alpha^b$-exceptional.

Lemma 2.5(ii) compares the two forms $\mathscr{E}_1$ and $\mathscr{E}_\alpha^b$ and shows that the closed domains $D(\mathscr{E}_1)$ and $D(\mathscr{E}_\alpha^b)$ coincide and that the norms $(\tilde{\mathscr{E}}_\alpha^b)^{1/2}$ and $\mathscr{E}_1^{1/2}$ are equivalent there. It follows that the notions of a nest, and hence of an exceptional set, are also the same for these two forms. In other words, it is without loss of generality that we take $b = 0$ and look only at the symmetric

Dirichlet form $(\mathscr{E}, D(\mathscr{E}))$, which is the closure of

$$\mathscr{E}(u,v) = \int \langle \nabla u(\mu), \nabla v(\mu)\rangle_\mu m(d\mu) = \int \Gamma(u,v)\, dm, \qquad u, v \in \mathscr{F}C_b^\infty.$$

LEMMA 6.2. *Suppose $(u_n)_{n\in\mathbb{N}}$ is a bounded sequence in $D(\mathscr{E})$, $(\tilde{u}_n)_{n\in\mathbb{N}}$ are $\mathscr{E}$-quasi-continuous m-versions of $(u_n)_{n\in\mathbb{N}}$ and $\tilde{u}_n \to u$ $\mathscr{E}$-q.e. on E. Then $u \in D(\mathscr{E})$ and u is $\mathscr{E}$-quasi-continuous.*

(i) *If, in addition, there is $h \in L^1(E; m)$ so that for some $v \in D(\mathscr{E})$, we have $\Gamma(u_n, v) \le h$, then $\Gamma(u, v) \le h$.*

(ii) *Alternatively, if $\Gamma(u_n, u_n) \le h$, then $\Gamma(u, u) \le h$.*

PROOF. Because $(u_n)_{n\in\mathbb{N}}$ is bounded in $D(\mathscr{E})$, there is a subsequence $(u_{n_k})_{k\in\mathbb{N}}$ whose averages $w_N = (1/N)\sum_{k=1}^N u_{n_k}$ converge strongly to some limit $w \in D(\mathscr{E})$, and so that $\mathscr{E}$-quasi-continuous m-versions $\tilde{w}_N$ of w_N converge $\mathscr{E}$-q.e. to an $\mathscr{E}$-quasi-continuous m-version $\tilde{w}$ of w. Because the functions $\tilde{u}_n$ are already $\mathscr{E}$-quasi-continuous, and because they converge $\mathscr{E}$-q.e. to u, it follows that u itself is an $\mathscr{E}$-quasi-continuous m-version of w.

In addition, in case (i) we have that, in $L^1(E; m)$,

$$\Gamma(u,v) = \lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^N \Gamma(u_{n_k}, v) \le \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^N h = h,$$

whereas in case (ii) we get

$$\begin{aligned}\Gamma(u,u) &= \lim_{N\to\infty}\Gamma\left(\frac{1}{N}\sum_{k=1}^N u_{n_k}, \frac{1}{N}\sum_{k=1}^N u_{n_k}\right)\\ &\le \lim_{N\to\infty}\frac{1}{N^2}\sum_{i,j=1}^N \Gamma(u_{n_i}, u_{n_j})\\ &\le \lim_{N\to\infty}\frac{1}{N^2}\sum_{i,j=1}^N \Gamma^{1/2}(u_{n_i})\Gamma^{1/2}(u_{n_j}) \le h. \qquad \square\end{aligned}$$

The following lemma shows that we can replace the bounded continuous functions f_i in (2.7) by bounded measurable functions, and the function u is still an $\mathscr{E}$-quasi-continuous member of $D(\mathscr{E})$.

LEMMA 6.3. *For any $\varphi \in C^\infty(\mathbb{R}^k)$ and $f_i \in \mathscr{B}_b(S)$, $i = 1,\dots,k$, the function*

$$u(\mu) = \varphi(\mu(f_1),\dots,\mu(f_k)) \tag{6.1}$$

is an $\mathscr{E}$-quasi-continuous member of $D(\mathscr{E})$ and for m-a.e. $\mu \in E$ we have

$$\Gamma(u)(\mu) = \sum_{i,j}\frac{\partial\varphi}{\partial y_i}(\mu(f_1),\dots,\mu(f_k))\frac{\partial\varphi}{\partial y_j}(\mu(f_1),\dots,\mu(f_k))\langle f_i, f_j\rangle_\mu.$$

PROOF. Let $\mathscr{H} = \{g \in \mathscr{B}_b(S) \mid \mu \mapsto u(\mu) := \mu(g)$ is an $\mathscr{E}$-quasi-continuous member of $D(\mathscr{E})$ and $\Gamma(u) = \langle g, g\rangle_\mu\}$. Then $\mathscr{H}$ is a linear space containing $C_b(S)$. Suppose $g_n \in \mathscr{H}$ and $g_n \uparrow g \in \mathscr{B}_b(S)$. Then $\mu(g_n) \uparrow \mu(g)$ at every $\mu \in E$ and so $u_n(\mu) \to u(\mu)$ boundedly pointwise and so in $L^2(E; m)$. Also, we have

$$\begin{aligned}\Gamma(u_n - u_m)(\mu) &= \langle g_n - g_m\rangle_\mu \\ &\le \int (g_n - g_m)^2(x)\mu(dx) \\ &\le 4\left(\|g\|_\infty^2 + \|g_1\|_\infty^2\right).\end{aligned}$$

Letting $n, m \to \infty$, we see that $\Gamma(u_n - u_m) \to 0$ pointwise boundedly, and so also in $L^1(E; m)$, so that $(u_n)_{n\in\mathbb{N}}$ is $\mathscr{E}_1$-Cauchy. By Lemma 6.2, $u(\mu)$ is an $\mathscr{E}$-quasi-continuous member of $D(\mathscr{E})$ and because $\Gamma(u_n) \to \Gamma(u)$ in $L^1(E; m)$ as $n \to \infty$, we find that $\Gamma(u) = \lim_n \langle g_n, g_n\rangle_\mu = \langle g, g\rangle_\mu$. Therefore, $g \in \mathscr{H}$ and so the monotone class theorem says that $\mathscr{H} = \mathscr{B}_b(S)$. The rest of the proof follows from general results on the calculus of square field operator Dirichlet forms; see, for example, [4] or [26]. □

REMARK. By Lemma 6.3, the definition of $(\mathscr{E}, D(\mathscr{E}))$ does not depend on the topology of S, but only on $\mathscr{B}(S)$.

Interpreted in terms of the process $(X_t)_{t\ge 0}$, Theorem 6.4 (in connection with Lemma 6.1) shows that if X_t does not charge a set F at fixed times, then the measure X_t will never charge F.

THEOREM 6.4. *Let $F \in \mathscr{B}(S)$. If $\mu(F) = 0$ for m-a.e. $\mu \in E$, then $\mu(F) = 0$ for $\mathscr{E}$-q.e. $\mu \in E$.*

PROOF. The result follows immediately from the fact that $\mu(F)$ is $\mathscr{E}$-quasi-continuous (cf. [20], IV.3.3(iii)). □

THEOREM 6.5. *If m-a.e. μ is purely atomic (i.e., $\sum_{x\in S}\mu(\{x\}) = 1$), then $\mathscr{E}$-q.e. μ is purely atomic.*

COMMENT. Applied to the symmetric Fleming–Viot process with bounded mutation operator, Theorem 6.5 shows that for $\mathscr{E}$-q.e. μ, X_t is purely atomic P_μ a.s. In this special case, this was first proved by Ethier and Kurtz (cf. [12], Chapter 10, Theorem 4.5 and [13], Theorem 7.2).

Before we can prove Theorem 6.5, we need some preparations.

DEFINITION 6.6. Let $\mathscr{A} = (A_j^n)_{j,n=1}^\infty$ be a collection of Borel subsets of S so that:

(i) For every n, $(A_j^n)_{j=1}^\infty$ is a partition of S, that is, $S = \bigcup_{j=1}^\infty A_j^n$ and $A_j^n \cap A_i^n = \varnothing$ for $i \ne j$.

(ii) $\operatorname{diam}(A_j^n) \le 1/n$ for all $j \ge 1$.

(iii) $(A_j^{n+1})_{j=1}^{\infty}$ is a refinement of $(A_j^n)_{j=1}^{\infty}$, that is, for each n, $j \ge 1$ there exists $i \ge 1$ so $A_j^{n+1} \subseteq A_i^n$.

LEMMA 6.7. *With $\mathscr{A}$ defined as above and μ any finite Borel measure on S, we have*

$$\inf_n \sup_j \mu(A_j^n) = \sup_{x \in S} \mu(\{x\}).$$

PROOF. Because $(A_j^n)_{j=1}^{\infty}$ covers the whole space for each n, it is clear that $\sup_j \mu(A_j^n) \ge \sup_{x \in S} \mu(\{x\})$ for every $n \ge 1$.

Now suppose $\inf_n \sup_j \mu(A_j^n) > c \ge 0$. For each $n \ge 1$, choose $j(n)$ so that $\mu(A_{j(n)}^n) \ge c$, and define $n^* = \sup\{m \mid A_{j(n)}^n \cap A_{j(m)}^m \ne \varnothing\}$. If $n^* < \infty$ for all n, then let $n_1 = 1$ and $n_{k+1} = n_k^* + 1$. However, the sets $(A_{j(n_k)}^{n_k})_{k \in \mathbb{N}}$ are disjoint and $\mu(A_{j(n_k)}^{n_k}) \ge c$ for every $k \in \mathbb{N}$. This is a contradiction, so n^* must be infinite for some n. In other words, there exists an increasing sequence $(n_i)_{i \in \mathbb{N}}$ so that $A_{j(n_i)}^{n_i} \supseteq A_{j(n_{i+1})}^{n_{i+1}}$ for all $i \ge 1$. By Definition 6.6(ii), the intersection contains at most one point, and because $\mu(A_{j(n_i)}^{n_i}) \ge c$ for all $i \ge 1$, we conclude that there exists $x \in S$ so that $\{x\} = \bigcap_i A_{j(n_i)}^{n_i}$ and $\mu(\{x\}) \ge c$. Taking the supremum over such c, now shows that $\inf_n \sup_j \mu(A_j^n) \le \sup_{x \in S} \mu(\{x\})$. $\square$

PROOF OF THEOREM 6.5. From Lemma 6.3 we recall that for any Borel set $B \subseteq S$, the function $\mu \mapsto \mu(B)$ is an $\mathscr{E}$-quasi-continuous member of $D(\mathscr{E})$ and $\Gamma(\mu(A), \mu(B)) = \mu(A \cap B) - \mu(A)\mu(B) \le \mu(A \cap B)$.

Now let $\mathscr{A}$ be the collection in Definition 6.6 and for any Borel set B define $\varphi_{n,B}^m(\mu) = \sup_{j=1}^n \mu(B \cap A_j^m)$, $\mu \in E$. Then $\varphi_{n,B}^m$ is an $\mathscr{E}$-quasi-continuous member of $D(\mathscr{E})$ and for every n_1, n_2, m_1, $m_2 \in \mathbb{N}$ and A, $B \in \mathscr{B}(S)$ we have, by [23], the proof of Lemma 3.2, that

$$\Gamma\big(\varphi_{n_1,A}^{m_1}, \varphi_{n_2,B}^{m_2}\big) \le \sup_{ij} \Gamma\big(\mu(A \cap A_i^{m_1}), \mu(B \cap A_j^{m_2})\big) \le \mu(A \cap B).$$

In particular, taking $A = B$, the sequence $(\varphi_{n,B}^m)_{n \in \mathbb{N}}$ satisfies $\mathscr{E}(\varphi_{n,B}^m) = \int \Gamma(\varphi_{n,B}^m)\, dm \le \int \mu(B)\, dm \le 1$, and because $\varphi_{n,B}^m$ are uniformly bounded pointwise by 1, we conclude that $(\varphi_{n,B}^m)_{n \in \mathbb{N}}$ is bounded in $D(\mathscr{E})$.

Also $\varphi_{n,B}^m \to \varphi_B^m := \sup_{j=1}^{\infty} \mu(B \cap A_j^m)$ pointwise everywhere, so we may apply Lemma 6.2 and conclude that the function φ_B^m is an $\mathscr{E}$-quasi-continuous member of $D(\mathscr{E})$ and $\Gamma(\varphi_B^{m_1}, \varphi_A^{m_2})(\mu) \le \mu(A \cap B)$ for any m_1, $m_2 \in \mathbb{N}$ and A, $B \in \mathscr{B}(S)$. Now applying Lemma 6.2 to the decreasing sequence $(\varphi_B^m)_{m \in \mathbb{N}}$ shows us that all of the functions $\varphi_B := \lim_m \varphi_B^m$ are also $\mathscr{E}$-quasi-continuous members of $D(\mathscr{E})$ and that $\Gamma(\varphi_B, \varphi_A)(\mu) \le \mu(A \cap B)$. We note, by Lemma 6.7, that $\varphi_B(\mu)$ measures the mass of the largest atom in B of μ, for small enough B.

Now define $u_n^m = \sum_{j=1}^m \varphi_{A_j^m}$. Then u_n^m is an $\mathscr{E}$-quasi-continuous member of

$D(\mathscr{E})$ and because $(A_j^m)_{j\in\mathbb{N}}$ is a partition,

$$\begin{aligned}\Gamma(u_n^m)(\mu) &= \sum_{j=1}^{n}\Gamma(\varphi_{A_j^m})(\mu) + \sum_{i\neq j}\Gamma(\varphi_{A_i^m}, \varphi_{A_j^m})\\ &\le \sum_{j=1}^{n}\mu(A_j^m) + 0\\ &\le 1.\end{aligned}$$

Applying Lemma 6.2 twice more, first letting $n\to\infty$ and then $m\to\infty$, we conclude that the limit function u is an $\mathscr{E}$-quasi-continuous member of $D(\mathscr{E})$. It is easy to check that

$$u(\mu) = \sum_{x\in S}\mu(\{x\}).$$

However, because we assumed that $u = 1$ m-a.e., it follows that $u = 1$ $\mathscr{E}$-q.e., which gives us the required result. □

We now have enough information about the process $(X_t)_{t\ge 0}$ to conclude that its sample paths are continuous in variation norm. In the special case of the Fleming–Viot process with bounded mutation operator this was proved by Shiga [28]. Before we can prove the result, we first require the following technical lemma.

LEMMA 6.8. *Let (S, d) be a Polish space and let $t\mapsto\mu_t$ be a map from $\mathbb{R}$ into $\mathscr{M}_1(S)$ satisfying the following conditions*:

1. *$t\mapsto\mu_t$ is continuous into $(\mathscr{M}_1(S)$, weak).*
2. *μ_t is purely atomic for all t.*
3. *There exist sequences, $(x_i)_{i\in\mathbb{N}}$ dense in S and $(r_j)_{j\in\mathbb{N}}$ dense in $(0,\infty)$, such that* (a) *$\mu_t(\bigcup_{i,j\in\mathbb{N}}\{y \mid d(x_i, y) = r_j\}) = 0$ for all t*; (b) *for every $i, j\in\mathbb{N}$, the map $t\mapsto\sup_{x\in B_{ij}}\mu_t(\{x\})$ is continuous, where $B_{ij} := \{y \mid d(x_i, y) < r_j\}$.*

Then $t\mapsto\mu_t$ is continuous into $(\mathscr{M}_1(S), \|\cdot\|_{\mathrm{var}})$.

PROOF. We first note that for any $y\in S$ and $\varepsilon>0$, we can find B_{ij} so that $y\in B_{ij}\subset\{x \mid d(y, x) < \varepsilon\}$. We also note that the boundary of B_{ij} is contained in $\{y \mid d(x_i, y) = r_j\}$ and so has μ_t-measure zero. Therefore, by weak continuity, the map $t\mapsto\mu_t(B_{ij})$ is continuous for all $i, j\in\mathbb{N}$.

Fix t and suppose that μ_t has the representation $\mu_t = \sum_{k=1}^{\infty}p_k\varepsilon_{y_k}$. Select k with $p_k>0$ and choose B_{ij} so that $y_k\in B_{ij}$ and $\mu_t(B_{ij}^c)\ge 1-(7p_k/6)$. Therefore, y_k is the unique largest atom of μ_t in B_{ij}.

Now take $\delta>0$ so small that if $s\in(t-\delta, t+\delta)$, then

$$|\mu_t(B_{ij}) - \mu_s(B_{ij})| \le p_k/6 \tag{6.2}$$

and

$$\left|\sup_{x\in B_{ij}}\mu_t(\{x\}) - \sup_{x\in B_{ij}}\mu_s(\{x\})\right| = \left|p_k - \sup_{x\in B_{ij}}\mu_s(\{x\})\right| \le p_k/6. \tag{6.3}$$

It follows that μ_s has an atom of mass greater than or equal to $5p_k/6$ in B_{ij}, and because $\mu_s(B_{ij}) \le 8p_k/6$, it is the unique largest atom of μ_s in B_{ij}. Let us denote it by $y(s)$, and we note that $\mu_s(B_{ij} \setminus \{y(s)\}) \le p_k/2$ for all $s \in (t-\delta, t+\delta)$.

We claim that $s \mapsto y(s)$ is continuous from $(t-\delta, t+\delta)$ into S. To show this, choose $s \in (t-\delta, t+\delta)$ and $\varepsilon > 0$. Pick $B'_{ij} \subseteq B_{ij}$ so that $y(s) \in B'_{ij} \subseteq \{x \mid d(x, y(s)) \le \varepsilon\}$ and take δ' so small that $(s-\delta', s+\delta') \subseteq (t-\delta, t+\delta)$, and if $s' \in (s-\delta', s+\delta')$, then $|\mu_s(B'_{ij}) - \mu_{s'}(B'_{ij})| \le p_k/6$. Then $\mu_{s'}(B'_{ij}) \ge 4p_k/6$, and because $\mu_{s'}(B'_{ij} \setminus \{y(s')\}) \le p_k/2$, we conclude that $y(s') \in B'_{ij}$ so $d(y(s), y(s')) \le \varepsilon$.

Now from the continuity and Lemma 6.8, condition 3(a), we see that $d(x_i, y(s)) = d(x_i, y(t))$ for all $i \in \mathbb{N}$ and $s \in (t-\delta, t+\delta)$, so by taking a subsequence of $(x_i)_{i\in\mathbb{N}}$ converging to $y(t)$, we conclude that $y(s) = y(t)$. In other words, the continuous map $s \mapsto y(s)$ must be constant. Therefore, $\sup_{x\in B_{ij}} \mu_s(\{x\}) = \mu_s(\{y_k\})$ for all $s \in (t-\delta, t+\delta)$.

Choose N so that $\mu_t(\{y_1, \dots, y_N\}) = \sum_{k=1}^N p_k \ge 1-\varepsilon$ and choose δ^* so that if $s \in (t-\delta^*, t+\delta^*)$, then $\sum_{k=1}^N |\mu_s(\{y_k\}) - \mu_t(\{y_k\})| \le \varepsilon$. For any Borel set B, we write

$$B = \bigcup_{k=1}^{N} (B \cap \{y_k\}) \cup (B \setminus \{y_1, \dots, y_N\}), \tag{6.4}$$

and so

$$\begin{aligned} &|\mu_s(B) - \mu_t(B)| \\ &\quad \le |\mu_s(B \setminus \{y_1, \dots, y_N\})| + |\mu_t(B \setminus \{y_1, \dots, y_N\})| \\ &\qquad + \sum_{k=1}^{N} |\mu_t(\{y_k\}) - \mu_s(\{y_k\})| \\ &\quad \le 4\varepsilon. \end{aligned}$$

Taking the supremum over $B \in \mathscr{B}(S)$ gives $\|\mu_s - \mu_t\|_{\mathrm{var}} \le 4\varepsilon$, which proves the continuity of the map $t \mapsto \mu_t$ at the point t. Because t was arbitrary, this gives the result. □

THEOREM 6.9. *If m-a.e. $\mu \in E$ is purely atomic, then*

$$P_\mu(t \to X_t \text{ is continuous in variation norm}) = 1$$

for $\mathscr{E}$-q.e. $\mu \in E$.

PROOF. From Theorem 3.5(ii), we already know that the process has continuous sample paths when $\mathscr{M}_1(S)$ is equipped with the weak topology. From Theorem 6.5, we know that

$$P_\mu(X_t \text{ is purely atomic for all } t \ge 0) = 1 \tag{6.5}$$

for $\mathscr{E}$-q.e. $\mu \in E$. Now let $(x_i)_{i\in\mathbb{N}}$ be any countable dense set in S and choose $(r_j)_{j\in\mathbb{N}}$ dense in $(0, \infty)$ so that $\nu_0(F) = 0$, where $F := \bigcup_{ij} \{y \mid d(x_i, y) = r_j\}$.

Here ν_0 is the mean measure of m; that is, $\nu_0 := \int \mu m(d\mu)$. By Theorem 6.4, we conclude that $P_\mu(X_t(F) = 0 \text{ for all } t \geq 0) = 1$, for $\mathscr{E}$-q.e. $\mu \in E$. Also, in the proof of Theorem 6.5, we have proved the $\mathscr{E}$-quasi-continuity of the map $\mu \mapsto \sup_{x \in B} \mu(\{x\})$ for any Borel set B, and so for every i, $y \in \mathbb{N}$ (by [19], IV.5.30),

$$P_\mu\Big(t \mapsto \sup_{x \in B_{ij}} X_t(\{x\}) \text{ is continuous}\Big) = 1 \tag{6.6}$$

for $\mathscr{E}$-q.e. $\mu \in E$, where B_{ij} is defined as in part 3(b) of the statement of Lemma 6.8.

Finding a common set of P_μ-probability 1 for $\mathscr{E}$-q.e. $\mu \in E$ on which all of these (countably many) conditions hold and applying Lemma 6.8 gives the result. □

7. Support properties in the perturbed Fleming–Viot case. We continue to explore the sample path properties of $(X_t)_{t \geq 0}$ by looking for $\mathscr{E}_\alpha^b$-exceptional sets, but specialize to the case where $m = \varphi^2 m^{\mathrm{FV}}$ (i.e., the Fleming–Viot process with interactive selection; cf. Section 5.2). We make the additional assumption that $c_1 \leq \varphi^2 \leq c_2$ for constants $c_2 \geq c_1 > 0$. This means that a set B is $\mathscr{E}_\alpha^b$-exceptional if and only if it is $\mathscr{E}$-exceptional, where $\mathscr{E}$ is the form associated with the unperturbed Fleming–Viot process. That is $(\mathscr{E}, D(\mathscr{E}))$ is the closure of

$$\mathscr{E}(u, v) = \int \langle \nabla u(\mu), \nabla v(\mu)\rangle_\mu m(d\mu) = \int \Gamma(u, v)\, dm, \qquad u, v \in \mathscr{F}C_b^\infty,$$

where $m = m^{\mathrm{FV}}$ is the symmetrizing probability measure for the Fleming–Viot process. When we want to emphasize the dependence on θ and ν_0 [cf. (5.1)] we will use the notation $\Pi_{\theta, \nu_0} := m$. Our ability to judge whether or not sets are $\mathscr{E}$-exceptional depends to a great extent on our understanding of the measure m. So let us begin with a few observations about the measure m. It is immediate that if $\nu_0(F) = 0$, then $\mu(F) = 0$ for m-a.e. $\mu \in E$, and also that m-a.e. μ is purely atomic. The quasi-everywhere versions of these statements now follow from Lemmas 6.4 and 6.5. In fact, because m only charges the set of purely atomic measures on S, all results from Section 6 apply here.

From Lemma 7.2 it follows that if $\nu_0(F) > 0$, then $\mu(F) > 0$ for m-a.e. $\mu \in E$. In addition, from the fact that $P(\rho_i > 0) = 1$ for all $i \in \mathbb{N}$, we see that if ν_0 is nonatomic, then m-a.e. μ has infinitely many atoms. In Proposition 7.6 and Corollary 7.10, we see that the corresponding quasi-everywhere statements are not true in general, but depend on the mutation intensity θ. If there is a lot of mutation (θ is large), then the measure X_t tends to be more spread out, whereas if there is very little mutation (θ is small), the support of the measure X_t occasionally collapses, in extreme cases even to a single point.

The following family of Dirichlet distributions, which are multidimensional analogs of the Beta distribution, will appear throughout our calculations in this section.

DEFINITION 7.1. If $\theta_1, \dots, \theta_n > 0$, then the Dirichlet$(\theta_1, \dots, \theta_n)$ distribution is the measure on

$$S_n := \left\{ x \in \mathbb{R}^n \mid x_i \geq 0 \; \forall i, \; \sum_{i=1}^{n} x_i = 1 \right\}$$

given by

$$\nu(dx) = \frac{\Gamma(\theta_1 + \cdots + \theta_n)}{\Gamma(\theta_1) \cdots \Gamma(\theta_n)} x_1^{\theta_1 - 1} \cdots x_n^{\theta_n - 1} \, dx_1 \cdots dx_{n-1}.$$

LEMMA 7.2. *If $(A_j)_{j=1}^n$ is a measurable partition of S, then under Π_{θ, ν_0} the random vector $(\mu(A_1), \dots, \mu(A_n))$ has a* Dirichlet$(\theta\nu_0(A_1), \dots, \theta\nu_0(A_n))$ *distribution on S_n. In addition, if $\nu_0(A) = 0$, then $\mu(A) = 0$ m-a.e. $\mu \in E$.*

PROOF. For $1 \leq i \leq n$, choose $x_i \in A_i$ and define $\tilde{\nu}_0 = \sum_{i=1}^n \nu_0(A_i)\varepsilon_{x_i}$. From the formula for the invariant measure, it is clear that $(\mu(A_1), \dots, \mu(A_n))$ has the same distribution under Π_{θ, ν_0} and $\Pi_{\theta, \tilde{\nu}_0}$. The result now follows from Lemma 2.2 of Ethier and Griffiths [11], where the lemma is proved to hold under $\Pi_{\theta, \tilde{\nu}_0}$.

For any Borel set A in S, we have

$$\begin{aligned} \int \mu(A) m(d\mu) &= E\Big(\sum_i \rho_i \varepsilon_{\xi_i}(A) \Big) = \sum_i E(\rho_i) E\big(\varepsilon_{\xi_i}(A)\big) \\ &= \sum_i E(\rho_i) P(\xi_i \in A) = \sum_i E(\rho_i) \nu_0(A) = \nu_0(A). \end{aligned}$$

Therefore, if $\nu_0(A) = 0$, then $\mu(A) = 0$ m-a.e. $\square$

LEMMA 7.3. *Suppose $U = (U_1, \dots, U_k) \in S_k$ has a* Dirichlet$(\theta_1, \dots, \theta_k)$ *distribution, that $V = (V_1, \dots, V_{n-k}) \in S_{n-k}$ has a* Dirichlet$(\theta_{k+1}, \dots, \theta_n)$ *distribution, R has a* Beta$(\theta_1 + \cdots + \theta_k, \theta_{k+1} + \cdots + \theta_n)$ *distribution on* $(0, 1)$ *and that U, V and R are independent. Then $W := (RU_1, RU_2, \dots, RU_k, (1 - R)V_1, \dots, (1 - R)V_{n-k}) \in S_n$ has a* Dirichlet$(\theta_1, \dots, \theta_n)$ *distribution.*

PROOF. Let $\tilde{\theta}_1 = \theta_1 + \cdots + \theta_k$ and $\tilde{\theta}_2 = \theta_{k+1} + \cdots + \theta_n$. Define two probability measures on $S := \{1, \dots, n\}$ by $\nu_1(i) = \theta_i / \tilde{\theta}_1$ for $1 \leq i \leq k$ (zero otherwise) and $\nu_2(i) = \theta_i / \tilde{\theta}_2$ for $k + 1 \leq i \leq n$ (zero otherwise). Let μ_1 and μ_2 be random measures with distributions $\Pi_{\tilde{\theta}_1, \nu_1}$ and $\Pi_{\tilde{\theta}_2, \nu_2}$, respectively, and independently of those let R have a Beta$(\tilde{\theta}_1, \tilde{\theta}_2)$ distribution. By Lemma 2.1 of Ethier and Griffiths [11], the random measure $\mu = R\mu_1 + (1 - R)\mu_2$ has a $\Pi_{\theta, \nu}$ distribution, where $\theta := = \tilde{\theta}_1 + \theta_2$ and $\nu(i) := \theta_i / \theta$ for $1 \leq i \leq n$. Applying Lemma 1, we have

$$\begin{aligned} (\mu_1(\{1\}), \dots, \mu_1(\{k\})) &\sim \text{Dirichlet}(\theta_1, \dots, \theta_k), \\ (\mu_2(\{k + 1\}), \dots, \mu_2(\{n\})) &\sim \text{Dirichlet}(\theta_{k+1}, \dots, \theta_n), \\ (\mu(\{1\}), \dots, \mu(\{n\})) &\sim \text{Dirichlet}(\theta_1, \dots, \theta_n), \end{aligned}$$

which gives the requires result. $\square$

LEMMA 7.4. *Suppose that $1 \le k \le n$, and let ν be the* Dirichlet$(\theta_1, \dots, \theta_k)$ *measure on S_k and $\tilde{\nu}$ be the* Dirichlet$(\theta_1, \dots, \theta_n)$ *measure on S_n. Then, for every $w \in C^\infty(\mathbb{R}^n)$.*

$$\frac{1}{2}\left(\frac{1-\tilde{\theta}_2}{2^{\tilde{\theta}_1}}\right)\frac{\Gamma(\tilde{\theta}_1+\tilde{\theta}_2)}{\Gamma(\tilde{\theta}_1)\Gamma(\tilde{\theta}_2)}\int_{S_k}\psi^2(u_1,\dots,u_k,0,\dots,0)\nu(du)$$
$$\le \int_B(\langle\nabla\psi(w), a(w)\nabla\psi(w)\rangle_{\mathbb{R}^n} + \psi^2(w))\tilde{\nu}(dw).$$

Here $\tilde{\theta}_1 = \theta_1 + \cdots + \theta_k$, $\tilde{\theta}_2 = \theta_{k+1} + \cdots + \theta_n$ and for $w \in S_n$, $a(w)$ is the $n \times n$ matrix with entries $a(w)_{ij} = (w_i\delta_{ij} - w_iw_j)$. Also $B := \{w \in S_n \mid \sum_{i=1}^k w_i > 1/2\}$.

PROOF. Without loss of generality we assume that $\tilde{\theta}_2 < 1$. Fix $u = (u_1, \dots, u_k) \in S_k$ and $v = (v_1, \dots, v_{n-k}) \in S_{n-k}$. For $r \in [0, 1]$ define $w := w(r) = (u_1r, \dots, u_kr, v_1(1-r), \dots, v_{n-k}(1-r)) \in S_n$ (cf. the statement of Lemma 7.3) and φ: $[0, 1] \to \mathbb{R}$ by $\varphi(r) = (2r-1)^+\psi(w(r))$. Because φ is absolutely continuous and vanishes for $r \le \frac{1}{2}$ we have $\varphi(1) = -\int_{1/2}^1 \varphi'(r)\,dr$ and so

$$\begin{aligned}
\varphi^2(1) &\le \left(\int_{1/2}^1 \varphi'(r)\,dr\right)^2 \\
&\le \left[\int_{1/2}^1 (\varphi'(r))^2 r(1-r)\frac{\Gamma(\tilde{\theta}_1+\tilde{\theta}_2)}{\Gamma(\tilde{\theta}_1)\Gamma(\tilde{\theta}_2)} r^{\tilde{\theta}_1-1}(1-r)^{\tilde{\theta}_2-1}\,dr\right] \\
&\quad\times\left[\frac{\Gamma(\tilde{\theta}_1)\Gamma(\tilde{\theta}_2)}{\Gamma(\tilde{\theta}_1+\tilde{\theta}_2)}\int_{1/2}^1 r^{-\tilde{\theta}_1}(1-r)^{-\tilde{\theta}_2}\,dr\right] \\
&\le \left[\int_{1/2}^1 (\varphi'(r))^2 r(1-r)f(r)\,dr\right]\left[\frac{\Gamma(\tilde{\theta}_1)\Gamma(\tilde{\theta}_2)}{\Gamma(\tilde{\theta}_1+\tilde{\theta}_2)}\frac{2^{\tilde{\theta}_1}}{1-\tilde{\theta}_2}\right],
\end{aligned} \tag{7.1}$$

where f is the density function for a Beta$(\tilde{\theta}_1, \tilde{\theta}_2)$ distribution. Now $\varphi(1) = \psi(u_1, \dots, u_k, 0, \dots, 0)$ and

$$\varphi'(r) = (2r-1)^+\langle\nabla\psi(w(r)), (u_1, \dots, u_k, -v_1, \dots, -v_{n-k})\rangle_{\mathbb{R}^n} + 2\,1_{(1/2,1]}(r)\psi(w(r)),$$

so that

$$(\varphi'(r))^2 \le 2\langle\nabla\psi(w(r)), (u_1, \dots, u_k, v_1, \dots, v_{n-k})\rangle_{\mathbb{R}^n}^2 + 8\,1_{(1/2,1]}(r)\psi^2(w(r)). \tag{7.2}$$

Letting e_k denote the vector in $\mathbb{R}^n$ whose first k entries are 1, and the remaining entries are 0, it is easy to check that

$$r(1-r)(u_1,\dots,u_k,-v_1,\dots,-v_{n-k}) = a(w(r))e_k,$$

and so

$$\begin{aligned}
&\langle \nabla\psi(w(r)), r(1-r)(u_1,\dots,u_k,-v_1,\dots,-v_{n-k})\rangle_{\mathbb{R}^n}^2\\
&\quad= \langle \nabla\psi(w(r)), a(w(r))e_k\rangle_{\mathbb{R}^n}^2\\
&\quad\le \langle \nabla\psi(w(r)), a(w(r))\nabla\psi(w(r))\rangle_{\mathbb{R}^n}\langle e_k, a(w(r))e_k\rangle_{\mathbb{R}^n}\\
&\quad= \langle \nabla\psi(w(r)), a(w(r))\nabla\psi(w(r))\rangle_{\mathbb{R}^n} r(1-r).
\end{aligned}$$

Dividing by $r(1-r)$ gives

$$\begin{aligned}
&r(1-r)\langle \nabla\psi(w(r)), (u_1,\dots,u_k,-v_1,\dots,-v_{n-k})\rangle_{\mathbb{R}^n}\\
&\qquad\le \langle \nabla\psi(w(r)), a(w(r))\nabla\psi(w(r))\rangle_{\mathbb{R}^n}.
\end{aligned}\tag{7.3}$$

Also, because $r(1-r)8\,1_{(1/2,1]}(r) \le 2$, using (7.2) and (7.3) we get

$$r(1-r)(\varphi'(r))^2 \le 2\langle \nabla\psi(w(r)), a(w(r))\nabla\psi(w(r))\rangle_{\mathbb{R}^n} + 2\psi^2(w(r)).$$

Now use the fact that $\varphi(1) = \psi(u_1,\dots,u_k,0,\dots,0)$ and plug back into (7.1) to obtain

$$\begin{aligned}
&\frac{1}{2}\left(\frac{1-\tilde{\theta}_2}{2^{\tilde{\theta}_1}}\right)\frac{\Gamma(\tilde{\theta}_1+\tilde{\theta}_2)}{\Gamma(\tilde{\theta}_1)\Gamma(\tilde{\theta}_2)}\psi^2(u_1,\dots,u_k,0,\dots,0)\\
&\qquad\le \int_{1/2}^1 (\langle \nabla\psi(w(r)), a(w(r))\nabla\psi(w(r))\rangle_{\mathbb{R}^n} + \psi^2(w(r)))f(r)\,dr.
\end{aligned}$$

Now integrating both sides over $u \in S_k$ with respect to ν and over $v \in S_{n-k}$ with respect to Dirichlet$(\theta_{k+1},\dots,\theta_n)$ measure, and using Lemma 7.3 gives the result. □

LEMMA 7.5. *Let D be a dense subspace of $D(\mathscr{E})$ that is closed under composition with smooth functions that vanish at the origin. Let $\tilde{D} := \{\tilde{u} \mid u \in D\}$ be a collection of $\mathscr{E}$-quasi-continuous m-versions of $u \in D$, such that if $u \ge 0$ m-a.e., then $\tilde{u}(\mu) \ge 0$ for all $\mu \in E$. If ν is a finite positive Borel measure on E so that*

$$u \mapsto \int \tilde{u}(\mu)\nu(d\mu) \tag{7.4}$$

is continuous from $(D, \mathscr{E}_1^{1/2})$ to $\mathbb{R}$, then there exists a unique Borel measure ν^* on E, which charges no $\mathscr{E}$-exceptional set, so that

$$\int \tilde{u}(\mu)\nu(d\mu) = \int \tilde{u}(\mu)\nu^*(d\mu) \tag{7.5}$$

for $\tilde{u} \in \tilde{D}$.

PROOF. The map (7.4) extends by continuity to a continuous linear functional on $(D(\mathscr{E}), \tilde{\mathscr{E}}_1^{1/2})$, which is represented by an element $v \in D(\mathscr{E})$; that is,

$$\mathscr{E}_1(u, v) = \int \tilde{u}(\mu)\nu(d\mu) \quad \text{for } u \in D.$$

We want to show that v is 1-excessive. Let φ_n be a sequence of positive smooth functions on $\mathbb{R}$, such that $\varphi_n(0) = 0$, $\varphi_n(x) \to |x|$ and $|\varphi_n'(x)| \le 1$. For $u \in D$ let $u_n := \varphi_n(u)$ so that $u_n \to |u|$ in $\mathscr{E}_1$ and thus

$$\mathscr{E}_1(|u|, v) = \lim_{n\to\infty} \mathscr{E}_1(u_n, v) = \int \tilde{u}_n(z)\nu(dz) \ge 0.$$

Now for an arbitrary $u \in D(\mathscr{E})$, let $u_n \in D$ so that $u_n \to u$ in $\tilde{\mathscr{E}}_1^{1/2}$-norm, which implies that $|u_n| \to |u|$ in $D(\mathscr{E})$, so

$$\mathscr{E}_1(|u|, v) = \lim_{n\to\infty} \mathscr{E}_1(|u_n|, v) \ge 0.$$

Therefore, v is 1-excessive. Let $\tilde{D}(\mathscr{E})$ denote the set of all $\mathscr{E}$-quasi-continuous m-versions of all elements in $D(\mathscr{E})$. Now by [20], VI.2.1, there is a unique σ-finite positive measure ν^* on $(E, \mathscr{B}(E))$, such that ν^* does not charge any $\mathscr{E}$-exceptional sets, $\tilde{D}(\mathscr{E}) \subset L^1(E, \nu^*)$ and

$$\mathscr{E}_1(u, v) = \int \tilde{u}(\mu)\nu^*(d\mu)$$

for all $u \in D(\mathscr{E})$. In the preceding integral, the choice of the $\mathscr{E}$-quasi-continuous version of u does not matter, so, in particular, for $u \in D$ we may choose the $\mathscr{E}$-quasi-continuous version $\tilde{u} \in \tilde{D}$ as in the statement of the lemma. Therefore, (7.5) holds for $\tilde{u} \in \tilde{D}$. □

PROPOSITION 7.6. *Let F be a Borel set in S so that $\nu_0(F) < 1$. Then $\mu(F) > 0$ for $\mathscr{E}$-q.e. $\mu \in E$ if and only if $\theta\nu_0(F) \ge 1$.*

PROOF. $\Leftarrow$: The random variable $\mu \to \mu(F)$ on $(E; m)$ has a Beta distribution with parameters $\theta\nu_0(F)$ and $\theta\nu_0(S \setminus F)$ (cf. Lemma 7.2). For $n \ge 1$ define $u_n \in D(\mathscr{E})$ by $u_n(\mu) = \varphi_n(\mu(F))$, where $\varphi_n\colon [0, \infty) \to \mathbb{R}$ satisfies $\varphi_n(x) = 1 - nx$ for $0 \le x \le 1/n$ and $\varphi_n(x) = 0$ otherwise. Hence by Lemma 6.3,

$$\Gamma(u_n, u_n) = \varphi_n'(\mu(F))^2(\mu(F) - \mu(F)^2),$$

and so

$$\begin{aligned}
\mathscr{E}(u_n) &= \int \Gamma(u_n, u_n)(\mu)m(d\mu) \\
&= \int n^2(\mu(F) - \mu^2(F))1_{\{0 \le \mu(F) \le 1/n\}} m(d\mu) \\
&= cn^2 \int_0^{1/n} (y - y^2) y^{\theta\nu_0(F)-1}(1-y)^{\theta\nu_0(S\setminus F)-1}\,dy \qquad (7.6)\\
&\le c'n^2(1/n)^{\theta\nu_0(F)+1} \\
&= c'n^{1-\theta\nu_0(F)},
\end{aligned}$$

where c, c' are positive constants that are independent of n. The functions u_n are uniformly bounded by 1, and because we have assumed $\theta\nu_0(F) \geq 1$, (7.6) shows us that $(u_n)_{n \in \mathbb{N}}$ is bounded in $\tilde{\mathscr{E}}_1^{1/2}$-norm. The functions u_n are also $\mathscr{E}$-quasi-continuous (cf. Lemma 6.3) and $u_n(\mu) \to u(\mu) := 1_{(\mu(F)=0)}$ pointwise on E, as $n \to \infty$. By Lemma 6.2, u is an $\mathscr{E}$-quasi-continuous member of $D(\mathscr{E})$. However, $u = 0$ m-a.e., so $u = 0$ $\mathscr{E}$-q.e., which gives the result.

$\Rightarrow$: We assume $\theta\nu_0(F) < 1$ and apply Lemma 7.5 to prove that the set $\{\mu \mid \mu(F) = 0\}$ is not $\mathscr{E}$-exceptional. We first need a few definitions. Define a probability measure on S by $\tilde{\nu}_0(A) = \nu_0(A \cap F^c)/\nu_0(F^c)$, and then define a related measure on E by $\nu := \Pi_{\theta\nu_0(F^c), \tilde{\nu}_0}$. Let $\mathscr{A} := \{(A_j)_{j=1}^n \mid (A_j)_{j=1}^n$ is a measurable partition of S for some n with $\nu_0(A_j) > 0 \ \forall j$, $F^c = A_1 \cup \cdots \cup A_k$ for some $1 \leq k \leq n\}$ and define

$$D := \Big\{u \in D(\mathscr{E}) \mid u(\mu) = \psi(\mu(A_1), \dots, \mu(A_n));$$
$$(A_j)_{j=1}^n \in \mathscr{A},\ \psi \in C^\infty(\mathbb{R}^n) \text{ for some } n \in \mathbb{N}\Big\}.$$

By Lemma 6.2, it is easy to see that D is dense in $D(\mathscr{E}_1)$. We let $\tilde{D}$ be the obvious set of $\mathscr{E}$-quasi-continuous m-versions of D by putting $\tilde{u}(\mu) = \psi(\mu(A_1), \dots, \mu(A_n))$ (cf. Lemma 6.3). By Lemma 7.2, under m if $(A_j)_{j=1}^n \in \mathscr{A}$, the random vector $(\mu(A_1), \dots, \mu(A_n))$ has a Dirichlet$(\theta\nu_0(A_1), \dots, \theta\nu_0(A_n))$ distribution, which has full support on S_n, so that if $u \geq 0$ m-a.e., then $\tilde{u} \geq 0$ everywhere on E. Under ν, the random vector $(\mu(A_1), \dots, \mu(A_k))$ has a Dirichlet$(\theta\nu_0(A_1), \dots, \theta\nu_0(A_k))$ distribution, whereas $\mu(A_j) = 0$ ν-a.e. for $k + 1 \leq j \leq n$. In particular, ν only charges the set $\{\mu \mid \mu(F) = 0\}$. Applying Lemma 7.4 (and Lemma 6.3) shows that for $u \in D$ we have

$$\frac{1}{2}\left(\frac{1 - \theta\nu_0(F)}{2^{\theta\nu_0(F^c)}}\right)\frac{\Gamma(\theta)}{\Gamma(\theta\nu_0(F))\Gamma(\theta\nu_0(F^c))}\int \tilde{u}^2(\mu)\nu(d\mu) \leq \mathscr{E}_1(u, u).$$

Because $\theta\nu_0(F) < 1$, this shows that the map $u \mapsto \int \tilde{u}\, d\nu$ is continuous in $\tilde{\mathscr{E}}_1^{1/2}$-norm and so we can apply Lemma 7.5 to find a measure ν^* on E, which does not charge $\mathscr{E}$-exceptional sets, and so $\int \tilde{u}\, d\nu = \int \tilde{u}\, d\nu^*$ for $\tilde{u} \in \tilde{D}$. Because $1 \in D$, we find that ν^* is a probability measure on E. Now let φ be a smooth function on $\mathbb{R}$ with $\varphi(0) = 1$ and $\varphi(x) < 1$ for $x \neq 0$. Then $1 = \int \varphi(\mu(F))\nu(d\mu) = \int \varphi(\mu(F))\nu^*(d\mu)$, which implies that the probability measure ν^* has all its mass on the set $\{\mu \mid \mu(F) = 0\}$. Because ν^* does not charge $\mathscr{E}$-exceptional sets, we conclude that $\{\mu \mid \mu(F) = 0\}$ is not $\mathscr{E}$-exceptional. □

COROLLARY 7.7. *For fixed $x_1, \dots, x_k$ in S, if $\nu_0(\{x_1, \dots, x_k\}) > 0$, then*

$$\left\{\mu \mid \mu = \sum_{i=1}^k p_i \varepsilon_{x_i},\ p_i \geq 0,\ \sum_{i=1}^k p_i = 1\right\}$$

is $\mathscr{E}$-exceptional if and only if

$$\theta\nu_0(S \setminus \{x_1, \dots, x_k\}) \geq 1.$$

PROOF. Apply Proposition 7.6 to set $F = S \setminus \{x_1, \dots, x_k\}$. □

The next main result. Theorem 7.10, tells us when the entire collection

$$\left\{ \mu \mid \mu = \sum_{i=1}^{k} p_i \varepsilon_{x_i},\ x_i \in S,\ p_i \geq 0,\ \sum_{i=1}^{k} p_i = 1 \right\}$$

is $\mathscr{E}$-exceptional. Here the points $x_1, \dots, x_k$ are allowed to range over all of S. Before we can prove Theorem 7.9 we need an additional lemma.

LEMMA 7.8. *Suppose that* $(X_1, \dots, X_n)$ *has a* Dirichlet$(\theta_1, \dots, \theta_n)$ *distribution. Define* $\theta = \theta_1 + \cdots + \theta_n$ *and for* $1 \leq k \leq n$, $\tilde{\theta}_k = \inf_{i_1, \dots, i_k} \theta - (\theta_{i_1} + \cdots + \theta_{i_k})$.

(i) *There exists a constant* $c(\theta)$, *depending on* θ *but not on* n, *so that when* $\varepsilon < 1/2$, *we have*

$$P\Big(\max_i X_i \geq 1 + \varepsilon\Big) \leq c(\theta) e^{\tilde{\theta}_1}. \tag{7.7}$$

(ii) *For* $1 < k \leq n$ *there exists a constant* $c(\theta, \delta, k)$, *not depending on* n, *so that if* $0 < \varepsilon \leq \delta \leq \frac{1}{2}$, *we have*

$$\begin{aligned} &P\Big(\max_{i_1, \dots, i_{k-1}} (X_{i_1} + \cdots X_{i_{k-1}}) < 1 - 2\delta, \\ &\qquad \max_{i_1, \dots, i_k} (X_{i_1} + \cdots + X_{i_k}) \geq 1 - \varepsilon\Big) \\ &\qquad \leq c(\theta, \delta, k) \varepsilon^{\tilde{\theta}_k}. \end{aligned} \tag{7.8}$$

PROOF. (i) First note that $P(\max_i X_i \geq 1 - \varepsilon) \leq \sum_{i=1}^{n} P(X_i \geq 1 - \varepsilon)$. Now X_i has a Beta$(\theta_i, \theta - \theta_i)$ distribution and so, using the fact that $ex\Gamma(x) \geq 1$ for all $x \geq 0$, we get

$$\begin{aligned} P(X_i \geq 1 - \varepsilon) &= \int_{1-\varepsilon}^{1} \frac{\Gamma(\theta)}{\Gamma(\theta_i)\Gamma(\theta - \theta_i)} x^{\theta_i - 1}(1-x)^{(\theta - \theta_i) - 1}\, dx \\ &\leq \frac{2\Gamma(\theta)}{\Gamma(\theta_i)\Gamma(\theta - \theta_i)} \frac{\varepsilon^{(\theta - \theta_i)}}{\theta - \theta_i} \\ &\leq 2\Gamma(\theta)\theta_i e^2 \varepsilon^{(\theta - \theta_i)} \\ &\leq 2\Gamma(\theta)\theta_i e^2 \varepsilon^{\tilde{\theta}_1}. \end{aligned}$$

Adding up over i gives $P(\max_i X_i \geq 1 - \varepsilon) \leq 2\Gamma(\theta)\theta e^2 \varepsilon^{\tilde{\theta}_i}$, which proves (i).

(ii) To prove (ii) we first note that

$$\begin{aligned} &P\Big(\max_{i_1, \dots, i_{k-1}} (X_{i_1} + \cdots + X_{i_{k-1}}) < 1 - 2\delta,\ \max_{i_1, \dots, i_k} (X_{i_1} + \cdots + X_{i_k}) \geq 1 - \varepsilon\Big) \\ &\quad \leq \sum_{i_1, \dots, i_k} P\big(X_{i_j} \geq \delta, 1 \leq j \leq k; X_{i_1} + \cdots + X_{i_k} \geq 1 - \varepsilon\big). \end{aligned}$$

Let $i_1,\dots,i_k$ be any set of k distinct indices between 1 and n, and set $(Y_1,\dots,Y_k)=(X_{i_1},\dots,X_{i_k})$. This random vector has a density on $\{y\in\mathbb{R}^k \mid y_i\ge 0,\ 1\le i\le k;\ \sum_{i=1}^k y_i\le 1\}$ given by

$$\tilde{c}y_1^{\theta_{i_1}-1}\cdots y_k^{\theta_{i_k}-1}(1-(y_1+\cdots-y_k))^{[\theta-(\theta_{i_1}+\cdots+\theta_{i_k})]-1},$$

where

$$\tilde{c}:=\frac{\Gamma(\theta)}{\Gamma(\theta_{i_1})\cdots\Gamma(\theta_{i_k})\Gamma(\theta-(\theta_{i_1}+\cdots+\theta_{i_k}))}.$$

Therefore, bounding each $y_j\theta^{i_j}$ by δ^{-1} and integrating, we find that

$$\begin{aligned}P\big(X_{i_j}\ge\delta,1\le j\le k;\ X_{i_1}+\cdots+X_{i_k}\ge 1-\varepsilon\big)&\\ \le\tilde{c}\delta^{-k}\frac{\varepsilon^{\theta-(\theta_{i_1}+\cdots+\theta_{i_k})}}{\big(\theta-(\theta_{i_1}+\cdots+\theta_{i_k})\big)}&\\ \le\Gamma(\theta)\delta^{-k}e^{k+1}(\theta_{i_1}\cdots\theta_{i_k})\varepsilon^{\tilde{\theta}_k}.&\end{aligned}$$

Adding over all k-tuples $(i_1,\dots,i_k)$ of distinct indices gives

$$\sum_{i_1,\dots,i_k}P\big(X_{i_j}\ge\delta,1\le j\le k;\ X_{i_1}+\cdots+X_{i_k}\ge 1-\varepsilon\big)\le\Gamma(\theta)\delta^{-k}\theta^k e^{k+1}\varepsilon^{\tilde{\theta}_k},$$

which gives the result. □

THEOREM 7.9. *For every $k\ge 1$, the set*

$$\left\{\mu\mid\mu=\sum_{i=1}^k p_i\varepsilon_{x_i};\ x_i\in S,\ p_i\ge 0,\ \sum_{i=1}^k p_i=1\right\}$$

is $\mathscr{E}$-exceptional if and only if

$$\theta\inf_{x_1,\dots,x_k}\nu_0(S\setminus\{x_1,\dots,x_k\})\ge 1.$$

PROOF. $\Rightarrow$: First we consider the case when $\theta<1$, and we show that the set $\{\varepsilon_x\mid x\in S\}$ is not $\mathscr{E}$-exceptional. The case when ν_0 itself is a point mass is trivial, so we assume $\nu_0\notin\{\varepsilon_x\mid x\in S\}$. Define a nontrivial measure ν on $\{\varepsilon_x\mid x\in S\}\subseteq E$ by letting ν be the image measure of ν_0 under the mapping from S to E, which sends x to ε_x. Similarly as in the proof of Proposition 7.6, let $\mathscr{A}=\{(A_j)_{j=1}^n\mid(A_j)_{j=1}^n$ is a measurable partition of S with $0<\nu_0(A_j)\le d$ for $1\le j\le n\}$, where $\sup_x\nu_0(\{x\})<d<1$, and define

$$D:=\Big\{u\in D(\mathscr{E})\mid u(\mu)=\psi(\mu(A_1),\dots,\mu(A_n)),\\ (A_j)_{j=1}^n\in\mathscr{A},\ \psi\in C_b^\infty(\mathbb{R}^n),\ n\in\mathbb{N}\Big\}.$$

For $u \in D$, we have

$$\int \tilde{u}(\mu)^2 \nu(d\mu) = \sum_{i=1}^{n} \nu_0(A_i)\psi^2(e_i), \tag{7.9}$$

where e_i is the point in S_n whose ith entry is equal to 1.

Under m, the random vector $(\mu(A_1), \dots, \mu(A_n))$ has the Dirichlet $(\theta\nu_0(A_1), \dots, \theta\nu_0(A_n))$ distribution on S_n; let us denote this measure by $\tilde{\nu}$. Then, letting $c := \sup_{0 \le x \le \theta} x\Gamma(x)$ and using the definition of the constant d, along with Lemma 7.4 (with $k = 1$ and ν the point mass at e_i), we get for each i,

$$\begin{aligned} &\frac{1}{2}\frac{(1-\theta)}{2^\theta}\nu_0(A_i)c^2(1-d)\Gamma(\theta)\psi^2(e_i) \\ &\quad \le \frac{1}{2}\frac{1-\theta\nu_0(S\setminus A_i)}{2^{\theta\nu_0(A_i)}}\frac{\Gamma(\theta)}{\Gamma(\theta\nu_0(A_i))\Gamma(\theta\nu_0(S\setminus A_i))}\psi^2(e_i) \\ &\quad \le \int_{S_n\cap\{w\mid w_i>1/2\}}(\langle\nabla\psi(w), a(w)\nabla\psi(w)\rangle_{\mathbb{R}^n} + \psi^2(w))\tilde{\nu}(dw). \end{aligned}$$

Because the sets $\{w \mid w_i > 1/2\}$ are disjoint, we may add these inequalities to obtain

$$\begin{aligned} &\frac{1}{2}\left(\frac{1-\theta}{2^\theta}\right)c^2(1-d)\Gamma(\theta)\sum_{i=1}^{n}\nu_0(A_i)\psi^2(e_i) \\ &\quad \le \int_{S_n}(\langle\nabla\psi(w), a(w)\nabla\psi(w)\rangle_{\mathbb{R}^n} + \psi^2(w))\tilde{\nu}(dw). \end{aligned}$$

Because $\theta < 1$, this means that there is a constant $c(\theta, d) \ge 0$ so that for $u \in D$ we have

$$\int \tilde{u}^2(\mu)\nu(d\mu) \le c(\theta, d)\mathscr{E}_1(u, u). \tag{7.10}$$

The rest of the proof of this case now follows as in Proposition 7.6.

Now suppose that $\theta \inf_{i_1,\dots,i_k} \nu_0(S\setminus\{x_{i_1},\dots,x_{i_k}\}) < 1$ and $\theta \ge 1$. Then there exist $\tilde{x}_1,\dots,\tilde{x}_k \in S$ so that $\nu_0(\{\tilde{x}_{i_1},\dots,\tilde{x}_{i_k}\}) > 0$ and $\theta\nu_0(S\setminus\{\tilde{x}_{i_1},\dots,\tilde{x}_{i_k}\}) < 1$. By Corollary 7.7 we find that the set

$$\left\{\mu \mid \mu = \sum_{i=1}^{k} p_i\varepsilon_{\tilde{x}_i};\ p_i \ge 0,\ \sum_{i=1}^{k} p_i = 1\right\}$$

is not $\mathscr{E}$-exceptional, and so the bigger set

$$\left\{\mu \mid \mu = \sum_{i=1}^{k} p_i\varepsilon_{x_1};\ x_i \in S,\ p_i \ge 0,\ \sum_{i=1}^{k} p_i = 1\right\}$$

is also not $\mathscr{E}$-exceptional.

$\Leftarrow$: Let $\mathscr{A} = (A_j^n)_{j,n\in\mathbb{N}}$ be a sequence of partitions as in Definition 6.6, and for $n, m \in \mathbb{N}$ and $1 \le k \le n$, define

$$u_{k,m}^n(\mu) := \sup_{i_1,\dots,i_k\le m} \mu(A_{i_1}^n) + \cdots + \mu(A_{i_k}^n).$$

As $m \to \infty$, $u^n_{k,m}$ converges pointwise boundedly to

$$u^n_k(\mu) := \sup_{i_1,\dots,i_k} \mu(A^n_{i_1}) \cdots + \cdots \mu(A^n_{i_k}),$$

and as $n \to \infty$, u^n_k converges pointwise boundedly to

$$u_k(\mu) := \sup_{i_1,\dots,i_k} \mu(\{x_{i_1},\dots,x_{i_k}\}).$$

This can be seen in Lemma 6.7. As in the proof of Lemma 6.2, one can show that u_k is an $\mathscr{E}$-quasi-continuous member of $D(\mathscr{E})$, with $\Gamma(u_k) \le 1$ m-a.e.

Next we need to know how likely the function u_k is to be close to 1, under m. We begin with the case $k = 1$ and get a bound on $m(u_1 \ge 1 - \varepsilon)$. Pick N so large that $\inf_i \nu_0(S \setminus A^N_i) > 0$. For $n \ge N$, choose m so large that

$$\nu_0\left(\bigcup_{j=1}^m A^n_j\right) \ge \inf_i \nu_0(S \setminus A^n_i).$$

Defining $(X_1,\dots,X_{m+1}) = (\mu(A^n_1),\dots,\mu(A^n_m), \mu(S \setminus \bigcup_{j=1}^m A^n_j))$ and applying Lemma 7.8(i) to the nontrivial subvector of $(X_1,\dots,X_{m+1})$, that is, taking only those i with $m(X_i = 0) = 0$, we get, for $\varepsilon < \frac{1}{2}$,

$$m(u^n_{1,m} \ge 1 - \varepsilon) \le c(\theta)\varepsilon^{\theta \inf_i : \nu_0(S \setminus A^n_i)}. \tag{7.11}$$

As $m \to \infty$, $m(u^n_{1,m} \ge 1 - \varepsilon) \to m(u^n_1 \ge 1 - \varepsilon)$ for all but countably many $\varepsilon < \frac{1}{2}$, and so

$$m(u_1 \ge 1 - \varepsilon) \le m(u^n_1 \ge 1 - \varepsilon) \le c(\theta)\varepsilon^{\theta \inf_i \nu_0(S \setminus A^n_i)}. \tag{7.12}$$

As $n \to \infty$, $\theta \inf_i \nu_0(S \setminus A^n_i) \to \theta \inf_i \nu_0(S \setminus \{x_i\})$, so we conclude that

$$m(u_1 \ge 1 - \varepsilon) \le c(\theta)\varepsilon^{\theta \inf_i \nu_0(S \setminus \{x_i\})}. \tag{7.13}$$

Similarly, using Lemma 7.8(ii) we can prove for $k > 1$.

$$\begin{aligned} &m(u_{k-1} < 1 - 2\delta, u_k \ge 1 - \varepsilon) \\ &\quad \le c(\theta, \delta, k)\varepsilon^{\theta \inf_{i_1,\dots,i_k} \nu_0(S \setminus \{x_{i_1},\dots,x_{i_k}\})}. \end{aligned} \tag{7.14}$$

Because our square field operator Γ satisfies the chain rule, it is not hard to show that for any $w_1,\dots,w_n \in D(\mathscr{E})$ we have

$$\Gamma(w_1 \vee \cdots \vee w_n) = \sum_{i=1}^n \Gamma(w_i) 1_{(w_i > \sup_{j \ne i} w_j)},$$

where $\Gamma(w_1 \vee \cdots \vee w_n) = 0$ m-a.e. wherever there is a "tie"' that is, on the set of $\mu \in E$ so that $w_i(\mu) = w_j(\mu)$ for some $i \ne j$. Also, we note that

$$\Gamma\big(\mu(A^n_{i_1}) + \cdots + \mu(A^n_{i_k})\big) = \mu\left(\bigcup_{j=1}^k A^n_{i_j}\right) - \mu^2\left(\bigcup_{j=1}^k A^n_{i_j}\right) \le \varepsilon$$

on the set $\{\mu \mid \mu(\bigcup_{j=1}^k A^n_{i_j}) \ge 1 - \varepsilon\}$. Now let φ be an absolutely continuous function on [0, 1] that vanishes for $x \le 1 - \varepsilon$, and $d = \sup_{0 \le x \le 1} |\varphi'(x)| < \infty$. Then $\varphi(u^n_{k,m})$ is an $\mathscr{E}$-quasi-continuous member of $D(\mathscr{E})$ with $\Gamma(\varphi(u^n_{k,m})) \le$

$d^2\varepsilon$. Applying Lemma 6.2, first to the sequence $(\varphi(u_{k,m}^n))_{m\in\mathbb{N}}$ and then to $(\varphi(u_k^n))_{n\in\mathbb{N}}$, we find that $\varphi(u_k)$ is an $\mathscr{E}$-quasi-continuous member of $D(\mathscr{E})$ and

$$\Gamma(\varphi(u_k)) \le d^2\varepsilon 1_{\{u_k \ge 1-\varepsilon\}}, \tag{7.15}$$

because by calculus for square field operator Dirichlet forms, we know that $\Gamma(\varphi(u_k)) = 0$ m-a.e. on the set $\{u_k \le 1 - \varepsilon\}$. We define the following absolutely continuous functions on $[0, 1]$; for $0 < \delta < 1/2$, let $\psi_\delta(x) := (1 - x/(1-\delta))^+$ and $\varphi_n(x) := (nx + (1-n))^+$.

We are now ready to prove the result. We will use induction on k, so let us begin with the case $k = 1$ and assume that $\theta \inf_i \nu_0(S \setminus \{x_i\}) \ge 1$. Now consider the sequence $v_n := \varphi_n(u_1)$ of $\mathscr{E}$-quasi-continuous members of $D(\mathscr{E})$. As $n \to \infty$, $v_n(\mu)$ converges pointwise boundedly to $v(\mu) := 1_{\{u_1(\mu)=1\}}$. From (7.15) we get

$$\Gamma(v_n) \le n^2(1/n)1_{\{u_1 \ge 1-1/n\}},$$

and by integrating over E with respect to m and using (7.13) we obtain

$$\mathscr{E}(v_n) \le nm\left(u_1 \ge 1 - \frac{1}{n}\right) \le c(\theta)n^{1-\theta \inf_i \nu_0(S\setminus\{x_i\})}.$$

Hence $(v_n)_{n\in\mathbb{N}}$ is bounded in $D(\mathscr{E})$ and so by Lemma 6.2 we conclude that v is $\mathscr{E}$-quasi-continuous. Because $v = 0$ m-a.e., it follows that $v = 0$ $\mathscr{E}$-q.e. and so $\{\mu \mid u_1(\mu) = 1\} = \{\mu \mid \mu = \varepsilon_x\colon x \in S\}$ is $\mathscr{E}$-exceptional.

Now let $k > 1$ and assume the result is true for $k-1$. Suppose $\theta \inf_{i_1,\dots,i_k} \nu_0(S \setminus \{x_{i_1}, \dots, x_{i_k}\}) \ge 1$. Then also $\theta \inf_{i_1,\dots,i_{k-1}} \nu_0(S \setminus \{x_{i_1}, \dots, x_{i_{k-1}}\}) \ge 1$, so that $\{\mu \mid u_{k-1}(\mu) = 1\}$ is $\mathscr{E}$-exceptional.

To prove that $\{\mu \mid u_k(\mu) = 1\}$ is $\mathscr{E}$-exceptional, it suffices for every fixed $\delta > 0$ to prove that $\{\mu \mid u_{k-1}(\mu) \le 1 - 2\delta,\ u_k(\mu) = 1\}$ is $\mathscr{E}$-exceptional because

$$\{\mu \mid u_k(\mu) = 1\} \subseteq \{\mu \mid u_{k-1}(\mu) = 1\} \cup \bigcup_{n=1}^{\infty}\left\{\mu \mid u_{k-1}(\mu) \le 1 - \frac{1}{n},\ u_k(\mu) = 1\right\}.$$

Now let $v_n := \psi_{2\delta}(u_{k-1})\varphi_n(u_k)$. As $n \to \infty$, $v_n(\mu)$ converges pointwise boundedly to $v(\mu) := 1_{\{u_k(\mu)=1\}}\psi_{2\delta}(u_{k-1})$. As in the case $k = 1$, calculating $\Gamma(v_n)$ shows that for some $\tilde{c} > 0$,

$$\Gamma(v_n) \le \tilde{c}n1_{\{u_{k-1} < 1-2\delta, u_k \ge 1-1/n\}}.$$

Integrating on E with respect to m and using (7.14), we obtain

$$\begin{aligned}\mathscr{E}(v_n) &\le \tilde{c}nm(u_{k-1} < 1 - 2\delta,\ u_k \ge 1 - 1/n)\\ &\le \tilde{c}c(\theta, \delta, k)n^{1-\theta \inf_{i_1,\dots,i_k} \nu_0(S\setminus\{x_{i_1},\dots,x_{i_k}\})}.\end{aligned}$$

Because we have assumed that $\theta \inf_{i_1,\dots,i_k} \nu_0(S \setminus \{x_{i_1}, \dots, x_{i_k}\}) \ge 1$, we see that $(v_n)_{n\in\mathbb{N}}$ is bounded in $D(\mathscr{E})$ and so we conclude that v is $\mathscr{E}$-quasi-continuous. Because $v = 0$ m-a.e., it follows that $v = 0$ $\mathscr{E}$-q.e. and so

$$\{\mu \mid u_{k-1}(\mu) < 1 - 2\delta, u_k(\mu) = 1\}$$

is $\mathscr{E}$-exceptional. This gives the result. □

COROLLARY 7.10. *If ν_0 is atomless, then $\mathscr{E}$-q.e. $\mu \in E$ has infinitely many atoms if and only if $\theta \geq 1$.*

Acknowledgment. This work was completed during a visit by B.S. to the University of Bonn. He gratefully acknowledges the hospitality of his hosts and the financial support of the Sonderforschungsbereich 256 for making this visit, and hence this work, possible.

REFERENCES

[1] ALBEVERIO, S., MA, Z. M. and RÖCKNER, M. (1993). Quasi-regular Dirichlet forms and Markov processes. *J. Funct. Anal.* **111** 118–154.

[2] ALBEVERIO, S., MA, Z. M. and RÖCKNER, M. (1994). Local property of Dirichlet forms and diffusions on general state space. *Math. Ann.* **296** 977–986.

[3] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

[4] BOULEAU, N. and HIRSCH, F. (1991). *Dirichlet Forms and Analysis on Wiener Space*. de Gruyter, Berlin.

[5] CARRILLO-MENENDEZ, S. (1975). Processus de Markov associé à une forme de Dirichlet non symétrique. *Z. Wahrsch. Verw. Gebiete* **33** 139–154.

[6] DAWSON, D. A. (1991). Measure-valued Markov processes. *Ècole d'Èté de Probabilités de Saint-Flour 1991. Lecture Notes in Math.* **1541.** Springer, Berlin.

[7] DAWSON, D. A. and MARCH, P. (1993). Resolvent estimates for Fleming–Viot operators and uniqueness of solutions to related martingale problems. Preprint.

[8] DEUSCHEL, J. D. and STROOCK, D. W. (1989). *Large Deviations*. Academic Press, Boston.

[9] DUDLEY, R. M. (1989). *Real Analysis and Probability*. Wadsworth & Brooks/Cole, Pacific Grove, CA.

[10] EBERLE, A. (1993). Diploma thesis, Univ. Bonn.

[11] ETHIER, S. N. and GRIFFITHS, R. C. (1992). The transition function of a Fleming–Viot process. Preprint.

[12] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes, Characterization and Convergence*. Wiley, New York.

[13] ETHIER, S. N. and KURTZ, T. G. (1993). Fleming–Viot processes in population genetics. *SIAM J. Control and Optim.* **31** 345–386.

[14] FUKUSHIMA, M. (1971). Dirichlet spaces and strong Markov processes. *Trans. Amer. Math. Soc.* **162** 185–224.

[15] FUKUSHIMA, M. (1980). *Dirichlet Forms and Markov Processes*. North-Holland, Amsterdam.

[16] FUKUSHIMA, M. (1982). A note on irreducibility and ergodicity of symmetric Markov processes. *Stochastic Processes in Quantum Theory and Statistical Physics: Proceedings of the International Workshop Held in Marseille 1981. Lecture Notes in Phys.* **173** 200–207. Springer, Berlin.

[17] FUKUSHIMA, M. and TAKEDA, M. (1984). A transformation of a symmetric Markov process and the Donsker–Varadhan theory. *Osaka J. Math.* **21** 311–326.

[18] LE JAN, Y. (1977). Balayage et forms de Dirichlet. *Z. Warsch. Verw. Gebiete* **37** 297–319.

[19] MA, Z. M., OVERBECK, L. and RÖCKNER, M. (1993). Markov processes associated with semi-Dirichlet forms. *Osaka J. Math.* To appear.

[20] MA, Z. M. and RÖCKNER, M. (1992). *Introduction to the Theory of (Non-Symmetric) Dirichlet Forms*. Springer, Berlin.

[21] MÜCK, S. (1993). Large deviations w.r.t. quasi-every starting point for symmetric right processes on general state spaces. *Probab. Theory Related Fields* **99** 87 527–548.

[22] RÖCKNER, M. and SCHMULAND, B. (1993). Quasi-regular Dirichlet forms: Examples and counterexamples. *Canad. J. Math.* To appear.

[23] RÖCKNER, M. and ZHANG, T. S. (1992). Uniqueness of generalized Schrödinger operators and applications. *J. Funct. Anal.* **105** 187–231.

[24] RÖCKNER, M. and ZHANG, T. S. (1992). Uniqueness of generalized Schrödinger operators—Part II. *J. Funct. Anal.* 119 456–467.

[25] SCHMULAND, B. (1991). A result on the infinitely many neutral alleles diffusion model. *J. Appl. Probab.* **28** 253–267.

[26] SCHMULAND, B. (1992). Dirichlet forms with polynomial domain. *Math. Japon* **37** 1015–1024.

[27] SILVERSTEIN, M. L. (1974). *Symmetric Markov Processes. Lecture Notes in Math.* **426.** Springer, Berlin.

[28] SHIGA, T. (1990). A stochastic equation based on a Poisson system for a class of measure-valued diffusions. *J. Math. Kyoto Univ.* **30** 245–279.

[29] TAKEDA, M. (1990). On Donsker–Varadhan's entropy and its applications. *Forum Math.* **2** 481–488.

L. OVERBECK
M. RÖCKNER
INSTITUT FÜR ANGEWANDTE MATHEMATIK
UNIVERSITÄT BONN
WEGELERSTRASSE 6
5300 BONN 1
GERMANY

B. SCHMULAND
DEPARTMENT OF STATISTICS
AND APPLIED PROBABILITY
UNIVERSITY OF ALBERTA
EDMONTON, ALBERTA
CANADA