# Balanced multicolour Pólya urns via smoothing systems analysis

## Cécile Mailler

Department of Mathematical Sciences,
University of Bath,
Claverton Down,
BA2 7AY Bath
United-Kingdom.
*E-mail address*: `c.mailler@bath.ac.uk`
*URL*: http://people.bath.ac.uk/cdm37/

**Abstract.** The present paper aims at describing in details the asymptotic composition of a class of $d$-colour Pólya urns: namely balanced, tenable and irreducible urns. We decompose the composition vector of such urns according to the Jordan decomposition of their replacement matrix. The projections of the composition vector onto the so-called *small* Jordan spaces are known to be asymptotically Gaussian, but the asymptotic behaviour of the projections onto the *large* Jordan spaces are not known in full details up to now and are described by a limit random variable called $W$, depending on the parameters of the urn.

We prove, via the study of smoothing systems, that the variable $W$ has a density and that it is moment-determined.

## 1. Introduction

A Pólya urn is a discrete time stochastic process which was originally introduced by Eggenberger and Pólya (1923) to model the spread of epidemics. Since then, they have been useful in many different areas of mathematics and theoretical computer science and are therefore broadly studied. We can for example cite applications to the analysis of random trees (AVL[1] Mahmoud, 1998, 2–3 trees Flajolet et al., 2005), to the analysis of the Bandit algorithm Lamberton et al. (2004), or to the reinforced random walks (see for example the survey of Pemantle, 2007).

The range of methods used to study this random object is also very large. Historically studied by enumerative combinatorics (see for example Bagchi and Pal,

---

[1]The AVL is a balanced search tree named after its inventors, Adel'son-Vel'skiĭ & Landis (see Adel'son-Vel'skiĭ and Landis, 1962).

1985), Pólya urns have been efficiently studied by embedding in continuous time (or *poissonization*) since the works of Athreya and Karlin (1968) (see for example Janson, 2004). In parallel, since the seminal paper by Flajolet, Dumas & Puyhaubert (Flajolet et al., 2006), the analytic combinatorics community successfully tackles the problem in the two-colour case.

A Pólya urn process is defined as follows: an urn contains balls of different colours, let us denote by $d$ the number of different colours available. Fix an initial composition $\boldsymbol{\alpha} = {}^t(\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$, meaning that there are $\alpha_i$ balls of colour $i$ at time zero in the urn, for all $i \in \{1, \ldots, d\}$. Fix a $d \times d$ matrix $R = (a_{i,j})_{1 \leqslant i,j \leqslant d}$ with integer coefficients. At each step of the process, pick a ball uniformly at random in the urn, denote its colour by $c$ ($c \in \{1, \ldots, d\}$), put this ball back in the urn and add into the urn $a_{c,i}$ balls of colour $i$ for all $1 \leqslant i \leqslant d$. The standard question asked is "how many balls of each colour are there in the urn" at time $n$? when $n$ tends to infinity?

Of course, the answer depends on the initial composition vector and on the replacement matrix chosen, and many different behaviours are exhibited in the literature. The aim of the present paper is to investigate into more detail the asymptotic behaviour of a very large class of Pólya urns: the $d$-colour, balanced, tenable and irreducible urns. This framework is very general and includes, among others, Pólya urns modelling m-ary trees (Chern and Hwang, 2001; Chauvin and Pouyanne, 2004), paged binary trees (Chern and Hwang, 2001; Mahmoud, 2002) or $B$-trees (Chauvin et al., 2016). Our precise assumptions are the following:

(T) The non-diagonal coefficients of $R$ are non-negative, and, for all $1 \leqslant i \leqslant d$, either $a_{i,i} \geqslant -1$ or $-a_{i,i}$ is the gcd of $\{a_{1,i}, \ldots, a_{d,i}, \alpha_i\}$.

(B) The urn is balanced, meaning that there exists an integer $S$, called the balance, such that, for all $c \in \{1, \ldots, d\}$, $\sum_{i=1}^{d} a_{c,i} = S$.

(I) The replacement matrix $R$ is irreducible, meaning that for all $1 \leqslant c, i \leqslant d$, there exists $n \geqslant 0$ such that $(R^n)_{c,i} > 0$.

These assumptions ensure us that $S$ is the eigenvalue of $R$ having maximal real part. We then consider the Jordan decomposition of the matrix $R$ and fix one Jordan block: this Jordan block is associated to a stable subspace $E$ (containing a unique eigenline) and to an eigenvalue $\lambda$. Note that several Jordan blocks can correspond to the same eigenvalue. We are interested in the behaviour of the projection of the urn composition vector on $E$, asymptotically when $n$ tends to infinity. When $\sigma := {}^{\mathrm{Re}\lambda}/S \leqslant {}^1/2$. This projection, appropriately renormalised, converges in law to a Gaussian distribution, independent of the initial composition $\boldsymbol{\alpha}$ of the urn. However in the case of "large" eigenspaces, i.e. if $\sigma > {}^1/2$, we observe a different behaviour, which we aim at describing better in the present paper.

The projection on a large eigenspace $E$ of the composition vector of the urn at time $n$, renormalised appropriately, converges almost surely and in all $L^p(p \geqslant 1)$ to a complex random variable $W_{\boldsymbol{\alpha}}^{DT}$ times an eigenvector $v$ associated to $E$. Moreover, if we embed the urn process in continuous time, the obtained Galton-Watson process $U^{CT}(t)$, projected onto $E$, and correctly renormalised, converges almost surely and in all $L^p(p \geqslant 1)$ to a complex random variable $W_{\boldsymbol{\alpha}}^{CT}$ times $v$. Our main goal is to gather information about the two random variables $W_{\boldsymbol{\alpha}}^{DT}$ and $W_{\boldsymbol{\alpha}}^{CT}$.

The discrete time and the continuous time process are closely related and one of our main tools will be to transport information from one setting to the other. The three main results of the paper are: for all initial composition $\boldsymbol{\alpha}$,

- the support of the random variables $W_{\boldsymbol{\alpha}}^{DT}$ and $W_{\boldsymbol{\alpha}}^{CT}$ is the whole complex plane if $\lambda \in \mathbb{C}\backslash\mathbb{R}$ and the whole real line if $\lambda \in \mathbb{R}$;
- $W_{\boldsymbol{\alpha}}^{DT}$ and $W_{\boldsymbol{\alpha}}^{CT}$ both admit a density on $\mathbb{C}$ (resp. $\mathbb{R}$ if $\lambda \in \mathbb{R}$);
- $W_{\boldsymbol{\alpha}}^{DT}$ and $W_{\boldsymbol{\alpha}}^{CT}$ are moment-determined (the Laplace transform of $W_{\boldsymbol{\alpha}}^{DT}$ converges on the whole plane).

These results are the first results in the literature about the variables $W$ induced by multi-colour urns.

To prove them, we first use the branching property of the urn process to prove that it is enough to consider the $d$ atomic initial compositions $\boldsymbol{e}_c$ ($c \in \{1, \ldots, d\}$) defined as follows: all the coefficients of $\boldsymbol{e}_c$ are zero except the $c$th which is equal to 1 if $a_{c,c} \geq 0$, and to $-a_{c,c}$ otherwise. The branching property of urn schemes is largely used in the literature; the novelty of our approach is that we can include the case of Pólya urn schemes with diagonal coefficients possibly less than $-1$. Using again the branching property, we prove that $(W_{\boldsymbol{e}_1}^{DT}, \ldots, W_{\boldsymbol{e}_d}^{DT})$ (resp. $(W_{\boldsymbol{e}_1}^{CT}, \ldots, W_{\boldsymbol{e}_d}^{CT})$) is solution of a system of smoothing equations.

To our knowledge, this article contains the first analysis of such a smoothing system: the literature contains examples of smoothing systems of two equations with real random variable solutions (see Chauvin et al., 2015 for an example) or one smoothing equation with complex solutions (see e.g. Chauvin and Pouyanne, 2004), but no smoothing system with complex solutions as in the present article. Such an analysis needed the development of new arguments and this approach would be useful in any other context where such smoothing systems would appear. In a very recent article (appeared on arxiv.org while the current paper was under review), Leckey (2016) gives a survey about smoothing systems and their applications, as well as general results about them.

From the smoothing system in discrete time, we deduce a system of equations verified by the Fourier transforms of the $W^{DT}$s, and prove that these Fourier transforms are integrable, inducing the existence of densities. We then show how to deduce the same result for the continuous time $W$s, and for any initial composition.

From the system in continuous time, we deduce an induction formula for the moments of the $W^{CT}$s. We then prove by induction that these moments are small enough to apply Carleman's criterion to conclude that they are moment-determined. We then transport the result in discrete time and for all initial initial composition.

**Plan of the paper:** We first describe in Section 2 our framework and the state of the art concerning the asymptotic behaviour of the multi-colour urns; at the end of Section 2, we also state our main results (namely Theorems 2.9 and 2.10). Section 3 is devoted to characterise the random variables $W^{CT}$ as a solution of a system of fixed point equations in law. In Section 4, we use these systems to study the moments of the $W$'s and prove that they are moment-determined, both in continuous and in discrete time. In Section 5, we imply, from the continuous time systems and from the moment study, that the random variables $W^{DT}$ are also characterised as a solution of a smoothing system. Finally, Section 6 contains the proof that the $W$'s admit a density, both in discrete and continuous time. It is very interesting to see how we will travel from discrete to continuous time all along the paper and how going from one world to the other is very fruitful. But first, let us discuss our set of hypothesis: (T), (B) and (I).

1.1. *Discussion of the hypothesis.* Different hypothesis are made in the literature, in order to control the behaviour of Pólya urns:

($T_{-1}$) The coefficients of $R$ are all non-negative except the diagonal coefficients which can be equal to $-1$.

(T) The non-diagonal coefficients of $R$ are non-negative, and, for all $1 \leqslant i \leqslant d$, either $a_{i,i} \geqslant -1$ or $-a_{i,i}$ is the gcd of $\{a_{1,i}, \ldots, a_{d,i}, \alpha_i\}$.

(B) The urn is balanced, meaning that there exists an integer $S$, called the balance, such that, for all $c \in \{1, \ldots, d\}$, $\sum_{i=1}^{d} a_{c,i} = S$.

Assuming ($T_{-1}$) or (T) permits to avoid non-tenable urns, i.e. urn schemes in which something impossible is asked: an example would be that you must subtract 3 balls of colour 1 from the urn while there is only 1 such ball in the urn. Allowing only non-negative coefficients in the replacement matrix and coefficients at least $-1$ on the diagonal permits to ensure that the urn is tenable. It is proven in Pouyanne (2008) that assumption (T) has the same effect although it is much weaker and allows to include a much wider class of Pólya urns into our framework. Some authors prefer working without such an assumption but then make all reasoning conditioned to tenability, i.e. conditioned on the event "no impossible configuration occur" (see for example Janson, 2004, Remark 4.2).

The balance assumption (B) is quite standard in the literature, though it is not always necessary (this assumption is not needed in Janson (2004), for example). This assumption implies that the total number of balls in the urn is a deterministic function of time and this property is the foundation of combinatorics approaches while continuous time analysis of Pólya urns can be done for non-balanced urns.

Finally, we will assume that the urn is irreducible, meaning that any colour can be produced from an initial composition with one unique ball: for a two–colour urn, being irreducible means having a non-triangular replacement matrix. The following definitions define the notion precisely.

**Definition 1.1** (see Janson, 2004, page 4)**.** Let $c, i \in \{1, \ldots, d\}$, we say that $c$ **dominates** $i$ if there exists $n \geqslant 1$ such that $(R^n)_{c,i} > 0$. A colour $c \in \{1, \ldots, d\}$ is said to be **dominating** if it dominates every other colour in $\{1, \ldots, d\}$.

**Definition 1.2** (see Janson, 2004, page 4)**.** We say that an urn of replacement matrix $R$ is **irreducible** if and only if every colour is a dominating colour.

(I) The replacement matrix $R$ is irreducible.

Note that it is sometimes enough (see, for example, Janson, 2004 and Pouyanne, 2008) to only assume the following weaker version of the irreducibility assumption. First note that the domination relation is transitive and reflexive and thus partitions the set of colours into some equivalence classes:

**Definition 1.3** (see Janson, 2004, page 4)**.** We say that two colours $i$ and $j$ are in the same class if $i$ dominates $j$ and $j$ dominates $i$. A **dominating class** is a class of dominating colours. An eigenvalue $\lambda$ belongs the one of the equivalence class $D$ if the restriction of $R$ to $D$ admits $\lambda$ as an eigenvalue.

We are now ready to assume this weaker version of irreducibility, which we call (S) for *simplicity*:

(S) The largest real eigenvalue $\lambda_{max}$ of $R$ is positive and is a simple eigenvalue of $R$. Furthermore, there is at least one ball of a dominating colour in the urn at time 0, and $\lambda_{max}$ belong to the dominating class.

Note that under (B), $S = \lambda_{max}$. We refer the reader to Janson (2004, page 5) for a discussion of this weaker assumption.

Cases of non-irreducible urns are studied in the literature: the diagonal case $R = SI_d$ is the original Pólya-Eggenberger process and its behaviour is well described in Athreya (1969); Blackwell and Kendall (1964); Johnson and Kotz (1977); Chauvin et al. (2015), and triangular urn schemes are developped in Janson (2006); Bose et al. (2009).

Note that a large two-colour urn cannot have negative diagonal coefficients, whereas there exists $d$–colour Pólya urns with possibly negative coefficients having large eigenvalues. We thus have to include such urns in our study, that is why we only assume (T) and not the more restrictive assumption (T$_{-1}$). In the present paper, we thus choose to assume (T), (B) and (S) (although it is necessary to assume (I) for Theorem 2.10): we can cite many examples of urn processes that fall in this framework (see for example m-ary trees (Chern and Hwang, 2001; Chauvin and Pouyanne, 2004), paged binary trees (Mahmoud, 2002; Chern and Hwang, 2001), $B$-trees (Chauvin et al., 2016) and we will see all along the paper how each of these assumptions is used in the proofs. We are interested in the asymptotic behaviour of an urn under these three conditions.

The present setting is different from Janson (2004)'s setting for $d$–colour Pólya urns ($d \geqslant 3$), where (T$_{-1}$) is assumed with further assumptions, but (B) is not: it is however mentioned in Janson (2004, Remark 4.2) that Janson's main results hold under our set of hypothesis. We will thus be able to apply Janson's results before going further in the study of the asymptotic behaviour of the urn.

The following section is devoted to summarising the results of the literature needed as preliminaries to state our main results.

## 2. Preliminaries and statement of the main results

The behaviour of the urn process is already quite well known: We recall hereby the main results of the literature (mainly by Athreya, 1969, Janson, 2004 and Pouyanne, 2008) and thereafter state our main results.

2.1. *Jordan decomposition.* In view of (B) and (S), the matrix $R$ admits $S$ as a simple eigenvalue, and every other eigenvalue $\lambda$ of $R$ verifies $\mathrm{Re}\lambda < S$. The matrix $R$ can be written on its Jordan normal form, meaning that it is similar to a diagonal of blocks $\mathtt{diag}(J_1, \ldots, J_r)$ where each $J_i$ is a matrix shaped as follows:

$$
J = \begin{pmatrix}
\lambda & 1 & 0 & \ldots & 0 \\
0 & \lambda & 1 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & 0 \\
\vdots & \ldots & \ddots & \lambda & 1 \\
0 & \ldots & \ldots & 0 & \lambda
\end{pmatrix},
$$

where $\lambda$ is an eigenvalue of $R$. Note that several Jordan blocks can be associated to the same eigenvalue. In the following, we chose a Jordan block and study the behaviour of the projection of the composition vector onto the subspace associated to this Jordan block. Note that the fact that $R$ is irreducible implies that $S$ is a single eigenvalue of $R$.

**Definition 2.1.** Let $\lambda$ be an eigenvalue for $R$ and $\sigma = {}^{\mathrm{Re}\lambda}/{S}$. We call $\lambda$ a **large eigenvalue** of $R$ if $1/2 < \sigma < 1$, or a **small eigenvalue** if $\sigma \leqslant 1/2$.

A **large Jordan block** is a Jordan block of $R$ associated with a large eigenvalue of $R$ and a **small Jordan block** is a Jordan block associated with a small eigenvalue of $R$.

We denote by $U_{\boldsymbol{\alpha}}(n) = U(n) \in \mathbb{N}^d$ the composition vector of the urn at time $n$ (the subscript $\boldsymbol{\alpha}$ is the initial composition of the urn): its $i^{\mathrm{th}}$ coordinate is by definition equal to the number of balls of colour $i$ at time $n$ in the urn. We are interested in the behaviour of $U(n)$ when $n$ tends to infinity. It is showed in the literature that $U(n)$ is easier to describe when decomposed according to the Jordan block decomposition of $R$. For every stable subspace $E$ associated to a Jordan block of $R$, we will denote by $\pi_E$ the projection on $E$ relative to the direct sum of all Jordan subspaces of $R$, and we will study separately each projection on a Jordan subspace $E$.

It is standard to embed urn processes in continuous time (see for example Athreya and Karlin, 1968): each ball is seen as a clock that rings after a random time with exponential law of parameter one, independently from other clocks in the urn. When a clock rings, it splits into $a_{i,j} + \delta_{i,j}$ balls of colour $j$ ($\forall j \in \{1, \ldots, d\}$) if the clock had colour $i$ (where we uses Kronecker's notation: $\delta_{i,j}$ equals 1 if $i = j$ and 0 otherwise). We denote by $\tau_n$ the time of the $n^{\mathrm{th}}$ ring in the urn and by $U^{CT}(t)$ the composition vector of the (continuous time) urn at time $t$. We have the following standard connection: almost surely,

$$(U(n))_{n \geqslant 0} = (U^{CT}(\tau_n))_{n \geqslant 0}. \tag{2.1}$$

In addition, the process $(U(n))_{n \geqslant 0}$ is independent of the sequence of stopping times $(\tau_n)_{n \geqslant 0}$.

The asymptotic behaviour of the different projections of $U(n)$ and $U^{CT}(t)$ is partially described in the literature:

- In continuous time (see Janson, 2004),
    - small projections have a Gaussian behaviour, and
    - renormalised large projections converge almost surely to a random variable $W^{CT}$.
- In discrete time,
    - if $R$ has only small eigenvalues apart from $S$, if $E$ is one of the largest Jordan block associated to the eigenvalue $\lambda$ realising the second highest real part (after $S$), and if $\sigma = {}^{\mathrm{Re}\lambda}/{S} = 1/2$ then projections onto $E$ have a Gaussian behaviour (see Janson, 2004, Theorems 3.22 et 3.23); and
    - renormalised large projections converge almost surely to a random variable $W^{DT}$ (see Pouyanne, 2008).

As one can see, the behaviour of small projections (i.e. projections on a small Jordan block) in discrete time is not known yet in full generality: Subsection 2.2 is devoted to describing the behaviour of $\pi_E(U(n))$ for all small Jordan space of $R$. However, the main aim of the paper is to describe the unexplored $W^{DT}$ and $W^{CT}$: Subsection 2.3 will state the results concerning the projections on large Jordan spaces (i.e. associated to a large Jordan block), as a preliminary to our main results.

2.2. *Projections on small Jordan spaces.* As explained above, the behaviour of the projections of the composition vector (in discrete time) onto the small Jordan blocks is not known yet in full generality. Although our main aim is to focus on the less understood projections onto large Jordan blocks, we sketch here a proof of a general result for small Jordan blocks in order to complete the theory of Pólya urns under (B), (T) and (S).

**Theorem 2.2.** *Under assumptions* (B), (T) *and* (S)*, if $E$ is a block of size $\nu + 1$ associated to a small eigenvalue $\lambda$ of $R$, then there exists a covariance matrix $\Sigma$ such that*

- *If $\operatorname{Re}\lambda = \frac{S}{2}$ then*

$$\frac{\pi_E(U_{\boldsymbol{\alpha}}(n))}{\sqrt{Sn \ln^{2\nu+1} n}} \to \mathcal{N}(0, \Sigma),$$

  *in distribution, asymptotically when $n$ tends to infinity.*
- *If $\operatorname{Re}\lambda < \frac{S}{2}$ then*

$$\frac{\pi_E(U_{\boldsymbol{\alpha}}(n))}{\sqrt{Sn}} \to \mathcal{N}(0, \Sigma),$$

  *in distribution, asymptotically when $n$ tends to infinity.*

*Moreover, $\Sigma$ does not depend on $\boldsymbol{\alpha}$.*

Let $E$ and $\lambda$ as in Theorem 2.2. The following result by Janson is the key of the proof of Theorem 2.2:

**Theorem 2.3** (Janson, 2004, Theorem 3.15 (i) and (ii))**.** *Under assumptions* (B), (T) *and* (S)*, for all vector $\boldsymbol{b} \in \mathbb{R}^d$, define*

$$\tau_{\boldsymbol{b}}(n) = \inf\{t \geqslant 0 \mid \langle \boldsymbol{b}, U_{\boldsymbol{\alpha}}^{CT}(t) \rangle \geqslant n\}.$$

*Then,*

  (i) *If $\operatorname{Re}\lambda = \frac{S}{2}$, then*

$$\frac{1}{\sqrt{Sn \ln^{2\nu+1} n}} \, \pi_E(U_{\boldsymbol{\alpha}}^{CT}(\tau_{\boldsymbol{b}}(n))) \to \mathcal{N}(0, \sigma),$$

  *in distribution, where $\sigma$ is a covariance matrix.*
 (ii) *If $\operatorname{Re}\lambda < \frac{S}{2}$, then*

$$\frac{1}{\sqrt{Sn}} \, \pi_E(U_{\boldsymbol{\alpha}}^{CT}(\tau_{\boldsymbol{b}}(n))) \to \mathcal{N}(0, \sigma),$$

  *n distribution, where $\sigma$ is a covariance matrix.*

Theorem 3.15 in Janson (2004) is slightly different than the above version. The above version corresponds to the special case $z = n$ in Janson's Theorem 3.15. The (*i*) of Janson (2004, Theorem 3.15) concerns the projection on the union of the small Jordan spaces, and it implies the (*i*) above by projection on a specified small Jordan space. The (*ii*) in Theorem 3.15 Janson (2004) is more general than the above version, which is the special case $k = \nu$ of Janson's result. The matrix $\sigma$ is given by Equations (3.11) and (2.15) in Janson (2004) for case (*i*) above, and by Equations (3.12) and (2.16) in Janson (2004) for case (*ii*) above. It is important to note that $\sigma$ does not depend on $\boldsymbol{\alpha}$.

Theorem 2.2 can be proved by using the *dummy* balls idea used in the proof of Theorems 3.21 and 3.22 in Janson (2004):

- Consider the continuous time urn with $d + 1$ colours, such that the $d + 1$ first colours evolve as the original $d$–colour process except that each time a ball splits, one ball of colour $d + 1$ is added to the process. When a ball of colour $d + 1$ splits, it splits into itself, adding no new balls in the process.
- Apply Theorem 2.3 to this $d + 1$–colour process (this process satisfies (S)).
- Go back to the original $d$–colour process by appropriate projection.

We do not develop the proof since no new idea is needed from there.

2.3. *Projections on large Jordan spaces.* Except for this digression on small eigenvalues, we are interested in the present paper in the behaviour of $U(n)$ along large Jordan blocks. **We will from now on fix $E$ a Jordan subspace of $R$ associated to a large eigenvalue $\lambda$.** We denote by $\nu + 1$ the size of its associated Jordan block (being also the dimension of $E$) and we denote by $v$ one eigenvector of $E$ associated to the eigenvalue $\lambda$.

2.3.1. *State of the art.* The asympotic behaviour of $U(n)$ projected onto the subspace $E$ is described by the following theorem:

**Theorem 2.4** (cf. Pouyanne, 2008). *Under* (B)*,* (T) *and* (S)*, if $\frac{1}{2} < \sigma < 1$, then,*

$$\lim_{n \to \infty} \frac{\pi_E(U(n))}{n^{\lambda/s} \ln^\nu n} = \frac{1}{\nu!} W^{DT} v,$$

*a.s. and in all $L^p$ ($p \geqslant 1$), where $\pi_E(U(n))$ is the projection of the composition vector at time $n$ onto $E$ (according to the Jordan decomposition of $R$).*

*Remark* 2.5. Note that different choices for $v$ are possible, and that the random variable $W^{DT}$ depends on this choice. The random variable $W^{DT}$ should actually be denoted by $W^{DT}_{E,v}$ since it depends on the Jordan subspace $E$ fixed and on the choice of $v$. For clarity's sake, we will stick to the ambiguous but simpler notation $W^{DT}$; there is no ambiguity since $E$ and $v$ are fixed all along the present paper. Note that one could also choose to include the $1/\nu!$ into the definition of $W^{DT}$; we choose to leave it outside as done in Pouyanne (2008).

In continuous time, the composition vector projected on a large stable subspace satisfies

**Theorem 2.6** (see Janson, 2004). *Under* (B)*,* (T) *and* (S)*, if $\frac{1}{2} < \sigma < 1$, then, almost surely and in all $L^p$ ($p \geqslant 1$),*

$$\lim_{t \to +\infty} \frac{\pi_E(U(t))}{t^\nu e^{\lambda t}} = \frac{1}{\nu!} W^{CT} v,$$

*where $\pi_E$ is the projection on the large stable subspace $E$. Moreover, the random variable $W^{CT}$ admits moments of all orders.*

*Remark* 2.7. Note that Theorem 2.6 is proven in Janson (2004) under (T$_{-1}$) and not (T); but Janson explains in his Remark 4.2 how to make it hold under (T).

We are interested in the two random variables $W^{DT}$ and $W^{CT}$ defined in Theorems 2.4 and 2.6. These random variables actually depend on the initial composition of the urn, denoted by $\boldsymbol{\alpha} =^t (\alpha_1, \ldots, \alpha_d)$, meaning that there are, for all $i \in \{1, \ldots, d\}$, $\alpha_i$ balls of type $i$ in the urn at time 0. It is thus more rigorous to denote by $W^{DT}_{\boldsymbol{\alpha}}$ (resp. $W^{CT}_{\boldsymbol{\alpha}}$) the random variable associated to the initial

composition $\boldsymbol{\alpha}$, emphasizing that we have to study two infinite families of random variables.

Connection (2.1) implies connections between the random variables $W$ induced by the discrete and continuous processes. We need the following result to deduce them:

$$\lim_{n \to \infty} n e^{-S\tau_n} = \xi, \tag{2.2}$$

almost surely, where $\xi$ is a random variable with Gamma law of parameter $(\alpha_1 + \ldots + \alpha_d)/S$. This result is shown for two–colour urns in Chauvin et al. (2011), and can be straightforwardly adapted to the present case, using the balanced hypothesis (B). We do not develop this proof, which is very standard in the study of Yule processes (see for example Athreya and Ney, 1972, page 120). We have

$$\frac{\pi_E\left(U^{CT}(\tau_n)\right)}{\tau_n^\nu e^{\lambda \tau_n}} = \frac{\pi_E\left(U^{DT}(n)\right)}{n^{\lambda/S} \ln^\nu n} \cdot \frac{n^{\lambda/S} \ln^\nu n}{\tau_n^\nu e^{\lambda \tau_n}}.$$

Moreover, Equation (2.2) implies $\frac{\ln n}{\tau_n} \to S$ when $n$ tends to $+\infty$,

$$\frac{n^{\lambda/S} \ln^\nu n}{\tau_n^\nu e^{\lambda \tau_n}} = \left(\frac{\ln n}{\tau_n}\right)^\nu \left(n e^{-S\tau_n}\right)^{\lambda/S} \to S^\nu \xi^{\lambda/S}.$$

Thus, for all initial composition $\boldsymbol{\alpha}$, we have (already mentioned in Janson, 2004):

$$W_{\boldsymbol{\alpha}}^{CT} \stackrel{(law)}{=} S^\nu \xi^{\lambda/S} W_{\boldsymbol{\alpha}}^{DT}, \tag{2.3}$$

where $\xi$ is a Gamma-distributed random variable with parameter $\frac{\alpha_1 + \ldots + \alpha_d}{S}$, and where $\xi$ and $W^{DT}$ are independent.

We also have that $(U^{DT}(n(t)))_{t \geq 0} = (U^{CT}(t))_{t \geq 0}$ almost surely, where $\sum_{i=1}^d \alpha_i + Sn(t)$ is the total number of balls in the continuous time urn at time $t$. It implies that, for all initial composition $\boldsymbol{\alpha}$,

$$W_{\boldsymbol{\alpha}}^{DT} \stackrel{(law)}{=} S^{-\nu} \xi^{-\lambda/S} W_{\boldsymbol{\alpha}}^{CT}, \tag{2.4}$$

where $\xi$ is a Gamma-distributed random variable with parameter $\frac{\alpha_1 + \ldots + \alpha_d}{S}$ but where $\xi$ and $W_{\boldsymbol{\alpha}}^{CT}$ are *not* independent, which can be verified via a covariance calculation.

2.3.2. *Statement of the main results.* The aim of the present paper is to gather information about $W_{\boldsymbol{\alpha}}^{DT}$ and $W_{\boldsymbol{\alpha}}^{CT}$. We will prove the two following theorems, which happen to be the first results on the variables $W$ induced by a multi–colour urn.

**Definition 2.8.** A complex-valued random variable $Z$ is moment-determined if for all random variable $Y$,

$$\mathbb{E}\left[Y^p \bar{Y}^q\right] = \mathbb{E}\left[Z^p \bar{Z}^q\right] \quad (\forall p, q \geq 1) \quad \Rightarrow \quad Y \stackrel{(law)}{=} Z.$$

**Theorem 2.9.** *Under assumptions* (B), (T) *and* (S), *for all initial composition* $\boldsymbol{\alpha}$,

   (i) *the random variable $W_{\boldsymbol{\alpha}}^{CT}$ is moment-determined.*
   (ii) *the Laplace series of $W_{\boldsymbol{\alpha}}^{DT}$ converges on the whole plane, which implies that $W_{\boldsymbol{\alpha}}^{DT}$ is moment-determined.*

Note that it is an open problem to determine whether the Laplace transform of $W^{CT}$ converges in a neighbourhood of 0. A similar result is already proved in Chauvin et al. (2015) in the two–colour case under assumptions (B), ($\mathtt{T_{-1}}$) and (I). Our proof is similar to the one developed there, although more involved due to the higher dimension and to the less restrictive tenability assumption: new higher–level arguments are needed here when simple calculations were sometimes enough in the two–colour case.

**Theorem 2.10.** *Under assumptions* (B), (T) *and* (I) *for all initial composition* $\boldsymbol{\alpha}$:
- *If* $\lambda \in \mathbb{C}\backslash\mathbb{R}$, *then the random variables* $W_{\boldsymbol{\alpha}}^{CT}$ *and* $W_{\boldsymbol{\alpha}}^{DT}$ *both admit a density on* $\mathbb{C}$, *and their support is the whole complex plane.*
- *If* $\lambda \in \mathbb{R}$, *then, the random variables* $W_{\boldsymbol{\alpha}}^{CT}$ *and* $W_{\boldsymbol{\alpha}}^{DT}$ *both admit a density on* $\mathbb{R}$, *and their support is the whole real line.*

A similar result is proved in Chauvin et al. (2011) or in Chauvin et al. (2015) for two–colours urns with the tenability condition ($\mathtt{T_{-1}}$) where the random variables $W^{CT}$ and $W^{DT}$ are real. Note that in dimension 2, an urn with negative entries in its replacement matrix admits no large Jordan block. Since the random variables $W^{CT}$ and $W^{DT}$ can be non-real in the multi–colour case, the proof of Theorem 2.10 needs additional input; for example, proving that the support of $W^{DT}$ is the whole complex plane becomes a non trivial step in the present article whereas the fact that the support was the whole real line it was straightforward in the real case.

## 3. Continuous time branching process – smoothing system

In this section, we focus on the continuous–time process and show how to use the tree–like structure of the process. It is the first step of the proofs of our two main results Theorems 2.9 and 2.10. As already mentioned, seeing urn schemes as branching processes is standard in the literature; the novelty of our approach is that we extend this analogy for urns that do not fall under the strong tenability condition ($\mathtt{T_{-1}}$) but only under the weaker (T).

First, we reduce the study to only $d$ initial compositions (instead of an infinite number), namely the initial compositions with a unique ball. Said differently, it is enough to study the random variables $W_{\boldsymbol{e_1}}, \ldots, W_{\boldsymbol{e_d}}$ where, for all $i \in \{1, \ldots, d\}$, $\boldsymbol{e_i}$ is the vector whose coordinates are all 0 except the $i^{\text{th}}$ which is 1 if $a_{i,i} \geqslant -1$ and $-a_{i,i}$ otherwise. We call $\boldsymbol{e_1}, \ldots, \boldsymbol{e_d}$ the **atomic initial compositions** of the urn. We then show, again using the tree-like structure of the process, that the random variables $W_{\boldsymbol{e_1}}, \ldots, W_{\boldsymbol{e_d}}$ satisfy a system of $d$ smoothing equations.

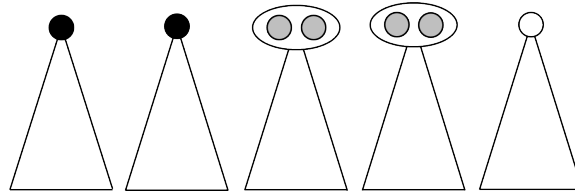First introduce further notations: For all $i \in \{1, \ldots, d\}$, let us denote

$$\tilde{\alpha}_i = \begin{cases} \alpha_i & \text{if } a_{i,i} \geqslant -1, \\[2mm] \dfrac{\alpha_i}{-a_{i,i}} & \text{otherwise,} \end{cases} \tag{3.1}$$

and for all $i, c \in \{1, \ldots, d\}$, let us denote

$$\tilde{a}_{c,i} = \begin{cases} a_{c,i} & \text{if } a_{i,i} \geqslant -1, \\[2mm] \dfrac{a_{c,i}}{-a_{i,i}} & \text{otherwise.} \end{cases} \tag{3.2}$$

In view of Assumption (T), for all $i, c \in \{1, \ldots, d\}$, $\tilde{a}_{c,i}$ and $\tilde{\alpha}_i$ are integers.

FIGURE 3.1. Decomposition of the urn process in continuous time
– example.



*Remark* 3.1. If we suppose further that $(\mathtt{T_{-1}})$ holds, then $\tilde{\alpha}_i = \alpha_i$ for all $i \in \{1, \dots, d\}$ and $\tilde{a}_{i,j} = a_{i,j}$ for all $i, j \in \{1, \dots, d\}$.

3.1. *Decomposition.* To explain how to decompose the continuous time urn process, we will focus on an example, before generalising to any urn process that satisfies $(\mathtt{B})$, $(\mathtt{S})$ and $(\mathtt{T})$. Assume for example that

$$R = \begin{pmatrix} 6 & 2 & 0 \\ 5 & -2 & 5 \\ 0 & 2 & 6 \end{pmatrix}.$$

One can verify that the urn process defined by $R$ satisfies $(\mathtt{B})$, $(\mathtt{I})$ and $(\mathtt{T})$. Moreover, its eigenvalues are 8, 6 and $-4$. In particular, 6 is a large eigenvalue which allows us to apply Section 2 and define $W_{\boldsymbol{\alpha}}^{CT}$ through Theorem 2.6 for any initial composition $\boldsymbol{\alpha}$. In the following, we will denote by $E_6$ the one-dimensional Jordan stable subspace associated to this eigenvalue 6 and by $\pi_6$ the Jordan projection onto it.

Note that we can decompose the multitype branching process as shown in Figure 3.1, which gives the following

$$U_{(2,4,1)}^{CT} \overset{(law)}{=} \sum_{k=1}^{2} U_{(1,0,0)}^{(k)} + \sum_{k=3}^{4} U_{(0,2,0)}^{(k)} + \sum_{k=5}^{5} U_{(0,0,1)}^{(k)},$$

where the $U^{(p)}$ are independent urn processes.

Let us quit the example and make the same reasoning as above in full generality under $(\mathtt{B})$, $(\mathtt{S})$ and $(\mathtt{T})$. Recall that, for all $c \in \{1, \dots, d\}$, $\boldsymbol{e_c}$ has all its coordinates equal to zero except for the $c^{\text{th}}$, which is equal to 1 if $a_{c,c} \geqslant 0$ and to $-a_{c,c}$ otherwise. We get

$$U_{\boldsymbol{\alpha}}^{CT}(t) \overset{(law)}{=} \sum_{c=1}^{d} \sum_{p=\beta_{c-1}+1}^{\beta_c} U_{\boldsymbol{e_c}}^{(p)}(t),$$

where $\beta_0 = 0$ and $\beta_c = \sum_{j \leqslant c} \tilde{\alpha}_j$, and where the $U_{\boldsymbol{e_c}}^{(p)}(t)$ are independent copies of $U_{\boldsymbol{e_c}}^{CT}(t)$, independent of each other.

Dividing this equality in law by $t^\nu e^{\lambda t}$, projecting onto the fixed large Jordan subspace $E$ via $\pi_E$, and applying Theorem 2.6 gives

**Proposition 3.2** (already mentioned in Janson, 2004, Remark 4.2). *For all replacement matrices $R$ and initial composition $\boldsymbol{\alpha}$ satisfying* (B), (S) *and* (T),

$$W_{\boldsymbol{\alpha}}^{CT} \overset{(law)}{=} \sum_{c=1}^{d} \sum_{p=\beta_{c-1}+1}^{\beta_c} W_{\boldsymbol{e_c}}^{(p)},$$

*where the $W_{\boldsymbol{e_c}}^{(p)}$ are independent copies of $W_{\boldsymbol{e_c}}^{CT}$, independent of each other and independent of $U$.*

Proposition 3.2 allows to reduce the study to only $d$ random variables, namely $(W_{\boldsymbol{e_1}}^{CT}, \ldots, W_{\boldsymbol{e_d}}^{CT})$ instead of having to study an infinite family of such variables. Any information gathered about those $d$ random variables will a priori give us some information about any $W_{\boldsymbol{\alpha}}^{CT}$.

3.2. *Dislocation.* In view of Proposition 3.2, it is enough to focus on the $d$ atomic initial compositions $\boldsymbol{e_1}, \ldots, \boldsymbol{e_d}$. Recall that for all $i \in \{1, \ldots, d\}$, $\boldsymbol{e_i}$ is the vector whose all components are zero, except the $i^{\text{th}}$ which is equal to 1 if $a_{i,i} \geqslant 0$ and to $-a_{i,i}$ otherwise. We will from now on denote by $\theta_i \in \{1, 2, \ldots\}$ the non zero component of $\boldsymbol{e_i}$:

$$\theta_i = \begin{cases} 1 & \text{if } a_{i,i} \geqslant 0 \\ -a_{i,i} & \text{otherwise.} \end{cases} \tag{3.3}$$

Let us again study first the particular example given by

$$R = \begin{pmatrix} 6 & 2 & 0 \\ 5 & -2 & 5 \\ 0 & 2 & 6 \end{pmatrix}.$$

The three atomic initial compositions are given by $\boldsymbol{e_1} = (1, 0, 0)$, $\boldsymbol{e_2} = (0, 2, 0)$ and $\boldsymbol{e_3} = (0, 0, 1)$, since $a_{1,1}, a_{3,3} \geqslant 0$ and $a_{2,2} = -2$. In all three cases, the first step is deterministic: we know the colour of the first ball to be drawn and we therefore know what is the composition of the urn after the first split time (cf. Figure 3.2). We therefore have that,

$$U_{\boldsymbol{e_1}}^{CT} \overset{(law)}{=} \sum_{k=1}^{7} U_{\boldsymbol{e_1}}^{(k)}(t - \tau^{(1)}) + \sum_{k=8}^{8} U_{\boldsymbol{e_2}}^{(k)}(t - \tau^{(1)}),$$

$$U_{\boldsymbol{e_2}}^{CT} \overset{(law)}{=} \sum_{k=1}^{5} U_{\boldsymbol{e_1}}^{(k)}(t - \tau^{(2)}) + \sum_{k=6}^{10} U_{\boldsymbol{e_3}}^{(k)}(t - \tau^{(2)}),$$

$$U_{\boldsymbol{e_3}}^{CT} \overset{(law)}{=} \sum_{k=1}^{1} U_{\boldsymbol{e_2}}^{(k)}(t - \tau^{(3)}) + \sum_{k=2}^{8} U_{\boldsymbol{e_3}}^{(k)}(t - \tau^{(3)}),$$

where the $U^{(k)}$ are independent continuous time urn processes with replacement matrix $R$, where $\tau^{(1)}$, $\tau^{(2)}$ and $\tau^{(3)}$ are independent random variables exponentially distributed of respective parameters 1, 2 and 1. The random variables $\tau^{(1)}$, $\tau^{(2)}$ and $\tau^{(3)}$ are the first split times that occur in urns of respective initial compositions $\boldsymbol{e_1}$, $\boldsymbol{e_2}$ and $\boldsymbol{e_3}$.

FIGURE 3.2. Dislocation of a continuous time urn process – the different atomic initial compositions and their composition after the first drawing.
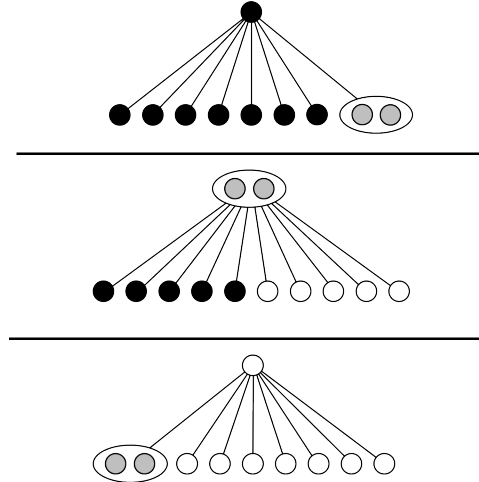
The same reasoning in full generality, for all replacement matrices $R$ satisfying (B), (S) and (T) gives that, for all $c \in \{1, \ldots, d\}$,

$$U_{\boldsymbol{e_c}}^{CT} \overset{(law)}{=} \sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} U_{\boldsymbol{e_i}}^{(k)}(t - \tau^{(c)}),$$

where $\gamma_0^{(c)} = 0$ and $\gamma_i^{(c)} = \sum_{j \leqslant i} \tilde{a}_{c,i} + \delta_{c,i}$, $\tau^{(c)}$ is an exponentially distributed random variable of parameter $\theta_c$ (which is a positive integer), and the $U^{(k)}$ are independent continuous time urn processes with replacement matrix $R$. We recall that $E$ is a fixed Jordan space of dimension $\nu + 1$ associated to a large eigenvalue $\lambda$. Dividing the previous equality in law by $t^\nu e^{\lambda t}$, projecting onto $E$ via $\pi_E$, and applying Theorem 2.6 gives

**Proposition 3.3** (already mentioned in Janson, 2004, Theorem 3.9). *Under assumptions* (B), (S) *and* (T), *for all $c \in \{1, \ldots, d\}$,*

$$W_{\boldsymbol{e_c}}^{CT} \overset{(law)}{=} U^{\lambda/\theta_c} \sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} W_{\boldsymbol{e_i}}^{(k)}, \qquad (3.4)$$

*where $\gamma_0^{(c)} = 0$, $\gamma_i^{(c)} = \sum_{j \leqslant i}(\tilde{a}_{c,j} + \delta_{c,j})$ (using Kronecker's notation $\delta_{c,j} = 1$ if $c = j$, 0 otherwise), where $U$ is a uniform random variable on $[0, 1]$, and where the $W_{\boldsymbol{e_i}}^{(k)}$ are independent copies of $W_{\boldsymbol{e_i}}^{CT}$, independent of each other and of $U$.*

*Remark* 3.4. If we assume $(T_{-1})$ instead of (T) in the result above, we get

$$W_{\boldsymbol{e_c}}^{CT} \overset{(law)}{=} U^\lambda \sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} W_{\boldsymbol{e_i}}^{(k)},$$

where $\gamma_0^{(c)} = 0$, $\gamma_i^{(c)} = \sum_{j \leqslant i}(a_{c,j} + \delta_{c,j})$, where $U$ is a uniform random variable on $[0,1]$, and where the $W_{\boldsymbol{e_i}}^{(k)}$ are independent copies of $W_{\boldsymbol{e_i}}^{CT}$, independent of each other and of $U$.

*Remark* 3.5. One can prove that the solution of System (3.4) is unique at fixed mean and under a condition of finite variance. A very similar proof in done in Janson (2004, Proof of Theorem 3.9(iii), page 232–233).

## 4. Moments – Proof of Theorem 2.9

This section is devoted to the proof of Theorem 2.9. We study here the moments of the random variables $(W_{\boldsymbol{e_i}}^{DT})_{i \in \{1,\dots,d\}}$ and $(W_{\boldsymbol{e_i}}^{CT})_{i \in \{1,\dots,d\}}$. First note that the convergence in all $L^p$ ($p \geqslant 1$) stated in Theorems 2.4 and 2.6 ensures us that those random variables admit moments of all orders. The first step in the proof is the following lemma, which concerns the continuous time process: Equation (2.4) will then allow us to infer results about the discrete time process.

**Lemma 4.1.** *Let* $(X_1, \dots, X_d)$ *be a solution of System* (3.4) *with moments of all orders. Then the sequences* $\left( \dfrac{\mathbb{E}|X_i|^p}{p! \ln^p p} \right)^{\frac{1}{p}}$, *for all* $i \in \{1, \dots, d\}$, *are bounded.*

*Proof*: Let $(X_1, \dots, X_d)$ be a solution of System (3.4), let $\varphi(p) := \ln^p(p+2)$ and let, for all $i \in \{1, \dots, d\}$,
$$u_p^{(i)} := \frac{\mathbb{E}|X_i|^p}{p! \varphi(p)}.$$
Let us prove by induction on $p \geqslant 1$ that, for all $i \in \{1, \dots, d\}$, the sequence $\left( \frac{\mathbb{E}|X_i|^p}{p! \varphi(p)} \right)^{\frac{1}{p}}$ is bounded. Raise the equations of System (3.4) to the power $p$ and apply the multinomial formula. Since $\mathbb{E}|U^{\mu p}| = \frac{1}{p \mathrm{Re}\mu + 1}$ for all $\mu \in \mathbb{C}$ (with $U$ uniformly distributed on $[0,1]$), for all $c \in \{1, \dots, d\}$,

$$\mathbb{E}|X_c|^p \leqslant \frac{1}{p\frac{\mathrm{Re}\lambda}{\theta_c} + 1} \Bigg( \sum_{i=1}^{d}(\tilde{a}_{c,i} + \delta_{c,i})\mathbb{E}|X_i|^p$$

$$+ \sum_{\substack{p_1 + \dots + p_{\gamma_d^{(c)}} = p \\ p_j \leqslant p-1}} \frac{p!}{\prod_{1 \leqslant j \leqslant \gamma_d^{(c)}} p_j!} \prod_{i=1}^{d} \prod_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \mathbb{E}|X_i|^{p_k} \Bigg),$$

which means

$$p\frac{\mathrm{Re}\lambda}{\theta_c}\mathbb{E}|X_c|^p \leqslant \sum_{i=1}^{d} \tilde{a}_{c,i}\mathbb{E}|X_i|^p + \sum_{\substack{p_1 + \dots + p_{\gamma_d^{(c)}} = p \\ p_j \leqslant p-1}} \frac{p!}{\prod_{1 \leqslant j \leqslant \gamma_d^{(c)}} p_j!} \prod_{i=1}^{d} \prod_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \mathbb{E}|X_i|^{p_k}.$$

It implies that, for all $c \in \{1, \dots, d\}$,

$$\frac{p\mathrm{Re}\lambda}{\theta_c}u_p^{(c)} \leqslant \sum_{i=1}^{d} \tilde{a}_{c,i}u_p^{(i)} + \sum_{\substack{p_1 + \dots + p_{\gamma_d^{(c)}} = p \\ p_j \leqslant p-1}} \frac{\prod_{1 \leqslant j \leqslant \gamma_d^{(c)}} \varphi(p_j)}{\varphi(p)} \prod_{i=1}^{d} \prod_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} u_{p_k}^{(i)}. \quad (4.1)$$

Let

$$\Phi_c(p) := \sum_{\substack{p_1 + \cdots + p_{\gamma_d^{(c)}} = p \\ p_j \leqslant p-1}} \frac{\prod_{1 \leqslant j \leqslant \gamma_d^{(c)}} \varphi(p_j)}{\varphi(p)}. \tag{4.2}$$

A slight generalisation of Chauvin et al. (2015, Lemma 1) ensures that

$$\Phi_c(p) \leqslant (1 + 8\ln(p+2))^{\gamma_d^{(c)}}, \tag{4.3}$$

for all $p \geqslant 2$, as soon as $\gamma_d^{(c)} \geqslant 1$; recall that $\gamma_d^{(c)} = \sum_{i=1}^d (\tilde{a}_{c,i} + \delta_{c,i}) \geqslant 1$.

Denote by $\Delta_p$ the determinant of $p\mathrm{Re}\lambda\Theta - \tilde{R}$ where $\Theta = \mathtt{diag}\{\theta_1^{-1}, \ldots, \theta_d^{-1}\}$ and $\tilde{R} = (\tilde{a}_{i,j})_{1 \leqslant i,j \leqslant d}$. This determinant is non-zero for all $p \geqslant 2$ since

$$\Delta_p = \mathtt{det}(p\mathrm{Re}\lambda\Theta - \tilde{R}) = \left(\prod_{i=1}^d \theta_i^{-1}\right) \mathtt{det}(p\mathrm{Re}\lambda I_d - R)$$

and since $p\mathrm{Re}\lambda > S$ (as, by assumption, $\mathrm{Re}\lambda > S/2$). Note that, under $(\mathtt{T_{-1}})$, we have $\Theta = I_d$.

Recall that $\theta_i \geqslant 1$ for all $i \in \{1, \ldots, d\}$ (see Equation (3.3)), and note that

$$\frac{\|p\mathrm{Re}\lambda\Theta - \tilde{R}\|_\infty}{p\mathrm{Re}\lambda} \to \frac{1}{\min_{1 \leqslant i \leqslant d} \theta_i}$$

when $p$ tends to $+\infty$, implying that there exists $p_0 \geqslant 1$ such that, for all $p \geqslant p_0$,

$$\frac{\|p\mathrm{Re}\lambda\Theta - \tilde{R}\|_\infty}{p\mathrm{Re}\lambda} \leqslant \frac{2}{\min \theta_i} =: \rho. \tag{4.4}$$

In addition, let us denote by $\Delta_p(j,i)$ the determinant of $p\mathrm{Re}\lambda\Theta - \tilde{R}$ in which the $i^{\mathrm{th}}$ column and the $j^{\mathrm{th}}$ line have been removed. For all $1 \leqslant i, j \leqslant d$, the polynomial $\Delta_p(j,i)$ has degree at most $d-1$ in $p$, which implies

$$\sup_{1 \leqslant i,j \leqslant d} \frac{|\Delta_p(j,i)|}{|\Delta_p|} = \mathcal{O}\left(\frac{1}{p}\right),$$

when $p$ goes to infinity, and there exists a constant $\eta > 0$ and an integer $p_1 \geqslant p_0$ such that, for all $p \geqslant p_1$,

$$\sup_{1 \leqslant i,j \leqslant d} \frac{|\Delta_p(j,i)|}{|\Delta_p|} \leqslant \frac{\eta}{p}.$$

Finally, let us denote by $\Delta_p(c)$ the determinant of the matrix $p\mathrm{Re}\lambda\Theta - \tilde{R}$ in which the $c^{\mathrm{th}}$ column has been replaced by a column of 1. We know that $\Delta_p$ has degree $d$ in $p$ whereas $\Delta_p(c)$ is a polynomial with degree at most $d-1$ in $p$. It implies that there exists an integer $p_2 \geqslant p_1$ such that, for all $p \geqslant p_2$, for all $c \in \{1, \ldots, d\}$,

$$\frac{\Delta_p(c)}{\Delta_p}(1 + 8\ln(p+2))^{\gamma_d^{(c)}} \leqslant \frac{1}{\rho\eta d^2 \mathrm{Re}\lambda}, \tag{4.5}$$

where the choice of the right-hand side constant will become clear later on.

Let us define

$$A := \max\{(u_q^{(i)})^{\frac{1}{q}}, 1 \leqslant q \leqslant p_2, 1 \leqslant i \leqslant d\},$$

and prove by induction on $p \geqslant p_2$ that, for all $q \leqslant p$ and $c \in \{1, \ldots, d\}$, $(u_q^{(c)})^{\frac{1}{q}} \leqslant A$. Fix $p > p_2$ and assume that the induction hypothesis is true for $p - 1$; then

Equations ([4.3](#)) and ([4.1](#)) imply

$$p\mathrm{Re}\lambda\Theta u_p^{(c)} \leqslant \sum_{i=1}^d \tilde{a}_{c,i} u_p^{(i)} + A^p \Phi_c(p).$$

Let $(v_1, \ldots, v_d)$ be the solution of the system

$$p\mathrm{Re}\lambda\Theta v_c = \sum_{i=1}^d \tilde{a}_{c,i} v_i + A^p \Phi_c(p).$$

We thus have that $v_c \geqslant u_p^{(c)}$ for all $1 \leqslant c \leqslant d$. Solving this Cramèr system, we get, in view of Equations ([4.3](#)) and ([4.5](#)),

$$v_c = A^p \Phi_c(p) \frac{\Delta_p(c)}{\Delta_p} \leqslant \frac{A^p}{\rho\eta d^2 \mathrm{Re}\lambda}.$$

For all $p \geqslant p_2$, using the fact that for all $d$-dimensional matrix $M$ and all vector $\boldsymbol{v}$, $\|M\boldsymbol{v}\|_\infty \leqslant d\|M\|_\infty \|\boldsymbol{v}\|_\infty$, we get

$$(p\mathrm{Re}\lambda\Theta - \tilde{R})\boldsymbol{u}^{(p)} \leqslant (p\mathrm{Re}\lambda\Theta - \tilde{R})\boldsymbol{v} \leqslant \|p\mathrm{Re}\lambda\Theta - \tilde{R}\|_\infty \frac{A^p}{\rho\eta d\mathrm{Re}\lambda}\boldsymbol{\omega},$$

where $\boldsymbol{u}^{(p)}$ and $\boldsymbol{v}$ denote the vectors of respective coordinates $(u_p^{(i)})_{1\leqslant i\leqslant d}$ and $(v_i)_{1\leqslant i\leqslant d}$, where $\boldsymbol{\omega}$ is the vector whose all coordinates are equal to 1, and where the sign $\leqslant$ between two vectors is to be read coordinate by coordinate. In particular, we have

$$\|(p\mathrm{Re}\lambda\Theta - \tilde{R})\boldsymbol{u}^{(p)}\|_\infty \leqslant A^p \frac{\|p\mathrm{Re}\lambda\Theta - \tilde{R}\|_\infty}{\rho\eta d\mathrm{Re}\lambda} \leqslant A^p \frac{p}{\eta d},$$

where we have used Equation ([4.4](#)). Let us denote $M = p\mathrm{Re}\lambda\Theta - \tilde{R}$. The coefficients of $M^{-1}$ are given by

$$(M^{-1})_{i,j} = (-1)^{i+j} \frac{\Delta_p(j,i)}{\Delta_p},$$

where $\Delta_p$ is the determinant of $M$, and $\Delta_p(j,i)$ is the determinant of the matrix $M$ in which the $j^{\mathrm{th}}$ line and the $i^{\mathrm{th}}$ column have been removed. By definition of $p_2$, for all $p \geqslant p_2$,

$$\|M^{-1}\|_\infty = \sup_{1\leqslant i,j\leqslant d} |(M^{-1})_{i,j}| \leqslant \frac{\eta}{p},$$

which implies that, for all $p \geqslant p_2$,

$$\|\boldsymbol{u}^{(p)}\|_\infty \leqslant \|M^{-1}\|_\infty A^p \frac{p}{\eta} \leqslant A^p.$$

Finally, for all $c \in \{1, \ldots, d\}$, $u_c^{(p)} \leqslant A^p$, which concludes the proof.    $\square$

*Proof of Theorem [2.9](#): (i)* To prove that a real-valued random variable $X$ is moment-determined, one can show that it satisfies Carleman's criterion:

$$\sum_{k=1}^\infty \mathbb{E}[X^{2k}]^{-1/2k} = \infty.$$

For a complex-valued random variable $Z$, one can for example apply Janson and Kaijser (2015, Theorem 10.3), which states that, if all moments of $|Z|$ are finite and if $|Z|$ satisfies Carleman's criterion, then $Z$ is moment-determined in the sense of Definition [2.8](#). Lemma [4.1](#) implies that, for all $i \in \{1, \ldots, d\}$, the random variable $|W_{\boldsymbol{e}_i}^{CT}|$, which admits moments of all orders in view of Theorem [2.6](#) satisfies

Carleman's criterion. Therefore, $W_{e_i}^{CT}$ is moment-determined. Proposition 3.2 eventually allows us to generalise Lemma 4.1 to any initial composition: For all initial composition $\boldsymbol{\alpha}$, $|W_{\boldsymbol{\alpha}}^{CT}|$ also satisfies the Carleman's criterion.

*(ii)* By Lemma 4.1, we have the following inequality: there exists a constant $C$ such that, for all integer $p \geqslant 2$, for any initial composition,

$$\frac{\mathbb{E}|W^{CT}|^p}{p!} \leqslant C^p \ln^p p.$$

It implies, via Equation (2.3), there exists a constant $D$ such that, for all integer $p \geqslant 2$,

$$\frac{\mathbb{E}|W^{DT}|^p}{p!} \leqslant D^p \frac{\ln^p p}{\Gamma\left(\frac{p\mathrm{Re}\lambda+1}{S}\right)},$$

where we recall that $\mathrm{Re}\lambda > S/2 > 0$. This implies that the Laplace series of $|W^{DT}|$ has an infinite radius of convergence. Since $\langle t, W^{DT}\rangle \leqslant 2|t|\|W^{DT}|$, this implies that the Laplace transform of $W^{DT}$ converges on the whole complex plane. $\qquad\square$

Theorem 2.9 gives an upper bound for the moments of $W^{CT}$ and $W^{DT}$; note that no lower bound is known up to now.

## 5. Discrete time urn process – Smoothing system

5.1. *Smoothing system in discrete time.* This subsection is devoted to deduce from Sections 3 and 4 that the random variable $(W_{e_1}^{DT}, \ldots, W_{e_d}^{DT})$ is a solution of a smoothing system:

**Proposition 5.1.** *Under assumptions* (B), (T) *and* (S), *for every colour* $1 \leqslant c \leqslant d$,

$$W_{e_c}^{DT} \stackrel{(law)}{=} \sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \left(V_k^{(c)}\right)^{\lambda/S} W_{e_i}^{(k)}, \tag{5.1}$$

*where* $\gamma_0 = 0$ *and* $\gamma_i^{(c)} = \sum_{j=1}^i (\tilde{a}_{c,j} + \delta_{c,j})$ *for all* $i \in \{1, \ldots, d\}$; *where the* $W_{e_i}^{(k)}$ *are independent copies of* $W_{e_i}^{DT}$, *independent of each other; and where* $V^{(c)} = (V_1, \ldots, V_{\gamma_d^{(c)}})$ *is a Dirichlet-distributed random vector independent of the* $W$ *and of parameter* $\boldsymbol{\pi}$, *given by*

$$\pi_k = \theta_i/S \quad if \quad \gamma_{i-1}^{(c)} < k \leqslant \gamma_i^{(c)}.$$

*Remark* 5.2. If we assume $(\mathtt{T_{-1}})$ instead of $(\mathtt{T})$ in the theorem above, we obtain the following system:

$$W_{e_c}^{DT} \stackrel{(law)}{=} \sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} V_k^{\lambda/S} W_{e_i}^{(k)},$$

where $\gamma_0 = 0$ and $\gamma_i^{(c)} = \sum_{j=1}^i (a_{c,j} + \delta_{c,j})$ for all $i \in \{1, \ldots, d\}$; where the $W_{e_i}^{(k)}$ are independent copies of $W_{e_i}^{DT}$, independent of each other; and where $V = (V_1, \ldots, V_{S+1})$ is a Dirichlet-distributed random vector of parameter $\left(\frac{1}{S}, \ldots, \frac{1}{S}\right)$, independent of the $W$.

*Proof*: We will present two proof for this statement: the first one is a moment proof, developed hereafter, the second one is the classical one based on the branching property of the urn process, we will detail this other proof in Subsection 5.4.

Let us prove that, for all $p, q \geqslant 1$,

$$\mathbb{E}\Big[(W_{e_c}^{DT})^p (\overline{W}_{e_c}^{DT})^q\Big] = \mathbb{E}\bigg[\bigg(\sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} V_k^{\lambda/S} W_{e_i}^{(k)}\bigg)^p \bigg(\sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} V_k^{\bar\lambda/S} \overline{W}_{e_i}^{(k)}\bigg)^q\bigg].$$

Since, $W_{e_c}^{DT}$ is moment-determined in view of Theorem 2.9, this will conclude the proof. Let us use Connection (2.1), which gives, for all $c \in \{1, \ldots, d\}$,

$$\mathbb{E}\big[(W_{e_c}^{CT})^p (\overline{W}_{e_c}^{CT})^q\big] = S^{(p+q)\nu} \mathbb{E}\Big[\xi_c^{\frac{p\lambda+q\bar\lambda}{S}}\Big] \mathbb{E}\big[(W_{e_c}^{DT})^p (\overline{W}_{e_c}^{DT})^q\big], \qquad (5.2)$$

where $\xi_c$ is a Gamma-distributed random variable, of parameter $\theta_c/s$.

Let $V_c$ be a Beta-distributed random variable of parameter $(\theta_c/s, 1)$. Note that if $U$ is uniformly distributed on $[0, 1]$, we have

$$V_c \stackrel{(law)}{=} U^{S/\theta_c}.$$

Let $(\zeta_{i,k})_{1 \leqslant i \leqslant d; k \geqslant 1}$ be a sequence of independent, Gamma-distributed random variables of parameter $\theta_c/s$. The random variable

$$\zeta_c = V_c \sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}}^{\gamma_i^{(c)}} \zeta_{i,k}$$

is Gamma-distributed with parameter $\theta_c/s$ (it can be verified by calculating its moments). Finally, let

$$V_{i,k}^{(c)} = \frac{\zeta_{i,k}}{\sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}}^{\gamma_i^{(c)}} \zeta_{i,k}}.$$

Then (see for example Bertoin, 2006, Lemma 2.2), the $\gamma_d^{(c)}$-dimensional random vector $V^{(c)}$ whose coordinates are given by

$$V_k^{(c)} = V_{i,k}^{(c)} \quad \text{if } \gamma_{i-1}^{(c)} < k \leqslant \gamma_i^{(c)}$$

is Dirichlet-distributed of parameter $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_{\gamma_d^{(c)}})$, where

$$\pi_k = \theta_i/s \quad \text{if} \quad \gamma_{i-1}^{(c)} < k \leqslant \gamma_i^{(c)},$$

and independent from $\zeta_c$. Thus, System (3.4) together with Equation (5.2), gives that, for all $c \in \{1, \ldots, d\}$,

$$S^{(p+q)\nu} \mathbb{E}\Big[\xi_c^{\frac{p\lambda + q\bar{\lambda}}{S}}\Big] \mathbb{E}\big[(W_{e_c}^{DT})^p (\overline{W}_{e_c}^{DT})^q\big] \tag{5.3}$$

$$= \mathbb{E}\Bigg[\Big(V_c^{\lambda/S} \sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} W_{e_i}^{CT,(k)}\Big)^p \Big(V_c^{\bar{\lambda}/S} \sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \overline{W}_{e_i}^{CT,(k)}\Big)^q\Bigg]$$

$$= \mathbb{E}\Bigg[\Big(\sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} V_c^{\lambda/S} S^\nu \zeta_{i,k}^{\lambda/S} W_{e_i}^{DT,(k)}\Big)^p \Big(\sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} V_c^{\bar{\lambda}/S} S^\nu \zeta_{i,k}^{\bar{\lambda}/S} \overline{W}_{e_i}^{DT,(k)}\Big)^q\Bigg]$$

$$= S^{(p+q)\nu} \mathbb{E}\Bigg[\Big(\sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} (\zeta_{i,k} V_c)^{\lambda/S} W_{e_i}^{DT,(k)}\Big)^p \Big(\sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} (\zeta_{i,k} V_c)^{\bar{\lambda}/S} \overline{W}_{e_i}^{DT,(k)}\Big)^q\Bigg].$$

$$\tag{5.4}$$

We also have that, by independence of $V^{(c)}$ and $\zeta_c$,

$$\mathbb{E}\Big[\zeta_c^{\frac{p\lambda + q\bar{\lambda}}{S}}\Big] \mathbb{E}\Bigg[\Big(\sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \big(V_{i,k}^{(c)}\big)^{\lambda/S} W_{e_i}^{DT,(k)}\Big)^p \Big(\sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \big(V_{i,k}^{(c)}\big)^{\bar{\lambda}/S} \overline{W}_{e_i}^{DT,(k)}\Big)^q\Bigg]$$

$$= \mathbb{E}\Bigg[\Big(\sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \big(\zeta_c V_{i,k}^{(c)}\big)^{\lambda/S} W_{e_i}^{DT,(k)}\Big)^p \Big(\sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \big(\zeta_c V_{i,k}^{(c)}\big)^{\bar{\lambda}/S} \overline{W}_{e_i}^{DT,(k)}\Big)^q\Bigg]$$

$$\mathbb{E}\Bigg[\Big(\sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} (\zeta_{i,k} V_c)^{\lambda/S} W_{e_i}^{DT,(k)}\Big)^p \Big(\sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} (\zeta_{i,k} V_c)^{\bar{\lambda}/S} \overline{W}_{e_i}^{DT,(k)}\Big)^q\Bigg]. \tag{5.5}$$

The result follows from Equations (5.4) and (5.5), using the fact that $W^{DT}$ is moment-determined. $\qquad\square$

5.2. *Unicity.* The main goal of this section is to prove that the solution of System (5.1) is unique. We therefore use the so-called contraction method. This method, presented for example in Neininger-Rüschendorf's survey (Neininger and Rüschendorf, 2006) consists in applying the Banach fixed point theorem in an appropriate complete Banach space. It has already been used in a Pólya urn context in the literature. In Knape and Neininger (2014) the contraction method is used as a new approach to prove an equivalent of Theorem 2.4 for large and small eigenvalues (for discrete time two-colour urns). In Chauvin et al. (2015), it is used as in the present paper, to prove the unicity of the solution of a two-equation system, in the study of large two–colour Pólya urns. In Janson (2004), it is also used to prove the unicity of the solution of system (3.4): therefore, we will only develop the proof for the discrete case. Similar proofs can be found in Knape and Neininger (2014) or Chauvin et al. (2015).

Let $\mathcal{M}_2$ be the space of complex-valued square integrable probability measures. For all $A \in \mathbb{C}$, let $\mathcal{M}_2^{\mathbb{C}}(A)$ be the subspace of measures in $\mathcal{M}_2$ with mean $A$. We

consider the Wasserstein distance as follows: for all $\mu$, $\nu$ two measures in $\mathcal{M}_2^{\mathbb{C}}(A)$,

$$d_W(\mu, \nu) = \inf_{X \sim \mu, Y \sim \nu} \|X - Y\|_2,$$

where $\|\cdot\|_2$ is the $L^2$-norm on $\mathbb{C}$.

For all $A_1, \ldots, A_d \in \mathbb{C}$. Let us denote by $\times_{i=1}^d \mathcal{M}_2^{\mathbb{C}}(A_i)$ the Cartesian product of the spaces $\mathcal{M}_2^{\mathbb{C}}(A_i)$. We define the Wasserstein distance on this space as follows: for all $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d)$ and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_d)$ two elements of $\times_{i=1}^d \mathcal{M}_2^{\mathbb{C}}(A_i)$,

$$d(\boldsymbol{\mu}, \boldsymbol{\nu}) = \max_{1 \leqslant i \leqslant d} \{d_W(\mu_i, \nu_i)\}.$$

We know that $(\mathcal{M}_2^{\mathbb{C}}(A), d_W)$ and thus $\times_{i=1}^d \mathcal{M}_2^{\mathbb{C}}(A_i)$ are complete metric spaces (see for example Dudley, 2002).

The random vector $(W_{\boldsymbol{e_1}}^{DT}, \ldots, W_{\boldsymbol{e_d}}^{DT})$ is a solution of System (5.1):

$$W_{\boldsymbol{e_c}} \overset{(law)}{=} \sum_{i=1}^d \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \left(V_k^{(c)}\right)^{\lambda/S} W_{\boldsymbol{e_i}}^{(k)}.$$

For all $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d) \in \times_{i=1}^d \mathcal{M}_2^{\mathbb{C}}(m_i)$, for all $c \in \{1, \ldots, d\}$, let

$$K_c(\boldsymbol{\mu}) = \mathcal{L}\left(\sum_{i=1}^d \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \left(V_k^{(c)}\right)^{\lambda/S} X_i^{(k)}\right),$$

where $\gamma_0 = 0$, $\gamma_i^{(c)} = \sum_{j \leqslant i}(\tilde{a}_{c,j} + \delta_{c,j})$ and for all $i \in \{1, \ldots, d\}$, the $(X_i^{(k)})_{1 \leqslant i \leqslant d}$ are independent random variables, independent of each other and of vector $V$, which is Dirichlet-distributed of parameter $\left(\pi_1^{(c)}, \ldots \pi_{\gamma_d^{(c)}}^{(c)}\right)$, and, for all $1 \leqslant i \leqslant d$ and $1 \leqslant k \leqslant \gamma_d^{(c)}$,

$$\gamma_{i-1}^{(c)} + 1 \leqslant k \leqslant \gamma_i^{(c)} \Rightarrow X_i^{(k)} \sim \mu_i \text{ and } \pi_k^{(c)} = \theta_i/S. \tag{5.6}$$

We define the function $K$ as

$$K(\boldsymbol{\mu}) = (K_1(\boldsymbol{\mu}), \ldots, K_d(\boldsymbol{\mu})),$$

and prove the following result:

**Proposition 5.3.** *For all large eigenvalue $\lambda$ of the replacement matrix $R$, for all $\boldsymbol{A} = (A_1, \ldots, A_d) \in \mathbb{C}^d$, denote by $\Psi(\boldsymbol{A})$ the vector whose coordinates are given by*

$$\Psi(\boldsymbol{A})_i = \frac{A_i}{\lambda + \theta_i} \quad \text{for all } i \in \{1, \ldots, d\},$$

*where $\lambda + \theta_i \neq 0$ since $\operatorname{Re}\lambda > S/2 > 0$ and $\theta_i \geqslant 1$.*

  (i)  *For all $\boldsymbol{A} = (A_1, \ldots, A_d) \in \mathbb{C}^d$ such that $\Psi(\boldsymbol{A}) \in \operatorname{Ker}(R - \lambda I_d)$, the function $K$ is an application from $\times_{i=1}^d \mathcal{M}_2^{\mathbb{C}}(A_i)$ into itself.*
  (ii) *Moreover, the law of $(W_{\boldsymbol{e_1}}^{DT}, \ldots, W_{\boldsymbol{e_d}}^{DT})$ is the unique square-integrable solution of (5.1) at fixed mean.*

*Remark* 5.4. If we assume $(\mathtt{T_{-1}})$ in addition, then, remark that $\Psi(\boldsymbol{A}) \in \operatorname{Ker}(R - \lambda I_d)$ if and only if $\boldsymbol{A} \in \operatorname{Ker}(R - \lambda I_d)$.

*Proof*: ($i$) First remark that

$$\mathbb{E}K_c(\boldsymbol{\mu}) := \int_{\mathbb{C}} x \, dK_c(\boldsymbol{\mu})(x) = \sum_{i=1}^{d} \sum_{p=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \mathbb{E}\left(V_p^{(c)}\right)^{\lambda/S} \mathbb{E}X_i^{(k)}$$

because $(V_1, \ldots, V_{S+1})$ is independent of $(X_i^{(1)}, \ldots, X_i^{(S+1)})_{1 \leqslant i \leqslant d}$. Since, for all $p \in \{1, \ldots, S+1\}$, for all $i \in \{1, \ldots, d\}$, $\mathbb{E}X_i^{(p)} = A_i$, we have

$$\mathbb{E}K_c(\boldsymbol{\mu}) = \sum_{i=1}^{d} A_i \sum_{p=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \mathbb{E}\left(V_p^{(c)}\right)^{\lambda/S}$$

$$= \sum_{i=1}^{d} A_i \frac{\gamma_i^{(c)} - \gamma_{i-1}^{(c)}}{1 + \lambda\theta_i^{-1}} = \sum_{i=1}^{d} A_i \frac{\tilde{a}_{c,i} + \delta_{c,i}}{1 + \lambda\theta_i^{-1}}$$

$$= \sum_{i=1}^{d} A_i \frac{a_{c,i} + \delta_{c,i}\theta_i}{\theta_i + \lambda} = \sum_{i=1}^{d} (a_{c,i} + \delta_{c,i}\theta_i)B_i,$$

where $\boldsymbol{B} = \Psi(\boldsymbol{A})$. The above calculations are true because $\left(V_p^{(c)}\right)^{\lambda/S} \stackrel{(law)}{=} U^{\lambda/\theta_c}$ for all $p \in \{1, \ldots, S+1\}$, where $U$ is a random variable uniformly distributed on $[0,1]$, and because $\mathrm{Re}\lambda > {}^{S}/_{2}$, which implies $\lambda \neq -\theta_c$ since $\theta_c \geqslant 1$ (see Equation (3.3)). Since $\lambda$ is an eigenvalue of $R$, and $\boldsymbol{B} = (B_1, \ldots, B_d) \in \mathrm{Ker}(R - \lambda I_d)$, we have

$$\sum_{i=1}^{d} a_{c,i}B_i = \lambda B_c$$

for all $1 \leqslant c \leqslant d$. It implies

$$\mathbb{E}K_c(\boldsymbol{\mu}) = (\lambda + \theta_c)B_c = A_c$$

for all $\boldsymbol{\mu} \in \bigtimes_{i=1}^{d} \mathcal{M}_2^{\mathbb{C}}(A_i)$. Moreover $K(\boldsymbol{\mu})$ is square-integrable, which implies that $K$ is indeed a function from $\bigtimes_{i=1}^{d} \mathcal{M}_2^{\mathbb{C}}(A_i)$ into itself, for all $\boldsymbol{A}$ such that $\Psi(\boldsymbol{A}) \in \mathrm{Ker}(R - \lambda I_d)$.

($ii$) Let $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_d) \in \bigtimes_{i=1}^{d} \mathcal{M}_2^{\mathbb{C}}(A_i)$ and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_d) \in \bigtimes_{i=1}^{d} \mathcal{M}_2^{\mathbb{C}}(A_i)$ be two solutions of System (5.1), meaning that

$$K\boldsymbol{\mu} = \boldsymbol{\mu} \quad \text{and} \quad K\boldsymbol{\nu} = \boldsymbol{\nu}.$$

Let us prove that $d(\boldsymbol{\mu}, \boldsymbol{\nu}) = 0$, using the total variance law: it is enough to prove that, for all $i \in \{1, \ldots, d\}$, $d_W(\mu_i, \nu_i) = 0$. Kantorovitch-Rubinstein's theorem implies that, for all $1 \leqslant i \leqslant d$, there exists two random variables $\mathcal{X}_i \sim \mu_i$ and $\mathcal{Y}_i \sim \nu_i$ such that $d_W(\mu_i, \nu_i) = \|\mathcal{X}_i - \mathcal{Y}_i\|_2$. Fix $1 \leqslant c \leqslant d$, and let $V$ be a Dirichlet-distributed random variable of parameter $\left(\pi_1^{(c)}, \ldots \pi_{\gamma_d^{(c)}}^{(c)}\right)$ (as in Equation (5.6)), $X = (X_1^{(c)}, \ldots, X_{\gamma_d^{(c)}}^{(c)})$ and $Y = (Y_1^{(c)}, \ldots, Y_{\gamma_d^{(c)}}^{(c)})$ be two sequences of independent random variables such that $X$ and $Y$ are both independent of $V$ (but not necessarily independent of each other), and, for all $1 \leqslant i \leqslant d$ and $1 \leqslant k \leqslant \gamma_d^{(c)}$,

$$\gamma_{i-1}^{(c)} + 1 \leqslant k \leqslant \gamma_i^{(c)} \Rightarrow \mathcal{L}(X_k^{(c)}, Y_k^{(c)}) = \mathcal{L}(\mathcal{X}_i, \mathcal{Y}_i) \text{ and } \pi_k^{(c)} = \theta_i/S.$$

We thus have, for all $c \in \{1, \dots, d\}$,

$$d_W(\mu_c, \nu_c)^2 = d_W(K_c(\boldsymbol{\mu}), K_c(\boldsymbol{\nu}))^2 \leqslant \left\| \sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \left(V_k^{(c)}\right)^{\lambda/S} (X_i^{(k)} - Y_i^{(k)}) \right\|_2^2$$

$$\leqslant \sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \mathbb{E}\left|\left(V_k^{(c)}\right)^{2\lambda/S}\right| \mathbb{E}|(X_i^{(k)} - Y_i^{(k)})|^2$$

$$= \sum_{i=1}^{d} \mathbb{E}|\mathcal{X}_i - \mathcal{Y}_i|^2 \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \mathbb{E}\left|\left(V_k^{(c)}\right)^{2\lambda/S}\right|,$$

implying that

$$d_W(K_c(\boldsymbol{\mu}), K_c(\boldsymbol{\nu}))^2 \leqslant \sum_{i=1}^{d} \frac{\gamma_i^{(c)} - \gamma_{i-1}^{(c)}}{2\mathrm{Re}\lambda\theta_i^{-1} + 1} d_W(\mu_i, \nu_i)^2 = \sum_{i=1}^{d} \frac{\tilde{a}_{c,i} + \delta_{c,i}}{2\mathrm{Re}\lambda\theta_i^{-1} + 1} d_W(\mu_i, \nu_i)^2.$$

If we let $\Delta_i = \frac{d_W(\mu_i, \nu_i)^2}{\theta_i + 2\mathrm{Re}\lambda}$, for all $i \in \{1, \dots, d\}$, we get that, for all $1 \leqslant c \leqslant d$,

$$2\mathrm{Re}\lambda\Delta_c \leqslant \sum_{i=1}^{d} a_{i,c}\Delta_i = (R\Delta)_c.$$

Let $\boldsymbol{v} = (v_1, \dots, v_d)$ be a horizontal vector with positive entries such that $\boldsymbol{v}R = S\boldsymbol{v}$, then

$$\boldsymbol{v}(R\Delta) = S\boldsymbol{v}\Delta.$$

The existence of such a $\boldsymbol{v}$ is a consequence of the irreducibility of the urn, namely hypothesis (S), as explained in Janson (2004). It implies that,

$$S\sum_{c=1}^{d} v_c\Delta_c = \sum_{c=1}^{d} v_c(R\Delta)_c \geqslant 2\mathrm{Re}\lambda \sum_{c=1}^{d} v_c\Delta_c.$$

Since $\mathrm{Re}\lambda/S > 1/2$, this last inequality implies that

$$\sum_{c=1}^{d} v_c\Delta_c = 0,$$

and thus, by positivity of the $v_c$ and non negativity of the $\Delta_c$, we get, for all $c \in \{1, \dots, d\}$, $\Delta_c = 0$. It thus implies that, for all $c \in \{1, \dots, d\}$,

$$\mu_c = \nu_c, \quad \text{and thus} \quad \boldsymbol{\mu} = \boldsymbol{\nu},$$

which concludes the proof.                                                                    $\square$

5.3. *Decomposition in discrete time.* The argument used to prove Proposition 5.1 can also be used to prove the following result from Proposition 3.2 and Theorem 2.9:

**Proposition 5.5.** *Under assumptions* (B), (T) *and* (S), *for all initial composition* $\boldsymbol{\alpha}$,

$$W_{\boldsymbol{\alpha}}^{DT} \overset{(law)}{=} \sum_{c=1}^{d} \sum_{p=\beta_{c-1}+1}^{\beta_c} Z_p^{\lambda/S} W_{\boldsymbol{e_c}}^{(p)},$$

*where $\beta_0 = 0$, $\beta_i = \sum_{j \leqslant i} \tilde{\alpha}_j$; where the $W_{e_c}^{(p)}$ are independent copies of $W_{e_c}^{DT}$, independent of each other; and where $Z = (Z_1, \ldots, Z_{\beta_d})$ is a Dirichlet-distributed random vector independent of the $W$ and of parameter $\boldsymbol{\eta}$ given by*

$$\boldsymbol{\eta}_k = \theta_i / S \quad \text{if } \beta_{i-1} < k \leqslant \beta_i.$$

*Remark* 5.6. If we assume $(\texttt{T}_{-1})$ instead of $(\texttt{T})$ in the theorem above, we obtain the following equation: for all initial composition $\boldsymbol{\alpha}$,

$$W_{\boldsymbol{\alpha}}^{DT} \stackrel{(law)}{=} \sum_{c=1}^{d} \sum_{p=\beta_{c-1}+1}^{\beta_c} Z_p^{\lambda/S} W_{e_c}^{(p)},$$

where $\beta_0 = 0$, $\beta_i = \sum_{j \leqslant i} \alpha_j$; where $Z = (Z_1, \ldots, Z_{\beta_d})$ is a Dirichlet-distributed random vector of parameter $\left(\frac{1}{S}, \ldots, \frac{1}{S}\right)$; and where the $W_{e_c}^{(p)}$ are independent copies of $W_{e_c}^{DT}$, independent of each other and of $Z$.

5.4. *Tree structure in discrete time.* Propositions 5.5 and 5.1 can be proven from scratch by an analogue of the proof of Propositions 3.2 and 3.3, using the underlying tree structure of the urn. The analysis of the tree structure in discrete-time is however more intricate since the subtrees of the considered forest are not independent. Such a tree decomposition in discrete-time is already proposed in Chauvin et al. (2015) (for two-colour urns) or Knape and Neininger (2014) but both papers assume (B), (I) and $(\texttt{T}_{-1})$. We will develop this alternative proof of Proposition 5.5 because it happens to be more complicated due to the possibly negative diagonal coefficient of the replacement matrix.

*Alternative proof of Proposition 5.5:* The discrete-time urn process can be seen as a forest whose leaves can be of $d$ different colours: we refer to Figure 5.3[2]. A time zero, the forest is composed of $\tilde{\alpha}_i$ roots of colour $i$ (for all $i \in \{1, \ldots, d\}$), each of these roots contains $\theta_i$ balls of colour $i$. At each step, we pick up uniformly at random a ball: this ball belongs to a leaf. The picked leaf then becomes an internal node, which has $\gamma_d^{(c)}$ children, amongst them $\tilde{a}_{i,j} + \delta_{i,j}$ contain $\theta_j$ balls of colour $j$ (for all $j \in \{1, \ldots, d\}$) if the picked up leaf was of colour $i$.

The composition of the urn is thus described by the set of leaves of the forest. Let us number the subtrees of the forest from trees rooted by colour 1 up to trees rooted by colour $d$. If we denote by $D_p(n)$ the number of balls in leaves of the $p^{\text{th}}$ subtree of the forest, then this $p^{\text{th}}$ subtree of the forest at time $n$ represents the composition vector of an urn process with inital composition of cardinal $\theta_c$ if $\beta_{c-1} < p \leqslant \beta_c$, taken at *internal time* $S_p(n) = \frac{D_p(n) - \theta_c}{S}$. Indeed, the *internal time* in the $p^{\text{th}}$ tree is the number of its internal nodes; the fact that the urn is balanced means that all internal nodes of the tree have given birth to $S + 1$ balls, which give the above relationship between leaves and internal nodes in the $p^{\text{th}}$ subtree.

We thus have

$$U_{\boldsymbol{\alpha}}(n) \stackrel{(law)}{=} \sum_{c=1}^{d} \sum_{p=\beta_{c-1}+1}^{\beta_c} U_{e_c}^{(p)} \left( \frac{D_p(n) - \omega_p}{S} \right), \tag{5.7}$$

---

[2] Figure 5.3 is an example of a two-colour urn. The example taken is chosen for it simplicity even if the chosen urn only has small Jordan blocks: it is called a *small* urn in the literature. It does not affect the arguments developed in the proof.

FIGURE 5.3. A realisation at time $n = 7$ of the forest associated to the urn process with initial composition ${}^t(2, 2)$ and with replacement matrix $R = \begin{pmatrix} -2 & 4 \\ 2 & 0 \end{pmatrix}$. Remark that $S_1(7) = 3$, $S_2(7) = 3$ and $S_3(7) = 1$. Moreover, $D_1(7) = 8$, $D_2(7) = 3$ and $D_3(7) = 1$.
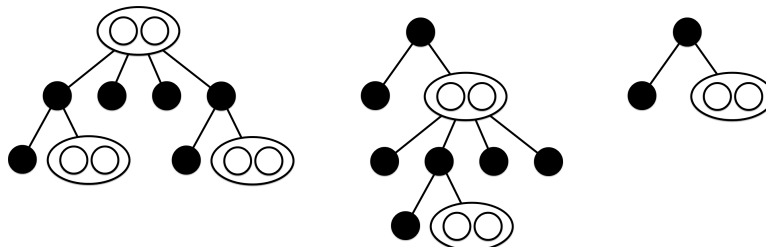


where $\beta_0 = 0$, where for all $c \geqslant 1$, $\beta_c = \sum_{i=1}^{c} \tilde{\alpha}_i$; where $\omega_p = \theta_i$ if $\beta_{i-1} < p \leqslant \beta_i$; and where the urn processes $U_{\boldsymbol{e_c}}^{(p)}$ are independent copies of the process $U_{\boldsymbol{e_c}}$, independent of each other. We are thus interested in the asymptotic behaviour of the vector $(D_1(n), \ldots, D_{\beta_d}(n))$ when $n$ grows to infinity. This vector happens to be the composition vector of a $\left(\beta_d = \sum_{i=1}^{d} \tilde{\alpha}_i\right)$–colour Pólya urn with initial composition $\boldsymbol{\omega} = {}^t(\omega_1, \ldots, \omega_{\beta_d})$, where

$$\omega_k = \theta_i \quad \text{if } \beta_{i-1} < k \leqslant \beta_i.$$

Indeed, forget the initial colouring of the leaves and colour the leaves of the $i$th subtree with colour $i$; at each step, a leaf of the forest picked up uniformly at random becomes an internal node and gives birth to $S + 1$ leaves of its same colour. Such a diagonal urn is called a Pólya-Eggenberger urn and has been long studied in the literature. We can for example cite this result by Athreya (1969) (for a complete proof, see Bertoin, 2006 or Chauvin et al., 2015):

**Theorem 5.7.** *Let $p$ and $K$ be two positive integers and $(D_1(n), \ldots, D_p(n))$ the composition vector at time $n$ of an urn process of initial composition $\boldsymbol{\nu} = {}^t(\nu_1, \ldots, \nu_p)$ and with replacement matrix $K I_p$, then, asymptotically when $n$ tends to infinity, almost surely,*

$$\frac{1}{nK} (D_1(n), \ldots, D_p(n)) \to \boldsymbol{Z} = (Z_1, \ldots, Z_p)$$

*where $\boldsymbol{Z}$ is a Dirichlet-distributed random vector of parameter $\left(\frac{\nu_1}{K}, \ldots, \frac{\nu_p}{K}\right)$.*

Thus, projecting Equation (5.7) onto $E$ via $\pi_E$, renormalising it by $n^{\lambda/s} \ln^{\nu} n$ and taking the limit when $n \to +\infty$ gives Proposition 5.5. $\qquad \square$

## 6. Densities – Proof of Theorem 2.10

In this section, we prove Theorem 2.10 via an analysis of the Fourier transforms of the random variables $W^{CT}$ and $W^{DT}$. We generalise the method developed by Liu (1998, 2001) for smoothing equations with positive solutions and refer to Chauvin et al. (2014) for a similar proof in the case of $m$-ary trees where a complex fixed point equation (but not a system) is studied. A similar result is also proved for

the two–colour case in Chauvin et al. (2015), but due to both the higher dimension and the weaker tenability condition, the present proof follows a different route.

The strategy of the proof is the following. First focus on the discrete time case, i.e. on $W^{DT}$:

- We prove that the support of $W^{DT}_{e_i}$ contains some non-discrete set.
- It implies that the Fourier transform of $W^{DT}$ is integrable and thus invertible (see Lemmas 6.4, 6.5 and 6.6), so that the variable $W^{DT}$ has a density on $\mathbb{C}$.
- We finally deduce from the existence of a density and from Lemma 6.2 that the support of $W^{DT}$ is the whole complex plane.

Via the martingale connection (2.3), one can consequently infer that, for all $c \in \{1, \ldots, d\}$ $W^{CT}_{e_c}$ has a density on $\mathbb{C}$. Propositions 5.5 and 3.2 permit to generalise to any initial composition.

We assume (B), (T) and (I) and recall that the eigenvalue $\lambda$, the Jordan subspace $E$, and its dimension $\nu + 1$ are fixed.

**Lemma 6.1.** *There exists $1 \leqslant c \leqslant d$ such that $\mathbb{E}W^{DT}_{e_c} \neq 0$.*

*Proof*: It is known (see for example Pouyanne, 2008) that, for all $c \in \{1, \ldots, d\}$,

$$\frac{1}{\nu!} W^{DT}_{e_c} \boldsymbol{v}$$

$$= \lim_{n \to +\infty} \left(1 + {}^t R_{|E}\right)^{-1} \left(1 + \frac{{}^t R_{|E}}{1 + S}\right)^{-1} \cdots \left(1 + \frac{{}^t R_{|E}}{1 + (n-1)S}\right)^{-1} \pi_E(U_{e_c}(n)),$$

as a martingale almost sure limit, where ${}^t R_{|E}$ stands for the restriction of the endomorphism induced by $R$ to the Jordan subspace $E$ (note that the unique eigenvalue of $R$ is thus $\lambda$ and since $\mathrm{Re}\lambda > S/2$, the inverses above are well defined). Thus, $\mathbb{E}W^{DT}_{e_c}\boldsymbol{v} = \nu!\pi_E(\boldsymbol{e_c})$. Note that, if $\mathbb{E}W^{DT}_{e_c}\boldsymbol{v} = \nu!\pi_E(\boldsymbol{e_c}) = 0$, for all $1 \leqslant c \leqslant d$, then, $\pi_E = 0$, which is impossible. $\qquad\square$

**Lemma 6.2.** *For all $\boldsymbol{z} = (z_1, \ldots, z_d) \in \bigtimes_{i=1}^{d} \mathrm{Supp}(W^{DT}_{e_i})$, for all $c \in \{1, \ldots, d\}$, for all $(v_1, \ldots, v_{\gamma_d^{(c)}})$ in the support of a Dirichlet- distributed random vector of parameter $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_{\gamma_d^{(c)}})$, where $\omega_i = {}^{\theta_i}/_S$ for all $i \in \{1, \ldots, c\}$, we have*

$$\sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} v_k^{\lambda/S} z_i \in \mathrm{Supp}(W^{DT}_{e_c}).$$

Note that the support of the Dirichlet distribution of parameter $\boldsymbol{\omega}$ is

$$\left\{(v_1, \ldots, v_{\gamma_d^{(c)}}) \in [0,1]^{\gamma_d^{(c)}} : \sum_{k=1}^{\gamma_d^{(c)}} v_k = 1\right\}.$$

*Proof*: Recall that for a given complex random variable $Z$, for all $z \in \mathbb{C}$,

$$z \in \mathrm{Supp}(Z) \Leftrightarrow \forall \varepsilon > 0, \mathbb{P}(|Z - z| < \varepsilon) > 0.$$

Let $\boldsymbol{z}$, $c$ and $(v_1, \ldots, v_{\gamma_d^{(c)}})$ satisfying the hypothesis of the lemma and let $\eta > 0$. We have

$$\left| W_{\boldsymbol{e_c}}^{DT} - \sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} v_k^{\lambda/s} z_i \right| \stackrel{(law)}{=} \left| \sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \left( V_k^{(c)} \right)^{\lambda/s} W_{\boldsymbol{e_i}}^{(k)} - \sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} v_k^{\lambda/s} z_i \right|$$

$$\stackrel{(law)}{=} \left| \sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \left( V_k^{(c)} \right)^{\lambda/s} (W_{\boldsymbol{e_i}}^{(k)} - z_i) - \sum_{i=1}^{d} \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \left( v_k^{\lambda/s} - \left( V_k^{(c)} \right)^{\lambda/s} \right) z_i \right|,$$

where $V = (V_1^{(c)}, \ldots, V_{\gamma_d^{(c)}}^{(c)})$ and the $W_{\boldsymbol{e_i}}^{(k)}$ are defined as in Theorem 5.1. Thus, since $(z \mapsto z^{\lambda/s})$ is continuous on $\mathbb{C}$, we have, with positive probability,

$$|W_{\boldsymbol{e_i}}^{(k)} - z_i| < \eta \quad \text{and} \quad \left| v_k^{\lambda/s} - \left( V_k^{(c)} \right)^{\lambda/s} \right| < \eta,$$

which imply

$$\left| W_{\boldsymbol{e_c}}^{DT} - \sum_{i=1}^{d} \sum_{p=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} v_p^{\lambda/s} z_i \right| \leqslant \eta \sum_{p=1}^{\gamma_d^{(c)}} \left| \left( V_p^{(c)} \right)^{\lambda/s} \right| + \eta \sum_{i=1}^{d} |z_i|$$

$$\leqslant (\gamma_d^{(c)} + \|z\|_1)\eta \leqslant (S + 1 + \|z\|_1)\eta,$$

where $\|z\|_1 = \sum_{i=1}^{d} |z_i|$. For all $\varepsilon > 0$, we fix $\eta = \frac{\varepsilon}{S+1+\|z\|_1}$ to conclude the proof. $\square$

**Lemma 6.3.** *There exists a non zero $z_0$ such that, for all $t \in (0,1)$,*

$$(t^{\lambda/s} + (1-t)^{\lambda/s})z_0 \in \bigcap_{i=1}^{d} \mathtt{Supp}(W_{\boldsymbol{e_i}}^{DT}).$$

*Proof*: Thanks to Lemma 6.1, there exists a colour $1 \leqslant c \leqslant d$ such that there exists $z_0 \neq 0 \in \mathtt{Supp}(W_{\boldsymbol{e_c}}^{DT})$. Applying Lemma 6.2, for all $i \in \{1, \ldots, d\}$ such that $a_{c,i} \neq 0$, $z_0 \in \mathtt{Supp}(W_{\boldsymbol{e_i}}^{DT})$. Iterating this argument allows us to conclude that, for all $i \in \{1, \ldots, d\}$ dominated by $c$, $z_0 \in \mathtt{Supp}(W_{\boldsymbol{e_i}}^{DT})$. By Assumption (I), all colours are dominating, and thus $c$ is dominating, meaning that

$$z_0 \in \bigcap_{i=1}^{d} \mathtt{Supp}(W_{\boldsymbol{e_i}}^{DT}).$$

Therefore, still by applying Lemma 6.2, for all $t \in [0,1]$,

$$(t^{\lambda/s} + (1-t)^{\lambda/s})z_0 \in \bigcap_{i=1}^{d} \mathtt{Supp}(W_{\boldsymbol{e_i}}^{DT}). \qquad \square$$

The three following lemmas are proven via very similar arguments as the ones developed in Chauvin et al. (2015). There is no additional idea to the proof here, except being careful to the slight changes induced by the higher $d \geqslant 3$ and by the weaker assumption (T) instead of (T$_{-1}$). For all $c \in \{1, \ldots, d\}$, let $\phi_c(t) = \mathbb{E}e^{i\langle t, W_{\boldsymbol{e_c}}^{DT} \rangle}$ for all $t \in \mathbb{C}$ and $\psi_c(r) = \sup_{|t|=r} |\phi_c(t)|$.

**Lemma 6.4.** *For all $c \in \{1, \ldots, d\}$, for all $r > 0$, $\psi_c(r) < 1$.*

*Proof*: We know that $\psi_c(0) = 1$ and that $\psi_c(r) \leqslant 1$ for all $r \geqslant 0$. Let us assume that there exists $r_c > 0$ such that $\psi_c(r_c) = 1$. Then, there exists $z_c \in \mathbb{C}$ and $\theta_c \in \mathbb{R}$ such that $|z_c| = r_c$ and

$$\mathbb{E}\mathbf{e}^{i\langle z_c, W_{e_c}^{DT}\rangle} = \mathbf{e}^{i\theta_c}.$$

Thus, the complex random variable $\mathbf{e}^{i\langle z_c, W_{e_c}^{DT}\rangle - i\theta_c}$ has mean 1 and takes its values in the unit disc. It is therefore almost surely equal to 1, implying that almost surely, $\langle z_c, W_{e_c}^{DT}\rangle \in \theta_c + 2\pi\mathbb{Z}$. Therefore, there exists $\varpi \in \mathbb{Z}$ such that, for all $t \in [0,1]$,

$$\langle z_c, (t^{\lambda/S} + (1-t)^{\lambda/S})z_0\rangle = \langle t^{\lambda/S} + (1-t)^{\lambda/S}, z_c\bar{z}_0\rangle = \theta_c + 2\pi\varpi.$$

Note that if $w \in \mathbb{C}\backslash\{0\}$ and $x \in \mathbb{R}$, then the set $\{z \in \mathbb{C}: \langle z, w\rangle = x\}$ is the line $x/w + iw\mathbb{R}$. Therefore, for all $t \in [0,1]$,

$$t^{\lambda/S} + (1-t)^{\lambda/S} \in \frac{\theta_c + 2\varpi\pi}{\bar{z}_0 z_c} + i\bar{z}_0 z_c\mathbb{R}. \tag{6.1}$$

If $\lambda \in \mathbb{R}$, then $t^{\lambda/S} + (1-t)^{\lambda/S} \in \mathbb{R}$ for all $t \in [0,1]$ which implies that both $(\theta_c + 2\varpi\pi)/(\bar{z}_0 z_c)$ and $i\bar{z}_0 z_c$ are real, which is impossible. If $\lambda \in \mathbb{C}\backslash\mathbb{R}$, then, the left-hand side of (6.1) is a spiral and thus contains at least three non-aligned points whereas the right-hand side is a line, which is impossible, and thus concludes the proof. $\square$

**Lemma 6.5.** *For all $c \in \{1, \ldots, d\}$, $\lim_{r\to\infty} \psi_c(r) = 0$.*

*Proof*: In view of Equation (5.1), for all $c \in \{1, \ldots, d\}$,

$$\phi_c(t) = \mathbb{E}\exp\left(i\Big\langle t, \sum_{i=1}^d \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \left(V_k^{(c)}\right)^{\lambda/S} W_{e_i}^{(k)}\Big\rangle\right)$$

$$= \mathbb{E}\left[\mathbb{E}\left[\exp\left(i\Big\langle t, \sum_{i=1}^d \sum_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \left(V_k^{(c)}\right)^{\lambda/S} W_{e_i}^{(k)}\Big\rangle\right)\Big| V_1, \ldots, V_{\gamma_d^{(c)}+1}\right]\right]$$

$$= \mathbb{E}\prod_{i=1}^d \prod_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \mathbb{E}\left[\mathbf{e}^{i\langle t, (V_k^{(c)})^{\lambda/S} W_{e_i}^{(k)}\rangle}\Big| V_1, \ldots, V_{\gamma_d^{(c)}+1}\right],$$

implying that

$$\psi_c(r) \leqslant \mathbb{E}\prod_{i=1}^d \prod_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \psi_i\left(\left(V_k^{(c)}\right)^{\mathrm{Re}\lambda/S} r\right).$$

By Fatou's Lemma, since $\mathbb{P}(V_k^{(c)} = 0) = 0$ for all $1 \leqslant k \leqslant \gamma_d^{(c)}$, it implies

$$\limsup_{r\to+\infty} \psi_c(r) \leqslant \mathbb{E}\prod_{i=1}^d \left(\limsup_{r\to+\infty} \psi_i(|V_i^{\lambda/S}|r)\right)^{\tilde{a}_{c,i}+\delta_{c,i}}.$$

Let us assume that there exists a colour $c$ such that $\limsup_{r\to\infty} \psi_c(r) = 1$. Let us define $\psi(r) = \max_{1\leqslant i\leqslant d} \psi_i(r)$, then $\limsup_{r\to\infty} \psi(r) = 1$. According to Lemma 6.4, for all $i \in \{1, \ldots, d\}$, $\psi_i(1) < 1$, and thus, $\psi(1) < 1$. Let $\varepsilon \in (0, 1-\psi(1))$, and define:

$$r_1(\varepsilon) = \max\{r \in (0,1), \psi(r) = 1 - \varepsilon\}$$
$$r_2(\varepsilon) = \min\{r > 1, \psi(r) = 1 - \varepsilon\}.$$

These definitions are legal since $\psi$ is continuous, $\psi(0) = 1$ and $\limsup_{r\to\infty} \psi(r) = 1$. Moreover, we have $\psi(r_1(\varepsilon)) = \psi(r_2(\varepsilon)) = 1 - \varepsilon$, and for all $r \in [r_1(\varepsilon), r_2(\varepsilon)]$, $\psi(r) \leqslant 1 - \varepsilon$.

We know that for all $c \in \{1, \ldots, d\}$

$$\psi_c(r) \leqslant \mathbb{E} \prod_{i=1}^{d} \prod_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \psi_i\left(\left(V_k^{(c)}\right)^{\mathrm{Re}\lambda/S} r\right).$$

In particular, for all $r \geqslant 0$,

$$\psi_c(r) \leqslant \mathbb{E}\psi_c\left(|V_c^{\lambda/S}|r\right),$$

where $V_c$ is Beta-distributed with parameter $(\theta_c/S, 1)$. For all $1 \leqslant i \leqslant d$, let $(A_k^{(i)})_{k\geqslant 1}$ be a sequence of i.i.d. random variables having the same law as $|V_i^{\lambda/S}|$. Iterating the last identity, we get

$$\psi_c(r) \leqslant \mathbb{E}\psi_c(rA_1^{(c)} \ldots A_n^{(c)}).$$

Let us define, for all $i \in \{1, \ldots, d\}$, for all integers $n$

$$\lambda_n^{(i)}(r, \varepsilon) = \mathbb{P}(r_1(\varepsilon) \leqslant rA_1^{(i)} \ldots A_n^{(i)} \leqslant r_2(\varepsilon)).$$

For all $r \geqslant 0$, for all $i \in \{1, \ldots, d\}$,

$$\psi_i(r) \leqslant \mathbb{E}\psi_c(rA_1^{(i)} \ldots A_n^{(i)}) \leqslant (1-\varepsilon)\lambda_n^{(i)}(r, \varepsilon) + 1 - \lambda_n^{(i)}(r, \varepsilon) \leqslant 1 - \varepsilon\lambda_n^{(i)}(r, \varepsilon).$$

For all $c \in \{1, \ldots, d\}$, we have,

$$\psi_c(r_2(\varepsilon)) \leqslant \mathbb{E} \prod_{i=1}^{d} \prod_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \psi_i\left(\left(V_k^{(c)}\right)^{\mathrm{Re}\lambda/S} r_2(\varepsilon)\right)$$

$$\leqslant \mathbb{E} \prod_{i=1}^{d} \prod_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \left(1 - \varepsilon\lambda_n^{(i)}\left(\left(V_k^{(c)}\right)^{\mathrm{Re}\lambda/S} r_2(\varepsilon), \varepsilon\right)\right).$$

In view of Lemma 6.4, $r_1(\varepsilon) \to 0$ when $\varepsilon \to 0$, implying that, for all $i \in \{1, \ldots, d\}$, when $\varepsilon$ goes to zero,

$$\lambda_n^{(i)}\left(\left(V_k^{(c)}\right)^{\mathrm{Re}\lambda/S} r_2(\varepsilon), \varepsilon\right) \to \mathbb{P}(0 \leqslant \left(V_k^{(c)}\right)^{\mathrm{Re}\lambda/S} A_1^{(i)} \ldots A_n^{(i)} \leqslant 1) = 1.$$

Thus, for all $c \in \{1, \ldots, d\}$,

$$\frac{1 - \mathbb{E} \prod_{i=1}^{d} \prod_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \left(1 - \varepsilon\lambda_n^{(i)}\left(\left(V_k^{(c)}\right)^{\mathrm{Re}\lambda/S} r_2(\varepsilon), \varepsilon\right)\right)}{\varepsilon} \to \sum_{i=1}^{d}(\tilde{a}_{c,i} + \delta_{c,i}) = \gamma_c^{(d)}.$$

Now note that

$$1 - \varepsilon = \psi(r_2(\varepsilon)) = \max_{c\in\{1,\ldots,d\}} \psi_c(r_2(\varepsilon))$$

$$\leqslant \max_{c\in\{1,\ldots,d\}} \mathbb{E} \prod_{i=1}^{d} \prod_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \left(1 - \varepsilon\lambda_n^{(i)}\left(\left(V_k^{(c)}\right)^{\mathrm{Re}\lambda/S} r_2(\varepsilon), \varepsilon\right)\right),$$

which implies that $\min_{c \in \{1,\dots,d\}} \gamma_d^{(c)} \leqslant 1$. Recall that, in view of Equation (3.3),

$$\gamma_d^{(c)} = 1 + \sum_{i=1}^{d} \tilde{a}_{c,i} = 1 + \sum_{i=1}^{d} \frac{a_{c,i}}{\theta_i} > 1,$$

since at least one of the $a_{c,i}$'s is non-zero as their sum is equal to $S \geqslant 1$. We have thus reached a contradiction, implying that for all $1 \leqslant c \leqslant d$,

$$\limsup_{r \to \infty} \psi_c(r) = 1. \qquad \square$$

**Lemma 6.6.** *For all $c \in \{1,\dots,d\}$, for all $\rho \in (0, \theta_c/\mathrm{Re}\lambda)$, asymptotically when $|t|$ tends to $+\infty$, $\phi_c(t) = \mathcal{O}(|t|^{-\rho})$.*

*Proof*: Let $\varepsilon > 0$. In view of Lemma 6.5, there exists $T > 0$ such that, for all $r \geqslant T$, for all $i \in \{1,\dots,d\}$, $\psi_i(r) \leqslant \varepsilon$. We already have proved that

$$\psi_c(r) \leqslant \mathbb{E} \prod_{i=1}^{d} \prod_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \psi_i\left(\left(V_k^{(c)}\right)^{\mathrm{Re}\lambda/S} r\right).$$

Thus, for all $r \geqslant T$,

$$\psi_c(r) \leqslant \varepsilon \mathbb{E} \psi_c\left(\left(V_{j(c)}^{(c)}\right)^{\mathrm{Re}\lambda/S} r\right) + \mathbb{P}\left(\left|\left(V_{\gamma_c^{(c)}}^{(c)}\right)^{\lambda/S} r\right| < T\right),$$

where $j(c)$ is any integer from $\{1,\dots,d\}\setminus\{\gamma_c^{(c)}\}$. Recall that $V_{\gamma_c^{(c)}}^{(c)} \overset{(law)}{=} U^{S/\theta_c}$ and $V_{j(c)}^{(c)} \overset{(law)}{=} U^{S/\theta_{j(c)}}$ where $U$ is a uniform random variable on $(0,1)$. Then, for all $r > T$,

$$\psi_c(r) \leqslant \varepsilon \mathbb{E} \psi_c\left(U^{\mathrm{Re}\lambda/\theta_{j(c)}} r\right) + \left(\frac{T}{r}\right)^{\theta_c/\mathrm{Re}\lambda}.$$

Said differently, there exists a positive constant $C$ such that, for all $r > T$,

$$\psi_c(r) \leqslant \varepsilon \mathbb{E} \psi_c\left(U^{\mathrm{Re}\lambda/\theta_{j(c)}} r\right) + C\left(\frac{1}{r}\right)^{\rho},$$

for all $\rho \in (0, \theta_c/\mathrm{Re}\lambda)$. For such $\rho$ (and actually for any $\rho$), $\mathbb{E} U^{-\rho \mathrm{Re}\lambda/\theta_{j(c)}} < +\infty$, and we can thus apply a Gronwall-type lemma (Liu, 1999, Lemma 4.1), which implies

$$\psi_c(r) \leqslant \frac{Cr^{-\rho}}{1 - \varepsilon \mathbb{E} U^{-\rho \mathrm{Re}\lambda \theta_{j(c)}}},$$

provided that $\varepsilon$ satisfies $1 - \varepsilon \mathbb{E} U^{-\rho \mathrm{Re}\lambda/\theta_{j(c)}} > 0$. $\qquad \square$

**Proposition 6.7.** *If $\lambda \in \mathbb{C}\setminus\mathbb{R}$, the distribution of $W_{e_c}^{DT}$ admits a density on $\mathbb{C}$ and its support is $\mathbb{C}$, for all $c \in \{1,\dots,d\}$.*

*If $\lambda \in \mathbb{R}$, the distribution of $W_{e_c}^{DT}$ admits a density on $\mathbb{R}$ and its support is $\mathbb{R}$, for all $c \in \{1,\dots,d\}$.*

*Proof*: Let us apply arguments already used in Chauvin et al. (2015, page 22) to prove that $\phi_c$ is integrable for all $c \in \{1,\dots,d\}$. Recall that, for all $r \geqslant 0$,

$$\psi_c(r) \leqslant \mathbb{E} \prod_{i=1}^{d} \prod_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \psi_i\left(\left(V_k^{(c)}\right)^{\mathrm{Re}\lambda/S} r\right).$$

In view of Lemma 6.6, there exists a constant $\kappa > 0$ such that, for all $i \in \{1, \ldots, d\}$, for all $\rho_i \in (0, \theta_i/\mathrm{Re}\lambda)$, for all $|t|$ large enough, $|\phi_i(t)| \leqslant \kappa|t|^{-\rho_i}$. Let $\eta_c = \sum_{i=1}^{d}(\tilde{a}_{c,i} + \delta_{c,i})\rho_i$, for all $r$ large enough, we have

$$\psi_c(r) \leqslant \frac{\kappa^{\gamma_d^{(c)}}}{r^{\eta_c}} \mathbb{E} \prod_{i=1}^{d} \prod_{k=\gamma_{i-1}^{(c)}+1}^{\gamma_i^{(c)}} \left(V_k^{(c)}\right)^{-\rho_i \mathrm{Re}\lambda/S} = \mathcal{O}\left(r^{-\eta_c}\right).$$

Note that, for all $1 \leqslant c \leqslant d$, in view of Equation (3.3),

$$\sum_{i=1}^{d}(\tilde{a}_{c,i} + \delta_{c,i})\frac{\theta_i}{S} = \sum_{i=1}^{d} \frac{a_{c,i} + \theta_i \delta_{c,i}}{S} = 1 + \frac{\theta_c}{S} > 1,$$

implying that we can choose $\rho_1, \ldots, \rho_d$ such that $\rho_i < \theta_i/S$ and $\eta_c > 1$. Therefore, $\phi_c$ is integrable and thus that $W_{\boldsymbol{e_c}}^{DT}$ admits a bounded and continuous density on $\mathbb{C}$, for all $c \in \{1, \ldots, d\}$.

Fix a dominating colour $c$ (actually, by Assumption (I), all colours are dominating, but for this argument, (S) would be enough). Since $W_{\boldsymbol{e_c}}^{DT}$ admits a continuous density, we know that there exists $z_0 \in \mathbb{C}$ and $\varepsilon > 0$ such that the opened ball centred in $z_0$ of radius $\varepsilon$, which we denote by $\mathcal{B}(z_0, \varepsilon)$, is contained in $\mathrm{Supp}(W_{\boldsymbol{e_c}}^{DT})$. In view of Lemma 6.2, for all colours $i$ dominated by $c$, and thus for all colours $i$ since $c$ is dominating,

$$\mathrm{Supp}(W_{\boldsymbol{e_c}}^{DT}) \subseteq \mathrm{Supp}(W_{\boldsymbol{e_i}}^{DT}),$$

implying that $\mathcal{B}(z_0, \varepsilon) \subseteq \bigcap_{i \in \{1,\ldots,d\}} \mathrm{Supp}(W_{\boldsymbol{e_i}}^{DT})$. By Lemma 6.3, we get that, for all $t \in [0, 1]$,

$$\left(t^{\lambda/S} + (1-t)^{\lambda/S}\right)\mathcal{B}(z_0, \varepsilon) \subseteq \bigcap_{i \in \{1,\ldots,d\}} \mathrm{Supp}(W_{\boldsymbol{e_i}}^{DT}). \tag{6.2}$$

Let us first assume that $\lambda \in \mathbb{C}\backslash\mathbb{R}$. Note that when $t$ is tending to zero, $t^{\lambda/S} + (1-t)^{\lambda/S} = 1 + t^{\lambda/S} + \mathcal{O}(t)$, where, for all complex functions $\eta(t)$, we denote $\eta(t) = \mathcal{O}(t)$, if $|\eta(t)| = \mathcal{O}(t)$. The function $t \mapsto t^{\lambda/S}$ maps any neighbourhood of 0 to an infinite spiral around 1 in the complex plane: the module of $t^{\lambda/S}$ goes to zero when $t$ goes to zero, while its argument, which is equal to $\mathrm{Im}\lambda \log t/S$ goes to $-\infty$. In particular, the trajectory of $t \mapsto 1 + t^{\lambda/S}$ intersects infinitely many times the set $\mathcal{B}(z_0, \varepsilon/2) \cap \{|z| = 1\}\backslash\{1\}$. Therefore, since $t^{\lambda/S} + (1-t)^{\lambda/S} = 1 + t^{\lambda/S} + \mathcal{O}(t)$, when $t$ goes to zero, we get that there exists $t_0 > 0$ such that

$$|t_0^{\lambda/S} + (1-t_0)^{\lambda/S}| = 1, \mathrm{arg}(t_0^{\lambda/S} + (1-t_0)^{\lambda/S}) \notin 2\pi\mathbb{Z} \text{ and } |t_0^{\lambda/S} + (1-t_0)^{\lambda/S} - 1| < \varepsilon/2.$$

Therefore, the application

$$z \mapsto \left(t_0^{\lambda/S} + (1-t_0)^{\lambda/S}\right)z$$

is a rotation of an angle sufficiently small so that $\mathcal{B}(z_0, \varepsilon)$ intersects its image. We also know, by Lemma 6.3 that $\bigcap_{i \in \{1,\ldots,d\}} \mathrm{Supp}(W_{\boldsymbol{e_i}}^{DT})$ is stable by this rotation. Therefore, if we iterate this application, we get that there exists $\tilde{\varepsilon} > 0$ $\{z \in \mathbb{C}: |z_0| - \tilde{\varepsilon} < |z| < |z_0| + \tilde{\varepsilon}\}$ is contained in $\bigcap_{i \in \{1,\ldots,d\}} \mathrm{Supp}(W_{\boldsymbol{e_i}}^{DT})$ (one can for example take $\tilde{\varepsilon} = \varepsilon/2$). In particular, $-z_0 \in \bigcap_{i \in \{1,\ldots,d\}} \mathrm{Supp}(W_{\boldsymbol{e_i}}^{DT})$. Using Lemma 6.2, we get that

$$\left(\frac{1}{2}\right)^{\lambda/S} (\mathcal{B}(z_0, \varepsilon) - z_0) = \mathcal{B}\left(0, (1/2)^{\mathrm{Re}\lambda/S} \varepsilon\right) \subseteq \bigcap_{i \in \{1,\ldots,d\}} \mathrm{Supp}(W_{\boldsymbol{e_i}}^{DT}).$$

Finally, recall that, in view of Lemma 6.3, $\bigcap_{i\in\{1,\dots,d\}}\mathrm{Supp}(W_{\boldsymbol{e_i}}^{DT})$ is stable by

$$z \mapsto (1/2)^{\lambda/S-1}z,$$

which sends any ball $\mathcal{B}(0,r)$ to the larger ball $\mathcal{B}\big(0,(1/2)^{\mathrm{Re}\lambda/S-1}r\big)$ (recall that $\mathrm{Re}\lambda < S$). Iterating this application thus sends $\mathcal{B}\big(0,(1/2)^{\mathrm{Re}\lambda/S}\varepsilon\big)$ to $\mathbb{C}$, implying that $\bigcap_{i\in\{1,\dots,d\}}\mathrm{Supp}(W_{\boldsymbol{e_i}}^{DT}) = \mathbb{C}$.

It remains to treat the case $\lambda \in \mathbb{R}$. From Proposition 5.1, we have, for all $c \in \{1,\dots,d\}$,

$$\sum_{i=1}^{d}(a_{c,i}+\delta_{c,i})\frac{\mathbb{E}W_{\boldsymbol{e_i}}^{DT}}{\theta_i+\lambda} = \lambda\frac{\mathbb{E}W_{\boldsymbol{e_c}}^{DT}}{\theta_c+\lambda}.$$

First assume that, for all $1 \leqslant i \leqslant d$, $\mathbb{E}W_{\boldsymbol{e_i}}^{DT} \geqslant 0$. Then, if we take $c$ such that

$$\frac{\mathbb{E}W_{\boldsymbol{e_c}}^{DT}}{\theta_c+\lambda} = \min_{1\leqslant i\leqslant d}\frac{\mathbb{E}W_{\boldsymbol{e_i}}^{DT}}{\theta_i+\lambda},$$

we get (see Equation (3.3))

$$(S+1)\frac{\mathbb{E}W_{\boldsymbol{e_c}}^{DT}}{\theta_c+\lambda} \leqslant \sum_{i=1}^{d}(a_{c,i}+\delta_{c,i})\frac{\mathbb{E}W_{\boldsymbol{e_i}}^{DT}}{\theta_i+\lambda} = \lambda\frac{\mathbb{E}W_{\boldsymbol{e_c}}^{DT}}{\theta_c+\lambda},$$

implying that $\mathbb{E}W_{\boldsymbol{e_c}}^{DT} = 0$ since $\lambda < S$ by assumption. The same reasoning holds if we assume that $\mathbb{E}W_{\boldsymbol{e_i}}^{DT} \leqslant 0$ for all $1 \leqslant i \leqslant d$. Therefore, either there exists $c$ such that $\mathbb{E}W_{\boldsymbol{e_c}}^{DT} = 0$, or there exist $i$ and $j$ such that $\mathbb{E}W_{\boldsymbol{e_i}}^{DT} < 0 < \mathbb{E}W_{\boldsymbol{e_j}}^{DT}$.

First assume that exists $c$ such that $\mathbb{E}W_{\boldsymbol{e_c}}^{DT} = 0$, and recall that, in view of Lemma 6.1, there exists a colour $i$ such that $\mathbb{E}W_{\boldsymbol{e_i}}^{DT} \neq 0$. Therefore, there exists $z_i \neq 0$ in the support of $W_{\boldsymbol{e_i}}^{DT}$. Since we have assumed that the urn is irreducible (assumption (I)), and thus, in particular that $i$ is a dominating colour, we get, by Lemma 6.2, that $z_i \in \bigcap_{j=1}^{d}\mathrm{Supp}W_{\boldsymbol{e_j}}^{DT}$. Similarly, the support of $W_{\boldsymbol{e_c}}^{DT}$, and thus $\bigcap_{j=1}^{d}\mathrm{Supp}W_{\boldsymbol{e_j}}^{DT}$, contain either 0 or some $z_c$ chosen such that $z_i$ and $z_c$ have opposite signs. In the first case, since $W_{\boldsymbol{e_j}}^{DT}$ admits a continuous density for all $1 \leqslant j \leqslant d$, there exists $\varepsilon > 0$ such that $[-\varepsilon,\varepsilon] \in \bigcap_{j=1}^{d}\mathrm{Supp}W_{\boldsymbol{e_j}}^{DT}$. In latter case, applying Lemma 6.2 implies that the segment $[z_i,z_c]$ (or $[z_c,z_i]$, depending on the signs) in contained in $\bigcap_{j=1}^{d}\mathrm{Supp}W_{\boldsymbol{e_j}}^{DT}$. This segment contains 0, and thus, in both cases, there exists $\varepsilon > 0$ such that

$$[-\varepsilon,\varepsilon] \subseteq \bigcap_{j=1}^{d}\mathrm{Supp}W_{\boldsymbol{e_j}}^{DT}.$$

Finally, by Lemma 6.2, $\bigcap_{j=1}^{d}\mathrm{Supp}W_{\boldsymbol{e_j}}^{DT}$ is stable by the application $z \mapsto (1/2)^{\lambda/S-1}z$, and iterating this application sends $[-\varepsilon,\varepsilon]$ to $\mathbb{R}$, which gives

$$\bigcap_{j=1}^{d}\mathrm{Supp}W_{\boldsymbol{e_j}}^{DT} = \mathbb{R},$$

as desired.

It now remains to treat the case when there exist $i$ and $j$ such that $\mathbb{E}W_{\boldsymbol{e_i}}^{DT} < 0 < \mathbb{E}W_{\boldsymbol{e_j}}^{DT}$. This is very similar to the previous case: there exists $z_i < 0$ in $\mathrm{Supp}W_{\boldsymbol{e_i}}^{DT}$ and $z_j > 0$ in $\mathrm{Supp}W_{\boldsymbol{e_j}}^{DT}$. By irreducibility, and applying Lemma 6.2,

$[z_i, z_j] \subseteq \bigcap_{j=1}^{d} \mathtt{Supp}W_{e_j}^{DT}$. This segment contains 0 in its interior and is therefore sent to $\mathbb{R}$ by iterating the application $z \mapsto (1/2)^{\lambda/S-1}z$, under which $\bigcap_{j=1}^{d} \mathtt{Supp}W_{e_j}^{DT}$ is stable. $\qquad\square$

*Proof of Theorem 2.10:* Connection (2.3) allows to transfer results stated in Proposition 6.7 to $W_{e_c}^{CT}$ for all $c \in \{1, \ldots, d\}$, since $\xi$ and $W_{e_c}^{DT}$ both admit a density and are independent. Finally, Propositions 5.5 and 3.2 allow to generalise to any initial composition $\boldsymbol{\alpha}$. $\qquad\square$

*Remark* 6.8. For two-colour Pólya urns, it is proven under (B), (I) and (T$_{-1}$) by Chauvin et al. (2011) that the density of $W_{\boldsymbol{\alpha}}^{CT}$ explodes in zero for all initial composition $\boldsymbol{\alpha}$. Is it possible to prove the same result for $d$-colour urns? We believe that the proof is similar to the one developed in Chauvin et al. (2011) if $\lambda$ is real, and under (T) instead of (T$_{-1}$), but what can be said about $W^{CT}$ in zero when $\lambda$ is complex?

## 7. Conclusion and open problems

This paper contains the first results concerning the $W$ random variables exhibited by the asymptotic behaviour of a multi–colour Pólya urn process projected along large Jordan blocks: we proved that $W$ is moments-determined and admits a density both in discrete and in continuous time.

However, many questions still remain about the random variables $W$: some of them are solved in the two–colour case (by methods that do not seem easy to generalise to multi–colour urns) but other are not proved even in the two–colour case. Can we find a lower bound for the moments of the $W$'s (Chauvin et al., 2011 for two–colour urns)? What is the exact order of their moments? Is the Laplace transform of $W^{CT}$ convergent (this is false in the two–colour case Chauvin et al. (2011))? Is the Fourier transform of $W^{CT}$ integrable (also false in the two–colour case Chauvin et al., 2011)? Is the density of $W^{CT}$ continuous (also false in two–colour Chauvin et al., 2011)? Do the $W$'s have heavy tails?

## Acknowledgements

## References

G. M. Adel'son-Vel'skiĭ and E. M. Landis. An algorithm for organization of information. *Dokl. Akad. Nauk SSSR* **146**, 263–266 (1962). MR0156719.

K. B. Athreya. On a characteristic property of Polya's urn. *Studia Sci. Math. Hungar.* **4**, 31–35 (1969). MR0247643.

K. B. Athreya and S. Karlin. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.* **39**, 1801–1817 (1968). MR0232455.

K. B. Athreya and P. E. Ney. *Branching processes.* Springer-Verlag, New York-Heidelberg (1972). MR0373040.

A. Bagchi and A. K. Pal. Asymptotic normality in the generalized Pólya-Eggenberger urn model, with an application to computer data structures. *SIAM J. Algebraic Discrete Methods* **6** (3), 394–405 (1985). MR791169.

J. Bertoin. *Random fragmentation and coagulation processes*, volume 102 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge (2006). ISBN 978-0-521-86728-3; 0-521-86728-2. MR2253162.

D. Blackwell and D. Kendall. The martin boundary of Pólya's urn scheme, and an application to stochastic population growth. *J. Appl. Probability* **1**, 284–296 (1964). MR0176518.

A. Bose, A. Dasgupta and K. Maulik. Strong laws for balanced triangular urns. *J. Appl. Probab.* **46** (2), 571–584 (2009). MR2535833.

B Chauvin, D. Gardy, N. Pouyanne and D.-H. Ton-That. B-urns. *ALEA Lat. Am. J. Probab. Math. Stat.* **13** (2), 605–634 (2016). MR3522821.

B. Chauvin, Q. Liu and N. Pouyanne. Limit distributions for multitype branching processes of $m$-ary search trees. *Ann. Inst. Henri Poincaré Probab. Stat.* **50** (2), 628–654 (2014). MR3189087.

B. Chauvin, C. Mailler and N. Pouyanne. Smoothing equations for large Pólya urns. *J. Theoret. Probab.* **28** (3), 923–957 (2015). MR3413961.

B. Chauvin and N. Pouyanne. $m$-ary search trees when $m \geqslant 27$: a strong asymptotics for the space requirements. *Random Structures Algorithms* **24** (2), 133–154 (2004). MR2035872.

B. Chauvin, N. Pouyanne and R. Sahnoun. Limit distributions for large Pólya urns. *Ann. Appl. Probab.* **21** (1), 1–32 (2011). MR2759195.

H.-H. Chern and H.-K. Hwang. Phase changes in random $m$-ary search trees and generalized quicksort. *Random Structures Algorithms* **19** (3-4), 316–358 (2001). MR1871558.

R. M. Dudley. *Real analysis and probability*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge (2002). ISBN 0-521-00754-2. MR1932358.

F. Eggenberger and G. Pólya. Über die statistik verketetter vorgäge. *Zeitschrift für Angewandte Mathematik und Mechanik* **1**, 279–289 (1923).

P. Flajolet, P. Dumas and V. Puyhaubert. Some exactly solvable models of urn process theory. In *Fourth Colloquium on Mathematics and Computer Science Algorithms, Trees, Combinatorics and Probabilities*, Discrete Math. Theor. Comput. Sci. Proc., AG, pages 59–118. Assoc. Discrete Math. Theor. Comput. Sci., Nancy (2006). MR2509623.

P. Flajolet, J. Gabarró and H. Pekari. Analytic urns. *Ann. Probab.* **33** (3), 1200–1233 (2005). MR2135318.

S. Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Process. Appl.* **110** (2), 177–245 (2004). MR2040966.

S. Janson. Limit theorems for triangular urn schemes. *Probab. Theory Related Fields* **134** (3), 417–452 (2006). MR2226887.

S. Janson and S. Kaijser. Higher moments of Banach space valued random variables. *Mem. Amer. Math. Soc.* **238** (1127), vii+110 (2015). MR3402381.

N. L. Johnson and S. Kotz. *Urn models and their application.* John Wiley & Sons, New York-London-Sydney (1977). MR0488211.

M. Knape and R. Neininger. Pólya urns via the contraction method. *Combin. Probab. Comput.* **23** (6), 1148–1186 (2014). MR3265841.

D. Lamberton, G. Pagès and P Tarrès. When can the two-armed bandit algorithm be trusted? *Ann. Appl. Probab.* **14** (3), 1424–1454 (2004). MR2071429.

K. Leckey. On densities for solutions to stochastic fixed point equations. *ArXiv Mathematics e-prints* (2016). arXiv: 1604.05787.

Q. Liu. Fixed points of a generalized smoothing transformation and applications to the branching random walk. *Adv. in Appl. Probab.* **30** (1), 85–112 (1998). MR1618888.

Q. Liu. Asymptotic properties of supercritical age-dependent branching processes and homogeneous branching random walks. *Stochastic Process. Appl.* **82** (1), 61–87 (1999). MR1695070.

Q. Liu. Asymptotic properties and absolute continuity of laws stable by random weighted mean. *Stochastic Process. Appl.* **95** (1), 83–107 (2001). MR1847093.

H. M. Mahmoud. On rotations in fringe-balanced binary trees. *Inform. Process. Lett.* **65** (1), 41–46 (1998). MR1606251.

H. M. Mahmoud. The size of random bucket trees via urn models. *Acta Inform.* **38** (11-12), 813–838 (2002). MR1943199.

R. Neininger and L. Rüschendorf. A survey of multivariate aspects of the contraction method. *Discrete Math. Theor. Comput. Sci.* **8** (1), 31–56 (2006). MR2247515.

R. Pemantle. A survey of random processes with reinforcement. *Probab. Surv.* **4**, 1–79 (2007). MR2282181.

N. Pouyanne. An algebraic approach to Pólya processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** (2), 293–323 (2008). MR2446325.