


A TRANSPORT PROCESS ON GRAPHS AND ITS LIMITING DISTRIBUTIONS

LEONARDO VIDELA ,* *Universidad de Santiago de Chile*

Abstract

Given a finite strongly connected directed graph $G = (V, E)$, we study a Markov chain taking values on the space of probability measures on V . The chain, motivated by biological applications in the context of stochastic population dynamics, is characterized by transitions between states that respect the structure superimposed by E : mass (probability) can only be moved between neighbors in G . We provide conditions for the ergodicity of the chain. In a simple, symmetric case, we fully characterize the invariant probability.

Keywords: Markov chains on the simplex; Dirichlet distribution

2020 Mathematics Subject Classification: Primary 60J05; 60B10
Secondary 92D25; 92D40

1. Introduction and motivation

1.1. Description

Let $G = (V, E)$ be a strongly connected directed graph without loops. This means that G is a graph with directed edges such that for any two different vertices there exists a directed path joining them. We endow G with a probability vector indexed by E , $\mathbf{p} := (p_{ij} : ij \in E)$. This means that $p_{ij} > 0$ if and only if $ij \in E$, and $\sum_{ij} p_{ij} = 1$. Hereafter, we refer to a pair (G, \mathbf{p}) as a *transport scheme on G* .

Assume that $|V| = m \geq 2$. Let Δ_V be the set of probability measures on V . We identify Δ_V with the $(m - 1)$ -dimensional simplex, and consequently we endow Δ_V with the Euclidean topology inherited from \mathbb{R}^{m-1} . Observe that since Δ_V is finite-dimensional, this topology and the weak topology of probability measures coincide in our case. In many situations we freely identify V with the set $[1, \dots, m]$, without further mention of this convention. We use the notation Leb_{m-1} to denote the Lebesgue measure on Δ_V . The space of continuous (resp. bounded Borel) functions on Δ_V will be denoted by $\mathcal{C}(\Delta_V)$ (resp. $\mathcal{B}(\Delta_V)$).

Given

$$\mu_0 = (\mu_0^{(1)}, \mu_0^{(2)}, \dots, \mu_0^{(m)}) \in \Delta_V,$$

in this paper we aim to study the following dynamics. For $n \geq 0$, given $\mu_n \in \Delta_V$, choose an edge $i_{n+1}j_{n+1} := e_{n+1} \in E$ with probability $p_{e_{n+1}}$ and independently draw

Received 24 July 2021; revision received 7 April 2022.

* Postal address: Departamento de Matemática y Ciencia de la Computación, Universidad de Santiago de Chile, Las Sophoras 173, Estación Central, Santiago, Chile. Email address: leonardo.videla@usach.cl

$U_{n+1} \sim \text{Uniform}(0, 1)$. Define μ_{n+1} as

$$\begin{aligned} \mu_{n+1}^{(k)} &= \mu_n^{(k)}, \quad k \neq i_{n+1}, j_{n+1}, \\ \mu_{n+1}^{(i_{n+1})} &= (1 - U_{n+1})\mu_n^{(i_{n+1})}, \\ \mu_{n+1}^{(j_{n+1})} &= \mu_n^{(j_{n+1})} + U_{n+1}\mu_n^{(i_{n+1})} \end{aligned} \tag{1}$$

In words: we move to j_{n+1} a chunk of size uniformly chosen from the heap at vertex i_{n+1} .

From now on, we refer to the process $(\mu_n : n \geq 0)$ as the *transport process on (G, \mathbf{p})* .

The description above has many points in common with other *mass transport* processes on graphs. For example, under the condition of symmetric weights (i.e. $p_{ij} = p_{ji}$), the transport process (or an easily derivable continuous-time version of it) can be regarded as a smooth version of the averaging process (see [1]), in that the expectations obey the same linear dynamics, up to a constant velocity factor (see the closing remarks at the end of this paper). It is known, however, that in the averaging process the limit distribution is the point mass at the uniform distribution, a situation that certainly does not hold in our model. Also, in [14], Smith considered general Gibbs samplers on graphs based on moves between adjacent vertices, and studied rates of convergence to the invariant measure. In his case, this invariant measure is always uniformly distributed. As the reader will see, the relevant distribution in our model (at least in the cases we study here) comes from the Dirichlet family. In [15], Smith studied convergence rates of a sampler with mass exchange on the set of vertices of the complete graph; although his model is strongly tuned to deal with this particular structure, he suggests that a slight modification to his model would yield the Dirichlet distribution as the limiting law. In contrast, our model allows for general connected graphs as the underlying structure for the allowed moves, and we are more interested in fundamental ergodic properties than estimates of convergence rates. In any case, as far as we know, the transport process, as described above, has not been studied in the previous literature.

1.2. Our results

Obviously, the transport process is a Δ_V -valued Markov chain. Let $\mathbb{P}_\mu, \mathbb{E}_\mu$ be the associated probability measure and expectation operator, respectively, when the chain is started from $\mu \in \Delta_V$. Let P be the transition kernel of the chain, that is, $P: \mathcal{B}(\Delta_V) \mapsto \mathcal{B}(\Delta_V)$ is the operator given by

$$PF(\mu) = \mathbb{E}_\mu(F(\mu_1)).$$

As usual, for $x \in \Delta_V$, we adopt the notation $P(x, \cdot)$ to denote the probability measure on Δ_V given by $\mathcal{B}(\Delta_V) \ni A \mapsto P\mathbf{1}_A(x)$. Let $\mathcal{P} := \mathcal{P}(\Delta_V)$ be the space of probability measures on $\mathcal{B}(\Delta_V)$. As is customary, we consider the action of the transition kernel P acting on \mathcal{P} from the right by definition: for $M \in \mathcal{P}$, MP is the unique element $M' \in \mathcal{P}$ such that for every $F \in \mathcal{B}(\Delta_V)$,

$$\int_{\Delta_V} PF(\mu)M(d\mu) = \int_{\Delta_V} F(\mu)M'(d\mu).$$

Recall that a probability measure M is said to be *invariant or stationary* for P if $MP = M$, and it is called a *limiting distribution* if for some initial probability measure π we have $\pi P^n \rightarrow M$ weakly as $n \rightarrow \infty$. In this case we say that π is in the basin of attraction of P . A chain (or its transition operator) is *ergodic* if it admits a unique invariant distribution whose basin of attraction is the whole of \mathcal{P} .

In this article we will focus on the problem of existence, uniqueness, and (at least in a simple case) the identification of the invariant distributions for P . The next results, whose proofs will be provided in the following pages, are the highlights of this article.

Theorem 1. *Assume $|V| \geq 2$. Then P is ergodic, and its unique invariant measure is absolutely continuous with respect to Lebesgue measure.*

Theorem 2. *Assume that $|V| = 2$. Let M be the unique stationary distribution of P . If $(X, 1 - X)$ is distributed as M , then X has the Beta distribution $\text{Beta}(p_{21}, p_{12})$.*

Theorem 3. *Assume that $|V| \geq 2$, and moreover assume that $p_{ij} = p_{ji}$ for every $ij \in E$. Then the unique stationary distribution is given by the Dirichlet distribution with concentration parameter $1/2$.*

1.3. A biological motivation

Although in this paper our presentation can be followed without mention of any particular application, some words about the motivation to study the transport process are in order. Indeed, the transport process arises as a first step in the agenda of characterizing the properties of a larger process regarding stochastic population dynamics. For example, consider a simple Lotka–Volterra model driven by Lévy noise, namely

$$dX_t = \text{diag}(X_t) \left((B + AX_t) dt + \Sigma dW_t + \int_{\mathbb{R}^n \setminus \{0\}} L(X_t, z) \tilde{N}(dt, dz) \right). \tag{2}$$

Here $B \in \mathbb{R}^n$ is a constant vector representing the rates of increase or decrease of the species, A, Σ are constant $n \times n$ real matrices, W is a standard n -dimensional Brownian motion, $L: \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}^n$ is a measurable, reasonable function, and \tilde{N} is a compensated Poisson measure of intensity measure $dt\nu(dz)$, where ν is a finite measure. Stochastic systems such as (2) have been previously studied in the literature (see e.g. [2], [9], and more recently [18]). Under certain conditions on \tilde{N} , it can be proved that the process X is an \mathbb{R}^n -valued right-continuous strong Markov–Feller process, with the vector X_t usually interpreted as the vector of abundances of the n species at time t . Consider the case where the process is intended to model a simple ecological network with one predator (species 1) and $m \geq 2$ prey species (species 2 through $m + 1$). In this situation the constant matrix A , called the interaction matrix, has components $A_{1,j} > 0$ and $A_{j,1} < 0$ for $j = 2, 3, \dots, m + 1$, and these values are interpreted as the average rate of increase (resp. decrease) of the fitness of the predator per encounter with prey j (resp. the fitness of the prey species j per encounter with the predator). Now it has been observed that some predators are able to modify their foraging strategies in accordance with the density of prey species. When the prey species j , say, is becoming less abundant than species k , the predator tends to change its hunting behavior, becoming increasingly specialized in hunting more of k and less of j . In turn, this shift in the interest of the predator modifies the fitness of the species j and k . This phenomenon, known in the ecology literature as prey-switching behavior (PSB: see e.g. [10]), has been studied previously in the mathematical biology literature; see [11] and [16] for a formulation based on deterministic dynamical systems, and, in a simple stochastic setting, [7]. To introduce this effect into our model, we impose dynamics on the interaction matrix, thus obtaining a new matrix-valued process A_t . This means that the interaction matrix, which was a constant in the model (2), now becomes a dynamic object.

Thus the fundamental object of study is now the pair $\Xi_t := (X_t, A_t)$ whose joint dynamics is governed by

$$dX_t = \text{diag}(X_t) \left((B + A_t X_t) dt + \Sigma dW_t + \int_{\mathbb{R}^n \setminus \{0\}} L(X_t, z) \tilde{N}(dt, dz) \right),$$

$$dA_t = \int_{[0,1]} \int_{[0,1]} \sum_{(j,k) \in S} \mathbf{1}_{u \in I_{j,k}(X_{t-})} Z(j, k, A_{t-}, z) N'(du, dt, dz).$$

Here S is the set of ordered pairs of prey species, $I_{j,k}(x)$ are pairwise disjoint sub-intervals of $(0, 1)$ whose endpoints depend continuously on x , and N' is an independent Poisson random measure with intensity measure $\text{Leb} \otimes \text{Leb} \otimes \mathfrak{m}$, with \mathfrak{m} a probability measure on $[0, 1]$. The length of the interval $I_{j,k}$, which depends on the state of the community, coincides with the rate at which a change in the foraging strategy involving species j and k is modified: the longer $I_{j,k}$ is, the more frequently we introduce a change in the coefficients of the interaction matrix through the application of the operator Z . This operator encodes the prey-switching behavior described above. More specifically, if $\eta = (u, T, z)$ is a point mass of the Poisson random measure N' with $u \in I_{j,k}(X_{T-})$, then the interaction matrix is modified as

$$A_T^{(1,j)} = A_{T-}^{(1,j)} - z A_{T-}^{(1,j)},$$

$$A_T^{(j,1)} = A_{T-}^{(j,1)} - z A_{T-}^{(j,1)},$$

$$A_T^{(1,k)} = A_{T-}^{(1,k)} + z A_{T-}^{(1,j)},$$

$$A_T^{(k,1)} = A_{T-}^{(k,1)} + z A_{T-}^{(j,1)},$$

with all the other entries unchanged. Observe that this modification preserves the sign structure of A , and moreover it preserves the total interaction strengths $\sum_j A_t^{(1,j)}$ and $\sum_j A_t^{(j,1)}$. Thus, if \mathfrak{m} is the Lebesgue measure on the unit interval, after normalization the action of Z on both the vector $(A^{(1,j)} : j = 2, \dots, m)$ and $(-A^{(j,1)} : j = 2, \dots, m)$ is precisely the transition prescribed for the transport process in (1). Moreover, it can be argued on a biological basis that the allowed changes in the foraging efforts are not arbitrary: when the preferred prey has reached low densities, the predator would generally tend to switch to a subset of the potential prey species, and this diet preference imposes certain constraints on the underlying structure of allowed *transport of foraging efforts between preys*. In the simplest case, i.e. if the predator is a generalist, the transport of foraging efforts from one prey to another can be modeled, in a first approach, through a transport scheme (G, \mathbf{p}) , with G a directed complete graph and \mathbf{p} a set of weights assigning equal probability to every possible pair of distinct species. In any case, once this underlying structure is fixed, the transport rules in the last display coincide with the transport process on the corresponding transport scheme.

It is well known that in a model such as (2), for any initial condition on the positive orthant, no extinction occurs in finite time a.s., and thus the fundamental question about coexistence translates mathematically into a question about stochastic persistence, namely that no extinction occurs asymptotically. The strongest version of this notion is called strong stochastic persistence, which can be defined as the fact that the system (2) possesses a unique invariant probability measure that does not charge the extinction set $\bigcup_i \{\mathbf{x} : x_i = 0 \text{ for at least one } i\}$. In recent years, much of the literature has successfully tackled the problem of finding operational conditions to ensure strong stochastic persistence of ecological models, at least in the

case of dynamics driven by Brownian motion; see e.g. the fundamental paper [6] and the aforementioned paper [18], which deals with the problem of finding conditions for strong stochastic persistence in a Lévy-driven framework. On the other hand, PSB can be considered as a special case of a Holling type III functional response (see [8]) in which predation is low at low densities but saturates quickly at a high value when prey is abundant (see e.g. [10]). As has been frequently recognized, prey-switching tends to relax or even remove predation pressure on highly impacted taxa, and thus allows them to recover from low densities (see e.g. [3]) as well as decrease prey competition due to a shared predator. Consequently, the ecological stability of natural ecosystems is enhanced [16, 17]. It is thus expected that PSB should *promote persistence* in any suitable model of prey–predator interaction (see [17]). This is the phenomenon we aim to quantify via a stochastic version of prey-switching as described above. Specifically, if PSB enhances stability, we wish to understand how much it facilitates the coexistence of species, in comparison with a situation where prey-switching is absent. As a fundamental step toward this end, we need to characterize, as much as possible, the convergence of the pair (X, A) for suitable initial conditions. In this program, a first step is to investigate the dynamics of the matrix component in the simpler, admittedly unrealistic case where the changes of the matrix are independent of the abundances. This means $I_{j,k}$ constant intervals. As the reader can easily convince herself, this amounts to studying *ergodic properties of the transport process as a Markov chain in its own right*.

This is the aim of the forthcoming sections.

2. Basic analytic properties

In this short section we provide some simple results that guarantee the existence of invariant measures.

For arbitrary functions $f: V \mapsto \mathbb{R}$, consider the Laplace functionals $\Delta_V \ni \mu \mapsto e^{\langle \mu, f \rangle} \in \mathbb{R}_+$. For $\mu \in \Delta_V$ we easily compute

$$\mathbb{E}_\mu(e^{\langle \mu, f \rangle}) = e^{\langle \mu, f \rangle} \left(\sum_{ij \in E} p_{ij} \frac{e^{\mu^{(i)}(f(j)-f(i))} - 1}{\mu^{(i)}(f(j) - f(i))} \right). \tag{3}$$

Observe that if $(\mu(k))_{k \geq 0}$ is a sequence such that $\langle \mu(k), f \rangle \rightarrow \langle \hat{\mu}, f \rangle$ for some probability measure $\hat{\mu}$ and for every f (which is simply weak convergence of probability measures when Δ_V is identified with $\mathcal{P}(V)$), then $\mathbb{E}_{\mu(k)}(e^{\langle \mu, f \rangle})$ converges to $\mathbb{E}_{\hat{\mu}}(e^{\langle \mu, f \rangle})$ for every f as well. On the other hand, since V is finite, the convergence $\langle \mu(k), f \rangle \rightarrow \langle \hat{\mu}, f \rangle$ for some probability measure $\hat{\mu}$ and for every f is equivalent to pointwise convergence of $\mu(k)$ to $\hat{\mu}$ when the elements of Δ_V are considered as vectors indexed by V .

Proposition 1. *The process $(\mu_n: n \geq 0)$ is a Feller chain.*

Proof. By the discussion above, $PG(\mu(k)) \rightarrow PG(\mu)$ whenever G is a Laplace functional and $\mu(k) \rightarrow \mu$ weakly. Now let $F \in \mathcal{C}(\Delta_V)$ be an arbitrary continuous function on Δ_V , fix $\varepsilon > 0$, and again consider a sequence $\mu(k)$ converging to μ weakly. Let G be a finite linear combination of functions of the form $\mu \mapsto e^{\langle \mu, f \rangle}$ that uniformly approximates the function F up to $\varepsilon/3$. Such a function exists by the Stone–Weierstrass theorem. Fix $k_0 > 0$ such that $|PG(\mu(k)) - PG(\mu)|$ for every $k \geq k_0$. Plainly

$$\begin{aligned} &|PF(\mu(k)) - PF(\mu)| \\ &\leq |PF(\mu(k)) - PG(\mu(k))| + |PG(\mu(k)) - PG(\mu)| + |PG(\mu) - PF(\mu)| \end{aligned}$$

$$\begin{aligned} &\leq \mathbb{E}_{\mu(k)}(|F(\mu_1) - G(\mu_1)|) + |PG(\mu(k)) - PG(\mu)| + \mathbb{E}_{\mu}(|G(\mu_1) - F(\mu_1)|) \\ &\leq \frac{\varepsilon}{3} + |PG(\mu(k)) - PG(\mu)| + \frac{\varepsilon}{3}, \end{aligned}$$

and we deduce that the last expression is no greater than ε whenever $k \geq k_0$. The result follows. □

Corollary 1. *P admits invariant measures.*

Proof. Δ_V is compact and, by Proposition 1, the chain is Feller. The result follows from the well-known Bogolyubov–Krylov theorem. □

3. The simplest case

In this section we specialize the previous discussion to the case $|V| = m = 2$, and for simplicity we first consider $p_{ij} = p_{ji} = \frac{1}{2}$ for $i \neq j$. Observe that in this case every probability measure on Δ_V is completely described by a random variable X supported on $[0, 1]$ via the prescription $\mu^{(1)} = X, \mu^{(2)} = 1 - X$. Indeed, it is easily seen that the description of the process $(\mu_n^{(1)} : n \geq 0)$ is equivalent to the next one. Let $(U_n : n \geq 1)$ be a family of independent and identically distributed (i.i.d.) Uniform $([0, 1])$ random variables, and let $(Y_n : n \geq 1)$ be an i.i.d. sequence of Bernoulli $(1/2)$ random variables, both defined on the same probability space (Ω, \mathbb{P}) . Given an initial condition X_0 , consider the discrete-parameter random dynamical systems

$$X_{n+1} = \begin{cases} X_n - X_n U_{n+1} & \text{if } Y_{n+1} = 0, \\ X_n + (1 - X_n) U_{n+1} & \text{if } Y_{n+1} = 1. \end{cases}$$

Then $(X_n : n \geq 0)$ has the same law as $(\mu_n^{(1)})$ for the same initial condition.

Proposition 2. *The unique invariant probability measure of the process $(X_n : n \geq 0)$ is given by the arcsine distribution, i.e. the absolutely continuous probability measure on $(0, 1)$ with density*

$$h(x) = \frac{1}{\pi} \frac{1}{\sqrt{x(1-x)}}, \quad x \in (0, 1).$$

Proof. To ease notation, write $X = X_0, U_1 = U$. For $t \in (0, 1)$, we compute

$$\begin{aligned} \mathbb{P}(X_1 \leq t) &= \frac{1}{2}(\mathbb{P}(X_1 \leq t \mid Y_1 = 0) + \mathbb{P}(X_1 \leq t \mid Y_1 = 1)) \\ &= \frac{1}{2}(\mathbb{P}(X(1 - U) \leq t) + \mathbb{P}(X + U(1 - X) \leq t)) \\ &= \frac{1}{2} \left(\mathbb{P}\left(U \geq \frac{X-t}{X}\right) + \mathbb{P}\left(U \leq \frac{t-X}{1-X}\right) \right) \\ &= \frac{1}{2} \left(\mathbb{P}\left(U \geq \frac{X-t}{X}, X > t\right) + \mathbb{P}(X \leq t) + \mathbb{P}\left(U \leq \frac{t-X}{1-X}, X \leq t\right) \right). \end{aligned} \tag{4}$$

Assume now that X has a density with respect to Lebesgue measure on the unit interval. Let g be such a density, and so $G' = g$ Lebesgue a.s. on $(0, 1)$. Then, by independence between X and U ,

$$\begin{aligned}
 \mathbb{P}(X_1 \leq t) &= \frac{1}{2} \left(\mathbb{P}(X \leq t) + \mathbb{P}\left(U \geq \frac{X-t}{X}, X > t\right) + \mathbb{P}\left(U \leq \frac{t-X}{1-X}, X \leq t\right) \right) \\
 &= \frac{1}{2} \left(G(t) + \int_t^1 \left(1 - \frac{x-t}{x}\right) g(x) \, dx + \int_0^t \left(\frac{t-1}{1-x} + 1\right) g(x) \, dx \right) \\
 &= \frac{1}{2} \left(G(t) + t \int_t^1 \frac{1}{x} g(x) \, dx - (1-t) \int_0^t \frac{1}{1-x} g(x) \, dx + G(t) \right) \\
 &= G(t) + \frac{1}{2} \left(t \int_t^1 \frac{1}{x} g(x) \, dx - (1-t) \int_0^t \frac{1}{1-x} g(x) \, dx \right).
 \end{aligned}$$

Let $\hat{h}(t)$ be the derivative of the left-hand side. Differentiation yields

$$\hat{h}(t) = g(t) + \frac{1}{2} \left(\int_t^1 \frac{g(x)}{x} \, dx - g(t) - g(t) + \int_0^t \frac{g(x)}{1-x} \, dx \right).$$

Thus, if $(X_n : n \geq 0)$ possesses a continuous invariant density, by the above equation it is differentiable and must thus satisfy

$$h' = \frac{h}{2} \left(\frac{1}{1-t} - \frac{1}{t} \right), \tag{5}$$

whose unique solution integrating up to 1 is given by the arcsine law. Thus h is the unique continuous invariant density for the dynamical system $(X_n : n \geq 0)$.

Now let P' be the transition kernel of the chain X_n . For a Borel set A and $x \in (0, 1)$, we easily compute

$$P' \mathbf{1}_A(x) = \frac{1}{2} \left(\frac{\text{Leb}(A \cap [0, x])}{x} + \frac{\text{Leb}(A \cap [x, 1])}{1-x} \right). \tag{6}$$

For small $h > 0$, we obtain the straightforward estimate

$$\left| \frac{\text{Leb}(A \cap [0, x])}{x} - \frac{\text{Leb}(A \cap [0, x+h])}{x+h} \right| \leq \frac{2h}{x+h}.$$

Thus the first term in (6) is the expression of a continuous function on $(0, 1)$. A similar estimate for the second term above shows that $P' \mathbf{1}_A$ is continuous on $(0, 1)$. By a standard approximation argument, $P' f$ is continuous for every bounded Borel function f . Thus P' possesses the strong Feller property. On the other hand, the above computation also shows that every point in $[0, 1]$ is accessible for the chain, in the sense that for every U_x open neighborhood of x , and for every $z \in [0, 1]$, there exists a $k \geq 1$ such that $P^k(z, U_x) > 0$. It is well known that every strong Feller Markov chain with at least one accessible point has at most one invariant measure (see e.g. [5, Proposition 2.7]). We deduce that the arcsine law is the unique invariant probability measure of the process. \square

The above proof is easily extended to the non-symmetric case. Indeed, if we put $p = p_{12}$, $q = p_{21}$, equation (4) above becomes

$$\mathbb{P}(X_1 \leq t) = p \left(\mathbb{P}\left(U \geq \frac{X-t}{X}, X > t\right) + \mathbb{P}(X \leq t) \right) + q \mathbb{P}\left(U \leq \frac{t-X}{1-X}, X \leq t\right),$$

and proceeding through the same steps detailed in the proof, equation (5) becomes

$$h' = h \left(\frac{q}{1-t} - \frac{p}{t} \right).$$

From this point on, the argument works in exactly the same way, and we arrive at the following result.

Proposition 3. *The unique invariant measure of the process $(\mu_k: k \geq 0)$ in the case $m = 2$ is the Beta distribution with density $\text{Beta}(p_{ij}, p_{ji})$, i.e. the probability measure on $[0, 1]$ with density*

$$\frac{1}{B(p_{ij}, p_{ji})} \frac{1}{x^{p_{ij}}(1-x)^{p_{ji}}}.$$

Still in the case $m = 2$, let π be the invariant measure identified above. We now state the geometric convergence of the transition operator to the measure π .

Theorem 4. *There exist constants $C > 0$, $\eta \in (0, 1)$ such that*

$$\|P^n f - \pi f\| \leq C \eta^n \|f - \pi f\|$$

for every bounded measurable function $f: [0, 1] \mapsto \mathbb{R}$.

Proof. Let $x, y \in [0, 1]$, and assume $x < y$. Then

$$\begin{aligned} P'(x, dz) - P'(y, dz) &= dz \left(p \mathbf{1}_{z < x} \frac{1}{x} + q \mathbf{1}_{z > x} \frac{1}{1-x} - p \mathbf{1}_{z < y} \frac{1}{y} - q \mathbf{1}_{z > y} \frac{1}{1-y} \right) \\ &= dz \left(p \mathbf{1}_{z < x} \left[\frac{1}{x} - \frac{1}{y} \right] + q \mathbf{1}_{z > y} \left[\frac{1}{1-x} - \frac{1}{1-y} \right] + \mathbf{1}_{z \in (x, y)} \left[\frac{q}{1-x} - \frac{p}{y} \right] \right). \end{aligned}$$

Let $\varphi: [0, 1] \mapsto \mathbb{R}$ be a measurable function with $|\varphi(x)| \leq 1$ -Lebesgue a.e. By virtue of the above computation,

$$\begin{aligned} &|P'\varphi(x) - P'\varphi(y)| \\ &\leq \frac{p(y-x)}{xy} \int_0^x |\varphi(z)| dz + \frac{q(y-x)}{(1-x)(1-y)} \int_y^1 |\varphi(z)| dz + \frac{|p-px-xy|}{(1-x)y} \int_x^y |\varphi(z)| dz \\ &= (y-x) \left(\frac{p}{y} + \frac{q}{1-x} + \frac{\max(p-px-xy, xy+px-p)}{(1-x)y} \right) \\ &= (y-x) \max \left(\frac{2p}{y}, \frac{2q}{1-x} \right) \\ &= \max \left(2p \left[1 - \frac{x}{y} \right], 2q \left[1 - \frac{1-y}{1-x} \right] \right) \\ &\leq 1. \end{aligned}$$

This is a sufficient condition to apply Harris's theorem in the compact state-space case (see [5]), and this concludes the proof. \square

4. The general case

Let us come back to the general model: $G = (V, E)$ is a directed, strongly connected graph, and \mathbf{p} is a probability vector indexed by the edges with $p_{ij} > 0$ if and only if $ij \in E$. Recall that P is the transition kernel of the chain. Observe that whenever $m > 2$, P is, in general, a degenerate transition operator, in the sense that it *does not* possess a density. Thus, to prove uniqueness of the invariant law, we cannot rely directly on P , and in fact we will need to prove some properties for higher-order iterates of P .

For this part we will need the following condition to hold.

Assumption 1. $uv \in E \Leftrightarrow vu \in E$.

The connectedness of G is the least we can ask from the underlying structure in order to guarantee uniqueness of the invariant measure. As for the above condition, we will need it in order to guarantee that any spanning tree of (the undirected version of) G can be regarded as a strongly connected subgraph of (the directed version of) G .

4.1. A transport algorithm with density

We introduce some extra notation. For any finite set V and $z > 0$, we write $z\Delta_V$ to denote the set of positive vectors indexed by V with $\sum_{v \in V} x_v = z$. The notation is consistent, since $\Delta_V = 1\Delta_V$. Let $G = (V, E)$ be a finite directed graph. For two points $x, y \in z\Delta_V$, we say that there exists a move that takes x into y if there exists an edge $uv \in E$ such that $y_w = x_w$ for every $w \neq u, v$ and $x_u \geq y_u$. In this case we write $x \rightarrow y$. The move is non-trivial if $x_u > y_u$, and in this case we write

$$x \xrightarrow{u, y_u - x_u, v} y.$$

Fix $x \in z\Delta_V$. We say that $y \in z\Delta_V$ is reachable from x if there exists a finite sequence of moves $x \rightarrow x_1 \rightarrow x_2 \rightarrow \dots \rightarrow x_{n-1} \rightarrow y$. Let \mathcal{R}_x be the set of points in $z\Delta_V$ reachable from x in one non-trivial move, and in general, for $j \geq 1$, let \mathcal{R}_x^j be the set of points that are reachable from x in j non-trivial moves, i.e. the points y such that there exist j non-trivial moves that take x into y . Clearly

$$\mathcal{R}_x^j = \bigcup_{y \in \mathcal{R}_x} \mathcal{R}_y^{j-1}.$$

Lemma 1. Let V be a set with $|V| = m$, and let $G = (V, E)$ be a finite strongly connected digraph that satisfies Assumption 1. Let $z > 0$, and fix $x \in z\Delta_V$. Then

$$\text{Leb}_{m-1} \{z\Delta_V \setminus \mathcal{R}_x^{m-1}\} = 0.$$

In words: almost every point is reachable in exactly $m - 1$ non-trivial moves.

Proof. Observe first that we can reduce the problem to the case where G is a tree, since the general situation follows by considering that the moves are constrained to be made on a fixed spanning tree of G .

So, let G be a tree. We will prove the following claim by induction on the number of vertices of G :

(P_m) For every tree $G = (V, E)$ on m vertices, for every $z > 0$, every $x \in z\Delta_V$ and almost every $y \in z\Delta_V$, there exists a sequence of $m - 1$ non-trivial moves that takes x into y .

In order to prove the claim, for $G, m, z, x,$ and y as above, we will define recursively a particular sequence of $m - 1$ non-trivial moves, hereafter denoted $t_{G,z}(x, y),$ that takes x into $y.$

If G is the tree on the vertices $\{1, 2\},$ the claim is obviously true: either $x_1 > y_1$ or $x_2 \geq y_2.$ In the first case let

$$t_{G,z}(x, y) = x \xrightarrow{1, x_1 - y_1, 2} y,$$

and in the second case define

$$t_{G,z}(x, y) = x \xrightarrow{2, x_2 - y_2, 1} y.$$

Observe that equality above arises only in a set of null Lebesgue measure.

Now suppose that (P_k) holds for $k = 2, 3, \dots, m - 1,$ and let G be a tree on m vertices. We proceed by *pruning* the tree of its leaves. Let v be a leaf of G and let u be its unique parent in $G.$ We distinguish two cases.

Case I: $x_v > y_v.$ Define the auxiliary configurations x' and y' on the $m - 1$ vertices of the tree $G \setminus \{v\}$ as follows: $x'_i = x_i$ for $i \neq u, v, x'_u = x_u + x_v - y_v, y'_i = y_i, i \neq v.$ Define $t_{G,z}(x, y)$ as the concatenation of

$$x \xrightarrow{v, x_v - y_v, u} x'$$

followed by $t_{G \setminus \{v\}, z - y_v}(x', y'),$ or with our notation

$$t_{G,z}(x, y) = x \xrightarrow{v, x_v - y_v, u} t_{G \setminus \{v\}, z - y_v}(x', y').$$

Case II: $x_v \leq y_v.$ In this case we define the auxiliary configurations x'' and y'' on the tree $G \setminus \{v\}$ as follows: $x''_i = x_i$ for $i \neq v, y''_i = y_i$ for $i \neq u, v,$ and $y''_u = y_u + y_v - x_v.$ The required sequence $t_{G,z}(x, y)$ is given by the concatenation of $t_{G \setminus \{v\}, z - x_v}(x'', y'')$ followed by the move

$$y'' \xrightarrow{u, y_v - x_v, v} y,$$

or

$$t_{G,z}(x, y) = t_{G \setminus \{v\}, z - x_v}(x'', y'') \xrightarrow{u, y_v - x_v, v} y.$$

In any case, by our induction hypothesis, we have proved that when G is a tree, the transport operation from configuration x to y can be performed in exactly $m - 1$ moves. Some of these moves, of course, may well be trivial moves, in that the second case in the description of the algorithm may occur with equality. We now show that for fixed x this can occur on a set of null Lebesgue measure. For every proper subset $J \subset V$ and constants $\epsilon_i, \eta_i, i \in J$ with $|\epsilon_i| = |\eta_i| = 1,$ define $\mathcal{N}_{J, \epsilon, \eta}$ as the set

$$\mathcal{N}_{J, \epsilon, \eta} := \left\{ y \in \Delta_V : \sum_{i \in J} \epsilon_i y_i = \sum_{j \in J} \eta_j x_j \right\}.$$

Then $\mathcal{N}_{J, \epsilon, \eta}$ is a proper linear sub-manifold of $\Delta_V,$ and thus has null Lebesgue measure. The same holds, of course, for the finite union $\mathcal{N} := \bigcup_{J, \epsilon, \eta} \mathcal{N}_{J, \epsilon, \eta}.$ Now the algorithm we have provided above shows that $t_{G,z}(x, y)$ is made up of less than $m - 1$ non-trivial moves on a y -set where Case II above, at some point in the recursive application of the algorithm, holds with equality. But this set is obviously contained in $\mathcal{N}.$ □

The proof of the above theorem indeed shows that P^{m-1} has a continuous component that has a density with respect to Lebesgue measure. Indeed, by the computation (6), this claim

holds true when G has two vertices. Assume it is true for graphs with at most $m - 1$ vertices for which Assumption 1 holds, and let G be a finite digraph on m vertices satisfying Assumption 1. Let T be a fixed spanning tree of G when considered as an undirected graph. Define the weight of T as $w_T := \prod_{e \in T} p_e$. Fix $x \in \Delta_V$. Let $f_{G,z}(x, dy)$ be the probability that x is taken into a small element $dy \subseteq \Delta_V$, conditioned on the event that the moves are performed on T and according to the sequence $t_{G,z}(x, y)$. Call this event $M(y)$. Clearly, for Lebesgue a.e. y ,

$$\mathbb{P}(M(y)) = \frac{w_T}{(m - 1)!} > 0.$$

Let $A \subseteq \Delta_V$ be a set with $\text{Leb}_{m-1}(A) > 0$. On one hand we have

$$P^{m-1}(x, A) \geq \frac{w_T}{(m - 1)!} \int_A f_{G,z}(x, dy).$$

On the other hand, with the notation of the above theorem, we have

$$\begin{aligned} & f_{G,z}(x, dy_1, dy_2, \dots, dy_{m-1}) \\ &= \mathbf{1}_{x_v > y_v} \frac{dy_v}{x_v} f_{G \setminus \{v\}, 1-y_v}(x', dy'_1, dy'_2, \dots, dy'_m) \\ &+ \mathbf{1}_{x_v \leq y_v} f_{G \setminus \{v\}, 1-x_v}(x', dy'_1, dy'_2, \dots, dy'_m) \frac{dy_v}{y_u + y_v - x_v}. \end{aligned}$$

By the induction hypothesis, both

$$f_{G \setminus \{v\}, 1-y_v}(x', dy'_1, dy'_2, \dots, dy'_m)$$

and

$$f_{G \setminus \{v\}, 1-x_v}(x', dy'_1, dy'_2, \dots, dy'_m)$$

have a density with respect to the measures $dy'_1 dy'_2 \dots dy'_m$ and $dy'_1 dy'_2 \dots dy'_m$, respectively (all the indices but v are present). On the other hand, both measures $dy'_1 dy'_2 \dots dy'_m$ and $dy'_1 dy'_2 \dots dy'_m$ are the image of $dy_1 dy_2 \dots dy_m$ (again, all the indices but v are present) under linear a.s. non-singular transformations, and thus have a density with respect to this measure. We have proved that $f_{G,1}(x, \cdot)$ has a density with respect to Leb_{m-1} . The following result is an immediate consequence.

Theorem 5. For every $A \subseteq \Delta_V$ with $\text{Leb}_{m-1}(A) > 0$, we have

$$P^{m-1}(x, A) > 0.$$

In particular, the chain is Leb_{m-1} -irreducible, that is, if τ_A is the first hitting time of the chain to the set A , then

$$\mathbb{P}(\tau_A < \infty) > 0.$$

We have all the ingredients to prove the following theorem.

Theorem 6. *There exists a unique invariant measure M for the chain $(\mu_n : n \geq 0)$, and for Leb_{m-1} a.e. $x \in \Delta_V$ we have*

$$\lim_{n \rightarrow \infty} \|P^n(x, \cdot) - M(\cdot)\|_{TV} = 0,$$

where $\|\cdot\|_{TV}$ denotes the total variation norm.

Proof. Let M be any invariant measure whose existence is guaranteed by Corollary 1. By the above theorem, P is Leb_{m-1} -irreducible, and thus by Lemma 3 of [12] we have $\text{Leb}_{m-1} \ll M$. Since the chain is obviously aperiodic, by application of the above theorem and Theorem 1 of [12], we obtain that for Leb_{m-1} a.e. $x \in \Delta_V$,

$$\lim_{n \rightarrow \infty} \|P^n(x, \cdot) - M(\cdot)\|_{TV} = 0.$$

On the other hand, since any invariant measure for P is invariant for P^{m-1} , the computation at the end of the last proof shows that every invariant measure has a density with respect to Lebesgue measure. In particular, any ergodic measure is equivalent to Lebesgue measure. Since two distinct ergodic measures must be mutually singular, this shows that M is the unique invariant measure for P . □

5. Identification of the invariant law

For every $i \in \{1, 2, \dots, m\}$, set

$$p_i^{\text{in}} = \sum_{j \neq i} p_{ji}, \quad p_i^{\text{out}} = \sum_{j \neq i} p_{ij}.$$

Also, let Q be the Q -matrix associated with the continuous-time random walk on G whose jump rate from state i to state j is p_{ij} . Recall that we are still under the running hypothesis that G is strongly connected, and thus the random walk associated with Q admits a unique stationary distribution, say $\nu := (\nu_i)_{i=1, \dots, m}$.

Theorem 7. *Let M be the unique invariant distribution for P . Then*

$$\mathbb{E}_M(\mu^{(i)}) = \nu_i, \quad i = 1, \dots, m. \tag{7}$$

Proof. We start from equation (3). To simplify notation, we write $Z_k = \mu^{(k)}$. For fixed $k \in \{1, 2, \dots, m\}$ and $t \in \mathbb{R}$, apply (3) to the function $f: V \mapsto \mathbb{R}$ given by $f(i) = 0$ for $i \neq k, f(k) = t$. Then integrate both sides with respect to the invariant measure M to obtain

$$\begin{aligned} \mathbb{E}_M(e^{Z_k t}) &= \mathbb{E}_M \left(e^{Z_k t} \sum_{i \neq k} \sum_{j \neq i, j \neq k} p_{ij} + \left[\frac{e^{Z_k t} - 1}{Z_k t} \right] p_k^{\text{out}} + \sum_{i \neq k} p_{ik} \left[\frac{e^{(Z_i + Z_k)t} - e^{Z_k t}}{Z_i t} \right] \right) \\ &= \mathbb{E}_M \left(e^{Z_k t} \left[1 - p_k^{\text{out}} - p_k^{\text{in}} \right] + \left[\frac{e^{Z_k t} - 1}{Z_k t} \right] p_k^{\text{out}} + \sum_{i \neq k} p_{ik} \left[\frac{e^{(Z_i + Z_k)t} - e^{Z_k t}}{Z_i t} \right] \right). \end{aligned}$$

Differentiate with respect to t on both sides. By virtue of the Lebesgue-dominated convergence theorem, after evaluation at $t = 0$ we obtain

$$\mathbb{E}_M(Z_k) = \mathbb{E}_M \left(Z_k \left[1 - p_k^{\text{out}} - p_k^{\text{in}} \right] + \frac{Z_k}{2} p_k^{\text{out}} + \sum_{i \neq k} p_{ik} \frac{Z_i}{2} + Z_k p_k^{\text{in}} \right),$$

and this yields

$$\mathbb{E}_M(\mu^{(i)}) = \frac{1}{p_i^{\text{out}}} \sum_{j \neq i} p_{ji} \mathbb{E}_M(\mu^{(j)}),$$

which is equation (7). □

After Proposition 2, intuition dictates that in the case $m \geq 2$, the invariant probability measure should be related in some way to the Dirichlet distribution. Indeed, this educated guess is exactly right in at least one simple case. For the rest of this paper, we will need the following condition to hold.

Assumption 2. *The transport probabilities are symmetric, namely*

$$p_{ij} = p_{ji}, \quad ij \in E.$$

Let M' be the symmetric $1/2$ -Dirichlet distribution. Assume that $X = (X_1, X_2, \dots, X_m) \sim M'$. It is well known that for any multi-index with positive integer components $\beta = (\beta_1, \dots, \beta_m)$, we have

$$\mathbb{E}(X^\beta) = \mathbb{E}\left(\prod_{i=1}^m X_i^{\beta_i}\right) = \frac{\Gamma(m/2)}{\Gamma(1/2)^m \Gamma(m/2 + \sum_{i=1}^m \beta_i)} \prod_{i=1}^m \Gamma(\beta_i + 1/2). \tag{8}$$

Let U be a Uniform $[0, 1]$ random variable independent of X . The action of the semigroup of the process ($\mu_n : n \geq 0$) on the monomials

$$f(x) := x^\beta = \prod_{i=1}^m x_i^{\beta_i}$$

when the initial distribution is M' can be written as

$$M'Pf = \sum_{ij \in E} p_{ij} \mathbb{E}\left((X_i(1 - U))^{\beta_i} (X_j + UX_i)^{\beta_j} \prod_{r \neq i,j} X_r^{\beta_r}\right). \tag{9}$$

Let S_{ij} be the general term of the sum above. Plainly

$$\begin{aligned} S_{ij} &= p_{ij} \mathbb{E}\left((X_i(1 - U))^{\beta_i} (X_j + UX_i)^{\beta_j} \prod_{r \neq i,j} X_r^{\beta_r}\right) \\ &= p_{ij} \sum_{k=0}^{\beta_j} \binom{\beta_j}{k} \mathbb{E}\left((1 - U)^{\beta_i} U^k X_i^{\beta_i} X_j^{\beta_j - k} X_i^k \prod_{r \neq i,j} X_r^{\beta_r}\right) \\ &= p_{ij} \sum_{k=0}^{\beta_j} \binom{\beta_j}{k} \mathbb{E}\left((1 - U)^{\beta_i} U^k\right) \mathbb{E}\left(X_i^{\beta_i + k} X_j^{\beta_j - k} \prod_{r \neq i,j} X_r^{\beta_r}\right), \end{aligned}$$

where the last line follows from the independence between X and U . On the one hand,

$$\mathbb{E}\left((1 - U)^{\beta_i} U^k\right) = B(\beta_i + 1, k + 1) = \frac{\Gamma(\beta_i + 1)\Gamma(k + 1)}{\Gamma(\beta_i + k + 2)}.$$

On the other hand, by (8),

$$\mathbb{E} \left(X_i^{\beta_i+k} X_j^{\beta_j-k} \prod_{r \neq i,j} X_r^{\beta_r} \right) = \frac{\Gamma(m/2)\Gamma(\beta_i+k+1/2)\Gamma(\beta_j-k+1/2)}{\Gamma(1/2)^m \Gamma(m/2 + \sum_{s=1}^m \beta_s)} \prod_{r \neq i,j} \Gamma(\beta_r + 1/2).$$

Thus

$$\begin{aligned} S_{ij} &= p_{ij} \frac{\Gamma(m/2)}{\Gamma(1/2)^m \Gamma(m/2 + \sum_{s=1}^m \beta_s)} \prod_{r \neq i,j} \Gamma(\beta_r + 1/2) \\ &\quad \times \sum_{k=0}^{\beta_j} \left(\frac{\Gamma(\beta_i+1)\Gamma(\beta_i+k+1/2)}{\Gamma(\beta_i+k+2)} \frac{\Gamma(\beta_j+1)\Gamma(\beta_j-k+1/2)}{\Gamma(\beta_j-k+1)} \right) \\ &= p_{ij} Mf \frac{\Gamma(\beta_i+1)\Gamma(\beta_j+1)}{\Gamma(\beta_i+1/2)\Gamma(\beta_j+1/2)} \\ &\quad \times \sum_{k=0}^{\beta_j} \left(\frac{\Gamma(\beta_i+k+1/2)}{\Gamma(\beta_i+k+2)} \frac{\Gamma(\beta_j-k+1/2)}{\Gamma(\beta_j-k+1)} \right). \end{aligned} \quad (10)$$

For a, b non-negative integer numbers, define

$$R(a, b) := \frac{\Gamma(a+1)\Gamma(b+1)}{\Gamma(a+1/2)\Gamma(b+1/2)} \sum_{k=0}^b \left(\frac{\Gamma(a+k+1/2)}{\Gamma(a+k+2)} \frac{\Gamma(b-k+1/2)}{\Gamma(b-k+1)} \right).$$

The following lemma establishes the crucial algebraic identity.

Lemma 2. For arbitrary $a, b \in \mathbb{Z}_+$,

$$R(a, b) + R(b, a) = 2.$$

Proof. A simple computation shows that

$$R(a, b) = \frac{1}{a+1} {}_3F_2 \left[\begin{matrix} 1, a + \frac{1}{2}, -b \\ a+2, -b + \frac{1}{2} \end{matrix}; 1 \right], \quad (11)$$

where ${}_3F_2[;]$ is the generalized (3, 2)-hypergeometric function. By the classic Pfaff–Saalschütz theorem (see [13, pages 48–49]), we have

$${}_3F_2 \left[\begin{matrix} 1, a + \frac{1}{2}, -b \\ a+2, -b + \frac{1}{2} \end{matrix}; 1 \right] = \frac{(a+1)_b \left(\frac{3}{2}\right)_b}{(a+2)_b \left(\frac{1}{2}\right)_b}, \quad (12)$$

where $(x)_n$ is the Pochhammer symbol:

$$(x)_n = x(x+1)(x+2) \cdots (x+n-1).$$

Substitute (12) into (11). After some simplifications, we get

$$R(a, b) = \frac{2b+1}{a+b+1},$$

and analogously

$$R(b, a) = \frac{2a + 1}{a + b + 1}.$$

The result follows. □

We have the ingredients to prove our main result.

Theorem 8. *Assume (G, \mathbf{p}) satisfies Assumption 2. The unique P -invariant probability measure is the symmetric Dirichlet distribution with concentration parameter $1/2$.*

Proof. By equations (9) and (10), for f a monomial, we have

$$\begin{aligned} M'Pf &= M'f \sum_{i < j} (p_{ij}R(\beta_i, \beta_j) + p_{ji}R(\beta_j, \beta_i)) \\ &= M'f \sum_{i < j} (2p_{ij}) \\ &= M'f, \end{aligned}$$

where the last line follows from Assumption 2. By finite additivity, the result is still true for polynomials. Since the class of polynomials is a point-separating algebra of continuous functions on Δ_V , which is a Polish space, in [4, Theorem 3.4.5, a)], we conclude that

$$M'P = M'$$

on $\mathcal{B}(\Delta_V)$. By Theorem 6, $M' = M$, the unique probability measure for P . □

6. Final remarks

We finish this article with two simple remarks that, although not connected with the general topic of this paper, show interesting possibilities for further study.

First, the process just described is, in some sense, related to the well-known averaging process (AP: see the beautiful survey by Aldous and Lanoue [1] for a simple description). Remember some easy facts about the AP. Given a finite connected graph $G = (V, E)$, we attach to every (undirected) edge $e \in E$ a Poisson process N_e of intensity ν_e . Initially, a real number $X_v(0)$ is given to every vertex $v \in V$. The AP is the continuous-time Markov chain $(X_u(t): u \in V, t \geq 0)$ with state-space \mathbb{R}^V that evolves according to the following rule: when the process N_e jumps, say at time T , then we update the values of the incident vertices at e averaging their values at time $T-$. More precisely, if $e = \{u, v\}$, we perform the change

$$X_u(T) = X_v(T) = \frac{X_u(T-) + X_v(T-)}{2}.$$

If the initial values are positive numbers adding up to 1, then the AP can be seen as a Δ_V -valued Markov process. It can be proved that the unique invariant probability measure for this process is the point mass at the Uniform distribution on V . Let Q be the generator of the continuous time random walk with rates $\nu_{(i,j)} = \nu_{(j,i)} = \nu_{\{i,j\}}$. It is straightforward to show that the averages obey the ODE

$$\frac{d}{dt} \mathbb{E}(X(t)) = \frac{1}{2} \mathbb{E}(X(t))Q, \quad t \geq 0. \tag{13}$$

Now the process $(\mu_n: n \geq 0)$ that we have studied in this paper can be put into the setting of a continuous-time process as follows. To every *directed* edge $e = (u, v)$ of G , attach a Poisson process N_e of rate p_{uv} . When the process $N_{(u,v)}$ jumps, transport a uniformly chosen chunk of mass from u to v . Observe that with this prescription for the jump rate of μ_t , the process is running at twice the speed of the Aldous–Lanoue AP. On the other hand, the reader can check that in this case the mean process $\mathbb{E}(\mu_t)$ obeys the same equation (13).

Finally, it would be interesting to devise a strategy allowing us to characterize the unique invariant measure in the non-symmetric case as well. Ideally, such a characterization should be based on the corresponding invariant measure of the underlying random walk. The reader can convince herself that the steps toward the proof of Theorem 8 do not generalize to the setting where the hypothesis of symmetric rates does not hold, and thus we need other tools to achieve the goal. On the other hand, the proof provided above is purely algebraic. It would be desirable to have a probabilistic insight of our results, presumably based on an argument of the stick-breaking type.

Funding information

This research was supported by ANID through Beca de Doctorado 2017-21170406, and partially funded by project ANID-FONDECYT 1200925.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

- [1] ALDOUS, D. AND LANOUE, D. (2012). A lecture on the averaging process. *Prob. Surveys* **9**, 90–102.
- [2] BAO, J., MAO, X., YIN, G. AND YUAN, C. (2012). Competitive Lotka–Volterra population dynamics with jumps. *Nonlinear Anal.* **74**, 6601–6616.
- [3] DUNNE, J. A., MASCHNER, H., BETTS, M. W., HUNTLY, N., RUSSELL, R., WILLIAMS, R. J. AND WOOD, S. A. (2016). The roles and impacts of human hunter-gatherers in North Pacific marine food webs. *Scientific Reports* **6**, 21179.
- [4] ETHIER, S. AND KURTZ, T. (1986). *Markov Processes: Characterization and Convergence*. John Wiley.
- [5] HAIRER, M. (2016) Convergence of Markov processes. Lecture notes, version of 13 January 2016. Available at <http://www.hairer.org/notes/Convergence.pdf>.
- [6] HENING, A. AND NGUYEN, D. H. (2018). Coexistence and extinction for stochastic Kolmogorov systems. *Ann. Appl. Prob.* **28**, 1893–1942.
- [7] HENING, A., NGUYEN, D. H. AND SCHREIBER, S. J. (2022). A classification of the dynamics of three-dimensional stochastic ecological systems. *Ann. Appl. Prob.* **32**, 893–931.
- [8] HOLLING, C. S. (1965). The functional response of predators to prey density and its role in mimicry and population regulation. *Mem. Entomol. Soc. Can.* **45**, 1–60.
- [9] LIU, M. AND WANG, K. (2014). Stochastic Lotka–Volterra systems with Lévy noise. *J. Math. Anal.* **410**, 750–763.
- [10] MURDOCH, W. W. (1969). Switching in general predators: experiments on prey specificity and stability of prey populations. *Ecol. Monog.* **39**, 335–354.
- [11] PILTZ, S., PORTER, M. AND MAINI, P. (2014). Prey switching with a linear preference trade-off. *SIAM J. Appl. Dynamical Systems* **13**, 658–682.
- [12] ROSENTHAL, S. (2001). A review of asymptotic convergence for general state space Markov chains. *Far East J. Theoret. Statist.* **5**, 37–50.
- [13] SLATER, L. (1966). *Generalized Hypergeometric Functions*. Cambridge University Press.
- [14] SMITH, A. (2013). Analysis of convergence rates of some Gibbs samplers on continuous state spaces. *Stoch. Process. Appl.* **123**, 3861–3876.
- [15] SMITH, A. (2014). A Gibbs sampler on the n -simplex. *Ann. Appl. Prob.* **24**, 114–130.

- [16] TERAMOTO, E., KAWASAKI, K. AND SHIGESADA, N. (1979). Switching effect of predation on competitive prey species. *J. Theoret. Biol.* **79**, 303–315.
- [17] VAN BAALEN, M., KRIVAN, V., VAN RIJN, P. AND SABELIS, M. (2001). Alternative food, switching predators, and the persistence of predator–prey systems. *Amer. Naturalist* **157**, 512–524.
- [18] VIDELA, L. (2022) Strong stochastic persistence of some Lévy-driven Lotka–Volterra systems. *J. Math. Biol.* **84**, 11.