On the asymptotic behavior of the Diaconis and Freedman's chain in a multidimensional simplex

Marc Peigné, Tat Dat Tran

On the asymptotic behavior of the Diaconis and Freedman's chain in a multidimensional simplex

Marc Peigné Tat Dat Tran

June 3, 2020

Abstract

In this paper, we give out a setting of an Diaconis and Freedman's chain in a multidimensional simplex and consider its asymptotic behavior. By using techniques in random iterated functions theory and quasi-compact operators theory, we first give out some sufficient conditions which ensure the existence and uniqueness of an invariant probability measure. In some particular cases, we give out explicit formulas of the invariant probability density. Moreover, we completely classify all behaviors of this chain in dimensional two. Eventually, some other settings of the chain are discussed.

MSC2000: 60J05, 60F05

Key words: Iterated function systems, quasi-compact linear operators, absorbing compact set, invariant probability measure, invariant probability density

Contents

1 Introduction

The main motivation in this paper is to propose a general setting for the so called "Diaconis and Freedman's chain" in $\mathbb{R}^d, d \geq 2$. First, we give out the most natural setting of this chain on a d-dimensional simplex and consider its asymptotic behavior by using techniques from random iterated functions theory and quasi-compact operators theory (see [Ladjimi and Peigné, 2019] for using these techniques in dimensional one). We have recently learnt that this multi-dimensional setting is also considered in [Nguyen and Volkov, 2019] where the authors used another approach and consider only the cases of ergodicity. Then, we also discuss some other possible extensions.

Markov chains generated by products of independent random iterated functions have been the object of numerous works for more than 60 years. We refer to [Harris, 1952], [Bush and Mosteller, 1953], [Karlin, 1953] for first models designed for analyzing data in learning, [Dubins and Freedman, 1966], [Guivarc'h and Raugi, 1986], [Letac, 1986], [Mirek, 2011] or [Stenflo, 2012] and references therein; see also [Peigné and Woess, 2011a] and [Peigné and Woess, 2011b] for such processes with weak contraction assumptions on the involved random functions.

In [Diaconis and Freedman, 1999], Diaconis and Freedman focus on the Markov chain $(Z_n)_{n\geq 0}$ on $[0, 1]$ randomly generated by the two families of maps $\mathcal{H} := \{h_t : [0, 1] \to [0, 1], x \mapsto tx\}_{t\in[0,1]}$ and $\mathcal{A} := \{a_t : [0, 1] \to [0, 1], x \mapsto tx + 1 - t\}_{t\in[0,1]}$; at each step, a map is randomly chosen with probability p in the set $\mathcal{H}$ and $q = 1-p$ in the set $\mathcal{A}$, then uniformly with respect to the parameter $t \in [0, 1]$. When the weight p is constant, the random maps (see Section 3 for a detail introduction) which control the transitions of this chain are i.i.d., otherwise the process $(Z_n)_{n\geq 0}$ is no longer in the framework of products of independent random functions. This class of such processes has been studied for a few decades, with various assumptions put on the state space (e.g. compactness) and the regularity of the weight functions. We refer to, for instance, [Kaijser, 1981], [Barnsley and Elton, 1988], [Barnsley et al., 1988], [Barnsley et al., 1989] with connections to image encoding a few years later, and [Kapica and Sleczka, 2017] more recently. All these works concern sufficient conditions for the unicity of the invariant measure and do not explore the case when there are several invariant measures. As far as we know, the coupling method does not seem to be relevant to study this type of Markov chains when there are further invariant measures, or, equivalently, when the space of harmonic functions is not reduced to constant.

For the Diaconis and Freedman' chain in dimension 1, a systematic approach has been developed in [Ladjimi and Peigné, 2019], based on the theory of quasi-compact operators (also described in [Peigné, 1993] and [Hennion and Hervé, 2001]); the authors describe completely the peripheral spectrum of the transition operator P of $(Z_n)_{n\geq 0}$ and use a precise control of the action of the family of functions generated by the sets $\mathcal{H}$ and $\mathcal{A}$ according to p and P. However, a multidimensional setting for such problems has not been touched; it is our aim to introduce and analyse it here.

The paper is organized as follows. In Section 2 we give out our setting of the Diaconis and Freedman's chain in dimension $d \geq 2$. Some properties of the transition operator and its dual operator have been considered and the uniqueness of the stationary density function has been shown (Corollary 4). In Section 3 we give out some results on uniqueness of invariant measures (Theorems 7 and 11) based on concepts and results from the iterated functions system theory. Some special cases where we can find the explicit formula of the unique invariant density are considered in Section 4. Section 5 contains our main result (Theorem 19) where we classify set of invariance probability measures and consider the asymptotic behavior of $(Z_n)_{n\geq 0}$. We discuss some future research directions in Section 6.

2 The Diaconis and Freedman's chain in dimension ≥ 2

In this section we consider a particular setting for the multi-dimensional problem of Diaconis and Freedman's chain. In fact, there are many ways to set which are based on different application models. Our setting here is fit for applications of robot controlling. Other interesting settings as well as their applications will be considered in details in somewhere else. Denote by

$$\Delta_d := \{\mathbf{x} = (x_i)_{1\leq i\leq d} \in \mathbb{R}^d_{\geq 0} : |\mathbf{x}| = x_1 + \cdots + x_d \leq 1\} = co\{\mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_d\}$$

a closed d-dimensional simplex with vertices $\mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_d$, where $\mathbf{e}_0 = (0, \ldots, 0)$ and $\mathbf{e}_i = (0, \ldots, \underbrace{1}_{i^{th}}, \ldots, 0)$ for $1 \leq i \leq d$. From now on, for any $\mathbf{x} \in \Delta_d$, we set $x_0 = 1 - |\mathbf{x}|$; it holds $\mathbf{x} = x_0\mathbf{e}_0 + x_1\mathbf{e}_1 + \ldots + x_d\mathbf{e}_d$ with $x_i \geq 0$ and $x_0 + \ldots + x_d = 1$.

We consider the Markov chain $(Z_n)_{n\geq 0}$ on the simplex Δ_d corresponding to the successive positions of a robot, according to the following rules:

- the robot is put randomly at a point Z_0 in Δ_d;

- if at time $n \geq 0$, it is located at $Z_n = \mathbf{x} \in \Delta_d$, then it chooses the vertex $\mathbf{e}_i, 0 \leq i \leq d$, with probability $p_i(\mathbf{x})$ for the next moving direction and uniformly randomly move to some point on the open line segment $(\mathbf{x}, \mathbf{e}_i) := \{t\mathbf{x} + (1-t)\mathbf{e}_i \mid t \in (0, 1)\}$.

We assume that the functions $p_i, 0 \leq i \leq d$, are continuous and non negative on Δ_d and satisfy $\sum_{i=0}^n p_i(\mathbf{x}) = 1$ for any $\mathbf{x} \in \Delta_d$.

Let us make this description more rigorous. For any $i = 0, \ldots, d$ and $\mathbf{x} \in \Delta_d$, denote by $\mu_i(\mathbf{x}, \cdot)$ the uniform distribution on $(\mathbf{x}, \mathbf{e}_i)$; it is defined on open intervals $(\mathbf{y}_1, \mathbf{y}_2) \in \mathcal{B}((\mathbf{x}, \mathbf{e}_i))$ as

$$\mu_i(\mathbf{x}, (\mathbf{y}_1, \mathbf{y}_2)) := |t(\mathbf{x}, \mathbf{y}_2, \mathbf{e}_i) - t(\mathbf{x}, \mathbf{y}_1, \mathbf{e}_i)|, \tag{1}$$

where the real number $t = t(\mathbf{x}, \mathbf{y}, \mathbf{e}_i) \in (0, 1)$ solves the equality $\mathbf{y} = t\mathbf{x} + (1-t)\mathbf{e}_i$. The *one-step transition probability function* P of $(Z_n)_{n\geq 0}$ is

$$P(\mathbf{x}, d\mathbf{y}) = \sum_{i=0}^{d} p_i(\mathbf{x})\mu_i(\mathbf{x}, d\mathbf{y} \cap (\mathbf{x}, \mathbf{e}_i)), \quad \mathbf{x} \in \Delta_d. \tag{2}$$

We illustrate this setting in Δ_2 in Figure 1.

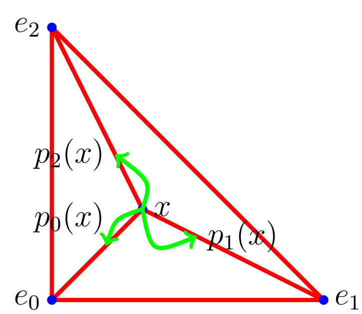

Figure 1: The Diaconis and Freedman's chain in Δ_2.

We want to classify the invariant probability measures of the chain $(Z_n)_{n\geq 0}$ and to describe its behavior as $n \to +\infty$. Our approach is based on the description of the spectrum, on some suitable space to specify, of the operator corresponding to the one-step transition probability function P, also denoted by P. Let us first introduce this transition operator.

We denote by $\mathbb{L}^\infty(\Delta_d, d\mathbf{x})$ the space of all bounded measurable functions $f : \Delta_d \to \mathbb{C}$ and $\mathbb{L}^1(\Delta_d, d\mathbf{y})$ the space of all integrable measurable functions $g : \Delta_d \to \mathbb{C}$; they are Banach spaces, endowed respectively with the norms $\|f\|_\infty := \sup_{\mathbf{x}\in\Delta_d} |f(\mathbf{x})|$ and $\|g\|_1 := \int_{\Delta_d} |g(\mathbf{y})| d\mathbf{y}$.

We also denote by $Den(\Delta_d, d\mathbf{y}) = \{g \in \mathbb{L}^1(\Delta_d, d\mathbf{y}) : g \geq 0 \text{ and } \int_{\Delta_d} g(\mathbf{y})d\mathbf{y} = 1\}$ the space of all probability densities on Δ_d with respect to the reference Lebesgue measure $d\mathbf{y}$. The set $(Den(\Delta_d, d\mathbf{y}), d)$ is a complete metric space for the distance $d(f, g) := \|f - g\|_1$; furthermore, $Den(\Delta_d, d\mathbf{y})$ is a nonempty closed convex subset of the Banach space $\mathbb{L}^1(\Delta_d, d\mathbf{y})$ and it contains the constant function $g(\mathbf{y}) \equiv d!$.

We drop the reference Lebesgue measure $d\mathbf{x}, d\mathbf{y}$ in our notations where no ambiguity arises.

The transition operator of the chain $(Z_n)_{n\geq 0}$ is defined by

$$\begin{aligned} P: \quad & \mathbb{L}^\infty(\Delta_d) \to \mathbb{L}^\infty(\Delta_d) \\ f \quad & \mapsto \left(Pf : \mathbf{x} \to \int_{\Delta_d} f(\mathbf{y})P(\mathbf{x}, d\mathbf{y}) \right). \end{aligned} \tag{3}$$

Its dual operator $P^* : \mathbb{L}^1(\Delta_d) \to \mathbb{L}^1(\Delta_d)$ is defined by

$$\int_{\Delta_d} Pf(\mathbf{x})g(\mathbf{x})d\mathbf{x} = \int_{\Delta_d} f(\mathbf{x})P^*g(\mathbf{x})d\mathbf{x}. \tag{4}$$

Let us explicit the form of these two operators.

Lemma 1. *Let P be the transition operator of $(Z_n)_{n\geq 0}$ and P^* its dual operator. Then*

$$Pf(\mathbf{x}) = \sum_{i=0}^{d} p_i(\mathbf{x}) \int_0^1 f(t\mathbf{x} + (1-t)\mathbf{e}_i)dt \tag{5}$$

and

$$P^*g(\mathbf{y}) = \sum_{i=0}^{d} \int_{1-y_i}^{1} t^{-d} G_i\left(\frac{1}{t}\mathbf{y} + \left(1 - \frac{1}{t}\right)\mathbf{e}_i\right) dt = \sum_{i=0}^{d} \int_1^{\frac{1}{1-y_i}} s^{d-2} G_i\left(s\mathbf{y} + \left(1 - s\right)\mathbf{e}_i\right) ds \tag{6}$$

where $G_i(\mathbf{y}) = g(\mathbf{y})p_i(\mathbf{y})$.

Proof. Equality Eq. (2) yields

$$\begin{aligned} Pf(\mathbf{x}) &= \int_{\Delta_d} f(\mathbf{y})P(\mathbf{x}, d\mathbf{y}) = \sum_{i=0}^{d} p_i(\mathbf{x}) \int_{\Delta_d} f(\mathbf{y})\mu_i(d\mathbf{y} \cap (\mathbf{x}, \mathbf{e}_i)) \\ &= \sum_{i=0}^{d} p_i(\mathbf{x}) \int_0^1 f(t\mathbf{x} + (1-t)\mathbf{e}_i)dt. \end{aligned}$$

For the computation of P^*, we assume $d = 2$; the same argument holds for any d. For all $f \in \mathbb{L}^\infty(\Delta_d)$ and $g \in \mathbb{L}^1(\Delta_d)$,

$$\begin{aligned} \int_{\Delta_2} f(\mathbf{x})P^*g(\mathbf{x})d\mathbf{x} &= \int_{\Delta_2} Pf(\mathbf{x})g(\mathbf{x})d\mathbf{x} \\ &= \sum_{i=0}^{2} \int_{\Delta_2} \left(p_i(\mathbf{x}) \int_0^1 f(t\mathbf{x} + (1-t)\mathbf{e}_i)dt \right) g(\mathbf{x})d\mathbf{x} \\ &= \sum_{i=0}^{2} \int_{\Delta_2} \left(G_i(\mathbf{x}) \int_0^1 f(t\mathbf{x} + (1-t)\mathbf{e}_i)dt \right) d\mathbf{x}. \end{aligned}$$

Let us detail the computation of the term $\int_{\Delta_2}\Big(G_0(\mathbf{x})\int_0^1 f(t\mathbf{x})dt\Big)d\mathbf{x}$; the same calculation holds for the other terms. Namely,

$$\begin{aligned}
\int_{\Delta_2}\Big(G_0(\mathbf{x})\int_0^1 f(t\mathbf{x})dt\Big)d\mathbf{x} &= \int_0^1\left[\int_0^{1-x_1}\Big(G_0(\mathbf{x})\int_0^1 f(t\mathbf{x})dt\Big)dx_2\right]dx_1\\
&= \int_0^1\left[\int_0^1\Big(\int_0^{1-x_1} G_0(\mathbf{x})f(t\mathbf{x})dx_2\Big)dt\right]dx_1\\
&= \int_0^1\left[\int_0^1\Big(\int_0^{1-x_1} G_0(\mathbf{x})f(t\mathbf{x})dx_2\Big)dx_1\right]dt\\
&\overset{y_2=tx_2}{=} \int_0^1\left[\int_0^1\Big(\int_0^{(1-x_1)t} G_0\Big(x_1,\frac{y_2}{t}\Big)f(tx_1,y_2)\frac{dy_2}{t}\Big)dx_1\right]dt\\
&\overset{y_1=tx_1}{=} \int_0^1\left[\int_0^t\Big(\int_0^{t-y_1} G_0\Big(\frac{y_1}{t},\frac{y_2}{t}\Big)f(y_1,y_2)\frac{dy_2}{t}\Big)\frac{dy_1}{t}\right]dt\\
&= \int_0^1\left[\int_{y_1}^1\Big(\int_0^{t-y_1} \frac{1}{t^2}G_0\Big(\frac{y_1}{t},\frac{y_2}{t}\Big)f(y_1,y_2)dy_2\Big)dt\right]dy_1\\
&= \int_0^1\left[\int_0^{1-y_1}\Big(\int_{y_1+y_2}^1 \frac{1}{t^2}G_0\Big(\frac{y_1}{t},\frac{y_2}{t}\Big)f(y_1,y_2)dt\Big)dy_2\right]dy_1\\
&= \int_{\Delta_2} f(\mathbf{y})\Big(\int_{1-y_0}^1 t^{-2}G_0\Big(\frac{1}{t}\mathbf{y}\Big)dt\Big)d\mathbf{y}.
\end{aligned}$$

Similarly $\int_{\Delta_2}\Big(G_i(\mathbf{x})\int_0^1 f(t\mathbf{x}+(1-t)\mathbf{e}_i)dt\Big)d\mathbf{x} = \int_1^{\frac{1}{1-y_i}} G_i\Big(s\mathbf{y}+\Big(1-s\Big)\mathbf{e}_i\Big)ds$ for $i=1,2$ and (6) follows. □

Remark 2. In dimension $d=1$, this is thus the expression of P^* given in [Ladjimi and Peigné, 2019]:

$$P^*g(y) = \int_{1-y}^1 t^{-1}G_1\left(\frac{1}{t}y+\Big(1-\frac{1}{t}\Big)\right)dt + \int_y^1 t^{-1}G_0\left(\frac{1}{t}y\right)dt = \int_0^y \frac{G_1(s)}{1-s}ds + \int_y^1 \frac{G_0(s)}{s}ds.$$

Let us summarize some simple properties of P and P^*.

Proposition 3. *1. The operator P is a Markov operator on $\mathbb{L}^\infty(\Delta_d, d\mathbf{x})$, i.e.*

(i) $Pf \geq 0$ whenever $f \in \mathbb{L}^\infty(\Delta_d, d\mathbf{x})$ and $f \geq 0$;

(ii) $P1 = 1$.

In particular, $\|Pf\|_\infty \leq \|f\|_\infty$ for any $f \in \mathbb{L}^\infty(\Delta_d, d\mathbf{x})$. Furthermore, P is a Feller operator on Δ_d, i.e. $Pf \in C(\Delta_d)$ for all $f \in C(\Delta_d)$.

2. P^* acts on $\mathbb{L}^1(\Delta_d, d\mathbf{y})$ and, for any non negative function $g \in \mathbb{L}^1(\Delta_d, d\mathbf{y})$,

$$P^*g \geq 0 \quad \text{and} \quad \|P^*g\|_1 = \|g\|_1.$$

Furthermore, P^* acts on $Den(\Delta_d, d\mathbf{y})$, i.e., $P^* : Den(\Delta_d, d\mathbf{y}) \to Den(\Delta_d, d\mathbf{y})$ and, for all $g_1 \neq g_2 \in Den(\Delta_d, d\mathbf{y})$,

$$\|P^*g_1 - P^*g_2\|_1 < \|g_1 - g_2\|_1. \tag{7}$$

Proof. The properties of P are quite obvious; in particular, the fact that P is a Feller operator is easily checked from the representation (5) of P. Similarly, the first properties of P^* follow from the definition.

To establish (7), we first recall that $|P^*h| \le P^*|h|$ for any $h \in \mathbb{L}^1(\Delta_d, d\mathbf{y})$, which yields

$$\|P^*h\|_1 \le \|(P^*|f|)\|_1 = \|h\|_1.$$

More precisely,

$$\begin{aligned}
|P^*h| = (P^*h)_+ + (P^*h)_- &= \max\{0, P^*h\} + \max\{0, -P^*h\} \\
&= \max\{0, P^*h_+ - P^*h_-\} + \max\{0, P^*h_- - P^*h_+\} \\
&\le \max\{0, P^*h_+\} + \max\{0, P^*h_-\} = P^*h_+ + P^*h_- = P^*|h|;
\end{aligned}$$

hence, equality $\|P^*h\|_1 = \|h\|_1$ holds if and only if $P^*h_- \equiv 0$ and $P^*h_+ \equiv 0$.

Now, we fix $g_1 \ne g_2 \in Den(\Delta_d, d\mathbf{y})$ and set $h = g_1 - g_2 \not\equiv 0$; it holds $\|P^*g_1 - P^*g_2\|_1 \le \|g_1 - g_2\|_1$. If $\|P^*g_1 - P^*g_2\|_1 = \|g_1 - g_2\|_1$ then $P^*h_- = P^*h_+ \equiv 0$ i.e. $P^*|h| \equiv 0$; therefore $\|h\|_1 = \|(P^*|h|)\|_1 = 0$, so that $h \equiv 0$, contradiction. □

As a direct consequence of (7), we may state the following corollary.

Corollary 4 (Uniqueness of the stationary density function)**.** *If there exists a stationary density function for the Markov chain $(Z_n)_{n\ge 0}$ then it is unique.*

Proof. Assume that there are two different stationary density functions $f \ne g \in Den(\Delta_d, d\mathbf{y})$, i.e. $P^*f = f$ and $P^*g = g$. This implies $d(P^*f, P^*g) = d(f, g)$, contradiction with (7). □

Remark 5.

1. Although $(Den(\Delta_d, d\mathbf{y}), d)$ is a complete metric space, the operator P^* is not uniformly contractive, i.e. there exists $q \in [0, 1)$ such that

$$d(P^*f, P^*g) \le q d(f, g) \quad \forall f, g \in Den(\Delta_d, d\mathbf{y})$$

therefore we can not apply the Banach fixed point theorem. In [Ramli and Leng, 2010, Proposition 2, p. 988-989], the authors applied the Banach fixed point theorem to prove the existence of the stationary density function but their argument does not work. A precise proof can be found in [Ladjimi and Peigné, 2019, Theorem 3.1] which covered all cases of $p_i(\mathbf{x})$ in dimension 1.

2. Although $Den(\Delta_d, d\mathbf{y})$ is a nonempty closed convex subset in a Banach space $\mathbb{L}^1(\Delta_d, d\mathbf{x})$, we can not apply the Browder fixed point theorem, because $\mathbb{L}^1(\Delta_d, d\mathbf{x})$ is not uniformly convex.

3. There are many cases of $p_i(\mathbf{x})$ such that there is no stationary density function for the $(Z_n)_{n\ge 0}$ even in dimension 1: see cases 2 and 3 in [Ladjimi and Peigné, 2019, Theorem 3.1] where the set of invariant probability measures consist of convex combinations of Dirac measures δ_0 and δ_1. It will be interesting to classify cases of $p_i(\mathbf{x})$ so that there exists (unique) a stationary density function. This is still an open question (see the last section of the present paper).

3 Uniqueness of invariant probability measure

In this section, we recall some concepts as well as well-known results of iterated function systems and apply them to our model.

3.1 Iterated function systems with place independent probabilities

Let (E, d) be a compact metric space and denote $\mathbb{L}\text{ip}(E, E)$ the space of Lipschitz continuous functions from E to E, i.e. of functions $T : E \to E$ such that

$$[T] := \sup_{\substack{x,y\in E\\ x\neq y}} \frac{d(T(x), T(y))}{d(x, y)} < \infty.$$

Let $(T_n)_{n\geq 1}$ be a sequence of i.i.d. random functions defined on a probability space $(\Omega, \mathcal{T}, \mathbb{P})$, with values in $\mathbb{L}\text{ip}(E, E)$ and common distribution μ. We consider the Markov chain $(X_n)_{n\geq 0}$ on E, defined by: for any $n \geq 0$,

$$X_{n+1} := T_{n+1}(X_n), \tag{8}$$

where X_0 is a fixed random variable with values in E. One says that the chain $(X_n)_{n\geq 0}$ is generated by the *iterated function system* $(T_n)_{n\geq 1}$. Its transition operator Q is defined by: for any bounded Borel function $\varphi : E \to \mathbb{C}$ and any $x \in E$

$$Q\varphi(x) := \int_{\mathbb{L}\text{ip}(E,E)} \varphi(T(x))\mu(\mathrm{d}T).$$

The chain $(X_n)_{n\geq 0}$ has the "Feller property", i.e. the operator Q acts on the space $C(E)$ of continuous functions from E to $\mathbb{C}$. The maps T_n being Lipschitz continuous on E, the operator Q acts also on the space of Lipschitz continuous functions from E to $\mathbb{C}$ and more generally on the space $\mathcal{H}_\alpha(E), 0 < \alpha \leq 1$, of α-Hölder continuous functions from E to $\mathbb{C}$, defined by

$$\mathcal{H}_\alpha(E) := \{f \in C(E) \mid \|f\|_\alpha := \|f\|_\infty + m_\alpha(f) < +\infty\}$$

where $m_\alpha(f) := \sup\limits_{\substack{x,y\in E\\ x\neq y}} \dfrac{|f(x) - f(y)|}{d(x, y)^\alpha} < \infty$. Endowed, with the norm $\|\cdot\|_\alpha$, the space $\mathcal{H}_\alpha(E)$ is a Banach space.

The behavior of the chain $(X_n)_{n\geq 0}$ is closely related to the spectrum of the restriction of Q to these spaces. Under some "contraction in mean" assumption on the T_n, the restriction of Q to $\mathcal{H}_\alpha(E)$ satisfies some spectral gap property. We first cite the following result in [Ladjimi and Peigné, 2019, Proposition 2.1].

Theorem 6 ([Ladjimi and Peigné, 2019])**.** *Assume that there exists $\alpha \in (0, 1]$ such that*

$$r := \sup_{\substack{x,y\in E\\ x\neq y}} \int_{\mathbb{L}\text{ip}(E,E)} \Big(\frac{d(T(x), T(y))}{d(x, y)}\Big)^\alpha \mu(\mathrm{d}T) < 1. \tag{9}$$

Then, there exists on E a unique Q-invariant probability measure ν. Furthermore, there exists constants $\kappa > 0$ and $\rho \in (0, 1)$ such that

$$\forall\varphi \in \mathcal{H}_\alpha(E),\ \forall x \in E \quad |Q^n\varphi(x) - \nu(\varphi)| \leq \kappa\rho^n. \tag{10}$$

Application to the Diaconis and Freedman's chain for p fixed in Δ_d. We assume $p_i(\mathbf{x}) = p_i$ for all $i = 0, \ldots, d$. We put the Diaconis and Freedman's chain into the framework of iterated random functions as follows. For each $i = 0, \ldots, d$ and $t \in [0, 1]$, we set $H_i(t, \cdot) : \Delta_d \to \Delta_d, \mathbf{x} \mapsto t\mathbf{x} + (1 - t)\mathbf{e}_i$ the affine transformation; these functions $H_i(t, \cdot)$ belong to the space $\mathbb{L}\mathrm{ip}(\Delta_d, \Delta_d)$ of Lipschitz continuous functions from Δ_d to Δ_d, with Lipschitz coefficient $m(H_i(t,\cdot)) = t$. Then, we consider the probability measure μ on $\mathbb{L}\mathrm{ip}(\Delta_d, \Delta_d)$ defined by

$$\mu(dT) := \sum_{i=0}^{d} p_i \int_0^1 \delta_{H_i(t,\cdot)}(dT)dt, \tag{11}$$

where δ_T is the Dirac mass at T. Eq. (5) may be rewritten as

$$\forall f \in \mathbb{L}^\infty(\Delta_d, d\mathbf{x}), \forall \mathbf{x} \in \Delta_d, \quad Qf(\mathbf{x}) = \int_{\mathbb{L}\mathrm{ip}(\Delta_d,\Delta_d)} f(T(\mathbf{x}))\mu(dT).$$

Hence, the Diaconis and Freedman's chain $(Z_n)_{n\geq 0}$ on Δ_d is generated by the *iterated function system* $(T_n)_{n\geq 1}$ in the sense of Eq. (8), where $(T_n)_{n\geq 1}$ be a sequence of i.i.d. random functions with common distribution μ defined by Eq. (11).

Theorem 7. *If $p_i(\mathbf{x}) = p_i$ for all $i = 0, \dots, d$, then the Diaconis and Freedman's chain in Δ_d admits a unique P-invariant probability measure $\nu \in \mathcal{P}(\Delta_d)$. Furthermore, there exists constants $\kappa > 0$ and $\rho \in (0,1)$ such that*

$$\forall \varphi \in \mathcal{H}_\alpha(\Delta_d),\ \forall \mathbf{x} \in \Delta_d \quad |Q^n\varphi(\mathbf{x}) - \nu(\varphi)| \leq \kappa\rho^n. \tag{12}$$

Proof. This is a direct consequence of Theorem 6 with

$$\begin{aligned} r = \sup_{\substack{\mathbf{x},\mathbf{y}\in\Delta_d \\ \mathbf{x}\neq\mathbf{y}}} \int_{\mathbb{L}\mathrm{ip}(\Delta_d,\Delta_d)} \Big(\frac{|T(\mathbf{x}) - T(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|}\Big)^\alpha \mu(\mathrm{d}T) &\leq \sum_{i=0}^{d} p_i \int_0^1 m(H_i(t,\cdot))^\alpha \mathrm{d}t \\ &= \int_0^1 t^\alpha \mathrm{d}t = \frac{1}{1+\alpha} < 1. \end{aligned}$$

□

Remark 8. The unique P-invariant probability measure ν is usually nothing but the Dirichlet distribution $Dir[\theta_0, \dots, \theta_k]$ as will be shown later. If $\theta_i > 0$ for all $i = 0, \dots, k$ we have a unique invariant probability density which is the Dirichlet density. If else, the Dirichlet distribution is singular and can be understood in the sense of [Ferguson, 1973, p. 211], [Ghosh and Ramamoorthi, 2003, Definition 3.1.1, p. 89], or [Jost et al., 2019, Definition 4.2].

3.2 Iterated function systems with place dependent probabilities

In this subsection, we extend the measure μ to a collection $(\mu_x)_{x\in E}$ of probability measures on E, depending continuously on x. We consider the Markov chain $(X_n)_{n\geq 0}$ on E whose transition operator Q is given by: for any bounded Borel function $\varphi : E \to \mathbb{C}$ and any $x \in E$,

$$Q\varphi(x) = \int_{\mathbb{L}\mathrm{ip}(E,E)} \varphi(T(x))\mu_x(\mathrm{d}T).$$

First, we introduce the following definition.

Definition 9. A sequence $(\xi_n)_{n\geq 0}$ of continuous functions from E to E is a contracting sequence if there exist $x_0 \in E$ such that

$$\forall x \in E \quad \lim_{n\to+\infty} \xi_n(x) = x_0.$$

We cite the following result in [Ladjimi and Peigné, 2019, Proposition 2.2].

Theorem 10 ([Ladjimi and Peigné, 2019]). *Assume that there exists $\alpha \in (0,1]$ such that*

H1. $r := \sup_{\substack{x,y\in E\\ x\neq y}} \int_{\mathbb{L}\mathrm{ip}(E,E)} \Big(\frac{d(T(x),T(y))}{d(x,y)}\Big)^\alpha \mu_x(\mathrm{d}T) < 1;$

H2. $R_\alpha := \sup_{\substack{x,y\in E\\ x\neq y}} \frac{|\mu_x - \mu_y|_{TV}}{d(x,y)^\alpha} < +\infty,$ *where $|\mu_x-\mu_y|_{TV}$ is the total variation distance between μ_x and μ_y;*

H3. *There exist $\delta > 0$ and a probability measure μ on E such that*

(i) $\forall x \in E \qquad \mu_x \geq \delta\mu;$

(ii) the closed semi-group T_μ generated by the support S_μ of μ possesses a contracting sequence.

Then, there exists on E a unique Q-invariant probability measure ν; furthermore, for some constants $\kappa > 0$ and $\rho \in (0,1)$, it holds

$$\forall \varphi \in \mathcal{H}_\alpha(E),\ \forall x \in E \quad |Q^n\varphi(x) - \nu(\varphi)| \leq \kappa\rho^n. \tag{13}$$

Let us now apply this statement to the Diaconis and Freedman's chain on Δ_d: for each $\mathbf{x} \in \Delta_d$, we define a space-dependent probability measure $\mu_{\mathbf{x}} \in \mathcal{P}(X)$ by

$$\mu_{\mathbf{x}}(dT) := \sum_{i=0}^d p_i(\mathbf{x}) \int_0^1 \delta_{H_i(t,\cdot)}(dT)dt,$$

where δ_T is the Dirac mass at T. With this collection $(\mu_{\mathbf{x}})_{\mathbf{x}\in\Delta_d}$ of probability measures, the Diaconis and Freedman's chain falls within the scope of iterated function systems with spacial dependent increments probabilities.

Theorem 11. *Assume that*

(1) for all $j = 0,\dots,d$, the functions p_j belong to $\mathcal{H}_\alpha(\Delta_d)$;

(2) there exists $i \in \{0,\dots,d\}$ such that $\delta_i := \min_{\mathbf{x}\in\Delta_d} p_i(\mathbf{x}) > 0$.

Then, the Diaconis and Freedman's chain in Δ_d has a unique P-invariant probability measure $\nu \in \mathcal{P}(\Delta_d)$. Furthermore, there exist constants $\kappa > 0$ and $\rho \in (0,1)$ such that

$$\forall \varphi \in \mathcal{H}_\alpha(\Delta_d),\ \forall \mathbf{x} \in \Delta_d \quad |P^n\varphi(\mathbf{x}) - \nu(\varphi)| \leq \kappa\rho^n. \tag{14}$$

Proof. This is a direct consequence of Theorem 10 since conditions $H1. - H3.$ hold in this context.

H1. For any $\mathbf{x} \neq \mathbf{y} \in \Delta_d$:

$$\int_{\mathbb{L}\mathrm{ip}(\Delta_d,\Delta_d)} \left(\frac{|T(\mathbf{x}) - T(\mathbf{y})|}{|\mathbf{x}-\mathbf{y}|}\right)^\alpha \mu_{\mathbf{x}}(dT) = \sum_{i=0}^d p_i(\mathbf{x}) \int_0^1 \Big(\frac{H_i(t,\mathbf{x}) - H_i(t,\mathbf{y})}{\mathbf{x}-\mathbf{y}}\Big)^\alpha dt = \frac{1}{1+\alpha} < 1;$$

H2. For any $\mathbf{x} \neq \mathbf{y} \in \Delta_d$ and any Borel set $A \subseteq \Delta_d$,

$$\frac{|\mu_{\mathbf{x}}(A) - \mu_{\mathbf{y}}(A)|}{|\mathbf{x}-\mathbf{y}|^\alpha} \leq \sum_{i=0}^d \frac{|p_i(\mathbf{x}) - p_i(\mathbf{y})|}{|\mathbf{x}-\mathbf{y}|^\alpha} \int_0^1 H_i(t,\mathbf{x})(A)dt \leq \sum_{i=0}^d m_\alpha(p_i) < \infty.$$

H3. Set $\mu(dT) := \int_0^1 \delta_{H_i(t,\cdot)}(dT)dt \in \mathcal{P}(X)$; it holds $\mu_{\mathbf{x}}(dT) \geq p_i(\mathbf{x}) \int_0^1 \delta_{H_i(t,\cdot)}(dT)dt \geq \delta\mu(dT)$ for all $\mathbf{x} \in \Delta_d$. Moreover, the constant function $\mathbf{x} \mapsto 0$ belongs to the support of μ so that the semigroup T_μ contains a contracting sequence with limit point 0.

□

4 Some explicit invariant probability densities

In this section we consider some special cases of weights for which it is possible to compute explicitly the unique invariant probability density. When $d = 1$, it has been known that, when both conditions $p_1(0) > 0$ and $p_0(1) > 0$ hold, there exists a unique invariant probability density of $(Z_n)_{n\geq 0}$ given by

$$g_\infty(y) = C \exp\left(\int_{1/2}^{y} \frac{p_1(t)}{1-t} dt - \int_{1/2}^{y} \frac{p_0(t)}{t} dt \right).$$

See for instance [Ramli and Leng, 2010] or [Ladjimi and Peigné, 2019]. We do not get such general result when $d \geq 2$, we can do it only in some specific cases. We would also like to emphasize that in [Nguyen and Volkov, 2019], based on Sethuraman's construction of the Dirichlet distributions (see, [Sethuraman, 1994]), the authors also gave out general results of the explicit formula of the stationary density in these special cases. Our approach is, however, very naturally and worth to be taken into account.

4.1 The case of constant weights

We first consider the case of constant weights, i.e., $p_i(\mathbf{x}) = p_i > 0$ for all $i = 0, \ldots, d$.

Theorem 12. *If $p_i(\mathbf{x}) = p_i > 0$ for all $i = 0, \ldots, d$ then the unique invariant probability density g_∞ of $(Z_n)_{n\geq 0}$ is the density of the Dirichlet distribution $Dir[p_0, \ldots, p_d]$, i.e.*

$$g_\infty(\mathbf{y}) = \frac{1}{\prod_{i=0}^{d} \Gamma(p_i)} \prod_{i=0}^{d} y_i^{p_i - 1} \mathbf{1}_{\mathring{\Delta}_d}(\mathbf{y}) \tag{15}$$

where $y_0 := 1 - y_1 - \cdots - y_d$.

Proof. It suffices to prove that $g_\infty(\mathbf{y}) = \prod_{i=0}^{d} y_i^{\alpha_i}$ with $\alpha_i = p_i - 1$ is the unique solution of the equation $P^* g(\mathbf{y}) = g(\mathbf{y})$. Indeed,

$$\begin{aligned}
P^* g_\infty(\mathbf{y}) &= \sum_{i=0}^{d} p_i \int_{1}^{\frac{1}{1-y_i}} s^{d-2} g_\infty\Big(s\mathbf{y} + \Big(1-s\Big)\mathbf{e}_i \Big) ds \\
&= \sum_{i=0}^{d} p_i \int_{1}^{\frac{1}{1-y_i}} s^{d-2} (sy_i + 1 - s)^{\alpha_i} \prod_{j\neq i} (sy_j)^{\alpha_j} ds \\
&= \sum_{i=0}^{d} p_i \int_{1}^{\frac{1}{1-y_i}} s^{d-2+\sum_{j\neq i}\alpha_j} (sy_i + 1 - s)^{\alpha_i} \frac{g_\infty(\mathbf{y})}{y_i^{\alpha_i}} ds \\
&= \sum_{i=0}^{d} \frac{p_i}{(1-y_i) y_i^{\alpha_i} (1-y_i)^{-2-\alpha_i}} \left(\int_0^{y_i} t^{\alpha_i} (1-t)^{-2-\alpha_i} dt \right) g_\infty(\mathbf{y})
\end{aligned} \tag{16}$$

where the last equality follows by using the change of variables $t = sy_i + 1 - s$ and the equality $\sum_{j=0}^{d} \alpha_j = -d$. Notice that, for $i = 0, \ldots, d$ and $\alpha_i > -1$,

$$(1 + \alpha_i) \int_0^{y_i} t^{\alpha_i} (1-t)^{-2-\alpha_i} dt = y_i^{1+\alpha_i} (1-y_i)^{-1-\alpha_i}.$$

As a matter of fact, the function $F : y_i \mapsto (1+\alpha_i)\int_0^{y_i} t^{\alpha_i}(1-t)^{-2-\alpha_i}dt - y_i^{1+\alpha_i}(1-y_i)^{-1-\alpha_i}$ satisfies $F(0) = 0$ and $F'(y_i) = 0$; therefore $F(y_i) \equiv 0$. Hence, equality (16) yields

$$P^* g_\infty(\mathbf{y}) = \sum_{i=0}^{d} y_i g_\infty(\mathbf{y}) = g_\infty(\mathbf{y}).$$

The uniqueness stems from Corollary 4. $\square$

Example 13. When $d = 2$, $p_1 = p_2 = p_0 = 1/3$, the unique invariant probability density is

$$g_\infty(\mathbf{y}) = \frac{1}{\Gamma(1/3)^3} y_1^{-\frac{2}{3}} y_2^{-\frac{2}{3}} (1 - y_1 - y_2)^{-\frac{2}{3}} \mathbf{1}_{\Delta_2^\circ}(\mathbf{y}). \qquad \text{(see Figure 2)}$$

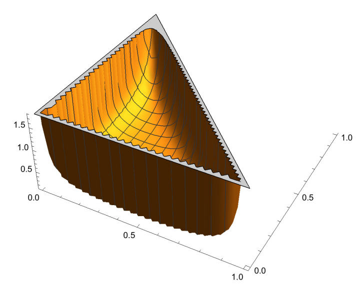

Figure 2: The density function of the Dirichlet distribution $Dir[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$.

Figure 3 represents random trajectories of $(Z_n)_{n\geq 0}$
- in $[0, 1]$, starting at $x_0 = 0.6$ with $p_1(x) = 0.2, p_0(\mathbf{x}) = 0.8$, in the left panel;
- in Δ_2, starting at $\mathbf{x}_0 = (0.3, 0.4)$ with $p_1(\mathbf{x}) = 0.5, p_2(\mathbf{x}) = 0.2, p_0(\mathbf{x}) = 0.3$ in the right panel.

They both illustrate the ergodic behavior of the chain.

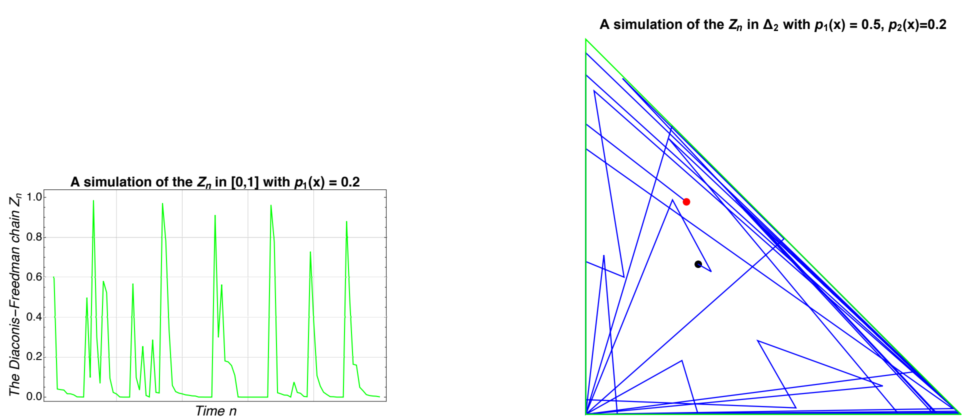

Figure 3: Left: Random trajectory of Z_n in $[0, 1]$ starting at $x_0 = 0.6$; Right: Random trajectory of Z_n in Δ_2 starting at $\mathbf{x}_0 = (0.3, 0.4)$.

4.2 The case of affine weights

In this section, we present a special case of non constant weight functions $p_i(\mathbf{y})$ which yield to the explicit form of the unique invariant density function.

Theorem 14. *Fix positive constants $\theta_0, \dots, \theta_d > 0$ with $|\boldsymbol{\theta}| = \theta_0 + \dots + \theta_d \leq 1$ and assume that, for any $i = 1, \dots, d$ and $\mathbf{y} = (y_1, \dots, y_d)$ in Δ_d,*

$$p_i(\mathbf{y}) = \tilde{p}_i(y_i) := \theta_i + (1 - |\boldsymbol{\theta}|)y_i$$

(which implicitly implies $p_0(\mathbf{y}) = \tilde{p}_0(y_0) := \theta_0 + (1 - |\boldsymbol{\theta}|)y_0$). Then, the unique invariant probability density g_∞ of $(Z_n)_{n \geq 0}$ is the Dirichlet distribution $Dir[\theta_0, \dots, \theta_d]$ given by

$$g_\infty(\mathbf{y}) = Dir[\theta_0, \dots, \theta_d](\mathbf{y}) = \frac{\Gamma(|\boldsymbol{\theta}|)}{\prod_{i=0}^d \Gamma(\theta_i)} \prod_{i=0}^d y_i^{\theta_i - 1} \mathbf{1}_{\mathring{\Delta}_d}(\mathbf{y}).$$

Proof. Using the same techniques as in the proof of Theorem 12, we only need to check that

$$\frac{\int_0^{y_i} p_i(t) t^{\alpha_i} (1-t)^{|\boldsymbol{\alpha}| - \alpha_i + d - 2} dt}{y_i^{\alpha_i}(1-y_i)^{|\boldsymbol{\alpha}| - \alpha_i + d - 1}} = y_i,$$

which can be easily done by a direct calculation. It completes the proof. □

Remark 15. (i) The case when $|\boldsymbol{\theta}| = 1$ corresponds to constant weights.

(ii) This result has a very closed connection to results studied in Wright-Fisher models with mutations (see for instance [Tran et al., 2015a], [Tran et al., 2015b], [Hofrichter et al., 2017]).

In Figure 4 we simulate random trajectories of $(Z_n)_{n \geq 0}$
- in $[0, 1]$, starting at $x_0 = 0.6$ with $p_1(x) = x, p_0(x) = 1 - x$ in the left panel;
- in Δ_2 starting at $\mathbf{x}_0 = (0.3, 0.4)$ with $p_1(\mathbf{x}) = x_1, p_2(\mathbf{x}) = x_2, p_0(\mathbf{x}) = 1 - x_1 - x_2$ in the right panel. They both illustrate the absorbing behavior of the chain.

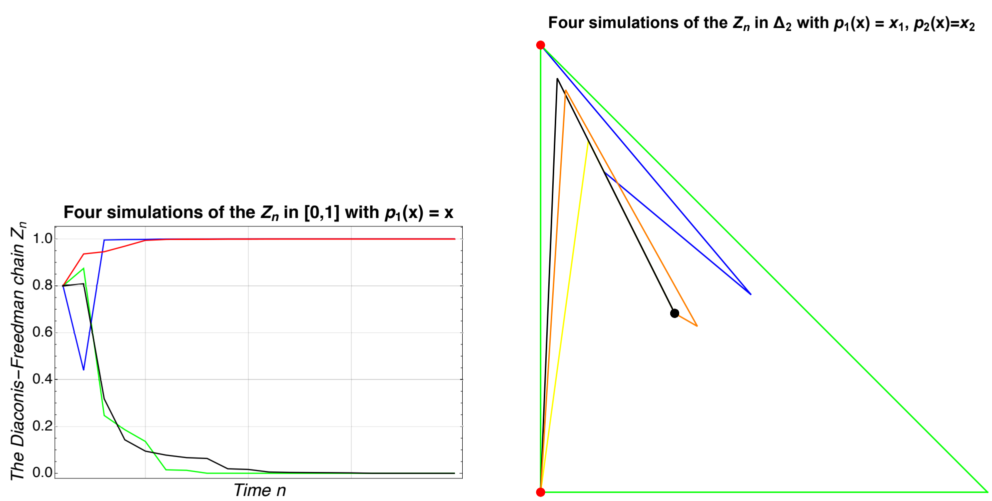

Figure 4: Left: Four random trajectories of Z_n in $[0, 1]$ starting at $x_0 = 0.8$, they will absorb in $\{0, 1\}$; Right: Four random trajectories of Z_n in Δ_2 starting at $\mathbf{x}_0 = (0.3, 0.4)$, they will absorb in $\{\mathbf{e}_0, \mathbf{e}_2\}$.

5 Asymptotic behavior of $(Z_n)_{n\geq 0}$

In this section, we describe the asymptotic behavior of $(Z_n)_{n\geq 0}$ using the notion of *minimal P-absorbing compact subsets*. First we establish some properties of the family $\mathcal{K}_m$ of these subsets and propose their classification. By a general results of [Hervé, 1994], this yields to the classification of the set of P- invariant probability measures as well as the description of the asymptotic behavior of $(Z_n)_{n\geq 0}$. The classification is complete in Δ_2 but partial in $\Delta_d, d > 2$.

5.1 The set $\mathcal{K}_m$ of minimal P-absorbing compact subsets

Definition 16. A non-empty compact subset $K \subseteq \Delta_d$ is said to be *P-absorbing* if for all $\mathbf{x} \in K$

$$P(\mathbf{x}, K^c) := P\mathbf{1}_{K^c}(\mathbf{x}) = \sum_{i=0}^{d} p_i(\mathbf{x}) \int_0^1 \mathbf{1}_{K^c}(t\mathbf{x} + (1-t)e_i)dt = 0,$$

where $K^c = \Delta_d \setminus K$. It is *minimal* when it does not contain any proper P-absorbing compact subset.

We denote by $\mathcal{K}_m$ is the set of all minimal P-absorbing compact subsets. For any $\mathbf{x}_0 \in \Delta_d$ and $\varepsilon > 0$, we set $B_\varepsilon(\mathbf{x}_0) = \{\mathbf{x} \in \Delta_d : |\mathbf{x} - \mathbf{x}_0| < \varepsilon\}$ and $B_\varepsilon = \cup_{i=0}^{d} B_\varepsilon(\mathbf{e}_i)$.

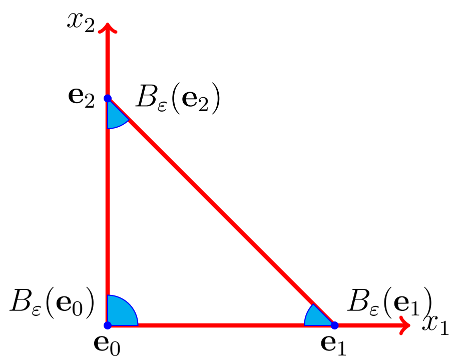

Figure 5: Domain $B_\varepsilon(\mathbf{e}_i)$

The following rules are useful to describe the minimal P-absorbing sets K.

Proposition 17. *(i) If $K \in \mathcal{K}_m$ then K contains at least one vertex.*

(ii) If $K \in \mathcal{K}_m$, $\mathbf{e}_i \in K$ and $p_i(\mathbf{e}_i) = 1$ then $K = \{\mathbf{e}_i\}$.

(iii) If $K \in \mathcal{K}_m$, $\mathbf{e}_i \in K$ and $p_j(\mathbf{e}_i) > 0$ for some $j \neq i$ then $[\mathbf{e}_i, \mathbf{e}_j] \subseteq K$.

(iv) If $K \in \mathcal{K}_m$ and $p_i(\mathbf{x}) > 0$ for some $\mathbf{x} \in K \setminus \{\mathbf{e}_i\}$ then $[\mathbf{e}_i, \mathbf{x}] \subseteq K$.

Proof. (i) Assume that $\mathbf{e}_i \notin K$ for all $i = 0, \ldots, d$. Since K^c is open, there exists $\varepsilon > 0$ such that $B_\varepsilon \subseteq K^c$. Therefore, for all $\mathbf{x} \in K$,

$$P(\mathbf{x}, K^c) = \sum_{i=0}^{d} p_i(\mathbf{x}) \int_0^1 \mathbf{1}_{K^c}(t\mathbf{x} + (1-t)e_i)dt \geq \sum_{i=0}^{d} p_i(\mathbf{x}) \int_0^1 \mathbf{1}_{B_\varepsilon}(t\mathbf{x} + (1-t)e_i)dt$$
$$= \sum_{i=0}^{d} p_i(\mathbf{x})\varepsilon = \varepsilon > 0.$$

This contradicts to the fact that K is P-absorbing.

(ii) It suffices to prove that $\{\mathbf{e}_i\} \in \mathcal{K}_m$. This is true because

$$P(\mathbf{e}_i, \{\mathbf{e}_i\}^c) = \sum_{j \neq i} p_j(\mathbf{e}_i) = 0.$$

(iii) If $[\mathbf{e}_i, \mathbf{e}_j] \cap K^c \neq \emptyset$, then there exist $\mathbf{x}_0 \in [\mathbf{e}_i, \mathbf{e}_j] \cap K^c$ and $\varepsilon > 0$ such that $B_\varepsilon(\mathbf{x}_0) \subseteq K^c$. Therefore $P(\mathbf{e}_i, K^c) \geq p_j(\mathbf{e}_i)\varepsilon > 0$, contradiction.

(iv) Again, if $[\mathbf{e}_i, \mathbf{e}_j] \cap K^c \neq \emptyset$, then there exist $\mathbf{x}_0 \in [\mathbf{e}_i, \mathbf{e}_j] \cap K^c$ and $\varepsilon > 0$ such that $B_\varepsilon(\mathbf{x}_0) \subseteq K^c$; hence,

$$P(\mathbf{x}, K^c) \geq p_i(\mathbf{x})\varepsilon > 0$$

which is a contradiction.

□

This Proposition (17) easily yields to the classification of $\mathcal{K}_m$ when$d = 1$ (see [Ladjimi and Peigné, 2019]):

(i) If $p_1(1) < 1$ and $p_0(0) < 1$ then $\mathcal{K}_m = \{[0, 1]\}$;

(ii) If $p_1(1) < 1$ and $p_0(0) = 1$ then $\mathcal{K}_m = \{\{1\}\}$;

(iii) If $p_1(1) = 1$ and $p_0(0) < 1$ then $\mathcal{K}_m = \{\{0\}\}$;

(iv) If $p_1(1) = 1$ and $p_0(0) = 1$ then $\mathcal{K}_m = \{\{0\}, \{1\}\}$.

In the following section, we describe the set $\mathcal{K}_m$ in Δ_2. Section 5.3 is devoted to the asymptotic behavior in distribution of $(Z_n)_{n\geq 0}$. The reader may be easily convinced that similar statements hold in higher dimension.

5.2 Classification of $\mathcal{K}_m$ in Δ_2

We assume $d = 2$ in this subsection. Unlike the case $d = 1$, for the case of $d = 2$ we need to classify the values of p_i not only on the vertices but also on the edges. We denote by $L_0 = \{\mathbf{x} \in [\mathbf{e}_1, \mathbf{e}_2] : p_0(\mathbf{x}) > 0\}$, $L_1 = \{\mathbf{x} \in [\mathbf{e}_0, \mathbf{e}_2] : p_1(\mathbf{x}) > 0\}$, and $L_2 = \{\mathbf{x} \in [\mathbf{e}_0, \mathbf{e}_1] : p_2(\mathbf{x}) > 0\}$. Let L_0^c (resp. L_1^c and L_2^c) be the complement of L_0 (resp. L_1 and L_2) in $[\mathbf{e}_1, \mathbf{e}_2]$ (resp. $[\mathbf{e}_0, \mathbf{e}_2]$ and $[\mathbf{e}_0, \mathbf{e}_1]$).

Let us fix $K \in \mathcal{K}_m$. There are several cases to consider.

1. $\underline{p_0(\mathbf{e}_0) = p_1(\mathbf{e}_1) = p_2(\mathbf{e}_2) = 1}$

   By Proposition 17 [i], either $\mathbf{e}_0 \in K$ or $\mathbf{e}_1 \in K$ or $\mathbf{e}_2 \in K$. When $\mathbf{e}_0 \in K$, Proposition 17 [ii] implies $K = \{\mathbf{e}_0\}$; similarly for $\mathbf{e}_1$ or $\mathbf{e}_2$. Finally $\mathcal{K}_m = \{\{\mathbf{e}_0\}, \{\mathbf{e}_1\}, \{\mathbf{e}_2\}\}$.

2. $\underline{p_0(\mathbf{e}_0) = p_1(\mathbf{e}_1) = 1 \text{ but } p_2(\mathbf{e}_2) < 1}$

   As above, if $\mathbf{e}_0 \in K$ (resp. $\mathbf{e}_1 \in K$) then $K = \{\mathbf{e}_0\}$ (resp. $\mathbf{e}_1 \in K$).

   Assume now $\mathbf{e}_2 \in K$. Proposition 17 [iii] implies $[\mathbf{e}_0, \mathbf{e}_2] \subseteq K$ when $p_2(\mathbf{e}_0) > 0$ and $[\mathbf{e}_1, \mathbf{e}_2] \subseteq K$ when $p_2(\mathbf{e}_1) > 0$. Therefore K contains $\mathbf{e}_1$ or $\mathbf{e}_2$, contradiction with the minimality of K. Finally $\mathcal{K}_m = \{\{\mathbf{e}_0\}, \{\mathbf{e}_1\}\}$.

   Similar statements hold when $p_0(\mathbf{e}_0) = p_2(\mathbf{e}_2) = 1$ but $p_1(\mathbf{e}_1) < 1$ or $p_1(\mathbf{e}_1) = p_2(\mathbf{e}_2) = 1$ but $p_0(\mathbf{e}_0) < 1$.

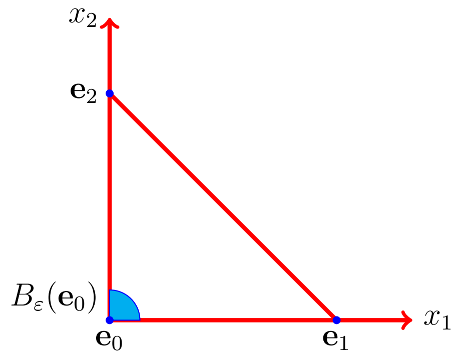

Figure 6: Domain $B_\varepsilon(\mathbf{e}_0)$

3. $p_0(\mathbf{e}_0) = 1$ but $p_1(\mathbf{e}_1), p_2(\mathbf{e}_2) < 1$

   Firstly, $\{\mathbf{e}_0\} \in \mathcal{K}_m$ and $K = \{\mathbf{e}_0\}$ as soon as $\mathbf{e}_0 \in K$.

   Assume now $\mathbf{e}_0 \notin K$ (thus $B_\varepsilon(\mathbf{e}_0) \subseteq K^c$ for some $\varepsilon > 0$) and suppose for instance $\mathbf{e}_1 \in K$ (the same argument holds with $\mathbf{e}_2$). Hence, $p_0(\mathbf{e}_1) = 0$; indeed, condition $p_0(\mathbf{e}_1) > 0$ implies $[\mathbf{e}_0, \mathbf{e}_1] \subseteq K$, contradiction. Consequently $p_2(\mathbf{e}_1) > 0$, which implies $[\mathbf{e}_1, \mathbf{e}_2] \subseteq K$, then $p_0(\mathbf{e}_2) = 0$ and $p_1(\mathbf{e}_2) > 0$. This readily implies that $L_0 = \emptyset$; otherwise, $p_0(\mathbf{x}_0) > 0$ for some $\mathbf{x}_0 \in [\mathbf{e}_1, \mathbf{e}_2]$, therefore $P(\mathbf{x}_0, K^c) \geq p_0(\mathbf{x}_0)\varepsilon > 0$, contradiction with the fact that $\mathbf{x}_0 \in K$ and K is invariant. The equality $L_0 = \emptyset$ yields $K = [\mathbf{e}_1, \mathbf{e}_2]$ and $\mathcal{K}_m = \{\{\mathbf{e}_0\}, [\mathbf{e}_1, \mathbf{e}_2]\}$.

   Finally

   $$\mathcal{K}_m = \begin{cases} \{\{\mathbf{e}_0\}, [\mathbf{e}_1, \mathbf{e}_2]\}, & \text{if } L_0 = \emptyset \\ \{\{\mathbf{e}_0\}\}, & \text{else.} \end{cases}$$

   Similar statements hold when $p_1(\mathbf{e}_1) = 1$ but $p_0(\mathbf{e}_0), p_2(\mathbf{e}_2) < 1$ or $p_2(\mathbf{e}_2) = 1$ but $p_0(\mathbf{e}_0), p_1(\mathbf{e}_1) < 1$.

4. $p_0(\mathbf{e}_0), p_1(\mathbf{e}_1), p_2(\mathbf{e}_2) < 1$ and $L_0 = \emptyset$

   By Proposition 17 [i], the set K contains at least one of the vertices. Assume for instance $\mathbf{e}_1 \in K$, thus $p_2(\mathbf{e}_1) > 0$ since $L_0 = \emptyset$, which implies $[\mathbf{e}_1, \mathbf{e}_2] \subseteq K$. The condition $L_0 = \emptyset$ also implies $P(\mathbf{x}, [\mathbf{e}_1, \mathbf{e}_2]^c) = p_0(\mathbf{x}) = 0$ for all $\mathbf{x} \in [\mathbf{e}_1, \mathbf{e}_2]$, finally $K = [\mathbf{e}_1, \mathbf{e}_2]$. The same conclusion holds when $\mathbf{e}_2 \in K$.

   Now, the set K cannot contain $\mathbf{e}_0$. Otherwise, the condition $p_0(\mathbf{e}_0) < 1$ implies either $[\mathbf{e}_0, \mathbf{e}_1] \subseteq K$ or $[\mathbf{e}_0, \mathbf{e}_2] \subseteq K$, thus $\mathbf{e}_1 \in K$ or $\mathbf{e}_2 \in K$. This yields $K = [\mathbf{e}_1, \mathbf{e}_2]$, contradiction.

   Similar statements hold when $L_1 = \emptyset$ or $L_2 = \emptyset$.

5. $p_0(\mathbf{e}_0), p_1(\mathbf{e}_1), p_2(\mathbf{e}_2) < 1$ and L_0, L_1, L_2 are nonempty

   In this case we always have $\{\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2\} \subseteq K$. Indeed, by Proposition 17 [i], the set K contains at least one vertex, say $\mathbf{e}_0 \in K$; since $p_0(\mathbf{e}_0) < 1$, it contains even one of the two sides $[\mathbf{e}_0, \mathbf{e}_1]$ or $[\mathbf{e}_0, \mathbf{e}_2]$. Assume $[\mathbf{e}_0, \mathbf{e}_1] \subset K$ (thus $\mathbf{e}_1 \in K$) and let us check that $\mathbf{e}_2 \in K$. Otherwise $B_\varepsilon(\mathbf{e}_2) \subseteq K^c$ for some $\varepsilon > 0$; since L_2 is a proper subset of $[\mathbf{e}_0, \mathbf{e}_1]$, there exists $\mathbf{x} \in [\mathbf{e}_0, \mathbf{e}_1] \subseteq K$ such that $P(\mathbf{x}, K^c) \geq p_2(\mathbf{x})\varepsilon > 0$, contradiction.

   Now, the inclusion $\{\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2\} \subseteq K$ combined with Proposition 17 [iv] yields

   $$co\{\mathbf{e}_0, L_0\} \subseteq K, co\{\mathbf{e}_1, L_1\} \subseteq K \quad \text{and} \quad co\{\mathbf{e}_2, L_2\} \subseteq K$$

So that, by the compactness of K,

$$K_0 := \overline{co\{\mathbf{e}_0, L_0\} \cup co\{\mathbf{e}_1, L_1\} \cup co\{\mathbf{e}_2, L_2\}} \subseteq K. \tag{17}$$

Next, we denote by

$$L_i(1) := \{\mathbf{x} \in K_0 : p_i(\mathbf{x}) > 0\}, \qquad \forall i \in \{0, 1, 2\}. \tag{18}$$

It is easy to see that $L_i(1) \supseteq L_i$ for all $i \in \{0, 1, 2\}$. Then, Proposition 17 [iv] and the compactness of K yield

$$K_1 := \overline{co\{\mathbf{e}_0, L_0(1)\} \cup co\{\mathbf{e}_1, L_1(1)\} \cup co\{\mathbf{e}_2, L_2(1)\}} \subseteq K. \tag{19}$$

We iteratively construct a sequence of increasing compact subsets $\{K_n\}_{n\geq 0} \subseteq K$ and consider two possibilities: if there is a finite n such that $K_n = \Delta_2$ then $K = \Delta_2$; otherwise, $K = K_\infty := \overline{\{\cup_{n\geq 0} K_n\}} \subset \Delta_2$. In this case $\mathcal{K}_m$ consists of a unique minimal P-absorbing compact set

$$\mathcal{K}_m = \begin{cases} \{\Delta_2\}, & \text{if there exists a finite } n \text{ such that } K_n = \Delta_2, \\ \{K_\infty\}, & \text{otherwise.} \end{cases}$$

We illustrate here two cases when $\mathcal{K}_m = \{K_0\}$ and $\mathcal{K}_m = \{K_1\}$.

(a) We assume that, for $i = 0, 1, 2$,

$$p_i(\mathbf{x}) = 0 \text{ for all } \mathbf{x} \in \Omega_i := \{\mathbf{y} \in K_0 : [\mathbf{y}, \mathbf{e}_i] \cap K_0^c \neq \emptyset\} \tag{20}$$

Then $K_1 = K_0$ and, iteratively, $K_\infty = K_0$. Moreover, from (20), $P(\mathbf{x}, K_0^c) = 0$ for all $\mathbf{x} \in K_0$, therefore $K = K_0$ by the minimality of K and (17).

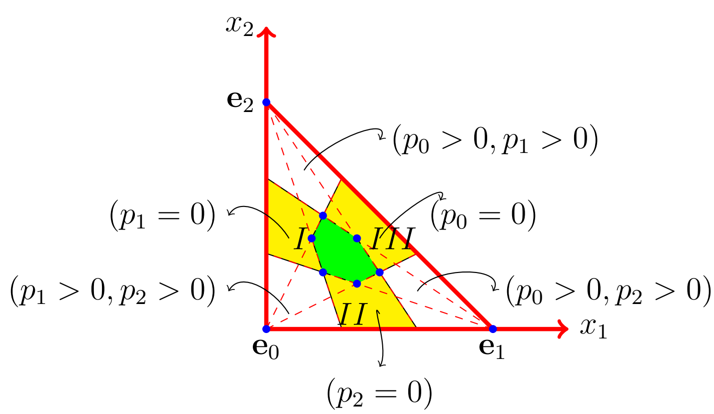

Figure 7: K_0 = white domains $\cup$ yellow domains $= \Delta_2 \setminus$ green domain; K_0^c is the green domain; $\Omega_1, \Omega_2, \Omega_0$ are yellow domains I, II, III correspondingly.

(b) We assume now, for $i = 0, 1, 2$,

$$p_i(\mathbf{x}) = 0 \text{ for all } \mathbf{x} \in \Omega_i(1) := \{i^{th} \text{ yellow region.}\} \tag{21}$$

Then, $K_1 \neq K_0$, $K_2 = K_1$ and iteratively $K_\infty = K_1$. Moreover, by (21), $P(\mathbf{x}, K_1^c) = 0$ for all $\mathbf{x} \in K_1$, therefore $K = K_1$ by the minimality of K and (19).

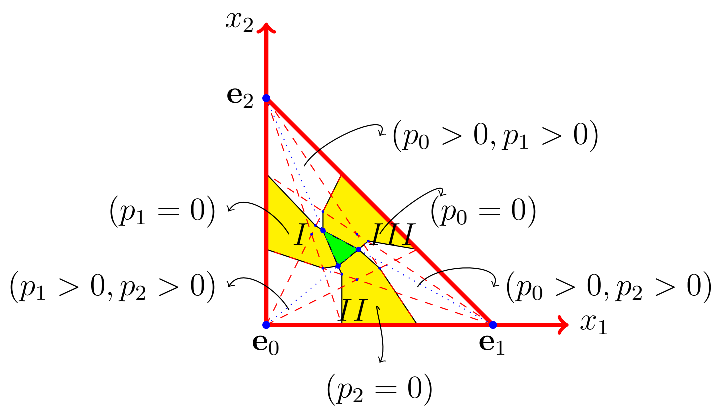

Figure 8: K_1 = white domains + yellow domains = Δ_2 - green domain; K_1^c is green domain; $\Omega_1(1), \Omega_2(1), \Omega_0(1)$ are yellow domains I, II, III correspondingly.

In summary, the complete classification of $\mathcal{K}_m$ in Δ_2 is as follows:

Theorem 18. *In Δ_2,*

1. *either $\mathcal{K}_m$ consists of 3 vertices;*
2. *or $\mathcal{K}_m$ consists of 2 vertices;*
3. *or $\mathcal{K}_m$ consists of 1 vertex;*
4. *or $\mathcal{K}_m$ consists of 1 edge;*
5. *or $\mathcal{K}_m$ consists of 1 vertex and 1 opposite edge;*
6. *or $\mathcal{K}_m$ consists of a compact subset $K_\infty \subseteq \Delta_2$ such that $K_\infty \cap \mathring{\Delta}_2 \neq \emptyset$. This set K_∞ may equal the whole set Δ_2.*

5.3 Asymptotic behavior of $(Z_n)_{n\geq 0}$ in Δ_2

Using Theorem 18, we may state the following Theorem in Δ_2.

Theorem 19. *Let $(Z_n)_{n\geq 0}$ be the Diaconis and Freedman's chain in Δ_2 with weight functions $p_i(\mathbf{x}) \in \mathcal{H}_\alpha(\Delta_2)$. Denote by $\mathcal{I}(P)$ the set of the invariant probability measures of $(Z_n)_{n\geq 0}$. Then, one of the following options holds.*

1. *If $\mathcal{K}_m = \{\{\mathbf{e}_0\}, \{\mathbf{e}_1\}, \{\mathbf{e}_2\}\}$ then $\mathcal{I}(P) = co\{\delta_{\mathbf{e}_0}, \delta_{\mathbf{e}_1}, \delta_{\mathbf{e}_2}\}$ and for any $\mathbf{x} \in \Delta_2$, the chain $(Z_n)_{n\geq 0}$ converges $\mathbb{P}_{\mathbf{x}}$-a.s. to a random variable Z_∞ with values in $\{\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2\}$ and distribution*
$$\mathbb{P}_{\mathbf{x}}(Z_\infty = \mathbf{e}_i) = h_i(\mathbf{x}), \quad i \in \{0, 1, 2\}$$
where h_i is a nonnegative function in $H_\alpha(\Delta_2)$ such that $Ph_i = h_i$, $h_0 + h_1 + h_2 \equiv 1$, and $h_i(\mathbf{e}_j) = 0$ for all $i \neq j \in \{0, 1, 2\}$. Moreover, there exist $\kappa > 0$ and $\rho \in [0, 1)$ such that
$$\forall \varphi \in \mathcal{H}_\alpha(\Delta_2),\ \forall \mathbf{x} \in \Delta_2 \quad \left| P^n\varphi(\mathbf{x}) - h_0(\mathbf{x})\varphi(\mathbf{e}_0) - h_1(\mathbf{x})\varphi(\mathbf{e}_1) - h_2(\mathbf{x})\varphi(\mathbf{e}_2) \right| \leq \kappa\rho^n \|\varphi\|_\alpha.$$

2. *If $\mathcal{K}_m = \{\{\mathbf{e}_0\}, \{\mathbf{e}_1\}\}$ then $\mathcal{I}(P) = co\{\delta_{\mathbf{e}_0}, \delta_{\mathbf{e}_1}\}$ and for any $\mathbf{x} \in \Delta_2$, the chain $(Z_n)_{n\geq 0}$ converges $\mathbb{P}_{\mathbf{x}}$-a.s. to a random variable Z_∞ with values in $\{\mathbf{e}_0, \mathbf{e}_1\}$ and distribution*

$$\mathbb{P}_{\mathbf{x}}(Z_\infty = \mathbf{e}_i) = h_i(\mathbf{x}), \quad i \in \{0, 1\}$$

where h_i is the unique function in $H_\alpha(\Delta_2)$ such that $Ph_i = h_i$, $h_0 + h_1 \equiv 1$, and $h_i(\mathbf{e}_j) = \delta_{ij}$ for all $i, j \in \{0, 1\}$. Moreover, there exist $\kappa > 0$ and $\rho \in [0, 1)$ such that

$$\forall \varphi \in \mathcal{H}_\alpha(\Delta_2),\ \forall \mathbf{x} \in \Delta_2 \quad \Big| P^n\varphi(\mathbf{x}) - h_0(\mathbf{x})\varphi(\mathbf{e}_0) - h_1(\mathbf{x})\varphi(\mathbf{e}_1) \Big| \leq \kappa\rho^n \|\varphi\|_\alpha.$$

Similar statements hold when $\mathcal{K}_m = \{\{\mathbf{e}_0\}, \{\mathbf{e}_2\}\}$ or $\mathcal{K}_m = \{\{\mathbf{e}_1\}, \{\mathbf{e}_2\}\}$.

3. *If $\mathcal{K}_m = \{\{\mathbf{e}_0\}\}$ then $\mathcal{I}(P) = \{\delta_{\mathbf{e}_0}\}$ and for any $\mathbf{x} \in \Delta_2$, the chain $(Z_n)_{n\geq 0}$ converges $\mathbb{P}_{\mathbf{x}}$-a.s. to $\mathbf{e}_0$. Moreover, there exist $\kappa > 0$ and $\rho \in [0, 1)$ such that*

$$\forall \varphi \in \mathcal{H}_\alpha(\Delta_2),\ \forall \mathbf{x} \in \Delta_2 \quad \Big| P^n\varphi(\mathbf{x}) - \varphi(\mathbf{e}_0) \Big| \leq \kappa\rho^n \|\varphi\|_\alpha.$$

Similar statements hold when $\mathcal{K}_m = \{\{\mathbf{e}_1\}\}$ or $\mathcal{K}_m = \{\{\mathbf{e}_2\}\}$.

4. *If $\mathcal{K}_m = \{[\mathbf{e}_1, \mathbf{e}_2]\}$ then $\mathcal{I}(P) = \{\mu^{12}_\infty(d\mathbf{x} \cap [\mathbf{e}_1, \mathbf{e}_2])\}$ where μ^{12}_∞ is the probability measure on $[\mathbf{e}_1, \mathbf{e}_2]$ with density*

$$g^{12}_\infty((t, 1-t), (s, 1-s)) := C \exp\left(\int_s^t \frac{p_1(u, 1-u)}{1-u} du + \int_t^s \frac{p_2(u, 1-u)}{u} du \right).$$

For any $\mathbf{x} \in \Delta_2$, the chain $(Z_n)_{n\geq 0}$ converges $\mathbb{P}_{\mathbf{x}}$-a.s. to a random variable Z_∞ with values on $[\mathbf{e}_1, \mathbf{e}_2]$ and distribution $\mu^{12}_\infty(d\mathbf{x} \cap [\mathbf{e}_1, \mathbf{e}_2])$. Moreover, there exist $\kappa > 0$ and $\rho \in [0, 1)$ such that

$$\forall \varphi \in \mathcal{H}_\alpha(\Delta_d),\ \forall \mathbf{x} \in \Delta_d \quad \Big| P^n\varphi(\mathbf{x}) - \mu^{12}_\infty(\varphi) \Big| \leq \kappa\rho^n \|\varphi\|_\alpha.$$

Similar statements hold when $\mathcal{K}_m = \{[\mathbf{e}_0, \mathbf{e}_1]\}$ or $\mathcal{K}_m = \{[\mathbf{e}_0, \mathbf{e}_2]\}$.

5. *If $\mathcal{K}_m = \{\{\mathbf{e}_0\}, [\mathbf{e}_1, \mathbf{e}_2]\}$ then $\mathcal{I}(P) = co\{\delta_{\mathbf{e}_0}, \chi_{[\mathbf{e}_1, \mathbf{e}_2]}\}$ and for any $\mathbf{x} \in \Delta_2$, the chain $(Z_n)_{n\geq 0}$ converges to $\mathbf{e}_0$ with probability $h_0(\mathbf{x})$ and to $[\mathbf{e}_1, \mathbf{e}_2]$ with probability $h_{12}(\mathbf{x})$. Moreover, there exist $\kappa > 0$ and $\rho \in [0, 1)$ such that $\forall \varphi \in \mathcal{H}_\alpha(\Delta_2), \forall \mathbf{x} \in \Delta_2$*

$$\Big| P^n\varphi(\mathbf{x}) - h_0(\mathbf{x})\varphi(\mathbf{e}_0) - h_{12}(\mathbf{x})\mu^{12}_\infty(\varphi) \Big| \leq \kappa\rho^n \|\varphi\|_\alpha.$$

6. *If $\mathcal{K}_m = \{K_\infty\}$ (possible equal Δ_2) then $\mathcal{I}(P) = \{\mu_\infty\}$ which is a probability measure on Δ_2 with support K_∞ (possible equal Δ_2).*

Proof. By a directed calculation, for all $\varphi \in H_\alpha(\Delta_2)$,

$$|P\varphi|_\alpha \leq \frac{1}{1+\alpha} |\varphi|_\alpha + \left(1 + \sum_{i=0}^{2} m_\alpha(p_i) \right) |\varphi|_\infty.$$

Hence, by [Hennion, 1993], the operator P is quasi-compact on $H_\alpha(\Delta_2)$. The operator P is Markov, so that $\mathbf{1} \in H_\alpha(\Delta_2)$ satisfies $P\mathbf{1} = \mathbf{1}$. Therefore, by using the [Hervé, 1994, Theorem 2.2], the eigenspace corresponding to eigenvalue 1 is nothing but $\ker(P - Id)$. All six above cases can be checked easily by following the [Hervé, 1994] (also see in [Ladjimi and Peigné, 2019] for a classification in dimension 1).

□

Remark 20. The cases considered in Section 4 where there is a unique invariant probability density all satisfy the case 6 where $\mathcal{K}_m = \{\Delta_2\}$, i.e. when $p_i(\mathbf{e}_i) < 1$ for all $i = 0, 1, 2$. The question of the existence (hence unicity) of an invariant probability density when $p_i(\mathbf{e}_i) < 1$ for all $i = 0, 1, \ldots, d$ is still open for $d \geq 2$ (it has been solved in $d = 1$ in [Ladjimi and Peigné, 2019]).

6 Discussion

We would like to briefly present here another interesting setting for the Diaconis and Freedman's chain in Δ_d. For any $i = 0, \ldots, d$ and $\mathbf{x} \in \Delta_d$, let $S_i(\mathbf{x})$ be the strict convex combination of $\mathbf{x}$ and all vertices $\mathbf{e}_j$ except $\mathbf{e}_i$, i.e. $S_i(\mathbf{x}) := \overset{\circ}{\Big(co\{\mathbf{x}, \{\mathbf{e}_j\}_{j\neq i}\}\Big)}$.

Assume that at time n, a walker $\mathbf{Z}$ is located at site $Z_n = \mathbf{x} \in \Delta_d$ and has probability $p_i(\mathbf{x})$ to move to the domain $S_i(\mathbf{x})$, the arrival point being chosen according to the uniform distribution $U_{S_i}(d\mathbf{x}))$ on this domain. In other words, the one-step transition probability function of the Markov chain generated by this walker is

$$P(\mathbf{x}, d\mathbf{y}) = \sum_{i=0}^{d} p_i(\mathbf{x}) \frac{1}{|S_i(\mathbf{x})|} \mathbf{1}_{S_i(\mathbf{x})}(\mathbf{y}) d\mathbf{y}, \quad \mathbf{x} \in \mathring{\Delta}_d.$$

We illustrate this setting in Δ_2 in Figure 9 but a such setting and its applications will be considered in details somewhere else.

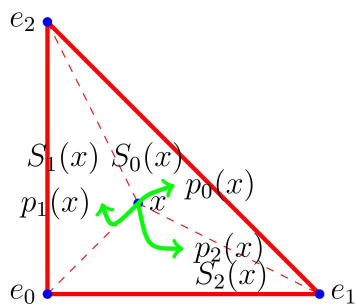

Figure 9: An alternative model

Acknowledgments

M. Peigné and T.D. Tran would like to thank warmly VIASM for financial support and hospitality where the first ideas of paper has been discussed: M. Peigné spent 6 months in 2017 and T.D. Tran spent three months in 2017 and two months in 2019 at VIASM, as visiting scientists. T.D. Tran would also like to thank Institut Denis Poisson for financial support and a warmly and friendly hospitality during his one-month visiting in 2018. We would like to thank Jürgen Jost for illuminating discussions.

References

[Barnsley et al., 1988] Barnsley, M. F., Demko, S. G., Elton, J. H., and Geronimo, J. S. (1988). Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities. *Ann. Inst. H. Poincaré Probab. Statist.*, 24(3):367–394.

[Barnsley and Elton, 1988] Barnsley, M. F. and Elton, J. H. (1988). A new class of Markov processes for image encoding. *Adv. in Appl. Probab.*, 20(1):14–32.

[Barnsley et al., 1989] Barnsley, M. F., Elton, J. H., and Hardin, D. P. (1989). Recurrent iterated function systems. *Constr. Approx.*, 5(1):3–31. Fractal approximation.

[Bush and Mosteller, 1953] Bush, R. R. and Mosteller, F. (1953). A stochastic model with applications to learning. *Ann. Math. Statist.*, 24(4):559–585.

[Diaconis and Freedman, 1999] Diaconis, P. and Freedman, D. (1999). Iterated random functions. *SIAM Rev.*, 41(1):45–76.

[Dubins and Freedman, 1966] Dubins, L. E. and Freedman, D. A. (1966). Invariant probabilities for certain markov processes. *Ann. Math. Statist.*, 37(4):837–848.

[Ferguson, 1973] Ferguson, T. S. (1973). A bayesian analysis of some nonparametric problems. *Ann. Statist.*, 1(2):209–230.

[Ghosh and Ramamoorthi, 2003] Ghosh, J. K. and Ramamoorthi, R. V. (2003). *Bayesian nonparametrics*. Springer Series in Statistics. Springer-Verlag, New York.

[Guivarc'h and Raugi, 1986] Guivarc'h, Y. and Raugi, A. (1986). Products of random matrices: convergence theorems. In *Random matrices and their applications (Brunswick, Maine, 1984)*, volume 50 of *Contemp. Math.*, pages 31–54. Amer. Math. Soc., Providence, RI.

[Harris, 1952] Harris, T. E. (1952). A method for limit theorems in markov chains. *Ann. Math. Statis*, 23:141.

[Hennion, 1993] Hennion, H. (1993). Sur un théorème spectral et son application aux noyaux lipchitziens. *Proc. Amer. Math. Soc.*, 118(2):627–634.

[Hennion and Hervé, 2001] Hennion, H. and Hervé, L. c. (2001). *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, volume 1766 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin.

[Hervé, 1994] Hervé, L. c. (1994). étude d'opérateurs quasi-compacts positifs. Applications aux opérateurs de transfert. *Ann. Inst. H. Poincaré Probab. Statist.*, 30(3):437–466.

[Hofrichter et al., 2017] Hofrichter, J., Jost, J., and Tran, T. D. (2017). *Information geometry and population genetics*. Understanding Complex Systems. Springer, Cham. The mathematical structure of the Wright-Fisher model.

[Jost et al., 2019] Jost, J., Le, H., Luu, H., and Tran, T. (2019). Probabilistic mappings and bayesian nonparametrics. arxiv1905.11448.

[Kaijser, 1981] Kaijser, T. (1981). On a new contraction condition for random systems with complete connections. *Rev. Roumaine Math. Pures Appl.*, 26(8):1075–1117.

[Kapica and Sleczka, 2017] Kapica, R. and Sleczka, M. (2017). Random iteration with place dependent probabilities. https://arxiv.org/abs/1107.0707.

[Karlin, 1953] Karlin, S. (1953). Some random walks arising in learning models. i. *Pacific J. Math.*, 3(4):725–756.

[Ladjimi and Peigné, 2019] Ladjimi, F. and Peigné, M. (2019). On the asymptotic behavior of the diaconis–freedman chain on [0,1]. *Statistics & Probability Letters*, 145:1 – 11.

[Letac, 1986] Letac, G. (1986). A contraction principle for certain Markov chains and its applications. In *Random matrices and their applications (Brunswick, Maine, 1984)*, volume 50 of *Contemp. Math.*, pages 263–273. Amer. Math. Soc., Providence, RI.

[Mirek, 2011] Mirek, M. (2011). Heavy tail phenomenon and convergence to stable laws for iterated Lipschitz maps. *Probab. Theory Related Fields*, 151(3-4):705–734.

[Nguyen and Volkov, 2019] Nguyen, T.-M. and Volkov, S. (2019). On a class of random walks in simplexes. arXiv:1709.00174.

[Peigné, 1993] Peigné, M. (1993). Iterated function systems and spectral decomposition of the associated Markov operator. In *Fascicule de probabilités*, volume 1993 of *Publ. Inst. Rech. Math. Rennes*, page 28. Univ. Rennes I, Rennes.

[Peigné and Woess, 2011a] Peigné, M. and Woess, W. (2011a). Stochastic dynamical systems with weak contractivity properties I. Strong and local contractivity. *Colloquium Mathematicum*, 125(1):31–54.

[Peigné and Woess, 2011b] Peigné, M. and Woess, W. (2011b). Stochastic dynamical systems with weak contractivity properties II. Iteration of Lipschitz mappings. *Colloquium Mathematicae*, 125(1):55–81.

[Ramli and Leng, 2010] Ramli, M. A. and Leng, G. (2010). The stationary probability density of a class of bounded Markov processes. *Advances in Applied Probability*, 42(4):986–993.

[Sethuraman, 1994] Sethuraman, J. (1994). A constructive definition of dirichlet priors. *Statistica Sinica*, 4(2):639–650.

[Stenflo, 2012] Stenflo, O. (2012). A survey of average contractive iterated function systems. *J. Difference Equ. Appl.*, 18(8):1355–1380.

[Tran et al., 2015a] Tran, T., Hofrichter, J., and Jost, J. (2015a). The free energy method and the wright-fisher model with 2 alleles. *Theory in Biosciences*, pages 1–10.

[Tran et al., 2015b] Tran, T., Hofrichter, J., and Jost, J. (2015b). The free energy method for the Fokker-Planck equation of the Wright-Fisher model. MIS-Preprint 29/2015.

Marc Peigné, Institut Denis Poisson UMR 7013, Université de Tours, Université d'Orléans, CNRS France.
E-mail address: peigne@univ-tours.fr

Tat Dat Tran, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Inselstrasse 22, D-04103 Leipzig, Germany
Mathematisches Institut, Universität Leipzig, Augustusplatz 10, D-04109 Leipzig, Germany
E-mail address: trandat@mis.mpg.de, tran@math-uni.leipzig.de