Species dynamics in the two-parameter Poisson-Dirichlet diffusion model

Matteo Ruggiero

© 2012 by Matteo Ruggiero. Any opinions expressed here are those of the authors and not those of the Collegio Carlo Alberto.
Species dynamics in the two-parameter Poisson-Dirichlet diffusion model

MATTEO RUGGIERO
University of Torino and Collegio Carlo Alberto
April 10, 2012

The recently introduced two-parameter infinite alleles model extends the celebrated one-parameter version, related to Kingman’s distribution, to diffusive two-parameter Poisson-Dirichlet frequencies. Here we investigate the dynamics driving the species heterogeneity underlying the two-parameter model. First we show that a suitable normalization of the number of species is driven by a critical continuous state branching process with immigration, or a squared Bessel process with dimension given by the immigration rate. Secondly, we identify an instance of the intensity rates driving the finite-dimensional mutation process that gives rise to the two-parameter model, which turn out to be inhomogeneous and unbounded. These, together with additional restrictions, allow to provide a finite-dimensional construction of the two-parameter model, obtained by means of a sequence of Feller diffusions of Wright-Fisher flavor which feature finitely-many types. Both results provide insight into the mathematical properties and biological interpretation of the two-parameter model, showing that it is structurally different from the one-parameter case in that the frequencies dynamics are driven by state-dependent rather than constant quantities.

Keywords: alpha diversity, infinite dimensional diffusion, infinite alleles model, mutation rate, Poisson-Dirichlet distribution.

MSC: 60J60, 60G57, 92D25.

1 Introduction and motivation

The two-parameter infinitely-many-neutral-alleles model (henceforth referred to as the two-parameter model) is a family of diffusion processes, introduced by Petrov (2009) and further investigated by Feng and Sun (2010), that extends the celebrated infinitely-many-neutral-alleles model (henceforth the one-parameter model). More specifically, let

\[ \nabla_\infty = \left\{ z \in [0,1]^\infty : z_1 \geq z_2 \geq \ldots \geq 0, \sum_{i=1}^\infty z_i \leq 1 \right\} \]

be the closure of the infinite-dimensional ordered simplex, and define, for constants \( 0 \leq \alpha < 1 \) and \( \theta > -\alpha \), the second-order differential operator

\[ B = \frac{1}{2} \sum_{i,j=1}^\infty z_i (\delta_{ij} - z_j) \frac{\partial^2}{\partial z_i \partial z_j} - \frac{1}{2} \sum_{i=1}^\infty (\theta z_i + \alpha) \frac{\partial}{\partial z_i}. \]
where \( \delta_{ij} \) denotes Kronecker delta, acting on a certain dense subalgebra of the space \( C(\nabla_\infty) \) of continuous functions on \( \nabla_\infty \). Then the closure of \( \mathcal{B} \) generates a strongly continuous semigroup of contractions on \( C(\nabla_\infty) \), and the sample paths of the associated process are almost surely continuous functions from \([0, \infty)\) to \( \nabla_\infty \). Such process can be thought of as describing the temporal evolution of the decreasingly ordered allelic frequencies \((z_1, z_2, \ldots)\) at a particular locus in an ideally infinite population with infinitely-many possible types or species. See Feng (2010) for a review of infinitely-many alleles models and Feng et al. (2011) for the transition function for the two-parameter case. See also Borodin and Olshanski (2009) for a general construction related to Petrov (2009), and Ruggiero, Walker and Favaro (2012) for a partially related model with diffusive parameter \( \theta \).

As shown by Petrov (2009) and Feng and Sun (2010), the two-parameter model is reversible and ergodic with respect to the Poisson-Dirichlet distribution with parameters \((\theta, \alpha)\), henceforth denoted \( \text{PD}(\theta, \alpha) \). Introduced by Pitman (1995) (see also Pitman and Yor (1997)), this extends the one-parameter version \( \text{PD}(\theta) := \text{PD}(\theta, 0) \) due to Kingman (1975), and has found numerous applications in several fields. See for example the monographs by Bertoin (2006) for fragmentation and coalescent theory, Pitman (2006) for excursion theory and combinatorics, Lijoi and Prünster (2009) for Bayesian inference, Teh and Jordan (2009) for machine learning and Feng (2010) for population genetics. Both these random discrete distributions arise as the decreasingly ordered weights of a Dirichlet process (Ferguson, 1973), when \( \alpha = 0 \), and of a two-parameter Poisson-Dirichlet process (Pitman, 1995) respectively, and can alternatively be constructed by means of the following so-called stick breaking procedure, also known as residual allocation model. Consider a sequence of random variables \((V_1, V_2, \ldots)\) obtained by setting

\[
V_1 = W_1, \quad V_n = W_n \prod_{i=1}^{n-1} (1 - W_i), \quad W_i \overset{ind}{\sim} \text{Beta}(1 - \alpha, \theta + i\alpha),
\]

with \( 0 \leq \alpha < 1 \) and \( \theta > -\alpha \). The vector \((V_1, V_2, \ldots)\) is said to have the GEM(\( \theta, \alpha \)) distribution, named after Griffiths, Engen and McCloskey, while the vector of descending order statistics \((V_1^{(1)}, V_2^{(2)}, \ldots)\) is said to have the PD(\( \theta, \alpha \)) distribution. See Feng and Wang (2007) for an infinite-dimensional diffusion process related to GEM distributions.

Besides sharing the above stick-breaking construction strategy, it is well known that the difference between these two random discrete distributions is structural and does not simply rely on a different parametrization. For example, the distribution \( \text{PD}(\theta) \) is obtained by ranking and normalizing the jumps of a Gamma process, whereas the \( \text{PD}(\theta, \alpha) \) is obtained by normalizing the jumps of a tilted stable subordinator of parameter \( \alpha \) (Pitman, 2003). Furthermore, the \( \text{PD}(\theta) \) distribution is obtained as weak limit of a Dirichlet distributed vector of frequencies (Kingman, 1975), while a similar construction for the two-parameter case is not
available. For what concerns their diffusive counterparts, the properties of the one-
parameter model, related to the PD(θ) distribution, are well understood, whereas
numerous are still the questions regarding the two-parameter model. In particu-
lar, given the above considerations, it is not surprising that a finite-dimensional
construction of the process with operator (2), in terms of a sequence of finite di-
mensional diffusion processes, is currently available only when α = 0. To be more
precise, consider the usual approximating diffusion for the Wright-Fisher discrete
genetic model with \( n \) selectively neutral alleles and symmetric mutation, which
corresponds to the operator

\[
B_n = \frac{1}{2} \sum_{i,j=1}^{n} z_i (\delta_{ij} - z_j) \frac{\partial^2}{\partial z_i \partial z_j} + \frac{1}{2} \sum_{i=1}^{n} b_i^{(n)}(z) \frac{\partial}{\partial z_i},
\]

acting on a suitable subspace of \( C^2(\nabla_n) \), with

\[
\nabla_n = \left\{ z \in \nabla_\infty : z_{n+1} = 0, \sum_{i=1}^{n} z_i = 1 \right\},
\]

and where the drift term \( b_i^{(n)}(z) \) is given by

\[
b_i^{(n)}(z) = \frac{\theta}{n - 1} (1 - z_i) - \theta z_i, \quad \theta > 0.
\]

Ethier and Kurtz (1981) formalized the conditions under which the sequence of
processes with operators defined by (3)-(4) converges in distribution to the one-
parameter model, with operator obtained by setting \( \alpha = 0 \) in (2). As anticipated,
a similar construction for the case \( 0 < \alpha < 1, \theta > -\alpha \) is currently unavailable.
It is to be said that two sequential constructions of the two-parameter model are
given in Petrov (2009) and Ruggiero and Walker (2009). In Section 3 we will ar-
gue that, despite offering interesting reads of the two parameter model, neither of
these offers particular insight into the interpretation of the species dynamics un-
derlying the infinite-dimensionality structure, in particular since both are based on
finitely-many items. The problem at hand could then be rephrased as that of un-
derstanding from which Wright-Fisher-type mechanism, if any, the two-parameter
model comes from. While the importance of providing a particle construction lies
in the fact that the individual dynamics are dealt with explicitly, the importance
of providing a sequential construction by means of Wright-Fisher-type diffusions
lies in the genetic interpretation one yields from the specification of the mutation
rates at the \( n \)-th step of the sequence. Such interpretation is clear in the case of
(4), whereby each type has the same chance of mutating (cf. also (18) below), but
is somewhat obscure for what regards the role of \( \alpha \) in (2), especially in terms of
its effect on finitely-many types. This role would be, at least partially, clarified by
identifying suitable mutation rates, which are basic building blocks of the model.
and give important information on the reproductive mechanism of the underlying population. Historically, the logic process has been the opposite, i.e. diffusion approximations were introduced for dealing mathematically in a simpler way with multi-type discrete models such as Wright-Fisher processes. But the recent advances, often stimulated by neighboring research fields, sometimes provide first what used to be the ultimate goal, as is the case with the two-parameter diffusion model, thus leaving a gap to be filled, both mathematically and in terms of interpretation.

Motivated by these considerations, the purpose of this paper is to investigate what lies underneath the infinite-dimensionality of the two-parameter model in terms of the forces driving the species dynamics. We pursue this task in two different ways. First we derive an $\alpha$-diversity diffusion for the two-parameter model. This is a continuous-time continuous-state extension of the corresponding notion for Poisson-Kingman models (Pitman, 2003), and describes the dynamics of the suitably normalized number of species in the underlying population. We show that such diffusion for the two-parameter model is a critical continuous-state branching process with immigration, or a squared Bessel process with dimension given by the immigration rate $\theta$. This is done in Section 2, where a corresponding quantity for the one-parameter case is also discussed. The second task is concerned with finding explicit transition rates for the mutation process which gives rise to the two-parameter model, and with providing a sequential construction for the limiting process in terms of finite-dimensional diffusions similar to those given by (3)-(4) for the one-parameter case. In Section 3 we collect some considerations on the problem of finding such mutation rates and on the fact that the existing sequential constructions do not provide insight from a genetical point of view. In Section 4 we identify an instance of such mutation rates, which turn out to be unbounded on the usual domain of the process and to depend on the current species abundances. By means of some additional restrictions, we formalize a sequential construction where each term of the sequence is a Feller diffusion on a finite-dimensional subspace of $\nabla_\infty$. Both approaches undertaken show that the one- and two-parameter model are structurally different, in that the species dynamics of the former are driven by constant terms, whereas those of the latter are driven by density-dependent quantities.

2 Heterogeneity in the two-parameter model

The notion of $\alpha$-diversity was introduced by Pitman (2003) for exchangeable partitions induced by random discrete distributions of Poisson-Kingman type. Let $\text{PK}(\varrho|t)$ denote the distribution of the weights $(P_i) = (J_i/T)$ determined by the ranked points $J_i$ of a Poisson process with Lévy density $\varrho$, given $T = t$. A Poisson-Kingman distribution with Lévy density $\varrho$ and mixing distribution $\gamma$ on $(0, \infty)$,
Species dynamics in the two-parameter diffusion model

For instance, the PD($\theta, \alpha$) distribution is obtained as a $PK(\varrho, \gamma)$ model, for $0 < \alpha < 1$ and $\theta > -\alpha$, where $\varrho_\alpha$ is the Lévy density of a stable subordinator of index $\alpha$, $\gamma_{\theta,\alpha}$ is

$$\gamma_{\theta,\alpha}(dt) = \frac{\Gamma(\theta + 1)}{\Gamma(\theta/\alpha + 1)} t^{-\theta} f_\alpha(t) dt,$$

and $f_\alpha(t)$ is the density of a positive stable random variable of index $\alpha$. Given an exchangeable random partition of $N$ induced by a Poisson-Kingman distribution, i.e. such that its ranked class frequencies have distribution $PK(\varrho, \gamma)$, this is said to have $\alpha$-diversity $S$ if and only if there exists a random variable $S$, with $0 < S < \infty$ almost surely, such that

$$\lim_{n \to \infty} \frac{K_n}{n^\alpha} = S \quad \text{a.s.}$$

where $K_n$ is the number of classes of the partition restricted to $\{1, \ldots, n\}$. For instance, in the case of a PD($\theta, \alpha$) partition, we have $S = T^{-\alpha}$ where $T$ has distribution $\gamma_{\theta,\alpha}$. See Pitman (2003), Proposition 13.

The idea of extending the concept of $\alpha$-diversity from random distributions on simplices to a continuous-time continuous-state framework has been formulated in Ruggiero, Walker and Favaro (2012), where a certain rescaled, inhomogeneous random walk on the integers, which tracks the dynamics of the number of species in a normalized inverse-Gaussian population, is shown to converge to a certain one dimensional diffusion process on $(0, \infty)$. Here we derive an $\alpha$-diversity diffusion for the two-parameter model, with the aim of providing insight into the species dynamics underlying the infinite-dimensional process. The derivation is based on the particle construction given in Ruggiero and Walker (2009), here briefly recalled for ease of the reader. Let $X^{(n)} = (X_1, \ldots, X_n)$ be a sample from a two-parameter Poisson-Dirichlet process, or equivalently (cf. Pitman, 1995) from the generalized Pólya urn scheme (also known as Pitman urn scheme) given by $X_1 \sim P_0$ and

$$P\{X_{i+1} \in \cdot | X_1, \ldots, X_i\} = \frac{\theta + \alpha K_i}{\theta + i} P_0(\cdot) + \frac{1}{\theta + i} \sum_{j=1}^{K_i} (n_j - \alpha) \delta_{X_j^*}(\cdot)$$

for $i = 2, \ldots, n - 1$. Here $P_0$ is a non atomic probability measure on the space of the observables (e.g. a Polish space), $K_i \leq i$ denotes the number of distinct elements $(X_1^*, \ldots, X_{K_i}^*)$ observed in $(X_1, \ldots, X_i)$ and $\delta_{X_j^*}$ is a point mass at $X_j^*$. A simple way to make the sample into a Markov chain is the following. Let
$X^{(n)}$ be updated at discrete times by replacing a uniformly chosen coordinate. Conditionally on $K_n = k$ at the current state, and exploiting the exchangeability of the sample from (6), the incoming particle will be of a new type with probability $(\theta + \alpha k r)/(\theta + n - 1)$, and will be a copy of one still in the vector after the removal with probability $(n - 1 - \alpha k r)/(\theta + n - 1)$, where $k_r$ is the value of $k$ after the removal.

The following proposition recalls the relation between the above described particle chain $\{X^{(n)}(m), m \in \mathbb{N}_0\}$ and the two-parameter model. For notational simplicity we omit here the details about the domain of the limiting operator.

**Proposition 2.1.** [Ruggiero and Walker, 2009] Let $Z(\cdot)$ be the two-parameter model, corresponding to the operator $B$ as in (2). Let also $\{X^{(n)}(m), m \in \mathbb{N}_0\}$ be the particle chain described above, and define $\{Y^{(n)}(t), t \geq 0\}$ by $Y^{(n)}(t) = \eta(X^{(n)}([n^2 t]))$, where $\eta(x^{(n)}) = (z_1, \ldots, z_n, 0, 0, \ldots)$ if $z_i$ is the relative size of the $i$-th largest cluster in $X^{(n)} = x^{(n)}$. Then

$$Y^{(n)}(\cdot) \Rightarrow Z(\cdot) \quad \text{in } C_{\nabla_\infty}([0, \infty))$$

as $n \to \infty$.

**Remark 2.2.** (a) Here “$\Rightarrow$” denotes weak convergence of the sequence of probability measures on the space $D_{\nabla_\infty}([0, \infty))$ of càdlàg functions from $[0, \infty)$ to $\nabla_\infty$. The convergence holds in the subspace $C_{\nabla_\infty}([0, \infty)) \subset D_{\nabla_\infty}([0, \infty))$ of continuous functions, since the limit probability measure is concentrated on $C_{\nabla_\infty}([0, \infty))$ and the Skorohod topology relativized to $C_{\nabla_\infty}([0, \infty))$ coincides with the topology on $C_{\nabla_\infty}([0, \infty))$.

(b) In Proposition 2.1 we use the expression “process corresponding to the operator” rather informally, only hinting at the notion of Feller semigroup generated by the closure of the operator and that of Markov process corresponding to such semigroup. In Section 4 (see Definition 4.1 below) we will treat these notions more formally.

(c) Proposition 2.1 is slightly different from the result stated in the quoted source, in particular due to the fact that this is a discrete parameter convergence. The proof of the above Proposition can be obtained by minor modifications of the original one, or alternatively the statements coincide if one superimposes to the Markov chain $\{X^{(n)}(m), m \in \mathbb{N}_0\}$ a Poisson clock of rate $n^2$ which governs the jump times of the chain. □

Denote now by $\{K_n(m), m \in \mathbb{N}_0\}$ the chain which keeps track of the number of distinct types in $X^{(n)}(m)$, and let $M_{1,n}$ denote the number of types in $X^{(n)}(m)$ with only one representative. The transition probabilities for $K_n(m)$, denoted for short

$$p(k, k') = \mathbb{P}\{K_n(m + 1) = k'|K_n(m) = k\}$$
Species dynamics in the two-parameter diffusion model

are given by

\[
p(k, k') = \begin{cases} 
  \left(1 - \frac{M_{1,n}}{n}\right) \frac{\theta + \alpha k}{\theta + n} & \text{if } k' = k + 1 \\
  \frac{M_{1,n}}{n(\theta + n)}(n - 1 - \alpha(k - 1)) & \text{if } k' = k - 1 \\
  1 - p(k, k + 1) - p(k, k - 1) & \text{if } k' = k \\
  0 & \text{else}
\end{cases}
\]

(7)

for \(1 \leq k \leq n\). Here \(M_{1,n}/n\) is the probability of removing a cluster of size one, and \(k = 1\) and \(k = n\) imply \(p(1, 0) = 0\) and \(p(n, n + 1) = 0\) respectively. Since (7) need not be markovian, we use an approximation of \(M_{1,n}\) based on the following asymptotic result. From (5) and Lemma 3.11 in Pitman (2006), we have that the number \(M_{1,n}\) of clusters of size one observed in the sample is such that

\[
M_{1,n}/n^{\alpha} \rightarrow \alpha S \quad \text{a.s.,}
\]

so that \(m_{1,n} \approx \alpha k\). This yields

\[
p(k, k') = \begin{cases} 
  \left(1 - \frac{\alpha k}{n}\right) \frac{\theta + \alpha k}{\theta + n} + o(n^{-1+\alpha}) & \text{if } k' = k + 1 \\
  \frac{\alpha k}{n(\theta + n)}(n - 1 - \alpha(k - 1)) + o(n^{-1+\alpha}) & \text{if } k' = k - 1 \\
  1 - p(k, k + 1) - p(k, k - 1) & \text{if } k' = k \\
  0 & \text{else}
\end{cases}
\]

(8)

The following theorem identifies the \(\alpha\)-diversity diffusion for the two-parameter model.

**Theorem 2.3.** Let \(\{K_n(m), m \in \mathbb{N}_0\}\) be a Markov chain on \(\mathbb{N}\) with transition probabilities as in (8), for \(0 < \alpha < 1\) and \(\theta > -\alpha\), and define \(\tilde{K}_n(t), t \geq 0\) by letting \(\tilde{K}_n(t) = K_n(\lfloor n^{1+\alpha} t \rfloor)/n^{\alpha}\). Let also \(\{S_{\theta,\alpha}(t), t \geq 0\}\) be a diffusion process on \([0, \infty)\) driven by the stochastic differential equation

\[
dS_{\theta,\alpha}(t) = \theta dt + \sqrt{2\alpha S_{\theta,\alpha}(t)} dB(t),
\]

where \(B(t)\) is a standard Brownian motion. If \(\tilde{K}_n(0) \Rightarrow S_{\theta,\alpha}(0)\), then

\[
\{\tilde{K}_n(t), t \geq 0\} \Rightarrow \{S_{\theta,\alpha}(t), t \geq 0\} \quad \text{in } C_{[0,\infty)}([0,\infty))
\]

(10)

and \(n \rightarrow \infty\).
Proof. Denote by $U_n$ the semigroup induced by (8). For notational brevity, here we do not distinguish between $(n, k)$ and $(n - 1, k - 1)$, since they are asymptotically equivalent. Then, for $f \in C_0([0, \infty))$, i.e. continuous functions vanishing at infinity, we can write

$$(U_n - I)f\left(\frac{k}{n^\alpha}\right) = \left[f\left(\frac{k+1}{n^\alpha}\right) - f\left(\frac{k}{n^\alpha}\right)\right] \left(1 - \frac{\alpha k}{n}\right) \frac{\theta + \alpha k}{\theta + n}$$

$$+ \left[f\left(\frac{k-1}{n^\alpha}\right) - f\left(\frac{k}{n^\alpha}\right)\right] \frac{\alpha k(n - \alpha k)}{n(\theta + n)} + O(n^{-1+\alpha}).$$

By means of a Taylor expansion we get

$$(U_n - I)f\left(\frac{k}{n^\alpha}\right) = \frac{1}{n^\alpha} f'\left(\frac{k}{n^\alpha}\right) \left[1 - \frac{\alpha k}{n}\right] \frac{\theta + \alpha k}{\theta + n} - \frac{\alpha k}{n} \left(n - \alpha k\right)$$

$$+ \frac{1}{2n^{2\alpha}} f''\left(\frac{k}{n^\alpha}\right) \left[1 - \frac{\alpha k}{n}\right] \frac{\theta + \alpha k}{\theta + n} + \frac{\alpha k}{n} \left(n - \alpha k\right) + o(n^{-1+\alpha}).$$

(11)

Since

$$\left(1 - \frac{\alpha k}{n}\right) \frac{\theta + \alpha k}{\theta + n} - \frac{\alpha k}{n} \left(n - \alpha k\right) = \frac{\theta}{\theta + n} + o(n^{-1})$$

and

$$\left(1 - \frac{\alpha k}{n}\right) \frac{\theta + \alpha k}{\theta + n} + \frac{\alpha k}{n} \left(n - \alpha k\right) = \frac{2\alpha k}{\theta + n} + o(n^{-1+\alpha})$$

we obtain

$$(U_n - I)f\left(\frac{k}{n^\alpha}\right) = \frac{1}{n^\alpha} f'\left(\frac{k}{n^\alpha}\right) \frac{\theta}{\theta + n} + \frac{1}{2n^{2\alpha}} f''\left(\frac{k}{n^\alpha}\right) \frac{2\alpha k}{\theta + n} + o(n^{-1+\alpha}).$$

From which, using (5), it follows that

$$(12) \sup_{s \in [0, \infty)} |Cf(s) - n^{1+\alpha}(U_n - I)f(s)| \to 0, \quad \text{as } n \to \infty,$$

where

$$Cf(s) = \theta f'(s) + 2\alpha s f''(s)$$

is the infinitesimal operator corresponding to (9). Here (12) holds for every $f$ belonging to an appropriate restriction $\mathcal{D}(C)$ of $C_0([0, \infty))$ (to be formalized in Proposition 2.4 below). Under these conditions, Theorem 1.6.5 in Ethier and Kurtz (1986) implies that

$$\sup_{s \in [0, \infty)} |U(t)f(s) - U_n([n^{1+\alpha}t])f(s)| \to 0, \quad f \in C_0([0, \infty)),$$

as $n \to \infty$ and for all $t \geq 0$, where $U$ is the semigroup operator corresponding to $C$. The assertion of the theorem with $C_{[0, \infty)}([0, \infty))$ replaced by $D_{[0, \infty)}([0, \infty))$
now follows from Theorem 4.2.6 in Ethier and Kurtz (1986), whereas the fact that the weak convergence holds in $C_{[0,\infty)}([0,\infty))$ follows from relativization of the Skorohod topology (cf. Remark 2.2-(a) above).

Hence dynamic heterogeneity of the two-parameter model is described by a non negative diffusion obtained with a space-time rescaling which depend on the parameter $\alpha$. Note that $S_{\theta,\alpha}(\cdot)$ in (9) can be seen as a critical continuous-state branching process with immigration (Kawazu and Watanabe, 1971; Li, 2006), or as the square of a $\theta$-dimensional Bessel process (see Revuz and Yor (1991), Chapter XI). Here $\theta > 0$ is interpreted as the immigration rate, whereas the case $\theta < 0$ has been treated in Göing-Jaeschke and Yor (2003).

The next proposition, which provides the complete boundary behaviour of the process driven by (9) and formalizes its well-definiteness, is essentially not new and included for formal completeness. Part of the statement can be found for example in Ikeda and Watanabe (1989), Example 4.8.2. Let $A_0$ be the second order differential operator

$$A_0 = \theta \frac{d}{ds} + \alpha s \frac{d^2}{ds^2}, \quad 0 < \alpha < 1, \quad \theta > -\alpha. \quad (13)$$

**Proposition 2.4.** The process $\{S_{\theta,\alpha}(t), t \geq 0\}$ driven by (9) has the following boundary behavior: the point $0$ is
- absorbing, for $\theta \leq 0$,
- instantaneously reflecting, for $0 < \theta < \alpha$,
- entrance, for $\theta \geq \alpha$,
and the point $\infty$ is
- natural and non attracting, for $\theta \leq \alpha$,
- natural and attracting, for $\theta > \alpha$.

Moreover, $S_{\theta,\alpha}(t)$ is null recurrent for $\theta = \alpha$ and transient for $\theta \neq \alpha$. For $A_0$ as in (13), define

$$\mathcal{D}_{\theta,\alpha}(A_0) = \left\{ \begin{array}{ll}
  f \in \mathcal{D}(A_0) & \text{if } \theta \geq \alpha \\
  f \in \mathcal{D}(A_0) : \lim_{s \to 0} A_0 f(s) = 0 & \text{if } \theta < \alpha
\end{array} \right\}
$$

with

$$\mathcal{D}(A_0) = \left\{ f \in C_0([0,\infty)) \cap C^2((0,\infty)) : A_0 f \in C_0([0,\infty)) \right\}.$$

Then $\{(f, A_0 f) : f \in \mathcal{D}_{\theta,\alpha}(A_0)\}$ generates a Feller semigroup on $C_0([0,\infty))$.

**Proof.** Let the drift and volatility of the process be $\mu(s) = \theta$ and $\sigma^2(s) = 2\alpha s$.

The scale function for the process equals

$$S(x) = \int_{x_1}^x s(y)dy \quad (14)$$
\[
= \int_{x_1}^{x} \exp \left\{ - \int_{x_0}^{y} \frac{2\mu(u)}{\sigma^2(u)} \, du \right\} \, dy \left\{ \begin{array}{ll}
x_0 \log(x/x_1) & \theta = \alpha \\
C x^{1-\theta/\alpha} & \frac{\alpha - \theta}{\alpha - \theta} \neq \alpha 
\end{array} \right.
\]
for some positive constant \( C \). Letting \( S[a, b] = S(b) - S(a) \), for \( 0 < a < b < \infty \), we have

\[
S(0, b) = \lim_{a \uparrow 0} S[a, b] \left\{ \begin{array}{ll}
< \infty & \theta < \alpha \\
= \infty & \theta \geq \alpha 
\end{array} \right.
\]

(15)

\[
S[a, \infty) = \lim_{b \downarrow \infty} S[a, b] \left\{ \begin{array}{ll}
= \infty & \theta \leq \alpha \\
< \infty & \theta > \alpha.
\end{array} \right.
\]

Moreover, for \( m(\xi) = (\sigma^2(\xi)s(\xi))^{-1} \), with \( \sigma^2(\xi) \) as above and \( s(\xi) \) as in (14), it can be shown that

\[
\Sigma(l) = \int_{l}^{x} S[l, \xi]m(\xi) \, d\xi \left\{ \begin{array}{ll}
(l - x \log(l))/2 & \theta = \alpha \\
x_0^{-\theta/\alpha} l (\alpha - \theta) - \alpha x^{\theta/\alpha} l^{1-\theta/\alpha} & 2\theta(\alpha - \theta) \neq \alpha 
\end{array} \right.
\]

for some constant \( C \), so that

\[
\Sigma(0) = \lim_{l \to 0} \Sigma(l) \left\{ \begin{array}{ll}
< \infty & \theta < \alpha \\
= \infty & \theta \geq \alpha 
\end{array} \right.
\]

while \( \Sigma(\infty) = \lim_{l \to \infty} \Sigma(r) = \infty \). Furthermore

\[
M[c, d] = \int_{c}^{d} m(x) \, dx = (2\theta)^{-1} x_0^{-\theta/\alpha} (d^{\theta/\alpha} - c^{\theta/\alpha})
\]

from which we have \( M[\{0\}] = 0 \) for \( \theta > 0 \), and

\[
M(0, d) = \lim_{c \to 0} M[c, d] \left\{ \begin{array}{ll}
= \infty & \theta < 0 \\
< \infty & \theta > 0 
\end{array} \right.
\]

\[
M[c, \infty) = \lim_{d \to \infty} M[c, d] \left\{ \begin{array}{ll}
< \infty & \theta < 0 \\
= \infty & \theta > 0.
\end{array} \right.
\]

Moreover

\[
N(r) = \frac{\theta r + x(\alpha - \theta) - \alpha r^{\theta/\alpha} x^{1-\theta/\alpha}}{2\theta(\alpha - \theta)}
\]

from which \( N(\infty) = \lim_{r \to \infty} N(r) = \infty \) and

\[
N(0) = \lim_{l \to 0} N(l) \left\{ \begin{array}{ll}
= \infty & \theta < 0 \\
< \infty & \theta > 0.
\end{array} \right.
\]

The first assertion then follows from Karlin and Taylor (1981), Section 15.6. The second assertion follows from (15) and, for example, Khasminskii (2012), Section 3.8. The third assertion follows from the first assertion together with Theorem 8.1.1 and Corollary 8.1.2 in Ethier and Kurtz (1986). \( \square \)
The following is an immediate consequence of the previous result, due to the transience or null recurrence of the process.

**Corollary 2.5.** Let \( \{S_{\theta,\alpha}(t), t \geq 0\} \) be as in (9). Then there exists no stationary density for the process.

We conclude the section with a brief discussion of the process corresponding to (9) for the one-parameter model. Although the notion of \( \alpha \)-diversity is given for Poisson-Kingman models with \( 0 < \alpha < 1 \) (cf. Pitman, 2003), a result analogous to Theorem 2.3 can be nonetheless derived for the one-parameter case. Note that the limit corresponding to (5) when \( \alpha = 0 \) is

\[
\lim_{n \to \infty} \frac{K_n}{\log n} = \theta \quad \text{a.s.;}
\]

see Korwar and Hollander (1973). Hence we expect the process for the normalized number of species to converge to a constant process, i.e.

\[
\{K_n(\lfloor c_n t \rfloor) / \log n, t \geq 0\} \overset{p}{\to} S(t) \equiv \theta,
\]

for some \( c_n \to \infty \). This can indeed be seen immediately from (11), which in this case becomes

\[
(U_n - I)f\left(\frac{k}{n^\alpha}\right) = \frac{1}{\log n} f'\left(\frac{k}{\log n}\right) \left[\left(1 - \frac{w}{n}\right) \frac{\theta}{\theta + n} - \frac{w}{n} \left(\frac{n}{\theta + n}\right)\right]
\]

\[
+ \frac{1}{2\log^2 n} f''\left(\frac{k}{\log n}\right) \left[\left(1 - \frac{w}{n}\right) \frac{\theta}{\theta + n} + \frac{w}{n} \left(\frac{n}{\theta + n}\right)\right] + o((n \log n)^{-1})
\]

where we have used the fact that

\[
M_{1,n} \overset{d}{\to} W \sim \text{Poisson}(\theta);
\]

see Arratia, Barbour and Tavarè (2003), Theorem 4.17. Hence \( c_n(U_n - I)f(k/n^\alpha) \), with \( c_n = n \log n \), converges to 0.

It follows that the dynamics of the number of species underlying the infinite alleles models are driven by a constant process in the one-parameter case, and by a diffusion process on \([0, \infty)\) with state-dependent volatility in the two-parameter case. This difference will be found again in Section 4 with a different approach.

## 3 Some considerations on mutation rates

In the Introduction it was mentioned that two different sequential constructions of the two-parameter model have been provided in Petrov (2009) and Ruggiero and Walker (2009). In this section we briefly outline why these offer only partial insight into the dynamics underlying the two-parameter model from a biological
perspective, motivating the need for further investigation. This is done by means of some considerations concerning the problem of finding the intensity rates for the mutation process which generates the model, which also prepare the ground for the construction detailed in the next section, developed in the spirit of that corresponding to (3)-(4).

The above-mentioned constructions consist respectively of a sequence taking values on the space of partitions of \( \mathbb{N} \), and of the Moran-type particle representation outlined in Section 2 above. Both cases are based on a dynamic system of finitely-many exchangeable particles and exhibit right-continuous sample paths, whereas (3) (with an appropriate domain) characterizes an \( n \)-dimensional diffusion process. Another notable feature of these constructions is the assumption that the mutant offspring distribution, i.e. the distribution which generates the mutant types, is non atomic, hence it selects types which appear for the first time with probability one. In the framework of Petrov (2009) this amounts to say that a new box is occupied with probability one. This in particular turns out to be the key point for proving the weak convergence of the sequences to the two-parameter model (see e.g. Ruggiero and Walker (2009) after Remark 3.1). Such assumption of non atomicity cannot be applied in a construction similar to (3)-(4), because the mass of the distribution must concentrate on the enumerated types, in order to keep the maximum amount of species constant in time. To be more specific, note first that the drift coefficient in Wright-Fisher diffusions as in (3) are determined as

\[
\begin{align*}
    b_i^{(n)}(z) &= \sum_{j \neq i} n q_{ji}^{(n)} z_j - \sum_{j \neq i} n q_{ij}^{(n)} z_i \\
    q_{ij}^{(n)} &= \frac{\theta}{n-1} , \quad i \neq j,
\end{align*}
\]

where \( q_{ij} \) is the intensity of a mutation from type \( i \) to type \( j \), and diagonal elements are \( q_{ii} = -\sum_{j \neq i} q_{ij} \), so that \( (q_{ij})_{i,j=1,...,n} \) is a square matrix with nonnegative off-diagonal elements and row sums equal to zero. The drift (4) for example is given by taking parent-independent symmetric mutations with rates

\[
q^{(n)}_{ij} = \frac{\theta}{n-1} , \quad i \neq j,
\]

whereby when a type \( i \) mutates, the new type will be any of the other \( n - 1 \) types with equal chances, and \( \theta \) controls how often on average a mutation occurs. Hence the mutant type is chosen with uniform probability, and the mutant distribution is discretely supported. The construction can be extended to have non-symmetric mutation (see for example Ethier and Kurtz (1981), Theorem 3.4), but the difference is not relevant for our purposes.

To sum up, the two existing constructions for the two-parameter model feature finitely-many objects potentially of infinitely-many types and a diffuse mutation offspring distribution, while the sought after construction should feature infinitely-
many objects of finitely-many types and a discretely supported mutation offspring distribution.

Before proceeding with the construction, it is worth dedicating a few additional considerations on such a problem from a mathematical point of view. The task is that of finding mutation rates $q_{ij}^{(n)}$ as in (17) which yield in the limit the drift term

$$\lim_{n \to \infty} b_i^{(n)}(z) = -\theta z_i - \alpha$$

uniformly on $z$, together with some boundary conditions which $b^{(n)}(z)$ must satisfy on

$$\Delta_n = \left\{ z \in [0, 1]^n : z_i \geq 0, \sum_{i=1}^{n} z_i = 1 \right\},$$

given by

$$b_i^{(n)}(z) \geq 0 \quad \text{if} \quad z_i = 0, \quad b_i^{(n)}(z) \leq 0 \quad \text{if} \quad z_i = 1.$$

The form of (19) suggests that we must let $q_{ij}^{(n)}$ depend on the current state $z$, so that (17) becomes

$$b_i^{(n)}(z) = \sum_{j \neq i} q_{ij}^{(n)}(z)z_j - \sum_{j \neq i} q_{ij}^{(n)}(z)z_i.$$

Suppose for the sake of the argument that $z_i > 0$, and consider for example the easiest modification of (18) with (19) in mind, that is

$$q_{ij}^{(n)}(z) = \frac{\theta}{n-1} + \frac{\alpha}{z_i(n-1)}, \quad i \neq j.$$

This, by means of (22) provides the drift component

$$b_i^{(n)}(z) = \frac{\theta}{n-1}(1 - z_i) - \theta z_i + \alpha - \alpha$$

giving (4) again. Consider instead

$$q_{ij}^{(n)}(z) = \frac{\theta}{n-1} + \frac{2\alpha j}{z_i n(n+1)}, \quad i \neq j,$$

which gives

$$b_i^{(n)}(z) = \frac{\theta}{n-1}(1 - z_i) - \theta z_i + \frac{2\alpha i(n-1)}{n(n+1)} - \alpha.$$
For this choice (19), holds uniformly, but (21) is not satisfied for all \( i = 1, \ldots, n \), since for \( z_i \to 0 \) we have that \( b_i^{(n)}(z) \geq 0 \) if and only if
\[
i \geq \frac{n(n+1)}{2\alpha(n-1)} \left( \alpha - \frac{\theta}{n-1} \right) \approx \frac{n}{2},
\]
and for \( z_i \to 1 \) we have that \( b_i^{(n)}(z) \leq 0 \) if and only if
\[
i \leq \frac{(\alpha + \theta)n(n+1)}{2\alpha(n-1)} \approx \frac{(1 + \theta)n}{2}.
\]
Finally, consider for some positive increasing sequence \( \{c_n\}_{n \in \mathbb{N}} \)
\[
q_{ij}^{(n)}(z) = \frac{\theta}{n-1} + \frac{\alpha c_n}{(n-1)(1 + c_n z_i)}, \quad i \neq j,
\]
which gives
\[
b_i^{(n)}(z) = \frac{\theta}{n-1} (1 - z_i) - \theta z_i + \frac{\alpha}{n-1} \sum_{j \neq i} \frac{c_n z_j}{1 + c_n z_j} - \frac{\alpha c_n z_i}{1 + c_n z_i}.
\]
This can be seen to satisfy (21), but (19) fails in some regions of \( \Delta_n \). Hence no intuitive choice seems to lead to the desired conditions, except that we expect the rates to be non constant and possibly state-dependent. It turns out that in order to identify at least an instance of the mutation rates which yield in the limit the two-parameter model we need to impose additional restrictions, mainly regarding the state space of the differential operator. These, together with the convergence argument, are formalized in the next section.

4 Sequential construction

Let \( n \geq 2 \) throughout the section, and let \( \Delta_n \) be as in (20). Consider a sequence of reals \( \{\varepsilon_n\}_{n \in \mathbb{N}} \) such that
\[
0 < \varepsilon_n < \frac{1}{n} \quad \forall n, \quad \varepsilon_n = \mathcal{O}(n^{-1}),
\]
and define the compact subspace of \( \Delta_n \) given by
\[
\Delta_{n,\varepsilon_n} = \left\{ z \in [0,1]^n : z_i \geq \varepsilon_n, \sum_{i=1}^n z_i = 1 \right\},
\]
where \( z \in \Delta_{n,\varepsilon_n} \) implies \( z_i \in [\varepsilon_n, 1 - (n-1)\varepsilon_n] \neq \emptyset \) for all \( i \). For \( z \in \Delta_{n,\varepsilon_n} \), consider the second order differential operator
\[
\mathcal{A}_n = \frac{1}{2} \sum_{i,j=1}^n a_{ij}^{(n)}(z) \frac{\partial^2}{\partial z_i \partial z_j} + \frac{1}{2} \sum_{i=1}^n b_i^{(n)}(z) \frac{\partial}{\partial z_i},
\]
with domain
\begin{equation}
\mathcal{D}(A_n) = \{ f : f \in C^2(\Delta_n, \varepsilon_n) \},
\end{equation}
where
\begin{equation}
C^2(\Delta_n, \varepsilon_n) = \{ f \in C(\Delta_n, \varepsilon_n) : \exists \tilde{f} \in C^2(\mathbb{R}^n), \tilde{f}|_{\Delta_n, \varepsilon_n} = f \}.
\end{equation}
The covariance components in (24) are specified to be
\begin{equation}
d_{ij}^{(n)}(z) = (z_i - \varepsilon_n)(\delta_{ij}(1 - n\varepsilon_n) - (z_j - \varepsilon_n))
= \begin{cases} 
(z_i - \varepsilon_n)(1 - (n - 1)\varepsilon_n - z_i) & \text{if } i = j \\
-(z_i - \varepsilon_n)(z_j - \varepsilon_n) & \text{if } i \neq j,
\end{cases}
\end{equation}
which can be seen as Wright-Fisher-type covariance terms restricted to $\Delta_n, \varepsilon_n$, since they vanish at $z_i = \varepsilon_n$ and $z_i = 1 - (n - 1)\varepsilon_n$. Consider now the state-dependent mutation rates given by
\begin{equation}
q_{ij}^{(n)}(z) = \frac{\theta}{n - 1} + \frac{2\alpha_j}{z_i n (n + 1)} \left[ 1 - \exp\left\{- (z_i - \varepsilon_n) e^{1/\varepsilon_n} \right\} \right], \quad i \neq j.
\end{equation}
Despite the cumbersome appearance of (27), some consideration can be done. A first, immediate remark, which is nonetheless crucial for the overall picture, is that (27) is strikingly different from (18), which is constant, whereas the contribution of $\alpha$ is heavily dependent on the current state of the process. Heuristically, this is consistent with the differences encountered between (16) and (10) in Section 2. Recall now that the limit operator (2) acts on functions defined on (1), where the frequencies have been ordered. The same ordering operation will be done before taking the limit of $A_n$, but the formal appearance of the operator will be unchanged. In light of this, the term $j$ at the numerator of $q_{ij}$ is an approximate indication of the size of the frequency $z_j$. Hence mutations from $i$ to $j$ occur more frequently if $z_j$ is relatively low, implying a redistributive effect. On the other hand, the behavior of the terms which depend on $z_i$ is not monotone. In particular, as $z_i$ decreases $[1 - \exp\{- (z_i - \varepsilon_n) e^{1/\varepsilon_n} \}]/z_i$ increases, except when $z_i$ is very close to $\varepsilon_n$, where it decreases abruptly to zero. Of these two behaviors, the latter is what ultimately keeps the process inside the state space. The former is instead related to a reinforcement feature noted in Lijoi, Mena and Prünster (2007), also possessed by the PD($\theta, \alpha$) distribution, which is intimately connected with the two-parameter diffusion model. It can indeed be observed that the probability that a further sample from (6) is an already observed species is not allocated proportionally to the current frequencies. The ratio of probabilities assigned to any pair of species $(i, j)$ is
\begin{equation}
r_{i,j} = \frac{n_i - \alpha}{n_j - \alpha}.
\end{equation}
When \( \alpha \to 0 \), the probability of sampling species \( i \) is proportional to the absolute frequency \( n_i \), or equivalently to \( z_i \), which in continuous time is reflected by a constant mutation rate as in (18). However, since \( (n_i - \alpha) / (n_j - \alpha) \) is increasing in \( \alpha \), a value of \( \alpha > 0 \) reallocates some probability mass from type \( j \) to type \( i \), so that, for example, for \( n_i = 2 \) and \( n_j = 1 \) we have \( r_{i,j} = 2, 3, 5 \) for \( \alpha = 0, 0.5, 0.75 \) respectively. Thus \( \alpha \) has a reinforcement effect on those species that have higher frequency. All together, these mechanisms imply higher volatility for lower frequencies, and vice versa.

The drift components in (24) are obtained from (27) by means of (22), yielding

\[
(28) \quad b_i^{(n)}(z) = \frac{\theta}{n-1}(1 - z_i) - \theta z_i + \frac{2\alpha i}{n(n+1)} \sum_{j \neq i} \left[ 1 - \exp \left\{ -(z_j - \varepsilon_n) e^{1/\varepsilon_n} \right\} \right] \\
- \alpha \left[ 1 - \exp \left\{ -(z_i - \varepsilon_n) e^{1/\varepsilon_n} \right\} \right].
\]

Note that the first two terms of \( b_i^{(n)}(z) \) equal

\[
\frac{\theta}{n-1}(1 - (n-1)\varepsilon_n - z_i) - \theta(z_i - \varepsilon_n),
\]

and that \( z_i = 1 - (n-1)\varepsilon_n \) implies \( z_j = \varepsilon_n \) for all \( j \neq i \). These two observations show that

\[
(29) \quad b_i^{(n)}(\varepsilon_n) > 0, \quad b_i^{(n)}(1 - (n-1)\varepsilon_n) < 0
\]

and thus \( b_i^{(n)} \) satisfies (21) restricted to \( \Delta_{n,\varepsilon_n} \).

Let \( K \) denote a compact space, and let \( ||f|| = \sup_{z \in K} f(z) \). The following definition recalls the notion of Feller semigroup. We refer the reader to the exhaustive textbook by Ethier and Kurtz (1986) for more details.

**Definition 4.1.** A one-parameter family of bounded linear operators \( \{T(t), t \geq 0\} \) on \( C(K) \) is said to be a Feller semigroup if \( T(t) \):

- has the semigroup property \( T(s + t) = T(s)T(t) \) for all \( s, t \geq 0 \);
- is strongly continuous, i.e.
  \[
  ||T(t)f - f|| \to 0 \quad \text{as} \quad t \to 0, \\
  f \in C(K);
  \]
- is a contraction, i.e. \( ||T(t)|| \leq 1 \) for all \( t \geq 0 \);
- is conservative, in the sense that \( T(t)1 = 1 \);
- is positive, in the sense that it preserves the cone of non negative functions.

Given a semigroup \( \{T(t)\} \) on \( C(K) \), a \( K \)-valued Markov process \( X(\cdot) = \{X(t), t \geq 0\} \) is said to correspond to \( \{T(t)\} \) if

\[
\mathbb{E}[f(X(t + s))|\mathcal{F}_t^X] = T(s)f(X(t))
\]

for all \( t, s \geq 0 \) and \( f \in C(K) \), where \( \mathcal{F}_t^X = \sigma(X(u), u \leq t) \). \( \square \)
Denote by \( C_K([0, \infty)) \) the space of continuous functions from \([0, \infty)\) to \( K \).

**Theorem 4.2.** Let \( \mathcal{A}_n \) be the operator defined by (24)-(25)-(26)-(28). Then the closure of \( \mathcal{A}_n \) in \( C(\Delta_{n,\varepsilon_n}) \) is single-valued and generates a Feller semigroup \( \{ T_n(t) \} \) on \( C(\Delta_{n,\varepsilon_n}) \). For each \( \nu_n \in \mathcal{P}(\Delta_{n,\varepsilon_n}) \), there exists a strong Markov process \( \mathcal{Z}^{(n)}(\cdot) = \{ Z^{(n)}(t), t \geq 0 \} \) corresponding to \( \{ T_n(t) \} \) with initial distribution \( \nu_n \) and such that

\[
\mathbb{P}\{ Z^{(n)}(\cdot) \in C(\Delta_{n,\varepsilon_n}([0, \infty])) \} = 1.
\]

**Proof.** It is easily seen that \( \mathcal{A}_n \) satisfies the positive maximum principle on \( \Delta_{n,\varepsilon_n} \), that is if \( f \in \mathcal{P}(\mathcal{A}_n) \) and \( z_0 \in \Delta_{n,\varepsilon_n} \) are such that \( f(z_0) = ||f|| \geq 0 \), then \( \mathcal{A}_n f(z_0) \leq 0 \). This is immediate in the interior of \( \Delta_{n,\varepsilon_n} \), while on the boundaries it follows from (29) and the fact that (26) vanishes at every boundary point. Denote now \( z^\sigma = z_1^{\sigma_1} \cdots z_n^{\sigma_n} \) and \( \sigma - \delta_i = (\sigma_1, \ldots, \sigma_i - 1, \ldots, \sigma_n) \) for \( \sigma_1, \ldots, \sigma_n \in \mathbb{N} \). Then

\[
\mathcal{A}_n z^\sigma = \sum_i b_i^{(n)}(z) z^{\sigma - \delta_i} + \sum_i \sigma_i(z_i - \varepsilon_n)(1 - (n - 1)\varepsilon_n - z_i)z^{\sigma - 2\delta_i}
- \sum_{i \neq j} \sigma_i \sigma_j (z_j - \varepsilon_n)(z_j - \varepsilon_n)z^{\sigma - \delta_i - \delta_j}
= \sum_i \left[ \begin{array}{c}
\frac{\theta}{n-1}z^{\sigma - \delta_i} - \frac{\theta}{n-1}z^{\sigma} - \theta z^{\sigma} + \frac{2\alpha i}{n(n+1)} \sum_{j \neq i} (1 - C_n e^{-z_j C''_n}) z^{\sigma - \delta_i} \\
- \alpha C_n z^{\sigma - \delta_i} + \alpha C_n e^{-z_j C''_n} z^{\sigma - \delta_i},
\end{array} \right]
+ \sigma_i \left[ \begin{array}{c}
1 - (n - 1)\varepsilon_n \right] z^{\sigma - \delta_i} - z^{\sigma} - \varepsilon_n[1 - (n - 1)\varepsilon_n]z^{\sigma - 2\delta_i} + \varepsilon_n z^{\sigma - \delta_i}
- \sum_{i \neq j} \sigma_i \sigma_j \left[ z^{\sigma} + \varepsilon_n z^{\sigma - \delta_j} + \varepsilon_n z^{\sigma - \delta_j} - \varepsilon_n^2 z^{\sigma - \delta_j - \delta_j} \right]
\]
where \( C_n = 1 + o(1), C'_n = \exp(-\varepsilon_n C''_n) \) and \( C''_n = \exp(1/\varepsilon_n) \). Hence, letting \( L_m \) denote the algebra of polynomials in \((z_1, \ldots, z_n)\) restricted to \( \Delta_{n,\varepsilon_n} \), with degree not greater than \( m \in \mathbb{N} \), the image of \( \mathcal{A}_n \) computed on \( L_m \) contains functions belonging to \( L_m \) and of type \( e^{-b_0 z_c} \), with \( b_0 \) fixed. Since for every \( g(x) \in C(K) \), with \( K \) compact, and \( f(x) = e^{b_0 x} g(x) \in C(K) \), there exists a sequence \( \{ p^{(k)} \} \) of polynomials on \( K \) such that \( ||f - p^{(k)}|| \to 0 \), so that \( ||e^{-b_0 z_c} - g|| \to 0 \), it follows that the image of \( \mathcal{A}_n \) is dense in \( C(\Delta_{n,\varepsilon_n}) \), and so is that of \( \lambda - \mathcal{A}_n \) for all \( \lambda > 0 \). Since \( \cup_m L_m \) is dense in \( C(\Delta_{n,\varepsilon_n}) \), Theorem 4.2.2 of Ethier and Kurtz (1986) now implies that the closure of \( \mathcal{A}_n \) on \( C(\Delta_{n,\varepsilon_n}) \) is single-valued and generates a strongly continuous, positive, contraction semigroup \( \{ T_n(t) \} \) on \( C(\Delta_{n,\varepsilon_n}) \). The fact that \((1,0)\) belongs to the domain of \( \mathcal{A}_n \) implies also that \( \{ T_n(t) \} \) is conservative. Note now that for every \( z_0 \in \Delta_{n,\varepsilon_n} \) and \( \delta > 0 \) there exists \( f \in \mathcal{P}(\mathcal{A}_n) \) such that

\[
\sup_{z \in B^c(z_0,\delta)} f(z) < f(z_0) = ||f|| \quad \text{and} \quad \mathcal{A}_n f(z_0) = 0,
\]
where $B^c(z_0, \delta)$ is a ball of radius $\delta$ centered at $z_0$. Take for example $f(z) = -C_\delta \sum_{i=1}^{n}(z_i - z_0)^4$ for an appropriate constant $C_\delta$ which depends on $\delta$. Then the second assertion follows from Theorem 4.2.7 together with Remark 4.2.10 in Ethier and Kurtz (1986).

Given $\nabla_\infty$ as in (1), define the subspaces

$$\nabla_\infty = \left\{ z \in \nabla_\infty : \sum_{i=1}^{\infty} z_i = 1 \right\}$$

and

$$\nabla_{n, \varepsilon_n} = \left\{ z \in \nabla_\infty : z_n \geq \varepsilon_n > z_{n+1} = 0 \right\},$$

and define the Borel measurable map $\rho_n : \Delta_n \to \nabla_\infty$ by

$$\rho_n(z) = (z(1), \ldots, z(n), 0, 0, \ldots), \quad z \in \Delta_n$$

where $z(i)$ are the decreasing order statistics of $z \in \Delta_n$. It is clear that $\rho_n$ maps $\Delta_{n, \varepsilon_n}$ into $\nabla_{n, \varepsilon_n}$. If $Z^{(n)}(\cdot)$ is the Markov process of Theorem 4.2, our aim is thus to show that

$$\rho_n(Z^{(n)}(\cdot)) \Rightarrow Z(\cdot)$$

in the sense of convergence in distribution in $C(\nabla_\infty([0, \infty)))$ as $n \to \infty$, where $Z(\cdot)$ is the diffusion process corresponding to the Feller semigroup generated by the closure of (2) in an appropriate space. To this end, consider the symmetric polynomials

$$\varphi_m(z) = \sum_{i \geq 1} z_i^m, \quad z \in \nabla_\infty, \quad m \geq 2,$$

and define

$$\mathcal{D}_0(\mathcal{B}) = \left\{ \text{subalgebra of } C(\nabla_\infty) \text{ generated by } 1, \varphi_2(z), \varphi_3(z), \ldots \right\}$$

and

$$\mathcal{D}(\mathcal{B}) = \left\{ \text{subalgebra of } C(\nabla_\infty) \text{ generated by } 1, \varphi_3(z), \varphi_4(z), \ldots \right\}.$$

**Lemma 4.3.** $\mathcal{D}(\mathcal{B})$ is dense in $C(\nabla_\infty)$.

**Proof.** In Ethier and Kurtz (1981) (see proof of Theorem 2.5) it is proved that the closure of $\mathcal{D}_0(\mathcal{B})$ equals $C(\nabla_\infty)$. Note now that

$$z_1 = \lim_{m \to \infty} \varphi_m(z)^{1/m}, \quad z_2 = \lim_{m \to \infty} (\varphi_m(z) - z_1)^{1/m}, \quad \ldots$$
from which

$$\varphi_2(z) = \lim_{m \to \infty} \varphi_m(z)^{2^m} + (\varphi_m(z) - z_1^{2^m} + \ldots$$

so that $\varphi_2 \in \mathcal{D}(\mathcal{B})$. It follows that $\mathcal{D}_0(\mathcal{B}) \equiv \mathcal{D}(\mathcal{B})$, from which the result follows.

Before providing the convergence argument, we recall the relevant theorems about the formal existence and the sample path properties of the process $Z(\cdot)$ appearing in (33).

**Theorem 4.4.** [Petrov, 2009] Let $\mathcal{B}$ be the operator defined by (2) and (35). The closure of $\mathcal{B}$ in $\mathcal{C}(\nabla_\infty)$ generates a Feller semigroup $\{\mathcal{T}(t)\}$ on $\mathcal{C}(\nabla_\infty)$, and for each $\nu \in \mathcal{P}(\nabla_\infty)$ there exists a strong Markov process $Z(\cdot) = \{Z(t), t \geq 0\}$ corresponding to $\{\mathcal{T}(t)\}$ with initial distribution $\nu$ and such that

$$\mathbb{P}\{Z(\cdot) \in C_{\nabla_\infty}([0, \infty))\} = 1.$$

Let $\nabla_\infty$ be as in (31). The following result shows that if the initial distribution of the Markov process $Z(\cdot)$ of Theorem 4.4 is PD($\theta, \alpha$), then the law of the process is concentrated on $C_{\nabla_\infty}([0, \infty))$.

**Theorem 4.5.** [Feng and Sun, 2010] Let $Z(\cdot) = \{Z(t), t \geq 0\}$ be the Markov process of Theorem 4.4, and assume $Z(0) \sim$ PD($\theta, \alpha$). Then

$$\mathbb{P}\{Z(t) \in \nabla_\infty, \forall t \geq 0\} = 1.$$

Denote now by $\mathcal{B}_n$ the operator (24) with coefficients defined by (26) and (28) when its domain is taken to be

$$\mathcal{D}(\mathcal{B}_n) = \{f \in C(\nabla_{n,\varepsilon_n}) : f \circ \rho_n \in C^2(\Delta_{n,\varepsilon_n})\}.$$

We are now ready to state the main result of the section.

**Theorem 4.6.** Let $Z^{(n)}$ and $Z(\cdot)$ be the Markov processes of Theorem 4.2 and Theorem 4.4, with initial distribution $\nu_n \in \mathcal{P}(\Delta_{n,\varepsilon_n})$ and $\nu$ in $\mathcal{P}(\nabla_\infty)$ respectively. If $\nu_n \circ \rho_n^{-1} \Rightarrow \nu$, then (33) holds in $C_{\nabla_\infty}([0, \infty))$. If in addition $\nu \in \mathcal{P}(\nabla_\infty)$, then (33) holds in $C_{\nabla_\infty}([0, \infty))$.

**Proof.** For $\rho_n$ as in (32), define $\pi_n : \mathcal{C}(\nabla_\infty) \to C(\Delta_n)$ by $\pi_n f = f \circ \rho_n$ and note that $\pi_n : \mathcal{D}(\mathcal{B}) \to \mathcal{D}(\mathcal{A}_n)$. Since for every $f \in \mathcal{D}(\mathcal{B})$ we have $\pi_n \mathcal{B} f = \mathcal{B}(f \circ \rho_n)$, for all such functions and $z \in \Delta_{n,\varepsilon_n}$ we have

$$\mathcal{A}_n \pi_n f(z) - \pi_n \mathcal{B} f(z) =$$

$$= \frac{1}{2} \sum_{i,j=1}^{n} \varepsilon_n \left[ z_i + z_j - \varepsilon_n - \delta_{ij} (1 + n(z_i - \varepsilon_n)) \right] \frac{\partial^2 f(\rho_n(z))}{\partial z_i \partial z_j}$$
For \( f \) bounded, since the inner sum converges for \( m \)
Here \( \sum_{n} \left( \frac{\theta}{n-1} \right) \exp \left\{ -(z_j - \varepsilon_n) e^{1/\varepsilon_n} \right\} \)
\( \alpha \exp \left\{ -(z_i - \varepsilon_n) e^{1/\varepsilon_n} \right\} \) \( \left| \frac{\partial f(\rho_n(z))}{\partial z_i} \right| \)
\( z_i + z_j - \varepsilon_n - \delta_{ij} (1 - n(z_i - \varepsilon_n)) = \begin{cases} 2z_i - \varepsilon_n - 1 - n(z_i - \varepsilon_n) & i = j \\ z_i + z_j - \varepsilon_n & i \neq j \end{cases} \)
whose absolute value is bounded above by \( n \), and also
\( \sum_{j \neq i} \left( 1 - \exp \left\{ -(z_j - \varepsilon_n) e^{1/\varepsilon_n} \right\} \right) \leq n - 1, \)
so that
\[
|A_n \pi_n f(z) - \pi_n B f(z)| \leq \frac{\varepsilon_n n}{2} \sum_{i,j=1}^{n} \left| \frac{\partial^2 f(\rho_n(z))}{\partial z_i \partial z_j} \right| + \frac{\theta}{2(n-1)} \sum_{i=1}^{n} \left| \frac{\partial f(\rho_n(z))}{\partial z_i} \right| \\
+ \frac{2\alpha}{n} \sum_{i=1}^{n} \left| \frac{\partial f(\rho_n(z))}{\partial z_i} \right| + \alpha \sum_{i=1}^{n} \exp \left\{ -(z_i - \varepsilon_n) e^{1/\varepsilon_n} \right\} \left| \frac{\partial f(\rho_n(z))}{\partial z_i} \right|.
\]
For \( f \in \mathcal{D}(\mathcal{B}) \) of type \( \varphi_{m_1} \times \cdots \times \varphi_{m_k} \), we have \( f(\rho_n(z)) = f(z) \) and
\[
\sum_{i=1}^{n} \left| \frac{\partial f(z)}{\partial z_i} \right| = \sum_{i=1}^{n} \sum_{j=1}^{k} m_j z_i^{m_j-1} \prod_{h \neq j} \varphi_{m_h} \leq \sum_{j=1}^{k} m_j \sum_{i=1}^{n} z_i^{m_j-1} \]
which is bounded above by \( \sum_{j=1}^{k} m_j < \infty \). The previous also implies
\[
\sum_{i=1}^{n} \exp \left\{ -(z_i - \varepsilon_n) e^{1/\varepsilon_n} \right\} \left| \frac{\partial f(z)}{\partial z_i} \right| \leq n \varepsilon_n^2 \sum_{j=1}^{k} m_j \to 0
\]
uniformly as \( n \to \infty \), given (23). Furthermore
\[
\sum_{i=1}^{n} \left| \frac{\partial f(z)}{\partial z_i} \right| \leq \sum_{j=1}^{k} m_j \sum_{i=1}^{n} z_i^{m_j-1}
\]
is bounded, since the inner sum converges for \( m_j \geq 3 \). Finally
\[
\sum_{i,j=1}^{n} \left| \frac{\partial^2 f(z)}{\partial z_i \partial z_j} \right| \leq \sum_{i,j=1}^{\infty} \left| \frac{\partial_{ij} \varphi_{m_i} \prod_{\ell \neq h} \varphi_{m_{\ell}} + \sum_{q \neq h} \partial_{i} \varphi_{m_h} \partial_{j} \varphi_{m_q} \prod_{\ell \neq h,q} \varphi_{m_{\ell}}}{\partial_{i} \varphi_{m_h} \partial_{j} \varphi_{m_q} \prod_{\ell \neq h,q} \varphi_{m_{\ell}}} \right|
\]
where the right-hand side is bounded. The above inequalities, together with (23), imply that
\begin{equation}
||A_n \pi_n f - \pi_n B f|| \to 0, \quad f \in D(B).
\end{equation}
The previous, together with Lemma 4.3 above and Theorem 1.6.1 of Ethier and Kurtz (1986), in turn implies the semigroup convergence
\begin{equation}
||T_n(t) \pi_n f - \pi_n T(t) f|| \to 0, \quad f \in C(\nabla_\infty),
\end{equation}
uniformly on bounded intervals, with $T_n$ as in Theorem 4.2 and $T$ as in Theorem 4.4. The first assertion with $C(\nabla_\infty([0, \infty)))$ replaced by $D(\nabla_\infty([0, \infty)))$ now follows from (40) and Corollary 4.8.7 in Ethier and Kurtz (1986) (here $\rho_n(Z^{(n)})$ obviously satisfies a compact containment condition). The first assertion with $C(\nabla_\infty([0, \infty)))$ follows from relativization to $C(\nabla_\infty([0, \infty)))$ of the topology on $D(\nabla_\infty([0, \infty)))$. The second assertion follows from Theorem 4.5 and from relativization to $C(\nabla_\infty([0, \infty)))$ of the topology on $C(\nabla_\infty([0, \infty)))$. $\square$

References


Notes, Monograph Series. IMS, Hayward. (References in the text are to the arXiv:math/0210396v1 version).


