# On the transition function of some time-dependent Dirichlet and gamma processes

Antonio Lijoi[*] Matteo Ruggiero[†] Dario Spanò[‡]

**Abstract**
We highlight a Bayesian interpretation of the transition functions of two classes of measure valued diffusions widely used in population genetics, given by Fleming–Viot and Dawson–Watanabe models, which describe time evolving random measures in the Dirichlet and gamma families respectively. We review some recent results on temporal conjugacy of these classes under certain assumptions on the data collection at discrete times, and discuss their interpretation in terms of dimensionality reduction of the transition function and of the associated forward propagation of the marginal measure of the process.

**Key Words:** Bayesian nonparametrics, Dawson–Watanabe process, Dirichlet process, Fleming–Viot process, gamma random measure.

## 1. Introduction

A currently active area of research in Bayesian nonparametrics is the construction of so-called dependent processes. These aim at accommodating weaker forms of dependence among the observables than exchangeability, such as partial exchangeability in the sense of de Finetti. Within this goal, the idea is to construct a family of random probability measures indexed by a covariate, so that the data are exchangeable conditionally on the covariate value but not overall exchangeable. This line of research has been inspired by MacEachern (1999, 2000) and has since generated a considerable amount of contributions. Most of these have concentrated their focus on extending to this dependent framework the milestone of Bayesian nonparametrics, the Dirichlet process (Ferguson, 1973). The great amount of activity around this idea is certainly due to the success of the Dirichlet process and associated hierarchical mixtures (Lo, 1982) as modelling approaches for Bayesian inference and its celebrated stick-breaking representation (Sethuraman, 1994), which lends itself to natural extensions to a non-exchangeable framework. Despite not being particularly easy to handle for deriving related analytical quantities, this has solidly proven to be extremely advantageous on the practical side, especially for implementing these models through computer aided posterior computation. The literature on this topic is by now quite vast, so we refer the reader to Hjort et al. (2010) and Müller and Mitra (2013) for reviews and references.

Many contributions have developed somewhat similar dependent structures for gamma random measures. These have been mainly instrumental towards applications which require a transformation of the random measure such as, for example, normalisation or tilting. A smaller amount of research effort has instead been devoted to developing and analysing the properties of dependent gamma random measures *per se*. Among the recent contributions in this respect, Ishwaran and Zarepour (2009) Spanò and Lijoi (2016) study the spectral properties of a dependent gamma model and build new general families of time-dependent gamma priors whose trajectories allow for possible discontinuities, and

[*]University of Pavia
[†]University of Torino and Collegio Carlo Alberto
[‡]University of Warwick

Papaspiliopoulos, Ruggiero and Spanò (2016) who show some analytical conjugacy properties of a class of gamma dependent processes with respect to Poisson point process data collected at discrete times.

In this note, we discuss some aspects of the transition functions of two dependent models for Dirichlet and gamma random measures: the Fleming–Viot and the Dawson–Watanabe process. See Ethier and Kurtz (1993); Dawson (1993); Etheridge (2000); Dawson (2010) for reviews. These have been formulated in the mathematical population genetics literature, but can be easily interpreted as time evolving Dirichlet and gamma random measures respectively, as discussed in the following Sections. Here we are particularly interested in detailing the connections between the transition functions of these models, and the conjugacy results obtained in Papaspiliopoulos, Ruggiero and Spanò (2016) for related hidden Markov measures. Specifically, we will outline an interpretation of such dependent processes from a Bayesian perspective, by relying on a construction strategy based on latent variables due to Pitt, Chatfield and Walker (2002); Pitt and Walker (2005), which makes them particularly appealing and relatively easy to interpret. Then, we will summarise how the transition mechanism for the two models can be reduced to a finite computation under certain assumptions on the collected data, and we will detail the connection between the two results.

## 2. Evolving Dirichlet measures

The above mentioned general recipe for constructing dependent processes using latent variables was used in Walker et al. (2007) for deriving a class of Fleming–Viot measure-valued diffusions, and can be quickly summarised as follows. Let $x$ be a Dirichlet process with base measure $\alpha = \theta P_0$, denoted as $x \sim \Pi_\alpha$, where $\theta > 0$ and $P_0$ is a nonatomic distribution on some space $\mathcal{Y}$ for the observables. For $m \in \mathbb{Z}_+$ consider sampling, given $x$, $m$ independent and identically distributed (*iid*) observations $(Y_1, \dots, Y_m)$ from $x$, and conditional on such observations sample a random measure from the associated posterior Dirichlet law. This can be written as the transition mechanism

$$P(x, \mathrm{d}x' \mid m) = \int_{\mathcal{Y}^m} \Pi_{\alpha+\sum_{i=1}^m \delta_{y_i}}(\mathrm{d}x')x^m(\mathrm{d}y_1, \dots, \mathrm{d}y_m)$$

where $x^m$ denotes the $m$-fold product measure $x \times \cdots \times x$. Randomising $m$, to be chosen with probability $d_m$, say, yields

$$P(x, \mathrm{d}x') = \sum_{m=0}^{\infty} d_m \int_{\mathcal{Y}^m} \Pi_{\alpha+\sum_{i=1}^m \delta_{y_i}}(\mathrm{d}x')x^m(\mathrm{d}y_1, \dots, \mathrm{d}y_m).$$

Let now the distribution of $m$ depend on time, i.e. $d_m = d_m(t)$, which formally gives

$$P_t(x, \mathrm{d}x') = \sum_{m=0}^{\infty} d_m(t) \int_{\mathcal{Y}^m} \Pi_{\alpha+\sum_{i=1}^m \delta_{y_i}}(\mathrm{d}x')x^m(\mathrm{d}y_1, \dots, \mathrm{d}y_m), \qquad t \geq 0. \tag{1}$$

If the probabilities $d_m(t)$ are such that (1) satisfies the Chapman–Kolmogorov equations, then $P_t$ is the transition function of a well-defined Markov process taking values in the space of discrete probability measures on $\mathcal{Y}$. It will also have Dirichlet marginals, by Corollary 1.1 in Antoniak (1974). Supposing in addition that $d_m(t)$ is the probability that a death process $M_t$, which starts at infinity and jumps from $m$ to $m-1$ at rate

$$\lambda_m = m(\theta + m - 1)/2, \tag{2}$$

is in $m$ at time $t$, then (1) is the transition function of a Fleming–Viot process. See Ethier and Griffiths (1993). The precise form of these probabilities has been described in Tavaré (1984), who computed that for $m \in \mathbb{N}$

$$d_m(t) = \begin{cases} 1 - \sum_{n=1}^{\infty}(-1)^{n-1}(\theta)_{(n-1)}n!^{-1}\gamma_{n,t,\theta} & m = 0, \\ \sum_{n=m}^{\infty}(-1)^{n-m}\binom{n}{m}(\theta+m)_{(n-1)}n!^{-1}\gamma_{n,t,\theta}, & m \in \mathbb{N}, \end{cases} \tag{3}$$

where $\gamma_{n,t,\theta} = (\theta + 2n - 1)e^{-\lambda_n t}$ and $a_{(n)} = a(a+1)\dots(a+n-1)$ for $n \in \mathbb{N}$ is the Pochhammer symbol, with $a_{(0)} = 1$. The interpretation of these $d_m(t)$ in the context of dependent random measures is that a larger $t$ implies sampling a lower amount of information from $x$ with higher probability, resulting in fewer atoms shared by $x$ and $x'$. The starting and arrival states therefore have correlation which decreases in $t$ and is controlled by $d_m(t)$. As $t \to 0$, infinitely many samples are drawn from $x$, and $x'$ will coincide with $x$. The trajectories of the process are continuous in total variation norm (Ethier and Kurtz, 1993). As $t \to \infty$, the fact that the death process which governs the probabilities $d_m(t)$ in (1) is eventually absorbed in 0 implies that $P_t(x, \mathrm{d}x') \to \Pi_\alpha$ as $t \to \infty$, so $x'$ is sampled from the prior $\Pi_\alpha$. Therefore this Fleming–Viot is stationary with respect to $\Pi_\alpha$ (in fact, it is also reversible). See also Jenkins and Spanò (2016) for exact simulation of these processes, where the difficulty lies on dealing with the $d_m(t)$ probabilities, Favaro et al. (2009); Ruggiero and Walker (2009a,b) for different constructions of Fleming–Viot related models, Mena and Ruggiero (2016); Mena et al. (2011) for different classes of measure-valued diffusions with applications to Bayesian nonparametrics.

There is a dynamic counterpart of the well known projective property of Dirichlet processes, which yield Dirichlet distributions, for these processes. Projecting a Fleming–Viot process $X_t$ onto a measurable partition $A_1, \dots, A_K$ of $\mathcal{Y}$ yields a $K$-dimensional Wright–Fisher diffusion, denoted here $\mathbf{X}_t$, which is reversible and stationary with respect to the Dirichlet distribution $\pi_{\boldsymbol{\alpha}} := \text{Dir}(\alpha_1, \dots, \alpha_K)$, for $\alpha_i = \theta P_0(A_i)$, $i = 1, \dots, K$. See (Dawson, 2010; Etheridge, 2009). Consistently, the transition function of a Wright–Fisher process is obtained by specialising (1) to

$$P_t(\mathbf{x}, \mathrm{d}\mathbf{x}') = \sum_{m=0}^{\infty} d_m(t) \sum_{\mathbf{m} \in \mathbb{Z}_+^K : \sum_i m_i = m} \binom{m}{\mathbf{m}} \mathbf{x}^{\mathbf{m}} \pi_{\boldsymbol{\alpha}+\mathbf{m}},$$

where $\mathbf{m} = (m_1, \dots, m_K)$ and $\pi_{\boldsymbol{\alpha}+\mathbf{m}} := \text{Dir}(\alpha_1 + m_1, \dots, \alpha_K + m_K)$. See Ethier and Griffiths (1993). The interpretation is analogous to that of (1) in a finite dimensional framework.

Let $x \sim \Pi_\alpha$ be the current state of the process, and suppose at time $t_0 = 0$, conditionally *iid* data $Y_1, \dots, Y_m$ are collected such that $Y_i | x \sim x$. Suppose also the data feature $K_m \le m$ distinct values $Y_1^*, \dots, Y_{K_m}^*$, where $Y_i^*$ has multiplicity $m_i$. By a well known result of (Ferguson, 1973), the posterior law is still Dirichlet with updated parameters, namely

$$x \mid y_1, \dots, y_m \sim \Pi_{\alpha + \sum_{i=1}^{K_m} m_i \delta_{y_i^*}}$$

Papaspiliopoulos, Ruggiero and Spanò (2016) showed that a Dirichlet dependent process $X_t$ of Fleming–Viot type is conjugate with respect to these type of data if $x$ above is the process state, i.e. $X_t = x$. Specifically, before additional data are collected, we have that the law of the process $X$ at time $t_0 + t$ can be written as

$$\psi_t\left(\Pi_{\alpha + \sum_{i=1}^{K_m} m_i \delta_{y_i^*}}\right) = \sum_{\mathbf{0} \le \mathbf{i} \le \mathbf{m}} p_{\mathbf{m}, \mathbf{m}-\mathbf{i}}(t) \Pi_{\alpha + \sum_{j=1}^{K_m} (m_j - i_j)\delta_{y_j^*}}. \tag{4}$$

Here we have denoted

$$\psi_t(\nu)(\mathrm{d}x') := \int_{\mathcal{X}} \nu(\mathrm{d}x) P_t(x, \mathrm{d}x')$$

the forward propagation of the law $\nu$ of $x$, obtained by means of the Fleming–Viot transition operator $P_t$. Furthermore, "$\leq$" denotes a partial ordering whereby for $\mathbf{i}, \mathbf{m} \in \mathbb{Z}_+^{K_m}$, $\mathbf{i} < \mathbf{m}$ if $i_j \leq m_j$ for all $j = 1, \dots, K_m$ and $i_j < m_j$ for some $j$. Denoting for short $|\mathbf{m}| = \sum_j m_j$, the weights $p_{\mathbf{m},\mathbf{m}-\mathbf{i}}(t)$ equal

$$p_{\mathbf{m},\mathbf{m}-\mathbf{i}}(t) = \begin{cases} e^{-\lambda_{|\mathbf{m}|}t}, & \mathbf{i} = \mathbf{0} \\ C_{|\mathbf{m}|,|\mathbf{m}|-|\mathbf{i}|}(t) p(\mathbf{i};\, \mathbf{m}, |\mathbf{i}|), & \mathbf{0} < \mathbf{i} \leq \mathbf{m}, \end{cases} \tag{5}$$

with

$$C_{|\mathbf{m}|,|\mathbf{m}|-|\mathbf{i}|}(t) = \left( \prod_{h=0}^{|\mathbf{i}|-1} \lambda_{|\mathbf{m}|-h} \right) (-1)^{|\mathbf{i}|} \sum_{k=0}^{|\mathbf{i}|} \frac{e^{-\lambda_{|\mathbf{m}|-k}t}}{\prod_{0 \leq h \leq |\mathbf{i}|, h \neq k} (\lambda_{|\mathbf{m}|-k} - \lambda_{|\mathbf{m}|-h})},$$

with $\lambda_n$ as in (2) and

$$p(\mathbf{i};\, \mathbf{m}, |\mathbf{i}|) = \binom{|\mathbf{m}|}{|\mathbf{i}|}^{-1} \prod_{j \geq 1} \binom{m_j}{i_j} \tag{6}$$

being the multivariate hypergeometric probability function, with parameters $(\mathbf{m}, |\mathbf{i}|)$, evaluated at $\mathbf{i}$. The important point to notice in (4) is that despite the transition function (1) of the Fleming–Viot process has an infinite series expansion, under the above assumptions the forward propagation of the prior Dirichlet law for $X_{t_0}$ to time $t_0 + t$, given data collected at time $t_0$, can be still expressed as a finite sum of Dirichlet measures. This in particular implies that alternating the operations of conditioning to the collected data, and propagating forward the current law of the random measure to the next data collection time, maintains the marginal law of the random measure inside the family of finite mixtures of Dirichlet processes.

The result in Papaspiliopoulos, Ruggiero and Spanò (2016) is obtained by exploiting a specific structure hidden in the time-reversal of the Fleming–Viot process, namely a dual process, which ultimately allows to drastically reduce the dimensionality of the transition operation. The opposite connection can be outlined as follows. Let $\mathbf{l} = \mathbf{m} - \mathbf{i}$. Then as $|\mathbf{m}| \to \infty$

$$C_{|\mathbf{m}|,|\mathbf{l}|}(t) \to d_{|\mathbf{l}|}(t)$$

with $d_m(t)$ as in (3). See Tavaré (1984), Section 6. Since (6) with $\mathbf{l}$ as above equals

$$\frac{\binom{m_1}{m_1 - l_1} \cdots \binom{m_K}{m_K - l_K}}{\binom{|\mathbf{m}|}{|\mathbf{m} - \mathbf{l}|}} = \frac{m_{1,[l_1]}}{m} \cdots \frac{m_{K,[l_k]}}{m - k + 1} \frac{|l|}{l_1! \cdots l_K!},$$

where $a_{[n]} = a(a+1)\dots(a+n-1)$ denotes the descending factorial and $a_{i,[n]}$ the same applied to $a_i$, it follows that by assuming $\mathbf{m}/|\mathbf{m}| \to \mathbf{x}$ we obtain

$$p_{\mathbf{m},\mathbf{l}}(t) \to d_{|\mathbf{l}|}(t) \binom{|\mathbf{l}|}{\mathbf{l}} \mathbf{x}^{\mathbf{l}}.$$

Defining now

$$x(\cdot) := \sum_{i=1}^{\infty} x_i \delta_{y_i^*}(\cdot)$$

one can see that $\Pi_{\alpha + \sum_{j=1}^{K_m} m_j \delta_{y_j^*}}$ degenerates to $x$ as $|\mathbf{m}| \to \infty$ and (4) converges to (1).

## 3. Evolving gamma measures

The class of Dawson–Watanabe processes can be considered, roughly speaking, as the gamma counterpart of Fleming–Viot processes, admitting therefore interpretation as dependent models for gamma random measures. More formally, they belong to the class of branching measure-valued diffusions taking values in the space of finite discrete measures. As in the Fleming–Viot case, they describe evolving discrete measures whose support varies with time and whose masses are each a positive diffusion and sum up to a finite quantity. See Dawson (1993); Li (2011) for reviews. Here we are interested in the special case of Dawson–Watanabe processes who admit the law of a gamma random measure as invariant measure, which correspond to subcritical branching with immigration. The transition function of a Dawson–Watanabe process admits a construction similar to that outlined for Fleming–Viot processes. Let $z$ on $\mathcal{Y}$ be a gamma random measure with shape measure $\alpha = \theta P_0$, with $\theta, P_0$ as in Section 2, and rate parameter $\beta > 0$, denoted $z \sim \Gamma_\alpha^\beta$, so that projections onto disjoint sets $A_1, \ldots, A_K \subset \mathcal{Y}$ yield independent gamma variables $z(A_i) \sim^{ind} \mathrm{Ga}(\alpha(A_i), \beta)$. Denoting $(z/|z|)^m$ the $m$-fold product of the normalised measure, where $|z| := z(\mathcal{Y})$ is the total mass of $z$, the transition function reads

$$P_t(z, \mathrm{d}z') = \sum_{m=0}^{\infty} d_m^{|z|,\beta}(t) \int_{\mathcal{Y}^m} \Gamma^{\beta+S_t^*}_{\alpha+\sum_{i=1}^m \delta_{y_i}}(\mathrm{d}z')(z/|z|)^m(\mathrm{d}y_1, \ldots, \mathrm{d}y_m). \tag{7}$$

Here, similarly to (1), a random sample size $m$ is chosen with time-dependent probability $d_m^{|z|,\beta}(t)$; conditional on $m$, data are collected as above from the normalised starting state, and given these, the arrival state is sampled from $\Gamma^{\beta+S_t^*}_{\alpha+\sum_{i=1}^m \delta_{y_i}}$. Unlike the Fleming–Viot case, not only the probability of $m$ but also a parameter of the gamma law depends on time. Specifically, we have where

$$d_m^{|z|,\beta}(t) = \mathrm{Po}\left(m \,\Big|\, \frac{\beta|z|}{e^{\beta t/2} - 1}\right) \quad \text{and} \quad S_t^* := \frac{\beta}{e^{\beta t/2} - 1}.$$

See Ethier and Griffiths (1993b). The main difference with respect to (1), apart from the different distributions involved, is that since in general $S_t^*$ is not integer–valued, the interpretation as sampling the arrival state $z'$ from a posterior gamma law is not formally correct under the usual Gamma-Poisson conjugate model. The sample size $m$ is chosen with probability $d_m^{|z|,\beta}(t)$, which is the probability that an $\mathbb{N}$-valued death process which starts at infinity at time 0 is in $m$ at time $t$, if it jumps from $m$ to $m-1$ at rate $(m\beta/2)(1-e^{\beta t/2})^{-1}$. Note that this death process admits representation as a time-changed Poisson process, whereby the above formulation. See Ethier and Griffiths (1993b) for details. So $z$ and $z'$ will share fewer atoms the farther they are apart in time. The Dawson–Watanabe process with the above transition is known to be stationary and reversible with respect to the law $\Gamma_\alpha^\beta$ of a gamma random measure. See Shiga (1990); Ethier and Griffiths (1993b).

The Dawson–Watanabe process satisfies a projective property similar to that seen in Section 2 for the Fleming–Viot process. Specifically, a projection of $Z_t$ with transition (7) onto a measurable partition $A_1, \ldots, A_K$ of $\mathcal{Y}$, yields the vector $(Z_t(A_1), \ldots, Z_t(A_K))$ of independent components $Z_t(A_i)$ each driven by a Cox–Ingersoll–Ross (CIR) diffusion (Cox, Ingersoll and Ross, 1985). These are reversible and ergodic with respect to a $\mathrm{Ga}(\alpha_i, \beta)$ distribution, with transition function

$$P_t^{(1)}(z_i, \mathrm{d}z_i') = \sum_{m_i=0}^{\infty} \mathrm{Po}\left(m_i \,\Big|\, \frac{z_i\beta}{e^{\beta t/2} - 1}\right) \mathrm{Ga}\left(\mathrm{d}z' \,\Big|\, \alpha_i + m_i, \beta + S_t^*\right). \tag{8}$$

This is not immediately clear as for (1)-(4). For every $K$ and every collection $(A_1,\ldots,A_K)$ of disjoint measurable sets of $\mathcal{Y}$, the evolution of the components $z_{i,t} = Z_t(A_i)$ for $t \geq 0$ and each $i = 1,\ldots,K$ can be read from (7) to be, for any $\mathbf{z} = (z_1,\ldots,z_K)$, $\mathbf{z}' = (z'_1,\ldots,z'_K)$,

$$
\begin{aligned}
P_t(\mathbf{z},\mathrm{d}\mathbf{z}') &= \sum_{m=0}^{\infty} d_m^{|\mathbf{z}|,\beta}(t)\int_{\mathcal{Y}^n}\Gamma^{\beta+S_t^*}_{\alpha+\sum_{i=1}^m\delta_{y_i}}(\mathrm{d}\mathbf{z}')(\mathbf{z}/|\mathbf{z}|)^m(\mathrm{d}y_1,\ldots,\mathrm{d}y_m).\\
&= \sum_{m=0}^{\infty}\mathrm{Po}\bigg(m\,\bigg|\,\frac{|\mathbf{z}|\beta}{e^{\beta t/2}-1}\bigg)\sum_{|\mathbf{m}|=m}\binom{m}{\mathbf{m}}\left(\frac{\mathbf{z}}{|\mathbf{z}|}\right)^{\mathbf{m}}\prod_{i=1}^{K}\mathrm{Ga}\Big(z'_i\,\Big|\,\alpha_i+m_i,\beta+S_t^*\Big)\\
&= \sum_{m_1+\cdots+m_d=0}^{\infty}\prod_{i=1}^{K}\mathrm{Po}\bigg(m_i\,\bigg|\,\frac{\beta}{e^{\beta t/2}-1}z_i\bigg)\mathrm{Ga}\Big(z'_i\,\Big|\,\alpha_i+m_i,\beta+S_t^*\Big)\\
&= \prod_{i=1}^{K}\sum_{m_i=0}^{\infty}\mathrm{Po}\bigg(m_i\,\bigg|\,\frac{z_i\beta}{e^{\beta t/2}-1}\bigg)\mathrm{Ga}\Big(z'_i\,\Big|\,\alpha_i+m_i,\beta+S_t^*\Big)\\
&= \prod_{i=1}^{K}P_t^{(1)}(z_i,\mathrm{d}z'_i),
\end{aligned}
$$

where $P_t^{(1)}$ is as in (8). See e.g. Spanò and Lijoi (2016) Sections 3 and 5.1 for further details.

Consider now, given the current process state $z$, Poisson data such that

$$
y_i \mid z, m \overset{iid}{\sim} z/|z|, \qquad m \mid z \sim \mathrm{Po}(|z|).
$$

By a well known result of (Lo, 1982), the posterior law is still gamma with updated parameters, namely

$$
z \mid y_1,\ldots,y_m \sim \Gamma^{\beta+1}_{\alpha+\sum_{i=1}^{K_m} m_i\delta_{y_i^*}},
$$

where the notation is as in Section 2. Papaspiliopoulos, Ruggiero and Spanò (2016) showed that a gamma dependent process $Z_t$ of Dawson–Watanabe type is conjugate with respect to these type of data if $z$ above is the process state, i.e. $Z_t = z$. Specifically, before additional data are collected, we have that the law of the process $Z$ at time $t_0 + t$ can be written as

$$
\psi_t\Big(\Gamma^{\beta+s}_{\alpha+\sum_{i=1}^{K_m} m_i\delta_{y_i^*}}\Big) = \sum_{\mathbf{0}\leq\mathbf{i}\leq\mathbf{m}}\tilde{p}_{\mathbf{m},\mathbf{m}-\mathbf{i}}(t)\Gamma^{\beta+S_t}_{\alpha+\sum_{i=1}^{K_m}(m_i-i_i)\delta_{y_i^*}}, \tag{9}
$$

where $\beta + s$ is the value at time $t$ of the rate parameter. Here $\psi_t$ is as in (4) with $P_t$ as in (7),

$$
\begin{aligned}
\tilde{p}_{\mathbf{m},\mathbf{m}-\mathbf{i}}(t) &= \mathrm{Bin}(|\mathbf{m}|-|\mathbf{i}|;\,|\mathbf{m}|,p(t))p(\mathbf{m}-\mathbf{i};\,\mathbf{m},|\mathbf{m}|-|\mathbf{i}|),\\
p(t) &= S_t/S_0, \quad S_t = \frac{\beta S_0}{(\beta+S_0)e^{\beta t/2}-S_0}, \quad S_0 = s,
\end{aligned}
$$

$p(\mathbf{m}-\mathbf{i};\,\mathbf{m},|\mathbf{m}|-|\mathbf{i}|)$ is as in (6) and $\mathrm{Bin}(|\mathbf{i}|;\,|\mathbf{m}|,p(t))$ is a Binomial pmf with parameters $(|\mathbf{m}|,p(t))$ evaluated at $|\mathbf{m}|-|\mathbf{n}|$. Note that as $t \to 0$, we have $p(t) \to 1$ which implies $\mathrm{Bin}(|\mathbf{m}|;\,|\mathbf{m}|,p(t)) \to 1$ and $p(\mathbf{m};\,\mathbf{m},|\mathbf{m}|) = 1$. That is, (9) puts mass one on the component $\Gamma^{\beta+s}_{\alpha+\sum_{i=1}^{K_m} m_i\delta_{y_i^*}}$ which coincides with the starting state. As $t \to \infty$ the opposite occurs, and as $s_t \to 0$, (9) puts mass one on the component $\Gamma^{\beta}_{\alpha}$ which coincides with the prior. See Papaspiliopoulos, Ruggiero and Spanò (2016), Section 1.3 for an illustration of this point.

A similar interpretation as that for the Fleming–Viot case holds here. Despite the transition function expansion of the Dawson–Watanabe process is in the form of an infinite series, under these assumptions on the data collection the propagation operation (5) can be reduced to a finite operation. The weights of the mixture have a different form, and in particular they also feature a deterministic component $S_t$ which has no analog in (5). The alternation of Bayesian updating and propagation leave therefore the marginal law of $Z_t$ inside the family of finite mixtures of gamma random measures.

In order to obtain the transition function from (9), as done for the Fleming–Viot process, let $\mathbf{l} := \mathbf{m} - \mathbf{i}$, and assume that, as $|\mathbf{m}| \to \infty$, $|\mathbf{m}|/s \to |z| > 0$. Then, since $S_t \to \beta\big(e^{\beta t/2} - 1\big)^{-1}$, we have

$$|\mathbf{m}|\frac{S_t}{s} \to \frac{\beta|z|}{e^{\beta t/2} - 1}.$$

The Poisson limit property of the binomial distribution now implies that

$$\text{Bin}(|\mathbf{l}|;\, |\mathbf{m}|, p(t)) \to \text{Po}\left(|\mathbf{l}| \,\Big|\, \frac{\beta|z|}{e^{\beta t/2} - 1}\right)$$

and one can see that (9) converges to the transition function (7) as expected.

## Acknowledgments

A. Lijoi and M. Ruggiero are supported by the European Research Council (ERC) through StG "N-BNP" 306406.

## REFERENCES

ANTONIAK, C.E. (1974). Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *Ann. Statist.* **2**, 1152–1174.

CARON, F. and TEH, Y.W. (2012). Bayesian nonparametric models for ranked data. *Neural Information Processing Systems (NIPS 2012)*, Lake Tahoe, USA, 2012.

COX, J.C., INGERSOLL, J.E. and ROSS, S.A. (1985). A theory of the term structure of interest rates. *Econometrica* **53**, 385–407.

DAWSON, D.A. (1993). *Measure-valued Markov processes*. Ecole d'Eté de Probabilités de Saint Flour XXI. Lecture Notes in Mathematics **1541**. Springer, Berlin.

DAWSON, D.A. (2010). *Introductory lectures on stochastic population systems*. Technical Report Series **451**, Laboratory for Research in Statistics and Probability, Carleton University.

ETHERIDGE, A.M. (2009). *Some mathematical models from population genetics*. École d'été de Probabilités de Saint-Flour XXXIX. Lecture Notes in Math. **2012**. Springer.

ETHERIDGE, A.M. (2000). *An introduction to superprocesses*. University Lecture Series, 20. American Mathematical Society, Providence, RI.

ETHIER, S.N. and GRIFFITHS, R.C. (1993). The transition function of a Fleming–Viot process. *Ann. Probab.* **21**, 1571–1590.

ETHIER, S.N. and GRIFFITHS, R.C. (1993b). The transition function of a measure-valued branching diffusion with immigration. In *Stochastic Processes. A Festschrift in Honour of Gopinath Kallianpur* (S. Cambanis, J. Ghosh, R.L. Karandikar and P.K. Sen, eds.), 71-79. Springer, New York.

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

FAVARO, S., RUGGIERO, M. and WALKER, S.G. (2009). On a Gibbs sampler based random process in Bayesian nonparametrics. *Electron. J. Statist.* **3**, 1556–1566.

FERGUSON, T.S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1**, 209–230.

HJORT, N.L., HOLMES, C.C., MÜLLER, P. AND WALKER, S.G., eds. (2010). *Bayesian Nonparametrics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge: Cambridge Univ. Press.

ISHWARAN, H. and ZAREPOUR, M. (2009). Series representations for multivariate generalized gamma processes via a scale invariance principle. *Statist. Sinica* **19**, 1665–1682.

JENKINS, P.A. and SPANÒ, D. (2016). Exact simulation of the Wright-Fisher diffusion. *Ann. Appl. Probab.*, to appear.

KAWAZU, K. and WATANABE, S. (1971). Branching processes with immigration and related limit theorems. *Theory Probab. Appl.* **16**, 36–54.

LI, Z. (2011). *Measure-valued branching Markov processes*. Springer, Heidelberg.

LO, A.Y. (1982). Bayesian nonparametric statistical inference for Poisson point process. *Z. Wahrsch. Verw. Gebiete* **59**, 55–66.

LO, A.Y. (1984). On a class of Bayesian nonparametric estimates: I. Density estimates. *Ann. Statist.* **12**, 351–357.

LO, A.Y. and WENG, C.-S. (1989). On a Class of Bayesian Nonparametric Estimates: II. Hazard rate Estimates. *Ann. Inst. Statist. Math.* **41**, 227–245.

MACEACHERN, S.N. (1999). Dependent Nonparametric Processes. In *ASA Proceedings of the Section on Bayesian Statistical Science*. American Statist. Assoc., Alexandria, VA.

MACEACHERN, S.N. (2000). Dependent Dirichlet processes. *Tech. Rep.*, Ohio State University.

MENA, R.H. and RUGGIERO, M. (2016). Dynamic density estimation with diffusive Dirichlet mixtures. *Bernoulli* **22**, 901–926.

MENA, R.H., RUGGIERO, M. and WALKER, S.G. (2011). Geometric stick-breaking processes for continuous-time Bayesian nonparametric modeling. *J. Statist. Plann. Inf.* **141**, 3217–3230.

MÜLLER, P. and MITRA, R. (2013). Bayesian nonparametric inference - why and how. *Bayesian Anal.* **8**, 269–302.

PAPASPILIOPOULOS, O. and RUGGIERO, M. (2014). Optimal filtering and the dual process. *Bernoulli* **20**, 1999–2019.

PAPASPILIOPOULOS, O., RUGGIERO, M. and SPANÒ, D. (2016). Conjugacy properties of time-evolving Dirichlet and gamma random measures. *Electron. J. Statist.*, to appear.

PITT, M.K., CHATFIELD, C. and WALKER, S.G. (2002). Constructing first order stationary autoregressive models via latent processes. *Scand. J. Statist.* **29**, 657–663.

PITT, M.K. and WALKER, S.G. (2005). Constructing stationary time series models using auxiliary variables with applications. *J. Amer. Statist. Assoc.* **100**, 554–564.

RUGGIERO, M. and WALKER, S.G. (2009a). Bayesian nonparametric construction of the Fleming–Viot process with fertility selection. *Statist. Sinica*, **19**, 707–720.

RUGGIERO, M. and WALKER, S.G. (2009b). Countable representation for infinite-dimensional diffusions derived from the two-parameter Poisson–Dirichlet process. *Elect. Comm. Probab.* **14**, 501–517.

SETHURAMAN, J. (1994). A constructive definition of the Dirichlet process prior. *Statist. Sinica* **2**, 639–650.

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

SPANÒ, D. and LIJOI, A. (2016). Canonical correlations for dependent gamma processes. arXiv:1601.06079.

TAVARÉ, S. (1984). Line-of-descent and genealogical processes, and their applications in population genetic models. *Theoret. Population Biol.* **26**, 119–164.

WALKER, S.G., HATJISPYROS S.J. and NICOLERIS, T. (2007). A Fleming–Viot process and Bayesian nonparametrics. *Ann. Appl. Probab.* **17**, 67–80.