Alpha-diversity processes and normalized inverse-Gaussian diffusions

Matteo Ruggiero
Stephen G. Walker
Stefano Favaro
Alpha-diversity processes and normalized inverse-Gaussian diffusions

MATTEO RUGGIERO
University of Turin and Collegio Carlo Alberto

STEPHENV G. WALKER
University of Kent

STEFANO FAVARO
University of Turin and Collegio Carlo Alberto

December 2011

The infinitely-many-neutral-alleles model has recently been extended to a class of diffusion processes associated with Gibbs partitions of two-parameter Poisson-Dirichlet type. This paper introduces a family of infinite-dimensional diffusions associated with a different subclass of Gibbs partitions, induced by normalized inverse-Gaussian random probability measures. Such diffusions describe the evolution of the frequencies of infinitely-many types together with the dynamics of the time-varying mutation rate, which is driven by an \( \alpha \)-diversity diffusion. Constructed as a dynamic version, relative to this framework, of the corresponding notion for Gibbs partitions, the latter is explicitly derived from an underlying population model and shown to coincide, in a special case, with the diffusion approximation of a critical Galton-Watson branching process. The class of infinite-dimensional processes is characterized in terms of its infinitesimal generator on an appropriate domain, and shown to be the limit in distribution of a certain sequence of Feller diffusions with finitely-many types. Moreover, a discrete representation is provided by means of appropriately transformed Moran-type particle processes, where the particles are samples from a normalized inverse-Gaussian random probability measure. The relationship between the limit diffusion and the two-parameter model is also discussed.

1 Introduction

Considerable attention has been devoted recently to a class of diffusion processes which extends the infinitely-many-neutral-alleles model to the case of two parameters. This family takes values in the space

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

namely the closure in \([0, 1]^{\infty}\) of the infinite dimensional ordered simplex, and is characterized, for constants \(0 \leq \alpha < 1\) and \(\theta > -\alpha\), by the second order differential operator

\[ \mathcal{L}^{\theta, \alpha} = \frac{1}{2} \sum_{i,j=1}^{\infty} z_i (\delta_{ij} - z_j) \frac{\partial^2}{\partial z_i \partial z_j} - \frac{1}{2} \sum_{i=1}^{\infty} (\theta z_i + \alpha) \frac{\partial}{\partial z_i} \]
acting on a certain dense sub-algebra of the space $C(\nabla_\infty)$ of continuous functions on $\nabla_\infty$ (throughout the paper $\delta_{ij}$ denotes Kronecker delta). The diffusion with operator (2) describes the evolution of the allelic frequencies at a particular locus in a large population subject to random genetic drift and mutation, where mutation is jointly driven by the parameters $(\theta, \alpha)$. Ethier and Kurtz (1981) characterized the corresponding process when $\alpha = 0$, whereas the two-parameter family was introduced by Petrov (2009) and further investigated by Ruggiero and Walker (2009) and Feng and Sun (2010). The latter is known to be stationary, reversible and ergodic with respect to the Poisson-Dirichlet distribution with parameters $(\theta, \alpha)$. This was introduced by Pitman (1995) (see also Pitman (1996) and Pitman and Yor (1997)) and extends the Poisson-Dirichlet distribution of Kingman (1975) as follows. Consider a random sequence $(V_1, V_2, \ldots)$ obtained by means of the so-called stick-breaking scheme

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

where $0 \leq \alpha < 1$ and $\theta > -\alpha$. The vector $(V_1, V_2, \ldots)$ is said to have the GEM distribution with parameters $(\theta, \alpha)$, while the vector of descending order statistics $(V_{(1)}, V_{(2)}, \ldots)$ is said to have the Poisson-Dirichlet distribution with parameters $(\theta, \alpha)$. The latter is also the law of the ranked frequencies of an infinite partition induced by a two-parameter Poisson-Dirichlet random probability measure, which generalizes the Dirichlet process introduced by Ferguson (1973). Two-parameter Poisson-Dirichlet models have found applications in several fields. See for example the monographs by Bertoin (2006) for fragmentation and coalescent theory, Pitman (2006) for excursion theory and combinatorics, Teh and Jordan (2009) for machine learning, Lijoi and Prünster (2009) for Bayesian inference and Feng (2010) for population genetics. See also Bertoin (2008), Handa (2009) and Favaro, Lijoi, Mena and Prünster (2009).

The Poisson-Dirichlet distribution and its two parameter extension in turn belong to a larger class of random discrete distributions induced by infinite partitions of Gibbs-type. These were introduced by Gnedin and Pitman (2005), and applications include fragmentation and coalescent theory (Bertoin, 2006; McCullagh, Pitman and Winkel, 2008; Goldschmidt, Martin and Spanò, 2008), excursion theory (Pitman, 2003), statistical physics (Berestycki and Pitman, 2007), and Bayesian nonparametric inference (Lijoi, Mena and Prünster, 2005; 2007a;b; Lijoi, Prünster and Walker, 2008a). See Pitman (2006) for a comprehensive account. See also Griffiths and Spanò (2007), Lijoi, Prünster and Walker (2008b) and Ho, James and Lau (2007).

This paper introduces a class of infinite-dimensional diffusions associated with a different subclass of Gibbs-type partitions, induced by normalized inverse-Gaussian random probabil-
ity measures. Such discrete distributions, recently investigated by Lijoi, Mena and Prünster (2005), are special cases of generalized gamma processes (Pitman, 2003; Lijoi, Mena and Prünster, 2007b), and their intersection with two-parameter Poisson-Dirichlet models is given by the sole case \((\theta, \alpha) = (0, 1/2)\), which corresponds to a normalized stable process with parameter 1/2 (Kingman, 1975). The class of diffusions studied in this paper is characterized in terms of the second order differential operator

\[
A = \frac{\beta}{s} \frac{\partial}{\partial s} + \frac{1}{2} s \frac{\partial^2}{\partial s^2} + \frac{1}{2} \sum_{i,j=1}^{\infty} z_i (\delta_{ij} - z_j) \frac{\partial^2}{\partial z_i \partial z_j} - \frac{1}{2} \sum_{i=1}^{\infty} \left( \frac{\beta}{s} z_i + \alpha \right) \frac{\partial}{\partial z_i}
\]

acting on a dense sub-algebra of \(C_0([0, \infty) \times \nabla_{\infty})\), the space of continuous functions on \([0, \infty) \times \nabla_{\infty}\) vanishing at infinity, for parameters \((\beta, \alpha)\), with \(\beta = a\tau^\alpha / \alpha\), \(a > 0\), \(\tau > 0\) and \(\alpha = 1/2\). By comparison with (2), it can be seen that the last two terms of (4) describe the time evolution of the frequencies of infinitely-many types. Common features between (2) and (4) are the variance-covariance terms \(z_i (\delta_{ij} - z_j)\) and the structure of the drift or mutation terms \(-[ (\beta/s) z_i + \alpha ]\), with a component acting proportionally to the frequency \(z_i\) and another acting regardless of the species abundance. In Section 2 it will be seen that this role of \(\alpha\) is a distinctive feature of Gibbs-type random probability measures with \(0 < \alpha < 1\). The additional feature with respect to (2) is given by the fact that the positive coefficient \(\theta_t = \beta / S_t\) varies in time, and is driven by what is termed here \(\alpha\)-diversity diffusion, whose operator is given by the first two terms of (4). Equivalently, \(S_t\) follows the stochastic differential equation

\[
dS_t = \frac{\beta}{S_t} dt + \sqrt{S_t} dB_t, \quad S_t \in [0, \infty)
\]

where \(B_t\) is a standard Brownian motion. This can be seen as a particular instance of a continuous-time analog of the notion of \(\alpha\)-diversity, introduced by Pitman (2003) for Poisson-Kingman models, which include Gibbs type partitions. An exchangeable random partition of \(\mathbb{N}\) 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, as \(n \to \infty\),

\[
K_n / n^\alpha \to S \quad \text{a.s.,}
\]

where \(K_n\) is the number of classes of the partition of \(\{1, \ldots, n\}\). The connection between (5) and (6) will become clear in Section 4, where the \(\alpha\)-diversity diffusion will be explicitly derived.

It is to be noted that (2) is not a special case of (4). Indeed the only way of making \(\theta_t = \beta / S_t\) constant is to impose null drift and volatility in (5), which implies \(\theta_t \equiv 0\).
Hence, consistently with the above recalled relation between normalized inverse-Gaussian and Poisson-Dirichlet random measures, (2) and (4) share only the case \((\theta, \alpha) \equiv (0, 1/2)\). Nonetheless, an interesting connection between these classes of diffusions can be stated. In particular it will be shown that performing the same conditioning operation in a pre-limit particle construction of normalized inverse-Gaussian diffusions yields a particular instance of the two-parameter model.

The paper is organized as follows. Section 2 recalls all relevant definitions, among which Gibbs-type partitions, the associated generalized Pólya-urn scheme and random probability measures of generalized gamma and normalized inverse-Gaussian type. Section 3 derives some new results on generalized gamma processes which are crucial for the construction. These are concerned with the convergence of the number of species represented only once in the observed sample and with the second order approximation of the weights of the generalized Pólya-urn scheme associated with normalized inverse-Gaussian processes. In Section 4, by postulating simple population dynamics underlying the time change of the species frequencies, we derive the \(\alpha\)-diversity diffusion for the normalized inverse-Gaussian case, by means of a time-varying analog of (6) with the limit intended in distribution, and highlight its main properties. In Section 5 normalized inverse-Gaussian diffusions are characterized in terms of the operator (4), whose closure is shown to generate a Feller semigroup on \(C_0([0, \infty) \times \nabla_{\infty})\), and the associated family of processes is shown to be the limit in distribution of certain Feller diffusions with finitely-many types. Section 6 provides a discrete representation of normalized inverse-Gaussian diffusions, which are obtained as limits in distribution of certain appropriately transformed Moran-type particle processes which model individuals explicitly, jointly with the varying population heterogeneity. Finally, Section 7 shows that conditioning on the \(\alpha\)-diversity process to be constant, i.e. \(S_t \equiv s\), in a pre-limit version of the particle construction yields, in the limit, the two-parameter model (2) with \((\theta, \alpha) = (s^2/4, 1/2)\).

2 Preliminaries

The Poisson-Dirichlet distribution and its two parameter extension belong to the class of random discrete distributions induced by infinite partitions of Gibbs-type, introduced by Gnedin and Pitman (2005). An exchangeable random partition of the set of natural numbers is said to have Gibbs form if for any \(1 \leq k \leq n\) and any \((n_1, \ldots, n_k)\) such that \(n_j \in \{1, \ldots, n\}\), for \(j = 1, \ldots, k\), and \(\sum_{j=1}^{k} n_j = n\), the law \(\Pi_k^{(n)}\) of the partition \((n_1, \ldots, n_k)\) can be written
as the product

\[ \Pi_k^{(n)}(n_1, \ldots, n_k) = V_{n,k} \prod_{j=1}^{k} (1 - \alpha)n_{j-1}. \]  

Here \( 0 \leq \alpha < 1, \)

\[ (a)_0 = 1, \quad (a)_m = a(a + 1) \cdots (a + m - 1), \quad m > 1, \]

is the Pochhammer symbol and the coefficients \( \{V_{n,k} : k = 1, \ldots, n; n \geq 1\} \) satisfy the recursive equation

\[ V_{n,k} = (n - \alpha k)V_{n+1,k} + V_{n+1,k+1}. \]  

The law of an exchangeable partition is uniquely determined by the function \( \Pi_k^{(n)}(n_1, \ldots, n_k), \) called exchangeable partition probability function, which satisfies certain consistency conditions, which imply invariance under permutations of \( \{1, \ldots, n\} \) and coherent marginalization over the \( (n + 1) \)-th item. Hence the law of a Gibbs partition is uniquely determined by the family \( \{V_{n,k} : k = 1, \ldots, n; n \geq 1\} \). Furthermore, a random discrete probability measure governing a sequence of exchangeable observations is said to be of Gibbs type if it induces a partition which can be expressed as in (7). These have associated predictive distributions which generalize the Blackwell and MacQueen (1973) Pólya-urn scheme to

\[ P\{X_{n+1} \in \cdot | X_1, \ldots, X_n\} = g_0(n, K_n)\nu_0(\cdot) + g_1(n, K_n)\sum_{j=1}^{K_n} (n_j - \alpha) \delta_{X_j^*}(\cdot) \]  

where \( \nu_0 \) is a non atomic probability measure, \( X_1^*, \ldots, X_{K_n}^* \) are the \( K_n \) distinct values observed in \( X_1, \ldots, X_n \) with absolute frequencies \( n_1, \ldots, n_{K_n} \), and the coefficients \( g_0 \) and \( g_1 \) are given by

\[ g_0(n, k) = \frac{V_{n+1,k+1}}{V_{n,k}}, \quad g_1(n, k) = \frac{V_{n+1,k}}{V_{n,k}}, \]

with \( \{V_{n,k} : k = 1, \ldots, n; n \geq 1\} \) as above. It will be of later use to note that integrating both sides of (10) yields

\[ g_0(n, K_n) + (n - \alpha K_n)g_1(n, K_n) = 1, \]

also obtained from (9) and (11). Examples of Gibbs-type random probability measures are the Dirichlet process (Ferguson, 1973), obtained for example from (10) by setting \( \theta > 0 \) and \( \alpha = 0 \) in

\[ g_0(n, k) = \frac{\theta + \alpha k}{\theta + n}, \quad g_1(n, k) = \frac{1}{\theta + n}, \]
the two-parameter Poisson-Dirichlet process (Pitman, 1995; 1996), obtained from (13) with $0 < \alpha < 1$ and $\theta > -\alpha$, the normalized stable process (Kingman, 1975), obtained from (13) with $0 < \alpha < 1$ and $\theta = 0$, the normalized inverse-Gaussian process (Lijoi, Mena and Prünster, 2005), and the normalized generalized gamma process (Pitman, 2003; Lijoi, Mena and Prünster, 2007b). See also Gnedin (2010) for a Gibbs-type model with finitely-many types.

The normalized generalized gamma process is a random probability measure with representation
\begin{equation}
\mu = \sum_{i=1}^{\infty} P_i \delta_{X_i}
\end{equation}
whose weights $\{P_i, i \in \mathbb{N}\}$ are obtained by means of the normalization
\begin{equation}
P_i = \frac{J_i}{\sum_{k=1}^{\infty} J_k}
\end{equation}
where $\{J_i, i \in \mathbb{N}\}$ are the points of a generalized gamma process, introduced by Brix (1999). This is obtained from a Poisson random process on $[0, \infty)$ with mean intensity
\[\lambda(ds) = \frac{1}{\Gamma(1 - \alpha)} \exp(-\tau s) s^{-(1+\alpha)} ds, \quad s \geq 0,\]
with $0 < \alpha < 1$ and $\tau \geq 0$, so that if $N(A)$ is the number of $J_i$’s which fall in $A \in \mathcal{B}([0, \infty))$, then $N(A)$ is Poisson distributed with mean $\lambda(A)$. Lijoi, Mena and Prünster (2007b) showed that a generalized gamma random measure defined via (14)-(15), denoted by $\text{GG}(\beta, \alpha)$, where $\beta = a\tau^\alpha / \alpha$ with $a > 0$ and $\tau > 0$, induces a random partition of Gibbs-type with coefficients $g_0(n, K_n)$ and $g_1(n, K_n)$ in (10) given by
\begin{equation}
\begin{aligned}
g_0(n, k) &= \frac{\alpha \sum_{i=0}^{n} \binom{n}{i} (-1)^i \beta^i / \alpha \Gamma(k + 1 - i / \alpha; \beta)}{\sum_{i=0}^{n-1} \binom{n-1}{i} (-1)^i \beta^i / \alpha \Gamma(k - i / \alpha; \beta)} \\
g_1(n, k) &= \frac{\sum_{i=0}^{n} \binom{n}{i} (-1)^i \beta^i / \alpha \Gamma(k - i / \alpha; \beta)}{\sum_{i=0}^{n-1} \binom{n-1}{i} (-1)^i \beta^i / \alpha \Gamma(k - i / \alpha; \beta)}
\end{aligned}
\end{equation}
where $\Gamma(c; x)$ denotes the upper incomplete gamma function
\begin{equation}
\Gamma(c; x) = \int_{x}^{\infty} s^{c-1} \exp(-s) ds.
\end{equation}
Special cases of a generalized gamma process with parameters \((\beta, \alpha)\) are the Dirichlet process, obtained by letting \(\tau = 1\) and \(\alpha \to 0\), the normalized stable process, obtained by setting \(\beta = 0\), and the normalized inverse-Gaussian process, obtained by setting \(\alpha = 1/2\).

We conclude the section with a brief discussion of the interpretation of \(\alpha\) in the context of species sampling with Gibbs-type partitions. Suppose \(K_n\) different species have been observed in the first \(n\) samples from (10). The probability that a further sample is an already observed species is \(g_1(n, K_n)(n - \alpha K_n)\), but this mass is not allocated proportionally to the current frequencies. The ratio of probabilities assigned to any pair of species \((i, j)\) is

\[
r_{i,j} = \frac{n_i - \alpha}{n_j - \alpha}.
\]

When \(\alpha \to 0\), the probability of sampling species \(i\) is proportional to the absolute frequency \(n_i\). 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. See Lijoi, Mena and Prünster (2007b) for a more detailed treatment of this aspect. Note that this is consistent with the role of \(\alpha\) in (2) and (4) highlighted in the Introduction, where its effect on the drift of each component can be seen as penalizing those frequencies which are relatively low.

## 3 Some results on generalized gamma random measures

In this section we investigate some properties of generalized gamma random measures which will be used in the subsequent constructions. In particular these regard the convergence of the number of species represented only once in the observed sample, and the second order approximation of the weights of the generalized Pólya-urn scheme associated with normalized inverse-Gaussian processes.

Let \(X_1, \ldots, X_n\) be an \(n\)-sized sample drawn from a generalized gamma process with parameters \((\beta, \alpha)\), let \(K_n\) denote the number of distinct species observed in the sample, and let \(N_n := (N_1, \ldots, N_{K_n})\) denote the vector of absolute frequencies associated with each observed species. The probability distribution of the random variable \((K_n, N_n)\), for any \(n \geq 1, k = 1, \ldots, n\) and frequencies \((n_1, \ldots, n_k)\) such that \(\sum_{i=1}^{k} n_i = n\), is provided by Lijoi, Mena and Prünster (2007b) and coincides with

\[
\mathbb{P}(K_n = k, N_n = (n_1, \ldots, n_{K_n})) = \quad (18)
\]
\[
\alpha^{k-1} e^\beta \prod_{j=1}^{k} (1 - \alpha)(n_j - 1) \frac{\Gamma(n)}{\Gamma(n_j)} \sum_{s=0}^{n-1} \binom{n-1}{s} (-1)^s \beta^s / \alpha \Gamma\left(k - \frac{s}{\alpha}; \beta\right)
\]

where \((1 - \alpha)(n_j - 1)\) and \(\Gamma\left(k - \frac{s}{\alpha}; \beta\right)\) are as in (8) and (17) respectively. Denote now by \(M_{j,n}\) the number of species represented \(j\) times in the sample. Then from Equation 1.52 in Pitman (2006) it follows that the distribution of \(M_n := (M_{1,n}, \ldots, M_{n,n})\) is given by

\[
P(M_n = (m_{1,n}, \ldots, m_{n,n})) = n! \alpha^{k-1} e^\beta \prod_{j=1}^{n} \frac{(1 - \alpha)(j-1)^{m_{j,n}}}{j!} \sum_{s=0}^{n-1} \binom{n-1}{s} (-1)^s \beta^s / \alpha \Gamma\left(k - \frac{s}{\alpha}; \beta\right)
\]

for any \(n \geq 1, k = 1, \ldots, n\) and vector \((m_{1,n}, \ldots, m_{n,n}) \in \mathcal{M}_{n,k}\), where

\[
\mathcal{M}_{n,k} = \left\{(m_{1,n}, \ldots, m_{n,n}) : m_{i,n} \geq 1, \sum_{i=1}^{n} m_{i,n} = k, \sum_{i=1}^{n} im_{i,n} = n\right\}.
\]

The following proposition identifies the speed of convergence of the number of species represented once in the sample. Denote by \(\mathcal{C}(n, k, \alpha)\) the generalized factorial coefficient

\[
\mathcal{C}(n, k, \alpha) = \frac{1}{k!} \sum_{j=0}^{k} (-1)^j \binom{k}{j} (-j\alpha)_n
\]

where \(\mathcal{C}(0, 0, \alpha) = 1\) and \(\mathcal{C}(n, 0, \alpha) = 0\). See Charalambides (2005), Chapter 2, for a complete account.

**Proposition 3.1.** Under the normalized generalized gamma process with parameters \((\beta, \alpha)\), one has

\[
P(M_{1,n} = m_{1,n}) = \alpha^{m_{1,n}-1} e^\beta \frac{n!}{\Gamma(n)m_{1,n}!} \sum_{s=0}^{n-1} \binom{n-1}{s} (-1)^s \beta^s / \alpha
\]

\[
\times \sum_{j=0}^{n-m_{1,n}} (-\alpha)^j \frac{(n - m_{1,n} - j + 1)(m_{1,n} + j)}{j!}
\]

\[
\times \sum_{k=0}^{n-m_{1,n} - j} \mathcal{C}(n - m_{1,n} - j, k, \alpha) \Gamma\left(k + m_{1,n} + j - \frac{s}{\alpha}; \beta\right).
\]

Moreover,

\[
\frac{M_{1,n}}{n^\alpha} \rightarrow \alpha S_\alpha \quad a.s.,
\]
where \( S_\alpha \) is a strictly positive and almost surely finite random variable with density function

\[
g_{S_\alpha}(s; \alpha, \beta) = e^{\beta - (\beta/s)^{1/\alpha}} f(s^{-1/\alpha}; \alpha) s^{1+1/\alpha}
\]

with \( f(\cdot; \alpha) \) being the density of a positive stable random variable with parameter \( \alpha \).

**Proof.** Denote \((x)_{[m]} = x(x-1) \ldots (x-m+1)\). From (19), for any \( r \geq 1 \) one has

\[
E[(M_{1,n})[r]] = 
\sum_{k=1}^{n} \sum_{M_{n,k}} n! \frac{\alpha^{k-1} e^{\beta}}{\Gamma(n)} \prod_{j=1}^{n} \left( \frac{(1-\alpha)(j-1)}{j!} \right) \frac{1}{m_{j,n}!} m_{j,n}[r] 
\times \sum_{s=0}^{n-1} \binom{n-1}{s} (-1)^s \frac{\beta^{s/\alpha}}{\Gamma(k-s/\alpha) \Gamma(k-\frac{s}{\alpha})} \beta^{s/\alpha} \Gamma(k-\frac{s}{\alpha}).
\]

In particular, by using the definition of generalized factorial coefficient in terms of sum over the set of partitions \( \mathcal{M}_{n,k} \) (see Charalambides (2005), eq. 2.62) we have

\[
\sum_{M_{n-r,k-r}} \prod_{j=1}^{n} \left( \frac{(1-\alpha)(j-1)}{j!} \right) \frac{1}{m_{j,n}!} m_{j,n} = \frac{(n)_{[r]}}{n! \alpha^{k-r}} \theta(n-r, k-r, \alpha).
\]

Therefore, we obtain

\[
E[(M_{1,n})[r]] = \sum_{k=1}^{n} \frac{\alpha^{r-1} (n)_{[r]} e^{\beta}}{\Gamma(n)} \theta(n-r, k-r, \alpha)
\times \sum_{s=0}^{n-1} \binom{n-1}{s} (-1)^s \frac{\beta^{s/\alpha}}{\Gamma(k-s/\alpha) \Gamma(k-\frac{s}{\alpha})} \beta^{s/\alpha} \Gamma(k-\frac{s}{\alpha}).
\]
In order to obtain the distribution of the random variable $M_{1,n}$ we can make use of the probability generating function of $M_{1,n}$, denoted $G_{(M_{1,n})}(t)$. From (23) we have

$$G_{(M_{1,n})}(t) = \sum_{r=0}^{\infty} \frac{\alpha^{r-1} e^{\beta}(n)[r]}{\Gamma(n)} \sum_{s=0}^{n-1} \left( \frac{n-1}{s} \right) (-1)^s \frac{\beta^s}{\alpha^n} \times \sum_{k=0}^{n} \mathcal{C}(n-r,k,\alpha) \frac{(t-1)^r}{r!}.$$ 

Therefore, the distribution of $M_{1,n}$ is given by

$$\mathbb{P}(M_{1,n} = m_{1,n}) = \frac{1}{m_{1,n}!} \sum_{j=0}^{\infty} \frac{\alpha^{m_{1,n}+j-1} e^{\beta}(n)[m_{1,n}+j]}{\Gamma(n)} \sum_{s=0}^{n-1} \left( \frac{n-1}{s} \right) (-1)^s \frac{\beta^s}{\alpha^n} \times \sum_{k=0}^{n} \mathcal{C}(n-m_{1,n}-j,k,\alpha) \frac{(t-1)^r}{r!}.$$ 

where the last identity is due to the fact that $\mathcal{C}(n,k,\alpha) = 0$ for any $k > n$. Proposition 3 in Lijoi, Mena and Prünster (2007b) shows that

$$K_n/n^\alpha \rightarrow S_\alpha$$

almost surely, where $S_\alpha$ is an almost surely positive and finite random variable with density function

$$g_{S_\alpha}(s;\alpha,\beta) = e^{\beta-(\beta/s)^{1/\alpha}} f(s^{-1/\alpha};\alpha)$$

with $f(\cdot;\alpha)$ being the density function of a positive stable random variable with parameter $\alpha$. In other terms, according to Definition 3.10 in Pitman (2006), an exchangeable partition.
of \( \mathbb{N} \) having EPPF (18) has \( \alpha \)-diversity \( S_\alpha \). A simple application of Lemma 3.11 in Pitman (2006) leads to (22).

A second aspect of generalized gamma random measures we need to address for later use is the approximate behavior of the coefficients in the generalized Pólya urn (10). It is well known that the first order behavior of (16) is that of a normalized stable process, i.e.

\[
g_0(n,k) \approx \alpha k/n \quad g_1(n,k) \approx 1/n,
\]

also implied by the next result. However it turns out that for the definition of the diffusion processes which are the object of the next two sections, it is crucial to know the second order approximation. The following proposition, whose proof is deferred to the Appendix, identifies such behaviour for the normalized inverse-Gaussian case \( \alpha = 1/2 \).

**Proposition 3.2.** Let \( g_0(n,k) \) and \( g_1(n,k) \) be as in (16). When \( \alpha = 1/2 \),

\[
g_0(n,k) = \frac{\alpha k}{n} + \frac{\beta/s_n}{n} + o(n^{-1})
\]

and

\[
g_1(n,k) = \frac{1}{n} - \frac{\beta/s_n}{n^2} + o(n^{-2})
\]

where \( s_n = k/n^\alpha \) and \( \beta = a\tau^\alpha/\alpha \).

### 4 Alpha-diversity processes

Making use of the results of the previous section, here we construct a one dimensional diffusion process which can be seen as a dynamic version of the notion of \( \alpha \)-diversity, recalled in (6), relative to the case of normalized inverse-Gaussian random probability measures. Such diffusion, which will be crucial for the construction of Section 5, is obtained as weak limit of an appropriately rescaled random walk on the integers, whose dynamics are driven by an underlying population process. This is briefly outlined here and will be formalized in Section 6. Consider \( n \) particles, denoted \( x^{(n)} = (x_1, \ldots, x_n) \) with \( x_i \in X \) for each \( i \), where \( X \) is a Polish space, and denote by \( K_n = K_n(x^{(n)}) \) the number of distinct values observed in \( (x_1, \ldots, x_n) \). Let the vector \( (x_1, \ldots, x_n) \) be updated at discrete times by replacing a uniformly chosen coordinate. Conditionally on \( K_n(x^{(n)}) = k \), the incoming particle will be a copy of one still in the vector, after the removal, with probability \( g_1(n - 1, k_r) \), and will be a new value with probability \( g_0(n - 1, k_r) \), where \( g_1(n - 1, k) \) and \( g_0(n - 1, k) \) are as in (16) and \( k_r \) is the value of
k after the removal. Denote by \( \{K_n(m), m \in \mathbb{N}_0\} \) the chain which keeps track of the number of distinct types in \((x_1, \ldots, x_n)\). Then, letting \( m_{1,n} \) be the number of clusters of size one in \((x_1, \ldots, x_n)\), the transition probabilities for \( K_n(m) \)

\[
p(k, k') = \mathbb{P}\{K_n(m + 1) = k' | K_n(m) = k\}
\]

can be written

\[
p(k, k') = \begin{cases} 
    (1 - \frac{m_{1,n}}{n}) g_0(n - 1, k) & \text{if } k' = k + 1 \\
    \frac{m_{1,n}}{n} g_1(n - 1, k - 1)(n - 1 - \alpha(k - 1)) & \text{if } k' = k - 1 \\
    1 - p(k, k + 1) - p(k, k - 1) & \text{if } k' = k \\
    0 & \text{else}
\end{cases}
\]

for \(1 \leq k \leq n\). That is, with probability \( m_{1,n}/n \) a cluster of size one is selected and removed, with probability \( g_0(n - 1, k) \) a new species appears and with probability \( g_1(n - 1, k)(n - 1 - \alpha(k - 1)) \) a survivor has an offspring. Note that \( k = 1 \) and \( k = n \) imply \( p(1, 0) = 0 \) and \( p(n, n + 1) = 0 \) respectively, since \( m_{1,n} \) equals 0 and \( n \) respectively, so that 1 and \( n \) act as barriers.

The following theorem finds the conditions under which the rescaled chain \( K_n(m)/n^\alpha \) converges to a diffusion process on \([0, \infty)\). Here we provide a sketch of the proof with the aim of favoring the intuition. The formalization of the result is contained in the proof of Theorem 6.1, while that of the fact that the limiting diffusion is well-defined, i.e. the corresponding operator generates a Feller semigroup on an appropriate subspace of \(C([0, \infty))\), is provided in Corollary 4.3 below.

Throughout the paper \( C_B(A) \) denotes the space of continuous functions from \( A \) to \( B \), while \( X_n \Rightarrow X \) denotes convergence in distribution.

**Theorem 4.1.** Let \( \{K_n(m), m \in \mathbb{N}_0\} \) be a Markov chain with transition probabilities as in (27) determined by a generalized gamma process with \( \beta \geq 0 \) and \( \alpha = 1/2 \), and define \( \{\tilde{K}_n(t), t \geq 0\} \) to be such that \( \tilde{K}_n(t) = K_n([n^{3/2}t]) / n^\alpha \). Let also \( \{S_t, t \geq 0\} \) be a diffusion process driven by the stochastic differential equation

\[
dS_t = \frac{\beta}{S_t} dt + \sqrt{S_t} dB_t, \quad S_t \geq 0,
\]

where \( B_t \) is a standard Brownian motion. If \( \tilde{K}_n(0) \Rightarrow S_0 \), then

\[
\{\tilde{K}_n(t), t \geq 0\} \Rightarrow \{S_t, t \geq 0\} \quad \text{in } C([0, \infty))([0, \infty)) \quad \text{as } n \to \infty.
\]
Proof. Let \( \alpha = 1/2 \). Recalling from (22) that \( m_{1,n} \approx \alpha s n^\alpha \approx \alpha k \), from Proposition 3.2 we can write (27) as follows (for ease of presentation we use \( n \) and \( k \) in place of \( n - 1 \) and \( k - 1 \) since it is asymptotically equivalent):

\[
p(k, k') = \begin{cases} 
1 - \frac{\alpha k}{n} \left( \frac{\alpha k}{n} + \frac{\beta/s_n}{n^2} \right) + o(n^{-1}) & \text{if } k' = k + 1 \\
\frac{\alpha k}{n} \left( \frac{1}{n} - \frac{\beta/s_n}{n^2} \right) (n - \alpha k) + o(n^{-3/2}) & \text{if } k' = k - 1 \\
1 - p(k, k + 1) - p(k, k - 1) + o(n^{-1}) & \text{if } k' = k \\
0 & \text{else}
\end{cases}
\]

when \( 1 < k < n \) and accordingly for the barriers 1 and \( n \). The conditional expected increment of the process \( \{K_n(m)/n^\alpha, m \in \mathbb{N}_0\} \) is

\[
\mathbb{E}\left( \frac{k'}{n^\alpha} - \frac{k}{n^\alpha} \middle| k \right) = \frac{1}{n^\alpha} \left[ \left( 1 - \frac{\alpha k}{n} \right) \left( \frac{\alpha k}{n} + \frac{\beta/s_n}{n^2} \right) - \frac{\alpha k}{n} \left( \frac{1}{n} - \frac{\beta/s_n}{n^2} \right) (n - \alpha k) \right] + o\left( \frac{1}{n^{1+\alpha}} \right) = \frac{\beta/s_n}{n^{1+\alpha}} + o\left( \frac{1}{n^{1+\alpha}} \right).
\]

Similarly, the conditional second moment of the increment is

\[
\mathbb{E}\left[ \left( \frac{k'}{n^\alpha} - \frac{k}{n^\alpha} \right)^2 \middle| k \right] = \frac{1}{n^{2\alpha}} \left[ \left( 1 - \frac{\alpha k}{n} \right) \left( \frac{\alpha k}{n} + \frac{\beta/s_n}{n^2} \right) + \frac{\alpha k}{n} \left( \frac{1}{n} - \frac{\beta/s_n}{n^2} \right) (n - \alpha k) \right] + o\left( \frac{1}{n^{1+2\alpha}} \right) = \frac{2\alpha k}{n^{1+2\alpha}} + o\left( \frac{1}{n^{1+2\alpha}} \right).
\]

Since \( k \approx sn^\alpha \), and recalling that \( s_n \to s \) almost surely, we have

\[n^{1+\alpha} \mathbb{E}\left( \frac{k'}{n^\alpha} - \frac{k}{n^\alpha} \middle| k \right) \to \beta/s\]

and

\[n^{1+\alpha} \mathbb{E}\left[ \left( \frac{k'}{n^\alpha} - \frac{k}{n^\alpha} \right)^2 \middle| k \right] \to 2\alpha s.
\]

It is easy to check that all conditional \( m \)-th moments of \( \Delta k/n^\alpha \) converge to zero for \( m \geq 3 \), whence it follows by standard theory (cf., e.g., Karlin and Taylor (1981)) that, as \( n \to \infty \), the process \( \tilde{K}_n(t) = K_n((n^{3/2}t))/n^\alpha \) converges in distribution to a diffusion process \( S_t \) on \([0, \infty)\) with drift \( \beta/S_t \) and diffusion coefficient \( \sqrt{2\alpha S_t} \). \( \square \)
As anticipated, the second order approximation of $g_0(n, k)$ is crucial for establishing the drift of the limiting diffusion, as the first order terms cancel. It is interesting to note that when $\beta = 0$, which yields the normalized stable case, the limiting diffusion reduces to the diffusion approximation of a critical Galton-Watson branching process, also known as the zero-drift Feller diffusion. See, e.g., Ethier and Kurtz (1986), Theorem 9.1.3. This also holds approximately for high values of $S_t$, in which case the drift becomes negligible.

In order to have some heuristics on the behaviour of the $\alpha$-diversity process, Figure 1 shows $3 \times 10^5$ steps of the random walk $\{K_n(m)/n^\alpha, m \in \mathbb{N}_0\}$ with dynamics as in Theorem 4.1, starting from $1/\sqrt{n}$ with $n = 200$. The three paths correspond to $\beta$ being equal to 0, 100 and 1000. It is apparent how $\beta$ influences the dynamic clustering structure in the population.

It is well known that when $\beta = 0$, the point 0 is an absorbing boundary for $S_t$. The next result provides the boundary classification, using Feller’s terminology, for the case $\beta > 0$.

**Proposition 4.2.** Let $S_t$ be as in Theorem 4.1 with $\beta > 0$. Then the points 0 and $\infty$ are respectively an entrance and a natural boundary.

**Proof.** The scale function for the process, defined as

$$
S(x) = \int_{x_0}^x s(y)dy, \quad 0 < x < \infty,
$$

where

$$
s(y) = \exp \left\{-\int_{y_0}^y \frac{2\mu(t)}{\sigma^2(t)} dt\right\}
$$

and $\mu(x)$ and $\sigma^2(x)$ denote drift and diffusion, equals

$$
S(x) = \int_{x_0}^x \exp \left\{-2\beta \left( \frac{1}{y_0} - \frac{1}{y} \right) \right\} dy
$$

$$
= e^{-2\beta/y_0} \left[ xe^{2\beta/x} - x_0e^{2\beta/x_0} - 2\beta \text{Ei}(2\beta/x) + \text{Ei}(2\beta/x_0) \right]
$$

where $\text{Ei}(z)$ is the exponential integral

$$
\text{Ei}(z) = -\int_{-z}^\infty t^{-1}e^{-t}dt.
$$

Letting $S[a, b] = S(b) - S(a)$, for $0 < a < b < \infty$, we have

$$
S(0, b) = \lim_{a \to 0} S[a, b] = \infty,
$$

(33)

$$
S[a, \infty) = \lim_{b \to \infty} S[a, b] = \infty.
$$
Figure 1: Three sample paths of the random walk \( \{K_n(m)/n^\alpha, m \in N_0\} \) with dynamics as in Theorem 4.1, starting from \( 1/\sqrt{n} \) with \( n = 200 \), for parameter values: (a) \( \beta = 0 \), (b) \( \beta = 100 \), (c) \( \beta = 1000 \). The figures show how \( \beta \) influences the dynamic clustering structure in the population.
Moreover the speed measure is given by

\[ M[c, d] = \int_c^d [\sigma^2(t)s(t)]^{-1} dt = e^{2\beta/y_0} [\text{Ei}(-2\beta/c) - \text{Ei}(-2\beta/d)] \]

from which \( M(0, d] = \lim_{c \downarrow 0} M[c, d] < \infty \) and

(34) \[ M[c, \infty) = \lim_{d \uparrow \infty} M[c, d] = \infty. \]

Now (33) implies that

\[ \Sigma(0) = \lim_{l \downarrow 0} \int_l^x S(l, y)dM(y) = \infty, \]

\[ \Sigma(\infty) = \lim_{l \uparrow \infty} \int_x^r S[y, r)dM(y) = \infty \]

and (34) implies

\[ N(\infty) = \lim_{r \uparrow \infty} \int_x^r S[x, y]dM(y) = \infty, \]

while

\[ N(0) = \lim_{l \downarrow 0} \int_l^x S[y, x]dM(y) = \lim_{l \downarrow 0} \int_l^x y^{-2\beta/y} e^{2\beta/y} - ye^{2\beta/y} + 2\beta \left( \text{Ei}\left( \frac{2\beta}{y} \right) - \text{Ei}\left( \frac{2\beta}{x} \right) \right) dy < \infty \]

since

\[ \lim_{y \downarrow 0} y^{-2\beta/y} e^{2\beta/y} \text{Ei}\left( \frac{2\beta}{y} \right) < \infty. \]

The theorem now follows from, e.g., Karlin and Taylor (1981), Section 15.6.

Hence when \( \beta > 0 \) neither boundary point is attainable from the interior of the state space, from which the actual state space is \([0, \infty)\) for \( \{S_t, t \geq 0\} \) and \((0, \infty)\) for \( \{S_t, t > 0\} \). The process can be made to start at 0, in which case it instantly moves towards the interior of the state space and never comes back. Consequently, we will use \((0, \infty)\) or \([0, \infty)\) as state space at convenience, with the agreement that \((0, \infty)\) is referred to \( \{S_t, t > 0\} \).

As a corollary, we formalize the well-definedness of the \( \alpha \)-diversity diffusion. Denote by \( C_0(K) \) the space of continuous functions vanishing at infinity on a locally compact set \( K \), and let \( \| \cdot \| \) be a norm which makes \( C_0(K) \) a Banach space. Recall that a Feller semigroup
on $C_0(K)$ is a one-parameter family of bounded linear operators $\{T(t), t \geq 0\}$ on $C_0(K)$ such that $T(t)$ has the semigroup property $T(s + t) = T(s)T(t)$ for all $s, t \geq 0$, is strongly continuous, i.e.

$$\|T(t)f - f\| \to 0 \quad \text{as} \quad t \to 0, \quad f \in C_0(K),$$

and, for all $t \geq 0$, $T(t)$ is a contraction, i.e. $\|T(t)\| \leq 1$, is conservative in the sense that $T(t)1 = 1$, and is positive in the sense that it preserves the cone of non negative functions.

**Corollary 4.3.** For $\beta \geq 0$, let $A_0$ be the second order differential operator

$$A_0 = \frac{\beta}{s} \frac{d}{ds} + \frac{1}{2s} \frac{d^2}{ds^2}$$

and define

$$\mathcal{D}(A_0) = \{f \in C_0([0, \infty)) \cap C^2((0, \infty)) : A_0f \in C_0([0, \infty))\}.$$ 

Then $\{(f, A_0f) : f \in \mathcal{D}(A_0)\}$ generates a Feller semigroup on $C_0([0, \infty))$.

**Proof.** The result follows from Proposition 4.2 together with Corollary 8.1.2 in Ethier and Kurtz (1986). \qed

An immediate question that arises is whether the $\alpha$-diversity diffusion is stationary. The following proposition, which concludes the section, provides a negative answer.

**Proposition 4.4.** Let $\{S_t, t \geq 0\}$ be as in Theorem 4.1. Then there exists no stationary density for the process.

**Proof.** A stationary density, if it exists, is given by

$$\psi(x) = m(x)[C_1S(x) + C_2], \quad x \geq 0,$$

where $m(x) = [s(x)\sigma^2(x)]^{-1}$, $s(x)$ and $S(x)$ are as in (32), and $C_1, C_2$ are constants determined in order to guarantee the non negativity and integrability to one of $\psi$. Here $s(x) = e^{-2\beta/x}$ and

$$S(x) = xe^{2\beta/x} - 2\beta \text{Ei}(2\beta/x)$$

so that

$$\psi(x) = C_1 - 2\beta C_1 x^{-1} e^{-2\beta/x} \text{Ei}(2\beta/x) + C_2 x^{-1} e^{-2\beta/x}.$$

The second term is not integrable in a neighborhood of infinity, since there exists an $x_0 > 0$ such that

$$-x^{-1} e^{-2\beta/x} \text{Ei}(2\beta/x) > x^{-1}, \quad \text{for all} \quad x > x_0,$$

hence $C_1$ must be zero, neither is the third, giving the result. \qed
5 Normalized inverse-Gaussian diffusions

The $\alpha$-diversity process constructed in the previous section is a key component in the definition of the class of normalized inverse-Gaussian diffusions. In this section we characterize such infinite-dimensional processes in terms of their infinitesimal generator, and show that they can be obtained as limit in distribution of a certain sequence of Feller diffusions with finitely-many types. The association of the limit family with the class of normalized inverse-Gaussian random probability measures will instead be shown in Section 6.

Consider the $(n-1)$-dimensional simplex

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

and the closed subspace of $\Delta_n$ given by

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

so that $\varepsilon_n \leq z_i \leq 1 - (n-1)\varepsilon_n$ for $z_i \in \tilde{\Delta}_n$, where $\{\varepsilon_n\} \subset \mathbb{R}_+$ is a non increasing sequence such that

$$(37) \quad 0 < \varepsilon_n < \frac{1}{n} \quad \forall n \geq 2, \quad n\varepsilon_n \downarrow 0.$$  

Define, for $(z_0,z_1,\ldots,z_n) \in (0,\infty) \times \tilde{\Delta}_n$, the differential operator

$$A_n = \frac{1}{2} \sum_{i,j=0}^n a_{ij}^{(n)}(z) \frac{\partial^2}{\partial z_i \partial z_j} + \frac{1}{2} \sum_{i=0}^n b_i^{(n)}(z) \frac{\partial}{\partial z_i}$$

where the covariance components $(a_{ij}^{(n)}(z))_{i,j=0,\ldots,n}$ are set to be

$$a_{ij}^{(n)}(z) = \begin{cases} z_0 & i = j = 0 \\ (z_i - \varepsilon_n)(\delta_{ij}(1-n\varepsilon_n) - (z_j - \varepsilon_n)) & 1 \leq i,j \leq n \\ 0 & \text{else,} \end{cases}$$

and, for $\beta > 0$, the drift components are

$$b_0^{(n)}(z) = \frac{\beta}{z_0},$$

$$b_i^{(n)}(z) = \frac{\beta}{z_0(n-1)}(1 - z_i) - \frac{\beta}{z_0} z_i - \alpha \left( 1 - \exp \left\{ -(z_i - \varepsilon_n) e^{1/\varepsilon_n} \right\} \right),$$

$$i = 1,\ldots,n.$$
Observe that \( a_{ij}^{(n)}(z) \), for \( 1 \leq i, j \leq n \), can be seen as a Wright-Fisher type covariance restricted to \([\varepsilon_n, 1 - (n - 1)\varepsilon_n]^n\), since for such indices \( i, j \)

\[
a_{ij}^{(n)}(z) = \begin{cases} (z_i - \varepsilon_n)(1 - (n - 1)\varepsilon_n - z_i) & i = j \\ -(z_i - \varepsilon_n)(z_j - \varepsilon_n) & i \neq j, \end{cases}
\]

(38)

and that the first two terms in \( b_i^{(n)}(z) \), \( i = 1, \ldots, n \), equal

\[
\frac{\beta}{z_0(n-1)}(1 - (n - 1)\varepsilon_n - z_i) - \frac{\beta}{z_0}(z_i - \varepsilon_n)
\]

from which the behavior at the boundary is clear. For ease of exposition and in analogy with the previous section we will at convenience denote \( z_0 \) by \( s \), so that \( A_n \) can be written more explicitly

\[
A_n = \frac{\beta}{s} \frac{\partial}{\partial s} + \frac{1}{2} s \frac{\partial^2}{\partial s^2} + \frac{1}{2} \sum_{i,j=1}^{n} a_{ij}^{(n)}(z) \frac{\partial^2}{\partial z_i \partial z_j} + \frac{1}{2} \sum_{i=1}^{n} b_i^{(n)}(z) \frac{\partial}{\partial z_i}
\]

(39)

with \( a_{ij}^{(n)}(z) \) and \( b_i^{(n)}(z) \) as above. The domain of \( A_n \) is taken to be

\[
\mathcal{D}(A_n) = \{ f : f = f_0 \times f_1, f_0 \in \mathcal{D}(A_0), f_1 \in C^2(\tilde{\Delta}_n) \},
\]

(40)

where \((f_0 \times f_1)(s, z) = f_0(s)f_1(z)\), \( \mathcal{D}(A_0) \) is (36), and

\[
C^2(\tilde{\Delta}_n) = \left\{ f \in C(\tilde{\Delta}_n) : \exists \tilde{f} \in C^2(\mathbb{R}^n), \tilde{f}|_{\tilde{\Delta}_n} = f \right\}.
\]

The operator \( A_n \) drives \( n + 1 \) components: those labeled from 1 to \( n \) can be seen as the frequencies associated to \( n \) species in a large population, bounded from below by \( \varepsilon_n \); the \( z_0 \) or \( s \) component is a positive real variable which evolves independently according to the \( \alpha \)-diversity diffusion (28) and contributes to drive the drift of the other \( n \) components.

Denote by \( C_0([0, \infty) \times \tilde{\Delta}_n) \) the Banach space of continuous functions on \([0, \infty) \times \tilde{\Delta}_n\) which vanish at infinity, equipped with the supremum norm \( ||f|| = \sup_{x \in [0, \infty) \times \tilde{\Delta}_n} f(x) \), and by \( \mathcal{P}(B) \) the set of Borel probability measures on \( B \). Recall that a Markov process \( \{X(t), t \geq 0\} \), taking values in a metric space \( E \), is said to correspond to a semigroup \( \{T(t)\} \), acting on a closed subspace \( L \) of the space of bounded functions on \( E \), if

\[
\mathbb{E}[f(X(t+s))|\mathcal{F}_t^X] = T(s)f(X(t)), \quad s, t \geq 0,
\]

for every \( f \in L \), where \( \mathcal{F}_t^X = \sigma(X(u), u \leq t) \).
Proposition 5.1. Let $\mathcal{A}_n$ be the operator defined in (39) and (40). The closure in $C_0([0, \infty) \times \tilde{\Delta}_n)$ of $\mathcal{A}_n$ generates a strongly continuous, positive, conservative, contraction semigroup $\{\mathcal{S}_n(t)\}$ on $C_0([0, \infty) \times \tilde{\Delta}_n)$. For every $\nu_n \in \mathcal{P}([0, \infty) \times \tilde{\Delta}_n)$ there exists a strong Markov process $Z^{(n)}(\cdot) = \{Z^{(n)}(t), t \geq 0\}$ corresponding to $\{\mathcal{S}_n(t)\}$ with initial distribution $\nu_n$ and sample paths in $C([0, \infty) \times \tilde{\Delta}_n)$ with probability one.

Proof. We proceed by verifying the hypothesis of the Hille-Yosida Theorem. Note first that $\mathcal{A}_n$ satisfies the positive maximum principle, that is for $f \in \mathcal{D}(\mathcal{A}_n)$ and $(s_*, z_*) \in (0, \infty) \times \tilde{\Delta}_n$ such that $\|f\| = f(s_*, z_*) \geq 0$ we have $\mathcal{A}_n f(s_*, z_*) \leq 0$. Indeed, writing $\mathcal{A}_n = \mathcal{A}_0 + \mathcal{A}_{n,1}$ to indicate the first two and last two terms in (39), it is immediate to check that $\mathcal{A}_0$ and $\mathcal{A}_{n,1}$ satisfy the positive maximum principle on $[0, \infty)$ and $\tilde{\Delta}_n$ respectively. If $f_0(s_*) \geq 0, f_1(z_*) \geq 0$, then $\mathcal{A}_0 f_0(s) \leq 0$ and $\mathcal{A}_{n,1} f_1(z) \leq 0$, while if $f_0(s_*) \leq 0, f_1(z_*) \leq 0$, then $\mathcal{A}_0 f_0(s_*) \geq 0$ and $\mathcal{A}_{n,1} f_1(z_*) \geq 0$. In both cases

$$\mathcal{A}_n f(s_*, z_*) = f_1(z_*) \mathcal{A}_0 f_0(s_*) + f_0(s_*) \mathcal{A}_{n,1} f_1(z_*) \leq 0.$$ 

Let now $L \subset \mathcal{D}(\mathcal{A}_n)$ be the algebra generated by functions $f = f_0 \times f_1$, with $f_0 \in \mathcal{D}(\mathcal{A}_0)$, $\mathcal{D}(\mathcal{A}_0)$ as in (36) and $f_1 = z^c = z_1^c \cdots z_n^c \in C^2(\tilde{\Delta}_n)$, $c_i \in \mathbb{N}_0$, so that $L$ is dense in $C_0([0, \infty) \times \tilde{\Delta}_n)$, and so is $\mathcal{D}(\mathcal{A}_n)$. Denoting $c + d_{(i)} = (c_0, \ldots, c_i + d, \ldots, c_n)$, for $f \in L$ we have

$$\mathcal{A}_n (f_0(s) \times z^c) = f_1(z) \mathcal{A}_0 f_0(s) + \frac{f_0(z_0)}{2} \left\{ \sum_{i=1}^{n} c_i (c_i - 1) \left( - z^c + (1 - (n - 2)\varepsilon_n)z^{c-1(i)} \right) \right.$$ 

$$- \varepsilon_n (1 - (n - 1)\varepsilon_n)z^{c-2(i)} \right.$$ 

$$+ \sum_{j \neq i} c_i c_j \left[ - z^c + \varepsilon_n (z^{c-1(j)} + z^{c-1(i)} - \varepsilon_n z^{c-1(i)-1(j)}) \right] \right.$$ 

$$+ \sum_{i=1}^{n} c_i \left[ \frac{\beta}{z_0(n - 1)} \left( z^{c-1(i)} - nz^c \right) \right.$$ 

$$- \alpha z^{c-1(i)} + \alpha e^{\varepsilon_n \exp(1/\varepsilon_n)} e^{-\exp(1/\varepsilon_n)z_i z^{c-\delta_i}} \right\},$$

so that the image of $\mathcal{A}_n$ contains functions of type $f_0 \times z^c$ and $f_0 \times e^{-b_0 z_i} z^c$, with $b_0$ fixed. 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 $\|e^{-b_0 x} p^{(k)} - g\| \to 0$. It follows that the image of $\mathcal{A}_n$ is dense in $C_0([0, \infty) \times \tilde{\Delta}_n)$, and so is that of $\lambda - \mathcal{A}_n$ for all but at most countably many $\lambda > 0$. The first assertion now follows from Theorem 4.2.2 of Ethier.
and Kurtz (1986) and by noting that $1 \in \mathcal{D}(\mathcal{A}_n)$ and $\mathcal{A}_n 1 = 0$, that is $\mathcal{A}_n$ is conservative. The second assertion with $D_{[0,\infty) \times \Delta_n}([0,\infty))$, the space of right-continuous functions with left limits, in place of $C_{[0,\infty) \times \Delta_n}([0,\infty))$, follows from Theorem 4.2.7 of Ethier and Kurtz (1986). To prove the almost sure continuity of sample paths it is enough to show that for every $z^* \in (0,\infty) \times \Delta_n$ and $\epsilon > 0$ there exists a function $f \in \mathcal{D}(\overline{\mathcal{A}}_n)$ such that

$$f(z^*) = ||f||, \quad \sup_{z \in B(z^*,\epsilon)^c} f(z) < f(z^*), \quad \overline{\mathcal{A}}_n f(z^*) = 0,$$

where $B(z^*,\epsilon)^c$ is the complement of an $\epsilon$-neighborhood of $z^*$ in the topology of coordinatewise convergence (cf. Ethier and Kurtz (1986), Remark 4.2.10). This can be done by means of a function $f \in \mathcal{D}(\mathcal{A}_n)$ which is flat in $z^*$ and rapidly decreasing away from $z^*$, e.g. of type $f(z) = c_1 - c_2 \sum_{i=0}^{n}(z_i - z_i^*)^4$ for appropriate constants $c_1, c_2$.

For $Z^{(n)}(\cdot)$ as in Proposition 5.1, consider now the mapping $\rho_n(Z^{(n)}(\cdot))$, where $\rho_n : [0,\infty) \times \Delta_n \rightarrow [0,\infty) \times \nabla_\infty$ is defined as

$$(42) \quad \rho_n(z) = (z_0, z_1, z_2, \ldots, z_n, 0, 0, \ldots),$$

$(z_1, \ldots, z_n)$ is the vector of decreasingly ordered statistics of $(z_1, \ldots, z_n) \in \Delta_n$, and $\nabla_\infty$ is the closure of the infinite-dimensional ordered simplex, defined in (1). The following Proposition states that $\rho_n(Z^{(n)}(\cdot))$ is still a well-defined Markov process. Define

$$\hat{\nabla}_n = \left\{ z \in \nabla_\infty : z_n \geq \varepsilon_n > z_{n+1} = 0, \sum_{i=1}^{n} z_i = 1 \right\}$$

and observe that $z \in \nabla_\infty$ satisfies $z_i \leq 1/i$ for all $i$, so that $\hat{\nabla}_n$ is non empty by (37).

**Proposition 5.2.** Let $\mathcal{A}_n$ be defined by the right hand side of (39), with domain

$$\mathcal{D}(\mathcal{A}_n) = \{ f : f = f_0 \times f_1, f_0 \in \mathcal{D}(\mathcal{A}_0), f_1 \in C^2_{\rho_n}(\hat{\nabla}_n) \}$$

where $\mathcal{D}(\mathcal{A}_0)$ is as in (36) and

$$C^2_{\rho_n}(\hat{\nabla}_n) = \{ f \in C(\hat{\nabla}_n) : f \circ \rho_n \in C^2(\Delta_n) \}.$$

Then the closure of $\mathcal{A}_n$ in $C_0([0,\infty) \times \hat{\nabla}_n)$ generates a strongly continuous, positive, conservative, contraction semigroup $\{ T_n(t) \}$ on $C_0([0,\infty) \times \hat{\nabla}_n)$. For every $\nu_n \in \mathcal{P}([0,\infty) \times \Delta_n)$, let $Z^{(n)}$ be as in Proposition 5.1. Then $\rho_n(Z^{(n)}(\cdot))$ is a strong Markov process corresponding to $\{ T_n(t) \}$ with initial distribution $\nu_n \circ \rho_n^{-1}$ and sample paths in $C([0,\infty) \times \hat{\nabla}_n([0,\infty))$ with probability one.

**Proof.** The result follows from Proposition 2.4 in Ethier and Kurtz (1981), with $\Delta_n, \nabla_n$ and $\rho_n$ there substituted by $[0,\infty) \times \Delta_n, [0,\infty) \times \hat{\nabla}_n$ and (42) respectively. \qed
We now turn the attention to the limit of $\rho_n(Z^{(n)}(\cdot))$ when the number of types goes to infinity. To this end, consider that $\mathcal{N}_{\infty}$ is a compact and metrizable metric space in the topology of coordinatewise convergence, and let $C_0([0, \infty) \times \mathcal{N}_{\infty})$ be the Banach space of continuous functions on $[0, \infty) \times \mathcal{N}_{\infty}$ which vanish at infinity, with the supremum norm $\|f\| = \sup_{z \in [0, \infty) \times \mathcal{N}_{\infty}} |f(z)|$. The key issue for showing that the closure of the differential operator $A$, defined in (4), generates a Feller diffusion on $C_0([0, \infty) \times \mathcal{N}_{\infty})$ is the choice of the domain of $A$. Here we adapt to the present framework a technique indicated by Ethier and Kurtz (1981). Consider polynomials $\varphi_m : \mathcal{N}_{\infty} \to [0, 1]$ defined as

\begin{equation}
\varphi_1(z) = 1, \quad \varphi_m(z) = \sum_{i=1}^{\infty} z_i^m, \quad m \geq 2.
\end{equation}

Since $z \in \mathcal{N}_{\infty}$ implies $z_i \leq i^{-1}$, functions $\varphi_m$ with $m \geq 2$ are uniformly convergent, and sums in (4) are assumed to be computed on

\begin{equation}
\mathcal{N}_{\infty} = \left\{ z = (z_1, z_2, \ldots) : z_1 \geq z_2 \geq \ldots \geq 0, \sum_{i=1}^{\infty} z_i = 1 \right\}
\end{equation}

and extended to $\mathcal{N}_{\infty}$ by continuity, so that for example

\[
\sum_{i=1}^{\infty} (1 + z_i) \frac{\partial}{\partial z_i} \varphi_2(z) = 2 + 2 \varphi_2(z)
\]

instead of $2 \sum_{i=1}^{\infty} z_i + 2 \varphi_2(z)$. Write

\begin{equation}
A = A_0 + A_1
\end{equation}

to indicate the first two and last two terms in (4), and denote

\begin{equation}
\mathcal{D}(A_1) = \left\{ \text{sub-algebra of } C(\mathcal{N}_{\infty}) \text{ generated by } \varphi_m \text{ as in (43)} \right\}.
\end{equation}

The domain $\mathcal{D}(A)$ of the operator (4) is then taken to be

\begin{equation}
\mathcal{D}(A) = \left\{ \text{sub-algebra of } C_0([0, \infty) \times \mathcal{N}_{\infty}) \text{ generated by } f = f_0 \times f_1 : f_0 \in \mathcal{D}(A_0), f_1 \in \mathcal{D}(A_1) \right\},
\end{equation}

with $\mathcal{D}(A_0)$ as in (36) and $\mathcal{D}(A_1)$ as above.

**Lemma 5.3.** The sub-algebra $\mathcal{D}(A_1) \subset C(\mathcal{N}_{\infty})$ is dense in $C(\mathcal{N}_{\infty})$. 
Proof. See the proof of Theorem 2.5 in Ethier and Kurtz (1981).

We also need the following lemma, which shows that the operator $A_1$ is triangulizable.

**Lemma 5.4.** Let $A_1$ be as in (45) and, for any $m \geq 2$, let $L_m$ be the algebra generated by polynomials as in (43), with degree not greater than $m$. Then $A_1 : L_m \to L_m$.

**Proof.** The assertion follows from equation (2.4) in Feng and Sun (2010), with $\theta$ replaced by $\beta/s$. □

Then we have the following result.

**Theorem 5.5.** Let $A$ be the operator defined by (4) and (47). The closure in $C_0([0, \infty) \times \nabla_\infty)$ of $A$ generates a strongly continuous, positive, conservative, contraction semigroup $\{T(t)\}$ on $C_0([0, \infty) \times \nabla_\infty)$. For every $\nu \in \mathcal{P}([0, \infty) \times \nabla_\infty)$, there exists a strong Markov process $Z(\cdot)$ corresponding to $\{T(t)\}$ with initial distribution $\nu$ and sample paths in $C([0, \infty) \times \tilde{\nabla}_\infty, [0, \infty))$ with probability one.

**Proof.** For every $g \in C_0([0, \infty) \times \nabla_\infty)$, define $r_n : C_0([0, \infty) \times \nabla_\infty) \to C_0([0, \infty) \times \tilde{\nabla}_n)$ to be the bounded linear map

$$r_n g = g|_{[0, \infty) \times \tilde{\nabla}_n}$$

given by the restriction of $g$ to $[0, \infty) \times \tilde{\nabla}_n$. Note that $r_n : \mathcal{D}(A) \to \mathcal{D}(\tilde{A}_n)$, with $\tilde{A}_n$ as in Proposition 5.2, and that

$$||r_n g - g|| \to 0, \quad g \in C_0([0, \infty) \times \nabla_\infty).$$

Then, for $g \in \mathcal{D}(A)$ and $z \in (0, \infty) \times \tilde{\nabla}_n$, we have

$$|\tilde{A}_n r_n g(z) - r_n A g(z)| =$$

$$= \frac{1}{2} \left| \sum_{i,j=1}^n \left( a_{ij}^{(n)}(z) - z_i (\delta_{ij} - z_j) \right) \frac{\partial^2 g(z)}{\partial z_i \partial z_j} \right. + \sum_{i=1}^n \left( \frac{\beta}{z_0(n-1)} (1 - z_i) - \alpha \left( \exp \left\{ - (z_i - \varepsilon_n) e^{1/\varepsilon_n} \right\} \right) \right) \frac{\partial g(z)}{\partial z_i} \right|$$

with $a_{ij}^{(n)}(z)$ as in (38). In particular

$$|a_{ij}^{(n)}(z) - z_i (\delta_{ij} - z_j)| = \begin{cases} 
\varepsilon_n [(z_i - \varepsilon_n)(n-1) + 1 - z_i] & i = j \\
\varepsilon_n [z_i + z_j - \varepsilon_n] & i \neq j,
\end{cases}$$

for $i, j = 1, \ldots, n$. □
which is bounded above by \( n\varepsilon_n \), from which

\[
|\tilde{A}_n r_n g(z) - r_n A g(z)| \leq n\varepsilon_n \sum_{i,j=1}^{n} \left| \frac{\partial^2 g(z)}{\partial z_i \partial z_j} \right|
+ \frac{\beta}{z_0(n-1)} \sum_{i=1}^{n} \left| \frac{\partial g(z)}{\partial z_i} \right| + \sum_{i=1}^{n} \exp \left\{ -(z_i - \varepsilon_n) e^{1/\varepsilon_n} \right\} \left| \frac{\partial g(z)}{\partial z_i} \right|.
\]

For \( g \in \mathcal{D}(A) \) of type \( g = f_0 \times f_1 \), with \( f_0 \in \mathcal{D}(A_0) \) and \( f_1 = \varphi_{m_1} \cdots \varphi_{m_k} \), we have

\[
\sum_{i=1}^{n} \left| \frac{\partial g(z)}{\partial z_i} \right| = |f_0(z_0)| \sum_{i=1}^{n} \sum_{j=1}^{k} m_j z_i^{m_j-1} \prod_{h \neq j} \varphi_{m_h} \leq |f_0(z_0)| \sum_{j=1}^{k} m_j \sum_{i=1}^{n} z_i^{m_j-1}
\]

so that

\[
\sum_{i=1}^{n} \exp \left\{ -(z_i - \varepsilon_n) e^{1/\varepsilon_n} \right\} \left| \frac{\partial g(z)}{\partial z_i} \right| \leq n\varepsilon_n |f_0(z_0)| \sum_{j=1}^{k} m_j \to 0
\]

uniformly as \( n \to \infty \) by (37). Furthermore

\[
\sum_{i,j=1}^{n} \left| \frac{\partial^2 g(z)}{\partial z_i \partial z_j} \right| \leq |f_0(z_0)| \sum_{i,j=1}^{\infty} \left[ \partial_{ij} \varphi_{m_h} \prod_{\ell \neq h} \varphi_{m_\ell} + \sum_{q \neq h} \partial_{i} \varphi_{m_h} \partial_j \varphi_{m_q} \prod_{\ell \neq h,q} \varphi_{m_\ell} \right]
\]

\[
= |f_0(z_0)| \left[ m_h (m_h - 1) \prod_{\ell \neq h} \varphi_{m_\ell} + \sum_{q \neq h} m_h m_q \varphi_{m_h + m_q - 2} \prod_{\ell \neq h,q} \varphi_{m_\ell} \right]
\]

whose right hand side is bounded. Since also the right hand side of (50) is bounded above by \( |f_0(z_0)| \sum_{j=1}^{k} m_j \), it follows by (37) that the right hand side of (49) goes to zero uniformly and, by means of (48), that

\[
||\tilde{A}_n r_n g - A g|| \to 0, \quad g \in \mathcal{D}(A).
\]

Proposition 5.2 implies that \( \tilde{A}_n \) is a dissipative operator for every \( n \geq 1 \), so that by (51) \( A \) is dissipative. Moreover, Lemma 5.3 and Lemma 5.4 respectively imply that \( \mathcal{D}(A) \) and the range of \( \lambda - A \), for all but at most countably many \( \lambda > 0 \), are dense in \( C_0([0, \infty) \times \nabla_\infty) \). The fact that the closure of \( A \) generates a strongly continuous contraction semigroup \( \{ T(t) \} \) on
Normalized inverse-Gaussian diffusions

\( C_0([0, \infty) \times \nabla_\infty) \) now follows from the Hille-Yosida Theorem (see Theorem 1.2.12 in Ethier and Kurtz (1986)). It is also immediate to check that \( A1 = 0 \), so that \((1, 0) \in (g, Ag)\) and \( \{T(t)\} \) is conservative. Finally, (51), together with Lemma 5.3 and Theorem 1.6.1 of Ethier and Kurtz (1986), implies the semigroup convergence

\[
||T_n(t)r_n g - T(t)g|| \longrightarrow 0, \quad g \in C_0([0, \infty) \times \nabla_\infty),
\]

uniformly on bounded intervals. From Proposition 5.2, \( \{T_n(t)\} \) is a positive operator for every \( n \geq 1 \), so that \( \{T(t)\} \) is in turn positive.

The second assertion of the Theorem, with \( D_{[0,\infty) \times \nabla_\infty ([0,\infty))} \) in place of \( C([0,\infty) \times \nabla_\infty ([0,\infty)) \), follows from Theorem 4.2.7 in Ethier and Kurtz (1986), while the continuity of sample paths follows from a similar argument to that used in the proof of Proposition 5.1.

\[ \square \]

**Corollary 5.6.** Let \( Z^{(n)}(\cdot) \) be as in Proposition 5.1 with initial distribution \( \nu_n \in \mathcal{P}([0, \infty) \times \hat{\Delta}_n) \), and let \( Z(\cdot) \) be as in Theorem 5.5 with initial distribution \( \nu \in \mathcal{P}([0, \infty) \times \nabla_\infty). \) If \( \nu_n \circ \rho_n^{-1} \Rightarrow \nu \) on \( \nabla_\infty \), then

\[
\rho_n(Z^{(n)}(\cdot)) \Rightarrow Z(\cdot) \quad \text{in} \ C([0, \infty) \times \nabla_\infty ([0, \infty))
\]

as \( n \to \infty. \)

**Proof.** The result with \( D_{[0,\infty) \times \nabla_\infty ([0,\infty))} \) in place of \( C([0,\infty) \times \nabla_\infty ([0,\infty)) \) follows from Proposition 5.2, together with (52) and Theorem 4.2.5 in Ethier and Kurtz (1986). The fact that the weak convergence holds in \( C([0,\infty) \times \nabla_\infty ([0,\infty)) \) follows from relativization of the Skorohod topology. \( \square \)

6 A population model for normalized inverse-Gaussian diffusions

By formalizing the population process briefly mentioned in Section 4 for constructing the \( \alpha \)-diversity diffusion, in this section we provide a discrete approximation, based on a countable number of particles, for the diffusion with operator (4). More specifically, this is obtained as limit in distribution of the process of frequencies of types associated with a set of particles sampled from a normalized inverse-Gaussian random probability measure, jointly with the normalized version of the diversity process.

In view of (10), the conditional distribution of the \( i \)-th component of an exchangeable sequence \( (X_1, \ldots, X_n) \) drawn from a random probability measure of Gibbs type can be written

\[
P\{X_i \in \cdot \mid X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n\} = \]
\[
\nu_0(n - 1, K_{n-1,i}) + g_1(n - 1, K_{n-1,i}) \sum_{j=1}^{K_{n-1,i}} (n_j - \alpha) \delta_{X_j^*}(\cdot)
\]

where \(\nu_0\) is a non atomic probability measure and \((X_1^*, \ldots, X_{K_{n-1,i}}^*)\) are the \(K_{n-1,i}\) distinct values in \((X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)\). For fixed \(n\), define a Markov chain \(\{X^{(n)}(m), m \geq 0\}\) on \(\mathbb{X}^n\) by means of the transition semigroup

\[
T_n f(x) = \int f(y) p_n(x, dy), \quad f \in C_0(\mathbb{X}^n)
\]

where \(x, y \in \mathbb{X}^n\), \(C_0(\mathbb{X}^n)\) is the space of Borel-measurable continuous functions on \(\mathbb{X}^n\) vanishing at infinity,

\[
p_n(x, dy) = \frac{1}{n} \sum_{i=1}^{n} \bar{p}_1(dy|x_{(-i)}) \prod_{k \neq i} \delta_{x_k}(dy_k),
\]

\(x_{(-i)} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)\) and \(\bar{p}_1(dy|x_{(-i)})\) is \((53)\). The interpretation is as follows. At each transition one component is selected at random with uniform probability, and is updated with a value sampled from \((53)\), conditional on all other components, which are left unchanged. Hence the incoming particle is either a new type (a mutant offspring) or a copy of an old type (a copied offspring). Embed now the chain in a pure jump Markov process on \(\mathbb{X}^n\) with exponentially distributed waiting times with intensity one, and denote the resulting process by \(X^{(n)}(\cdot) = \{X^{(n)}(t), t \geq 0\}\). The infinitesimal generator of \(X^{(n)}(\cdot)\) is given by

\[
B_n f(x) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{K_{n-1,i}} g_0^{(n-1,i)} [P_i f(x) - f(x)]
\]

\[
+ \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{K_{n-1,i}} g_1^{(n-1,i)} (n_j - \alpha) \Phi_j^* f(x) - f(x)]
\]

with domain

\[
\mathcal{D}(B_n) = \{f : f \in C_0(\mathbb{X}^n)\}.
\]

Here \(\Phi_j^*: C_0(\mathbb{X}^n) \to C_0(\mathbb{X}^{n-1})\) is defined as

\[
\Phi_j^* f(x_1, \ldots, x_n) = f(x_1, \ldots, x_{i-1}, x_j^*, x_{i+1}, \ldots, x_n)
\]

for \(x_j^* \in (x_1^*, \ldots, x_{K_{n-1,i}}^*)\), \(P\) is the transition semigroup

\[
P g(z) = \int g(y)p_1(z, dy), \quad g \in C_0(\mathbb{X}),
\]
where \( p_1(z, dy) \) is given by
\[
(59) \quad p_1(z, dy) = \nu_0(dy),
\]
\( P_i f \) denotes \( P \) acting on the \( i \)-th coordinate of \( f \), and we have set for brevity
\[
(60) \quad g_j^{(n-1,i)} = g_j(n-1, K_{n-1,i}), \quad j = 0, 1.
\]
Defining (58) and (59) separately is somewhat redundant, but will allow us to provide a general expression for the global mutation rate in this particle representation before making the assumptions of non atomicity and parent independence as in (59). See (A.80) below.

Define now the map \( w : \mathbb{X}^n \to \nabla_\infty \) by
\[
(61) \quad w(x) = w(x^{(n)}) = (z_1, \ldots, z_{K_n}, 0, 0, \ldots)
\]
where \( z_j \) and \( K_n \) respectively denote the relative frequency of the \( j \)-th most abundant type and the number of types in \( X^{(n)} \). Let also \( A \) be as in (4). The next theorem states that
\[
(62) \quad [K_n(\cdot)/n^\alpha, w(X^{(n)}(\cdot))] = \{[K_n(t)/n^\alpha, w(X^{(n)}(t))], t \geq 0\}
\]
if appropriately rescaled in time, converges in distribution to the process with generator \( \overline{A} \). The proof is deferred to the Appendix and contains, as a byproduct, a more formal derivation of Theorem 4.1.

**Theorem 6.1.** Let \( X^{(n)}(\cdot) \) be the \( \mathbb{X}^n \)-valued process with generator (55)-(56), \( w : \mathbb{X}^n \to \nabla_\infty \) as in (61) and \( Z(\cdot) \) as in Theorem 5.5. If
\[
(63) \quad [K_n(0)/n^\alpha, w(X^{(n)}(0))] \Rightarrow Z(0)
\]
then
\[
(64) \quad [K_n(n^{3/2}t)/n^\alpha, w(X^{(n)}(n^{2}t/2))] \Rightarrow Z(t)
\]
in \( C_{[0,\infty) \times \nabla_\infty}([0,\infty)) \).

We conclude the Section by showing the reversibility of the particle process. Denote the joint distribution of an \( n \)-sized sequence from the generalized Pólya urn scheme (10) by
\[
\mathcal{M}_n(dx_1, \ldots, dx_n)
= \nu_0(dx_1) \prod_{i=1}^{n-1} \left[ g_0(i, K_i)\nu_0(dx_{i+1}) + g_1(i, K_i) \sum_{j=1}^{K_i} (n_j - \alpha) \delta_{x_j^*}(dx_{i+1}) \right].
\]
Proposition 6.2. Let $X^{(n)}(\cdot)$ be the $\mathbb{X}^n$-valued process with generator given by (55)-(56). Then $X^{(n)}(\cdot)$ is reversible with respect to $\mathcal{M}_n$.

Proof. Let $q_n(x,dy)$ denote the infinitesimal transition kernel on $\mathbb{X}^n \times \mathcal{B}(\mathbb{X}^n)$ of $X^{(n)}$. Since $X^{(n)}$ has discontinuities at rate $\lambda_n$, recalling (54) we have

\[
\mathcal{M}_n(dx)q_n(x,dy) = \mathcal{M}_n(dx)\lambda_n \frac{1}{n} \sum_{i=1}^{n} p_1(dy_i|x_{(-i)}) \prod_{k \neq i} \delta_{x_k}(dy_k)
\]

\[
= \frac{\lambda_n}{n} \sum_{i=1}^{n} \mathcal{M}_{n-1}(dx_{(-i)}) p_1(dx_{i}|x_{(-i)}) p_1(dy_{i}|x_{(-i)}) \prod_{k \neq i} \delta_{x_k}(y_k)
\]

\[
= \frac{\lambda_n}{n} \sum_{i=1}^{n} \mathcal{M}_{n-1}(dy_{(-i)}) p_1(dx_{i}|y_{(-i)}) p_1(dy_{i}|y_{(-i)}) \prod_{k \neq i} \delta_{y_k}(x_k)
\]

\[
= \mathcal{M}_n(dy) \frac{1}{n} \sum_{i=1}^{n} \lambda_n p_n(dx_{i}|y_{(-i)}) \prod_{k \neq i} \delta_{y_k}(x_k) = \mathcal{M}_n(dy)q_n(y,dx)
\]

giving the result. \hfill \square

7 Conditioning on the alpha-diversity

We conclude by discussing an interesting connection with the two-parameter model (2). In the Introduction it was observed that conditioning on the $\alpha$-diversity diffusion $S_t$ to be constant in the operator (4) only yields the special case $\beta = 0$ and $\alpha = 1/2$, consistently with the associated random probability measures. It turns out that performing the same conditioning operation in the particle construction of the previous section, before taking the limit for $n \to \infty$, yields a particular instance of the two-parameter model. The following proposition states that under this pre-limit conditioning with $S_t \equiv s$, the normalized inverse-Gaussian model with operator (4) reduces to the two-parameter model with $(\theta, \alpha) = ((\alpha s)^2, \alpha)$ and $\alpha = 1/2$.

Proposition 7.1. Denote by $\tilde{Z}^{(n)}(\cdot)$ the process obtained from (61), and by $V^{\theta,\alpha}(\cdot)$ the process corresponding to $L^{\theta,\alpha}$ as in (2). Then

$$
\tilde{Z}^{(n)}(\cdot) \bigg| S_t \equiv s \Rightarrow V^{(\alpha s)^2,\alpha}(\cdot)
$$

in $C_{[0,\infty) \times [0,\infty)}$ as $n \to \infty$. 
Proof. In the pre-limit version of the process of frequencies derived from the particle process, i.e. (61), conditioning on \( S_t \equiv s \) means conditioning on \( K_n(\cdot) \) being constant over time, hence with zero conditional first and second moment. Denote

\[
z - \varepsilon_i = (z_1, \ldots, z_{i-1}, z_i - 1, z_{i+1}, \ldots)
\]

and assume \( z \) has \( k \) non null components obtained from \( n \) particles. Then, as in Section 6, when a particle is removed we have the change of frequency

\[
z \mapsto z - \frac{\varepsilon_i}{n} \quad \text{w.p. } z_i
\]

where \( z - \varepsilon_i/n \) has

1. \( k \) non-null components w.p. \( 1 - \frac{m_{1,n}}{n} \)
2. \( k - 1 \) non-null components w.p. \( \frac{m_{1,n}}{n} \).

Conditional on case (1), the number of non null components remains \( k \) if the incoming particle is a copy of an existing type, i.e. we observe either of

\[
z - \frac{\varepsilon_i}{n} \mapsto z - \frac{\varepsilon_i}{n} + \frac{\varepsilon_i}{n} \quad \text{w.p. } g_1^{(n,k)}(n_i - 1 - \alpha)/(1 - g_0^{(n,k)})
\]

\[
z - \frac{\varepsilon_i}{n} \mapsto z - \frac{\varepsilon_i}{n} + \frac{\varepsilon_j}{n} \quad \text{w.p. } g_1^{(n,k)}(n_j - \alpha)/(1 - g_0^{(n,k)})
\]

where \( g_0^{(n,k)} \) and \( g_1^{(n,k)} \) are as in (60), while conditional on case (2) we observe

\[
z - \frac{\varepsilon_i}{n} \mapsto z - \frac{\varepsilon_i}{n} + \frac{\varepsilon_{k+1}}{n} \quad \text{w.p. } 1.
\]

If the events occur at rate \( \lambda_n \), the generator of the process \( Z^{(n)}(\cdot) \) in this case can be written

\[
B_{n,1} f_1(z) = \lim_{\delta t \to 0} \frac{1}{\delta t} \left\{ \lambda_n \delta t \sum_{i=1}^k z_i \left[ f_1(z) \left( 1 - \frac{m_{1,n}}{n} \right) \frac{g_1^{(n,k)}(n_i - 1 - \alpha)}{1 - g_0^{(n,k)}} + \sum_{j \neq i} f_1 \left( z - \frac{\varepsilon_i}{n} + \frac{\varepsilon_j}{n} \right) \left( 1 - \frac{m_{1,n}}{n} \right) \frac{g_1^{(n,k)}(n_j - \alpha)}{1 - g_0^{(n,k)}} \\
+ f_1 \left( z - \frac{\varepsilon_i}{n} + \frac{\varepsilon_{k+1}}{n} \right) \frac{m_{1,n}}{n} \right] \right. \\
+ \left. (1 - \lambda_n \delta t) f_1(z) + O((\delta t)^2) - f_1(z) \right\}
\]
for \( f_1 \in C^2(\nabla_n) \),

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

and \( \nabla_\infty \) as in (44). Exploiting the relation

\[
\left(1 - \frac{m_{1,n}}{n}\right) g_1 \sum_{j=1}^k \frac{(n_j - \alpha)}{1 - g_0} + \frac{m_{1,n}}{n} = 1
\]

we can write

\[
B_{n,1}f_1(z) = \lambda_n \sum_{i=1}^k \left\{ f_1(z) - f_1(z') \right\} \left(1 - \frac{m_{1,n}}{n}\right) \frac{g_1^{(n,k)}(n_i - \alpha)}{1 - g_0^{(n,k)}}
\]

\[
+ \sum_{j \neq i} \left[ f_1 \left(z - \frac{\varepsilon_i}{n} + \frac{\varepsilon_j}{n}\right) - f_1(z) \right] \left(1 - \frac{m_{1,n}}{n}\right) \frac{g_1^{(n,k)}(n_j - \alpha)}{1 - g_0^{(n,k)}}
\]

\[
+ \left[ f_1 \left(z - \frac{\varepsilon_i}{n} + \frac{\varepsilon_{k+1}}{n}\right) - f_1(z) \right] \frac{m_{1,n}}{n}
\]

\[
= \lambda_n \sum_{i,j} z_i \left[ f_1 \left(z - \frac{\varepsilon_i}{n} + \frac{\varepsilon_j}{n}\right) - f_1(z) \right] \left(1 - \frac{m_{1,n}}{n}\right) \frac{g_1^{(n,k)}(n_j - \alpha)}{1 - g_0^{(n,k)}}
\]

By making use of Taylor’s Theorem, it can be easily verified that the following three relations holds:

\[
\sum_{i,j} z_i z_j \left[ f_1 \left(z - \frac{\varepsilon_i}{n} + \frac{\varepsilon_j}{n}\right) - f_1(z) \right] = \frac{1}{n^2} \sum_{i,j} z_i (\delta_{ij} - z_j) \frac{\partial^2 f_1}{\partial z_i \partial z_j} + o(n^{-2}),
\]

\[
\sum_{i,j} z_i \left[ f_1 \left(z - \frac{\varepsilon_i}{n} + \frac{\varepsilon_j}{n}\right) - f_1(z) \right] = \frac{1}{n} \sum_i \frac{\partial f_1}{\partial z_i} - \frac{k}{n} \sum_i z_i \frac{\partial f_1}{\partial z_i} + O(n^{-2}),
\]

\[
\sum_i z_i \left[ f_1 \left(z - \frac{\varepsilon_i}{n} + \frac{\varepsilon_{k+1}}{n}\right) - f_1(z) \right] = -\frac{1}{n} \sum_i z_i \frac{\partial f_1}{\partial z_i} + O(n^{-2}).
\]

By means of the last three expressions we can write (66) as

\[
B_{n,1}f_1(z) = \frac{\lambda_n}{n^2} \sum_{i,j} z_i (\delta_{ij} - z_j) \frac{\partial^2 f_1(z)}{\partial z_i \partial z_j} + B_{n,1}^{\text{dir}}f_1(z) + o(\lambda_n n^{-2})
\]
where $B_{n,1}^{dr}f_1(z)$ is the drift term, given by

$$B_{n,1}^{dr}f_1(z) = \frac{\lambda_n}{n} \left[ \left( \frac{\alpha kg_1}{1 - g_0} \left( 1 - \frac{m_{1,n}}{n} \right) - \frac{m_{1,n}}{n} \right) \sum_i z_i \frac{\partial f_1}{\partial z_i} ight. \\
\left. - \frac{\alpha g_1}{1 - g_0} \left( 1 - \frac{m_{1,n}}{n} \right) \sum_i \frac{\partial f_1}{\partial z_i} \right].$$

Using (22) and Proposition 3.2 it can be seen that

$$\frac{\alpha kg_1}{1 - g_0} \left( 1 - \frac{m_{1,n}}{n} \right) - \frac{m_{1,n}}{n} \approx -\frac{\alpha km_{1,n}}{n^2} \approx -\frac{\alpha^2 s^2}{n}$$

and

$$\frac{\alpha g_1}{1 - g_0} \left( 1 - \frac{m_{1,n}}{n} \right) \approx \frac{\alpha}{n},$$

yielding

$$B_{n,1}f_1(z) = \frac{\lambda_n}{n^2} \sum_{i,j} z_i (\delta_{ij} - z_j) \frac{\partial^2 f_1(z)}{\partial z_i \partial z_j} - \frac{\lambda_n}{n^2} \sum_i ((\alpha s^2) z_i + \alpha) \frac{\partial f_1(z)}{\partial z_i} + o(\lambda_n n^{-2}).$$

For $f \in C(\nabla_\infty)$, define $\tilde{r}_n : C(\nabla_\infty) \to C(\nabla_n)$ to be

$$(67) \quad \tilde{r}_n f = f|_{\nabla_n}$$

namely the restriction of $f$ to $\tilde{\nabla}_n$ and note that

$$||\tilde{r}_n f - f|| \rightarrow 0, \quad f \in C(\nabla_\infty).$$

Choosing $\lambda_n = n^2/2$ implies that for $f$ as in (46) and $L^{\theta,\alpha}$ as in (2)

$$||L^{(\alpha s^2, \alpha)} f - B_{n,1}^{\tilde{r}_n, f}|| \rightarrow 0.$$

The strong convergence of the corresponding semigroups on $C(\nabla_\infty)$, similar to (52), and the statement of the Proposition now follow from an application of Theorems 1.6.1 and 4.2.11 in Ethier and Kurtz (1986), together with the relativization of the Skorohod topology to $C_{[0,\infty) \times \nabla_\infty}([0,\infty))$. \qed

Acknowledgements

The first author is grateful to Antonio Lijoi, Bertrand Lods and Igor Prünster for several useful discussions.
Proof of Proposition 3.2

Consider first that $V_{n,k}$ appearing in (11), in the case of generalized gamma processes, can be written (cf. Lijoi, Mena and Prünster (2007b))

$$V_{n,k} = \frac{a^k}{\Gamma(n)} \int_0^\infty x^n \exp \left\{ -\frac{a}{\alpha} \left[ (\tau + x)^\alpha - \tau^\alpha \right] \right\} (\tau + x)^{ak-n} dx.$$  

Together with (9) this leads to writing

$$g_0(n,k) = V_{n,k} - \frac{(n - \alpha k) V_{n+1,k}}{V_{n,k}} = 1 - (1 - \alpha k/n) w(n,k)$$

where

$$w(n,k) = \frac{\int_0^\infty x^n \exp \left\{ -\frac{\alpha}{a} \left[ (\tau + x)^\alpha - \tau^\alpha \right] \right\} (\tau + x)^{ak-n-1} dx}{\int_0^\infty x^{n-1} \exp \left\{ -\frac{\alpha}{a} \left[ (\tau + x)^\alpha - \tau^\alpha \right] \right\} (\tau + x)^{ak-n} dx}.$$  

Denote by $f(x)$ the integrand of the denominator of $w(n,k)$, so

$$w(n,k) = \frac{\int_0^\infty \frac{x}{\tau + x} f(x) dx}{\int_0^\infty f(x) dx}.$$  

Since $f(x)$ is unimodal, by means of the Laplace method one can approximate $f(x)$ with the kernel of a normal density with mean given by

$$x^* = \arg \max_{x>0} x^{n-1} \exp \left\{ -\frac{\alpha}{a} \left[ (\tau + x)^\alpha - \tau^\alpha \right] \right\} (\tau + x)^{ak-n}$$

and variance given by $-\left[ f''(x) \right]^{-1} |_{x=x^*}$. It follows that

$$w(n,k) \approx \frac{f_N(x_N^*) C(x_N^*, -\left[ f''_N(x) \right]^{-1} |_{x=x^*})}{f(x_D^*) C(x_D^*, -\left[ f''(x) \right]^{-1} |_{x=x^*})}$$

where $f_N$ denotes the integrand of the numerator, $x_N^*$ and $x_D^*$ the modes of the integrands of numerator and denominator respectively, and $C(x,y)$ is the normalizing constant of a normal kernel with mean $x$ and variance $y$, yielding

$$w(n,k) \approx \frac{f_N(x_N^*)}{f(x_D^*)} \left( \frac{f''(x_D^*)}{f''_N(x_N^*)} \right)^{1/2}.$$  

\[\text{(A.68)}\]
From (A.67), the mode $x_D^*$ is the only positive real root of the equation

\[(n - 1)x^{-1} + (\alpha k - n)(\tau + x)^{-1} - a(\tau + x)^{n-1} = 0\]

which, for $\alpha \neq 1/2, 1/3$, involves finding roots of polynomials of degree greater than 4. When $\alpha = 1/2$ we have

\[
x_D^* = \frac{(k - 2)^2}{12a^2} - \frac{\tau}{3} + \frac{48a^2 \tau (n - 1)(k - 2) + [4a^2 \tau - (k - 2)^2]^2}{6 \cdot 2^{1/3}a^2p_{1,D}(a, \tau, n, k)} + \frac{p_{1,D}(a, \tau, n, k)}{12 \cdot 2^{1/3}a^2}
\]

where

\[
p_{1,D}(a, \tau, n, k) = \left\{ p_{2,D} + \left[ p_{2,D} + 4 \left( -48a^2 \tau (n - 1)(k - 2) - \left(4a^2 \tau - (k - 2)^2\right)^2\right)^{1/3} \right]^{1/3} \right\}
\]

and $p_{2,D} = p_{2,D}(a, \tau, n, k)$ with

\[
p_{2,D}(a, \tau, n, k) = 2(k - 2)^3[(k - 2)^3 - 12a^2 \tau (k + 4 - 6n)] + 96a^4 \tau^2(k + 2) + 10 - 6n(k + 4) + 18n^2 - 128a^6 \tau^3.
\]

Similarly one finds that

\[
x_N^* = \frac{(k - 2)^2}{12a^2} - \frac{\tau}{3} + \frac{48a^2 \tau n(k - 2) + [4a^2 \tau - (k - 2)^2]^2}{6 \cdot 2^{2/3}a^2p_{1,N}(a, \tau, n, k)} + \frac{p_{1,N}(a, \tau, n, k)}{12 \cdot 2^{1/3}a^2}
\]

where

\[
p_{1,N}(a, \tau, n, k) = [p_{2,N} + \left( p_{3,N} \right)^{1/3}]^{1/3}
\]

with

\[
p_{2,N}(a, \tau, n, k) = p_{2,D}(a, \tau, n, k) - 144a^2 \tau [4a^2 \tau (k + 1 - 6n) - (k - 2)^3]
\]

and

\[
p_{3,N}(a, \tau, n, k) = -33210a^6n^2 \tau^3 \left[(k - 2)^4 - 2n(k - 2)^3 - 4a^2 \tau [8 + 2k^2 + 9n(3n + 4) - 2k(9n + 4)] + 16a^4 \tau^2 \right].
\]

When $k \approx sn^\alpha$ (cf. (24) above) and $\alpha = 1/2$, it can be checked that

\[
n^{-1}p_{1,i}(a, \tau, n, k) \to 2^{1/3}s^2, \quad i = N, D
\]
from which
\[(A.71) \quad n^{-1}x_i^* \to (s/2a)^2, \quad i = N, D.\]

Using this fact one finds that
$$\frac{f_N(x_N^*)}{f(x_D^*)} \approx \frac{1}{1 + \tau(s/2a)^{-2}}.$$ Computing also the ratio of the two second derivatives it can be seen that
$$w(n,k) \to 1,$$ which by means of (A.66) and (12) implies (25). In order to find the speed at which $w(n,k)$ goes to 1, consider
$$1 - w(n,k) = \int_0^\infty \frac{\tau}{\tau + x} f(x) dx / \int_0^\infty f(x) dx.$$

The denominator is unchanged, while the mode of $\tau/(\tau + x)f(x)$ is
$$x_N^* = \frac{(k - 4)^2}{12a^2} - \frac{t}{3} - \frac{-48a^2t(n - 1)(k - 4) - [4a^2t - (k - 4)^2]^2}{6 \cdot 2^{2/3}a^2 p'_1(a, \tau, n, k)}$$
$$+ \frac{p'_1(a, \tau, n, k)}{12 \cdot 2^{1/3}a^2}$$
where
$$p'_1(a, \tau, n, k) = \left\{ p_{2,D} + \left[ p_{2,D}^2 + 4 \left( -48a^2t(n - 1)(k - 4) - \left( 4a^2t - (k - 4)^2 \right)^2 \right) \right]^{\frac{1}{2}} \right\}^{\frac{1}{3}}$$
and $p_{2,D} = p_{2,D}(a, \tau, n, k)$ with
$$p'_{2,D}(a, \tau, n, k) = 2(k - 4)^3[(k - 4)^3 - 12a^2\tau(k + 2 - 6n)] + 96a^4\tau^2(k - 2) + 10 - 6n(k + 2) + 18n^2 - 64a^6\tau^3.$$

Moreover, $p'_{2,D}$ satisfies (A.70), and (A.71) follows. Unfortunately the fact that the two modes grow with an asymptotically equivalent rate is too rough an approximation for our purposes here, which ignores how far apart they are if this is negligible with respect to the growth speed. Indeed it turns out that
$$n^{-1/2}(x_D^* - x_N^*) \to s/a^2.$$

Using this information in the Laplace approximation for $w(n,k) - 1$ yields
$$n(1 - w(n,k)) \to 2a\sqrt{\tau}/s = \beta/s.$$
with $\beta$ as in the statement of the Proposition. From (A.66) it is now easy to see that
\[ ng_0(n, k) - \alpha k w(n, k) = n(1 - w(n, k)) \]
which provides the second order approximation for $g_0(n, k)$,
\[ g_0(n, k) = \frac{\alpha k}{n} + \frac{\beta / s_n}{n} + o(n^{-1}) \]
where the first term is of order $n^{-1/2}$, yielding immediately, by means of (12), the second order term for $g_1(n, k)$, i.e.
\[ g_1(n, k) = \frac{1}{n} - \frac{\beta / s_n}{n^2} + o(n^{-2}). \]
\[ \square \]

**Proof of Theorem 6.1**

We can write the generator of (62) as
\[ (A.72) \quad B_n(f_0 \times f_1) = f_1 B_{n,0} f_0 + f_0 B_{n,1} f_1, \]
where $B_{n,0}$ and $B_{n,1}$ drive $K_n(\cdot)/n^\alpha$ and $w(X^{(n)}(\cdot))$ respectively and $f_0 \in \mathcal{D}(A_0)$, with $\mathcal{D}(A_0)$ as in (36), $f_1 \in C^2(\nabla_n)$ and $\nabla_n$ is as in (65). From (27) we can write $B_{n,0} f_0$ as
\[ B_{n,0} f_0 \left( \frac{k}{n^\alpha} \right) = \]
\[ = \lim_{\delta t \downarrow 0} \frac{1}{\delta t} \left\{ \delta t \left[ f_0 \left( \frac{k+1}{n^\alpha} \right) \left( 1 - \frac{m_{1,n}}{n} \right) g_0^{(n-1,k)} \right. \right. \]
\[ + f_0 \left( \frac{k-1}{n^\alpha} \right) m_{1,n} g_1^{(n-1,k-1)} (n - 1 - \alpha(k - 1)) \]
\[ + f_0 \left( \frac{k}{n^\alpha} \right) \left( 1 - \left( 1 - \frac{m_{1,n}}{n} \right) g_0^{(n-1,k)} \right. \]
\[ - \frac{m_{1,n}}{n} g_1^{(n-1,k-1)} (n - 1 - \alpha(k - 1)) \left. \right]\]
\[ \left. \right] + (1 - \delta t) f_0 \left( \frac{k}{n^\alpha} \right) + O((\delta t)^2) - f_0 \left( \frac{k}{n^\alpha} \right) \}
\[ = \left( 1 - \frac{m_{1,n}}{n} \right) g_0^{(n-1,k)} \left[ f_0 \left( \frac{k+1}{n^\alpha} \right) - f_0 \left( \frac{k}{n^\alpha} \right) \right] \]
An application of Taylor’s Theorem yields

\[ B_{n,0} f_0 \left( \frac{k}{n^\alpha} \right) = (1 - \frac{m_{1,n}}{n}) g_0^{(n-1,k)} \]

\[ = \left[ f_0 \left( \frac{k}{n^\alpha} \right) + \frac{1}{n^\alpha} f'_0 \left( \frac{k}{n^\alpha} \right) + \frac{1}{2n^{2\alpha}} f''_0 \left( \frac{k}{n^\alpha} \right) + O(n^{-3\alpha}) - f_0 \left( \frac{k}{n^\alpha} \right) \right] \]

\[ + \frac{m_{1,n}}{n} g_i^{(n-1,k-1)} (n-1 - \alpha(k-1)) \]

\[ = \frac{1}{n^{\alpha}} f_0 \left( \frac{k}{n^\alpha} \right) \left[ 1 - \frac{m_{1,n}}{n} g_0^{(n-1,k)} \right] \]

\[ - \frac{m_{1,n}}{n} g_i^{(n-1,k-1)} (n-1 - \alpha(k-1)) \]

\[ + \frac{1}{2n^{2\alpha}} f''_0 \left( \frac{k}{n^\alpha} \right) \left[ \left( 1 - \frac{m_{1,n}}{n} \right) g_0^{(n-1,k)} \right] \]

\[ + \frac{m_{1,n}}{n} g_i^{(n-1,k-1)} (n-1 - \alpha(k-1)) \right] + O(n^{-4\alpha}). \]

From (30)-(31) we have

\[ B_{n,0} f_0 \left( \frac{k}{n^\alpha} \right) = \frac{1}{n^{\alpha}} f_0 \left( \frac{k}{n^\alpha} \right) \left[ \frac{\beta/s_n}{n} + o(n^{1}) \right] \]

\[ + \frac{1}{2n^{2\alpha}} f''_0 \left( \frac{k}{n^\alpha} \right) \left[ \frac{s_n}{n^{2\alpha}} + O(n^{-1-2\alpha}) \right] + O(n^{-4\alpha}) \]

\[ = \frac{\beta/s_n}{n^{1+\alpha}} f'_0 \left( \frac{k}{n^\alpha} \right) + \frac{s_n}{2n^{3\alpha}} f''_0 \left( \frac{k}{n^\alpha} \right) + o(n^{-1-\alpha}) \]

from which it follows that

(A.73) \[ ||A_0 f_0 - n^{3/2} B_{n,0} f_0|| \to 0 \]

as \( n \to \infty \), with \( A_0 \) as in (35). Since (63) implies \( K_n(0)/n^\alpha \Rightarrow S_0 \), the previous expression, together with Theorems 1.6.1 and 4.2.5 in Ethier and Kurtz (1986), implies (29) with \( C_{[0,\infty)}([0,\infty)) \) replaced by \( D_{[0,\infty)}([0,\infty)) \), while (29) follows from relativization of the Skorohod topology to \( C_{[0,\infty)}([0,\infty)) \).
In order to describe $B_n$ in (A.72), define
\[
\phi_n(\mu) = f_1(\langle h_1, \mu \rangle, \ldots, \langle h_n, \mu \rangle) \quad f_1 \in C_0^2(\mathbb{R}^n), \ h_i \in C(X)
\]
and
\[
\mu_n(t) = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i(t)}, \quad t \geq 0,
\]
Then the generator of the $\mathcal{P}(X)$-valued process $\mu_n(\cdot) = \{\mu_n(t), t \geq 0\}$ can be written
\[
B_n \phi_n(\mu) = \frac{1}{n} \sum_{i=1}^{n} g_0^{(n-1,i)}[\langle Ph_i, \mu \rangle - \langle h_i, \mu \rangle] \frac{\partial f_1}{\partial z_i}
\]
\[
- \frac{\alpha}{n} \sum_{i=1}^{n} g_1^{(n-1,i)} K_{n-1,i} [\langle Q^{(n-1,i)} h_i, \mu \rangle - \langle h_i, \mu \rangle] \frac{\partial f_1}{\partial z_i}
\]
\[
+ \frac{1}{n} \sum_{1 \leq k \neq i \leq n} g_1^{(n-1,i)} [\langle h_i h_j, \mu \rangle - \langle h_i, \mu \rangle \langle h_j, \mu \rangle] \frac{\partial^2 f_1}{\partial z_i \partial z_j}
\]
where $g_{z_i}$ is the derivative of $g$ with respect to its $i$-th argument and $Q^{(n-1,i)}$ is defined as
\[
Q^{(n-1,i)} g(z) = \int g(y)p_{n-1,i}^*(z, dy), \quad g \in C_0(X),
\]
with
\[
p_{n-1,i}^*(z, dy) = \frac{1}{K_{n-1,i}} \sum_{j=1}^{K_{n-1,i}} \delta_{x_j^*}(dy).
\]
Unlabel now the model by choosing $\phi_n(\mu)$ as in (A.74) with $h_j(\cdot)$ being the indicator function of the $j$-th largest atom in $\mu$, so that $\langle h_j, \mu \rangle = z_j$ is the relative frequency associated with the $j$-th most abundant species. Note that some arguments of $f_1(z_1, \ldots, z_n)$ can be null since $K_{n-1,i} \leq K_n \leq n$. With this choice we have $\langle h_i h_j, \mu \rangle - \langle h_i, \mu \rangle \langle h_j, \mu \rangle = z_i \delta_{ij} - z_i z_j$, where $\delta_{ij}$ is Kronecker delta, and, under (58) and (A.77),
\[
\langle Ph_i, \mu \rangle - \langle h_i, \mu \rangle = \sum_{j=1}^{n} p_1(x_j^*, dx_j^*) z_j - z_i = \sum_{j=1}^{n} [p_1(x_j^*, dx_j^*) - \delta_{ij}] z_j
\]
and
\[
\langle Q^{(n-1,i)} h_i, \mu \rangle - \langle h_i, \mu \rangle = \sum_{j=1}^{n} [p_{n-1,i}^*(x_j^*, dx_j^*) - \delta_{ij}] z_j.
\]
It follows that $\mathbb{B}_n$ reduces to

$$
(A.79) \quad \mathcal{B}_{n,1} f_1 = \frac{1}{n} \sum_{i=1}^{n} g_0^{(n-1,i)} \left( \sum_{j=1}^{n} \left[ p_1(x_j^*, dx_i^*) - \delta_{ij} \right] z_j \right) \frac{\partial f_1}{\partial z_i} 
$$

$$
- \frac{\alpha}{n} \sum_{i=1}^{n} g_1^{(n-1,i)} K_{n-1,i} \left( \sum_{j=1}^{n} \left[ p_{n-1,i}^*(x_j^*, dx_i^*) - \delta_{ij} \right] z_j \right) \frac{\partial f_1}{\partial z_i} 
$$

$$
+ \frac{1}{n} \sum_{i,j=1}^{n} g_1^{(n-1,i)} z_i (\delta_{ij} - z_j) \frac{\partial^2 f_1}{\partial z_i \partial z_j}.
$$

Here a mutation from type $i$ to type $j$ occurs at rate $q_{ij} = q_{ij}(z)$ given by

$$
(A.80) \quad q_{ij} = \frac{1}{n} \left[ g_0^{(n-1,i)} \left[ p_1(i, \{j\}) - \delta_{ij} \right] - \alpha g_1^{(n-1,i)} K_{n-1,i} \left[ p_{n-1,i}^*(i, \{j\}) - \delta_{ij} \right] \right]
$$

where $p_1(i, \{j\})$ and $p_{n-1,i}^*(i, \{j\})$ stand for $p_1(x_j^*, dx_i^*)$ and $p_{n-1,i}^*(x_i^*, dx_j^*)$. When (59) holds, from the nonatomicity of $\nu_0$ we have $p_1(z, dy) = 0$ for every $y \in \mathbb{X}$, and when (A.78) holds, we have $p_{n-1,i}^*(z, dy) = K_{n-1,i}^*$, from which

$$
(A.81) \quad q_{ij} = \frac{1}{n} \left[ - \delta_{ij} g_0^{(n-1,i)} - \alpha g_1^{(n-1,i)} K_{n-1,i} \delta_{ij} \right]
$$

$$
= \frac{1}{n} \left[ - \delta_{ij} (1 - (n-1)g_1^{(n-1,i)}) - \alpha g_1^{(n-1,i)} \right]
$$

$$
= \frac{1}{n} \left[ - \delta_{ij} \left( 1 - (n-1) \left( \frac{1}{n-1} - \frac{\beta/s}{(n-1)^2} \right) \right) \right.
$$

$$
- \alpha \left( \frac{1}{n-1} - \frac{\beta/s}{(n-1)^2} \right) + o(n^{-2})
$$

$$
= - \frac{\delta_{ij} \beta/s}{n(n-1)} - \frac{\alpha}{n(n-1)} + o(n^{-2})
$$

where the second equality follows from (12) and the third from (26). Once again it is clear that the key point for determining the limiting behaviour of the diffusion is the second order approximation of the predictive weights, as obtained in Proposition 3.2. Hence we have

$$
\sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} z_j = - \sum_{i=1}^{n} \left[ \frac{z_i \beta/s}{n(n-1)} + \frac{\alpha}{n(n-1)} + o(n^{-2}) \right]
$$

from which (A.79), substituting (26) in the third term, reduces to

$$
\mathcal{B}_{n,1} f_1 = \sum_{i,j=1}^{n} \left[ n^{-2} - O(n^{-3}) \right] z_i (\delta_{ij} - z_j) \frac{\partial^2 f_1}{\partial z_i \partial z_j}
$$
\[
- \sum_{i=1}^{n} \left[ \frac{z_i \beta}{n(n-1)} + \frac{\alpha}{n(n-1)} + o(n^{-2}) \right] \frac{\partial f_1}{\partial z_i}.
\]

which in turn implies that
\[
\|A_1 f_1 - (n^2/2)B_{n,1} \tilde{r}_n f_1\| \to 0 \tag{A.82}
\]
with \(A_1\) as in (45), \(f_1 \in \mathcal{D}(A_1)\) as in (46) and \(\tilde{r}_n\) as in (67). From (A.73) and (A.82) it follows that
\[
\|A(f_0 \times f_1) - (\tilde{r}_n f_1)n^3/2B_{n,0} f_0 - f_0 n^2B_{n,1} \tilde{r}_n f_1/2\| \to 0,
\]
with \(A\) as in (4) and \((f_0 \times f_1)\) as in (47). The fact that (64) holds in \(D_{[0,\infty) \times \nabla_\infty}([0,\infty))\) follows from the density of (47) in \(C_0([0,\infty) \times \nabla_\infty)\), together with Theorems 1.6.1 and 4.2.11 in Ethier and Kurtz (1986), which imply respectively the strong convergence of the associated semigroups on \(C_0([0,\infty) \times \nabla_\infty)\), similarly to (52), and the weak convergence of the probability measures induced on \(D_{[0,\infty) \times \nabla_\infty}([0,\infty))\). The same assertion with \(C_{[0,\infty) \times \nabla_\infty}([0,\infty))\) in place of \(D_{[0,\infty) \times \nabla_\infty}([0,\infty))\) now follows from relativization of the Skorohod topology. \(\square\)

References


