



**HAL**  
open science

# A fixed-point equation approach for the superdiffusive elephant random walk

Hélène Guérin, Lucile Laulin, Kilian Raschel

► **To cite this version:**

Hélène Guérin, Lucile Laulin, Kilian Raschel. A fixed-point equation approach for the superdiffusive elephant random walk. 2024. hal-04190858v2

**HAL Id: hal-04190858**

**<https://hal.science/hal-04190858v2>**

Preprint submitted on 18 Apr 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A fixed-point equation approach for the superdiffusive elephant random walk

H el ene Gu erin\*, Lucile Laulin† and Kilian Raschel‡

April 18, 2024

## Abstract

We study the elephant random walk in arbitrary dimension  $d \geq 1$ . Our main focus is the limiting random variable appearing in the superdiffusive regime. Building on a link between the elephant random walk and P olya-type urn models, we prove a fixed-point equation (or system in dimension two and larger) for the limiting variable. Based on this, we deduce several properties of the limit distribution, such as the existence of a density with support on  $\mathbb{R}^d$  for  $d \in \{1, 2, 3\}$ , and we bring evidence for a similar result for  $d \geq 4$ . We also investigate the moment-generating function of the limit and give, in dimension 1, a non-linear recurrence relation for the moments.

**Keywords:** Elephant random walk, Multi-dimensional elephant random walk, P olya-type urn models, Asymptotic distribution, Fixed-point equations, Krylov subspaces

**AMS MSC 2020:** 60E05, 60E10, 60J10, 60G50

## 1 Introduction and main results

**Context** The one-dimensional elephant random walk (ERW) was introduced by Sch utz and Trimper [43] in 2004, in order to see how memory could induce subdiffusion in random walk processes. It turned out that the ERW is, in fact, always at least diffusive. Nevertheless, the easy definition of the process together with the underlying deepness of the results has led to a great interest from mathematicians over the last two decades.

The process  $(S_n)_{n \geq 0}$  is defined as follows. We denote by  $(X_n)_{n \geq 0}$  its successive steps. The elephant starts at the origin at time zero:  $S_0 = 0$ . For the first step  $X_1$ , the elephant moves one step to the right with probability  $q$  or one step to the left with probability  $1 - q$ , for some  $q$  in  $[0, 1]$ . The next steps are performed by choosing uniformly at random an integer  $k$  among the previous times. Then the elephant moves exactly in the same direction as at

---

\*Universit e du Qu ebec  a Montr eal (UQAM), D epartement de Math ematiques, Canada. Funded by Natural Sciences and Engineering Research Council of Canada (NSERC) discovery grant (RGPIN-2020-07239); [guerin.helene@uqam.ca](mailto:guerin.helene@uqam.ca)

†Nantes Universit e, Laboratoire de Math ematiques Jean Leray, France. Funded by Centre Henri Lebesgue, programme ANR-11-LABX-0020-01; [lucile.laulin@math.cnrs.fr](mailto:lucile.laulin@math.cnrs.fr)

‡Universit e d'Angers, CNRS, Laboratoire Angevin de Recherche en Math ematiques, France. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No. 759702 and from Centre Henri Lebesgue, programme ANR-11-LABX-0020-01; [raschel@math.cnrs.fr](mailto:raschel@math.cnrs.fr)

time  $k$  with probability  $p \in [0, 1]$ , or in the opposite direction with probability  $1 - p$ . In other words, defining for all  $n \geq 1$ ,

$$X_{n+1} = \begin{cases} +X_k & \text{with probability } p, \\ -X_k & \text{with probability } 1 - p, \end{cases} \quad (1.1)$$

with  $k \sim \mathcal{U}\{1, \dots, n\}$ , the position of the ERW at time  $n + 1$  is given by

$$S_{n+1} = S_n + X_{n+1}. \quad (1.2)$$

The probability  $q$  is called the first step parameter and  $p$  the memory parameter of the ERW. An ERW trajectory is sampled in Figure 1.

A wide range of literature is now available on the ERW and its extensions in dimension  $d = 1$ , see for instance [3, 16, 20, 19, 22, 34]. In dimension  $d \geq 1$ , the multi-dimensional ERW (MERW) was later introduced by Bercu and Laulin [6]. Back to dimension 1, the behavior of the process, in particular its dependency on the value of  $p$  with respect to the critical value  $3/4$ , is now well understood. In the diffusive regime  $p < 3/4$  and the critical regime  $p = 3/4$ , a strong law of large numbers and a central limit theorem for the position, properly normalized, were established, see [3, 16, 17, 43] and the more recent contributions [5, 18, 25, 30, 41, 42, 45]. The main changes between the two regimes are the rate of the associated convergences.

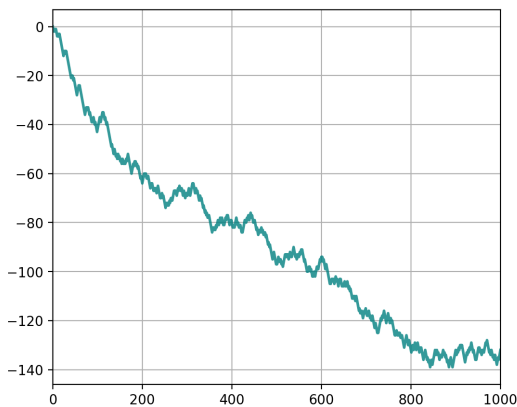


Figure 1: A trajectory of the Elephant random walk in the superdiffusive regime until time  $n = 1000$ , with  $p = 0.87$  and  $q = 0.9$ .

The superdiffusive regime  $p > 3/4$  is even more intriguing. Introduce

$$a := 2p - 1.$$

It has been established that

$$\lim_{n \rightarrow \infty} \frac{S_n}{n^a} = L_q \quad \text{a.s.}, \quad (1.3)$$

where  $L_q$  is a non-degenerate, non-Gaussian random variable, see [3, 4, 16]. This result has been extended to MERW in [6], with a rate of convergence depending on the dimension  $d$ . In dimension 1, it has been proved in [33] that the fluctuations of the ERW around its limit  $L_q$  are Gaussian, as indeed

$$\sqrt{n^{2a-1}} \left( \frac{S_n}{n^a} - L_q \right) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N} \left( 0, \frac{1}{2a-1} \right). \quad (1.4)$$

See [7] for a similar result in dimension  $d \geq 2$ . These results were established thanks to a martingale approach.

Little is known about the limit variable  $L_q$ , although it is a natural question to try to find out more about its law. Indeed, among the open questions concerning the (M)ERW over the last two decades, two stood out: the law of the random variable  $L_q$  arising in the superdiffusive regime, and the recurrence and transience of the MERW, depending on the regime. In this work we give a sharp answer to the first question, as detailed below. The second question was first addressed by Bertoin in dimension  $d = 1$  in [9], establishing that the ERW is recurrent in the diffusive regime and transient in the superdiffusive regime. However, until recently, this remained unresolved for dimensions  $d \geq 2$ . Qin in [40] recently provided an answer, showing that in dimension  $d = 2$ , the MERW behaves similarly to the ERW, whereas in dimensions  $d \geq 3$ , it is always transient, see also [21] for an alternative proof of the recurrence in the diffusive regime for  $d = 2$ .

**Our contributions** In this article, we are interested in studying the distribution of the asymptotics of the superdiffusive ERW, first in dimension 1 (the variable  $L_q$  appearing in (1.3)), then in higher dimension. Indeed, before writing this paper, a lot of natural questions regarding this limiting random variable (such as the existence of a density, an explicit formula for the moments, the moment problem, the finiteness of the moment-generating function, etc.) were still open. While it is generally very difficult to obtain results for the ERW in dimension  $d \geq 2$ , an interesting feature of the present work is that the techniques work in all dimensions, even though we don't have a complete proof in dimension  $d \geq 4$ : our result is proved in dimensions 2 and 3, and we give a method for  $d \geq 4$  based on a (linear algebra) analysis of a certain Krylov space. The main idea of this paper is to establish fixed-point equations related to the limit distribution of the ERW.

**Dimension 1** Using a connection between the ERW and Pólya-type urns, we first establish in Section 2 that in dimension  $d = 1$ , if we impose  $q = 1$ , which means that the elephant first goes to the right, then the random variable  $L_1$  satisfies the following fixed-point equation, from which everything will follow:

**Theorem 1.1.** *Assume  $a = 2p - 1 > \frac{1}{2}$ . Let  $L_1$  be the limit variable appearing in (1.3) with  $q = 1$ . The distribution of  $L_1$  satisfies the fixed-point equation*

$$L_1 \stackrel{\mathcal{L}}{=} V^a L_1^{(1)} + (2\xi_p - 1)(1 - V)^a L_1^{(2)}, \quad (1.5)$$

where

- $V$  is a uniformly distributed random variable on  $[0, 1]$ ;
- $\xi_p$  is a Bernoulli distributed random variable with parameter  $p$ ;
- $L_1^{(1)}$  and  $L_1^{(2)}$  have the same distribution as  $L_1$ ;
- all the variables are independent.

The proof of Theorem 1.1 is detailed in Section 2.2. As we shall see in Remark 2.3, one has

$$L_q \stackrel{\mathcal{L}}{=} (2\xi_q - 1)L_1, \quad (1.6)$$

where as above  $\xi_q$  denotes a Bernoulli random variable with parameter  $q$ , independent of  $L_1$ .

The connection between ERW and Pólya-type urns was first used by Baur and Bertoin [3] in order to obtain functional convergence for the ERW, thanks to the work of Janson [32] on generalized Pólya urns. The originality of our work lies in the use of tree decomposition:

this connection between the ERW and urn models is different from what has been done until now. This point of view together with the initial urn connection has enabled us to establish the fixed-point equation (1.5).

Equations such as (1.5) have already been used to study the limit of generalized Pólya urn processes with deterministic replacement, see [35, 36, 11, 12]. In the same spirit, in our random replacement context, we will deduce various properties of the limit distribution from the fixed-point equation that they satisfy. We mention that there is a wide literature on the study of solutions of general fixed-point equations, see e.g. [24, 1, 31, 37] and the references therein.

The next result, proved in Section 2.2, ensures the existence and uniqueness of the solution to (1.5), with constraints on the first and second moments (we already know by [32], recalled here in Theorem 2.1, that  $\mathbb{E}[L_1] = 1/\Gamma(1+a)$ ).

**Theorem 1.2.** *Let  $a \in (\frac{1}{2}, 1)$ . The random variable  $L_1$  is the unique solution of the distributional equation (1.5) having  $\frac{1}{\Gamma(1+a)}$  as expectation and finite second moment.*

Further, in Section 2.3, we solve a conjecture on the existence of a density of the limit random variable:

**Theorem 1.3.** *For any  $q \in [0, 1]$ ,  $a \in (\frac{1}{2}, 1)$ , the random variable  $L_q$  is absolutely continuous with respect to Lebesgue measure. Its density is a positive, bounded, smooth function on  $\mathbb{R}$ .*

Few illustrations of that density function are presented in Figure 4, Section 2.5.

We also prove in Section 2.4 that the random variable  $L_1$  has finite moments of all order and that its probability distribution is determined by its moments (see Theorem 2.8). Another interesting consequence of the fixed-point equation (1.5) is to give a simple way to compute all moments of  $L_1$  by induction. So far, only the first four moments were known, see [4].

**Theorem 1.4.** *Let  $a \in (\frac{1}{2}, 1)$ . The moments of  $L_1$  are given by the following recursive equation. Let  $(m_k)_{k \geq 1}$  be defined by  $m_1 = 1$  and, for  $k \geq 2$ ,*

$$m_k = \frac{1}{ka - c_k} \sum_{j=1}^{k-1} c_j m_j m_{k-j}, \quad (1.7)$$

where  $c_k = 1$  for even  $k$  and  $c_k = a$  for odd  $k$ . Then, for any  $k \geq 1$ ,

$$\mathbb{E}[L_1^k] = \frac{(k-1)!}{a\Gamma(ka)} m_k, \quad (1.8)$$

and the moment-generating function of  $L_1$  is given by, for  $t \in \mathbb{R}$ ,

$$\mathbb{E}[e^{tL_1}] = \sum_{k \geq 0} \frac{m_k}{\Gamma(ka+1)} t^k.$$

This theorem is proved at the end of Section 2.4. Using (1.6), we immediately deduce the moments of  $L_q$ :  $\mathbb{E}[L_q^k] = (q + (-1)^k(1-q))\mathbb{E}[L_1^k]$ . Finally, we prove that the moment-generating function of  $L_1$  and  $L_1^2$  are both finite everywhere on  $\mathbb{R}$ , see Corollaries 2.11 and 2.12.

At the end of the study in dimension 1, we give in Section 2.6 a series representation (2.16) of the random variable  $L_q$  using classical distributions, which a direct consequence of the interpretation of ERW in terms of random trees (see [2, 10] for more details on the subject).

**Dimension  $d \geq 2$**  Section 3 is devoted to the multidimensional case  $d \geq 2$ , for which we derive similar results as in dimension 1, up to some peculiarities and subtleties. A great feature of our way of using the link between the elephant random walk and the Pólya-type urns model in dimension 1 is that it generalizes nicely to higher dimensions, whereas most work on MERW has had to use stronger tools to break out of dimension 1. We first recall the definition of the multidimensional ERW in Section 3.1, in particular the result [6] giving the existence of a limit variable  $\mathbf{L}$  in the superdiffusive regime.

In Section 3.2, similarly to the dimension 1 case, we associate to the MERW a Pólya-type urn process with  $2d$  different colors (representing the  $2d$  possible directions of the MERW). In the superdiffusive limit, the asymptotic Pólya urn process obtained by [32] (recalled here in Theorem 3.4) can be characterized by two different fixed-point systems, given by (3.13) and (3.17), analogous to Theorem 1.1 in dimension 1, which will both be used in our study. These fixed-point equations were more delicate to obtain than in dimension 1. It is also important to remark that as soon as the ERW is concerned, we only need a partial information from the urn reformulation, as the MERW has natural dimension  $d$ , while the urn model has dimension  $2d$ . The link between the asymptotic  $\mathbf{W}$  of the urn process and the limit MERW  $\mathbf{L}$  is formalized in Section 3.3.

Existence and uniqueness of the solution of the first fixed-point system (3.13) is studied in Section 3.4. In Section 3.5, working with the second fixed-point system (3.17), we give various properties of the limit variable  $\mathbf{W}$  in the urn interpretation, when its support is the full space. In particular we prove existence of a smooth density in Theorem 3.8. We conjecture that our hypothesis on the support is systematically satisfied, and in Section 3.6 we show that indeed it holds in dimension 2 and 3. We further give a method to check this hypothesis in dimension  $d \geq 4$ , based on the analysis of certain Krylov spaces.

In Section 3.7 we show how to transfer from the urn limit variable  $\mathbf{W}$  new properties on the limit variable  $\mathbf{L}$  of the MERW, and we deduce the following theorem in dimension  $d = 2$  and  $d = 3$ .

**Theorem 1.5.** *We assume that  $d = 2$  or  $d = 3$ . Let  $\frac{1}{2} < a < 1$  with  $a = \frac{2dp-1}{2d-1}$ . The limit  $\mathbf{L}$  of a superdiffusive elephant random walk in dimension  $d$  has a smooth bounded density on  $\mathbb{R}^d$ , and its support is  $\mathbb{R}^d$ .*

Finally, in Section 3.8 we prove that the limit MERW variable has finite moments of all order, that its probability distribution is determined by its moments, and that in addition the moment-generating function of  $\mathbf{L}$  and  $\|\mathbf{L}\|^2$  are globally well defined.

## 2 One-dimensional superdiffusive limit using Pólya-type urns

### 2.1 Connection between the ERW and Pólya-type urns

We now present the urn approach of the ERW in dimension 1, but everything can be extended to dimension  $d$  to obtain a more general fixed-point equation. The higher dimensional case is studied in full details in Section 3.

Let  $(\mathbf{U}(n))_{n \geq 0}$  be a discrete-time urn with balls of two colors, say red and blue. The composition of the urn at time  $n \geq 0$  is given by a vector  $\mathbf{U}(n) = (R_n, B_n)^T$ , where  $R_n$  (resp.  $B_n$ ) stands for the number of red (resp. blue) balls at time  $n$ . The initial composition of the urn is  $(1, 0)$  with probability  $q$  or  $(0, 1)$  with probability  $1 - q$ . The composition of urn then evolves as follows. At any time  $n \geq 1$ , a ball is drawn uniformly at random, its color is observed, then it is returned to the urn together with a ball of the same color with probability  $p$ , or with a ball of the other color with probability  $1 - p$ .

Hence, the replacement matrix  $A$  is defined by

$$A = \begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{with probability } p, \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \text{with probability } 1 - p. \end{cases} \quad (2.1)$$

The composition of the urn at time  $n + 1$  satisfies

$$\mathbf{U}(n + 1) = \mathbf{U}(n) + A_{n+1} \begin{pmatrix} \varepsilon_{n+1} \\ 1 - \varepsilon_{n+1} \end{pmatrix},$$

where  $\varepsilon_{n+1}$  is equal to 1 if a red ball is picked and to 0 otherwise, and  $(A_n)_{n \geq 1}$  is a sequence of i.i.d. random matrices with same distribution as  $A$ .

The behavior of the process is deeply linked to the spectral decomposition of the mean replacement matrix  $\mathbb{E}[A]$ , given by

$$\mathbb{E}[A] = \begin{pmatrix} p & 1 - p \\ 1 - p & p \end{pmatrix}.$$

The eigenvalues of  $\mathbb{E}[A]$  are  $\lambda_1 = 1$  and  $\lambda_2 = 2p - 1$ , and the corresponding unit eigenvectors are

$$\mathbf{v}_1 = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{v}_2 = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

It is well known, see [13, 26, 27, 32], that the asymptotics of the urn depends on the ratio  $a = \lambda_2/\lambda_1 = 2p - 1$  with respect to  $1/2$ . In particular, applied to our special case we have the following large urn convergence result [32, 3].

**Theorem 2.1** ([32, 3]). *Let  $U(0) = \alpha = (\alpha_1, \alpha_2) \in \{0, 1\}^2$  be such that  $\alpha_1 + \alpha_2 = 1$ . If  $a = 2p - 1 > 1/2$ , then it holds*

$$\lim_{n \rightarrow \infty} \frac{\mathbf{U}(n) - n\mathbf{v}_1}{n^a} = W_\alpha \mathbf{v}_2 \quad a.s.,$$

where  $W_\alpha$  is a non-degenerate random variable such that

$$\mathbb{E}[W_\alpha] = \frac{\alpha_1 - \alpha_2}{\Gamma(1 + a)} \quad \text{and} \quad \mathbb{E}[W_\alpha^2] = \frac{1}{(2a - 1)\Gamma(2a)}.$$

In particular,

$$\lim_{n \rightarrow \infty} \frac{R_n - B_n}{n^a} = W_\alpha \quad a.s.$$

The connection to the ERW model is straightforward. Let  $(S_n)_{n \geq 0}$  denote the ERW started from  $S_0 = 0$  and such that  $S_1 = R_1 - B_1$ ; then for every  $n \geq 1$ ,

$$S_n \stackrel{\mathcal{L}}{=} R_n - B_n.$$

In other words, the difference between the number of red and blue balls in the urn behaves like an ERW with first step equal to  $R_1 - B_1$  (the red balls corresponding to steps to the right, and the blue ball corresponding to left jumps). Hence, if  $W_\alpha$  denotes the limiting random variable for the urn process started from the composition vector  $\alpha$ , we have the equality in distribution

$$L_q \stackrel{\mathcal{L}}{=} \xi_q W_{(1,0)} + (1 - \xi_q) W_{(0,1)}, \quad (2.2)$$

where  $\xi_q$  is a Bernoulli variable with parameter  $q$ , and  $\xi_q$ ,  $W_{(1,0)}$  and  $W_{(0,1)}$  are mutually independent. In particular, when the elephant starts with a right jump, we have  $L_1 \stackrel{\mathcal{L}}{=} W_{(1,0)}$ . This explains how the asymptotic behavior of the ERW is determined by the spectral decomposition of the (mean) replacement matrix of the corresponding urn.

Finally, we briefly recall a nice result on  $L_q$ , obtained by using a connection between the ERW and random recursive trees on which a Bernoulli bond percolation has been performed (see [34, 10] for more details on the subject). It turns out that results on cluster sizes after percolation, e.g. from [2, Lem. 3.3], can lead to the following decomposition of the random variable  $L_q$ :

$$L_q = C_1 Z_1^{(q)} + \sum_{j=2}^{+\infty} C_j Z_j \quad \text{a.s.}, \quad (2.3)$$

where  $Z_1^{(q)}$  has Rademacher distribution with parameter  $q$  (denoted by  $\mathcal{R}(q)$ ),  $(Z_j)_{j \geq 2}$  are i.i.d. with  $\mathcal{R}(1/2)$  distribution,  $C_1$  has a Mittag-Leffler distribution with parameter  $a$ , and  $(C_j)_{j \geq 2}$  is a sequence of dependent, almost surely positive random variables (their explicit expression is given in Proposition 2.15, Equation (2.16)). This decomposition ensures that the random variable  $L_q$  has a continuous distribution (no atoms). Although we find this expression very elegant, we will not use it because we have not been able to prove the existence of a density on  $\mathbb{R}$  for the distribution of  $L_q$  from it. We will instead focus on the urn-connection and the underlying fixed-point equations to obtain properties of the superdiffusive limit.

## 2.2 Tree-structure and fixed-point equations

We will prove that both random variables  $W_{(1,0)}$  and  $W_{(0,1)}$ , introduced in Theorem 2.1, admit a density, by showing that their characteristic functions are integrable. Our method is strongly inspired by the work of Chauvin et al. [12] for urn processes with deterministic replacement.

The tree structure of the urn process is as follows: a sequence of trees  $(T_n)_{n \geq 0}$  grows at each drawing from the urn. At time 0, there is a red ball in the urn with probability  $q$ , or a blue ball with probability  $1 - q$ . Then the tree starts from a red node with probability  $q$  or a blue node with probability  $1 - q$ :

with probability  $q$  ● ● with probability  $1 - q$

At time  $n$ , each leaf in the tree represents a ball in the urn. When a leaf is chosen (i.e., a ball is drawn), it becomes an internal node and gives birth to 2 balls of the same color with probability  $p$ , or 1 ball of each color with probability  $1 - p$ . Due to the urn process correspondence, leaves are uniformly chosen among the leaves of the tree

At time 1, the tree is decomposed into two subtrees (or a forest with two trees); the three possible decompositions of the tree into two subtrees at time 1 are illustrated on Figure 2. Two

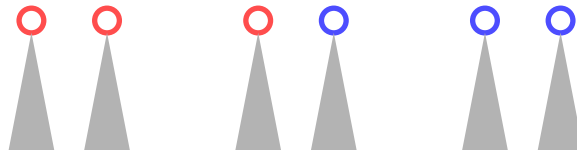


Figure 2: The three possible subtrees at time  $n = 1$ .

possibilities for the tree decomposition at time 4 are displayed on Figure 3 below, depending on the first ball added to the urn (depending on the color of the root node). The balls that have been picked are represented by the empty discs, while the balls that can be picked at



next times are plain discs (as already said, they correspond to the leaves of the tree). On Figure 3, after time 4, there are 5 balls in the urn.

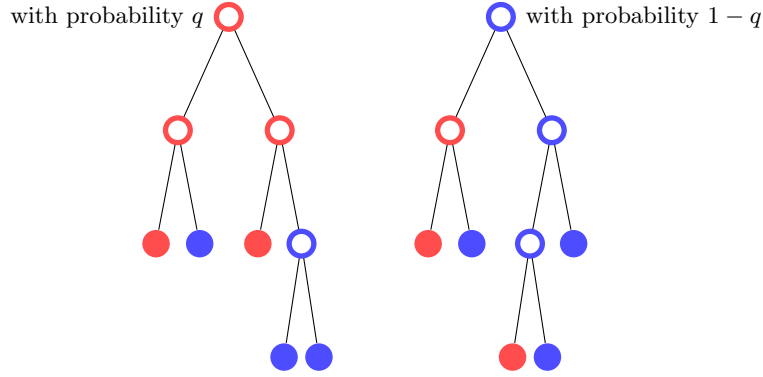


Figure 3: Two examples of tree at time 4.

For any  $n \geq 1$ , denote by  $D_1(n)$  the number of leaves at time  $n$  of the first subtree (starting on the left node at time 1), and similarly  $D_2(n)$  the number of leaves at time  $n$  of the second subtree (starting on the right node at time 1). At time  $n$ , we note that  $D_1(n) + D_2(n) = n + 1$ , and the number of drawings in the  $k$ -th subtree is  $D_k(n) - 1$ . These numbers represent the time inside each subtree.

We now describe the evolution of  $D(n)$ . Remember that the balls of the whole urn are uniformly drawn at any time and notice that at each drawing in the  $k$ -th subtree,  $D_k(n)$  increases by 1. The random vector  $\mathbf{D}(n) = (D_1(n), D_2(n))$  can indeed be seen as a classical 2-color urn process (each color modeling each subtree), with  $I_2$  as (deterministic) replacement matrix and  $(1, 1)$  as initial composition vector.

We recall that  $\mathbf{U}_{(1,0)}$  is the urn process starting from a red ball. Then the tree will start from a red node. The number of red/blue leaves in the whole tree is equal to the sum of the number of red/blue leaves in the first subtree and the number of red/blue leaves in the second subtree. We also easily observe that the first subtree of  $\mathbf{U}_{(1,0)}$  always starts from a red node, and the second subtree starts from a red node with probability  $p$  and a blue node with probability  $1 - p$ .

Gathering these arguments, the distribution of the urn process  $\mathbf{U}_{(1,0)}$  can be described the following way: consider simultaneously

- an original 2-color urn process  $\mathbf{D} = (D_1, D_2)$  having the identity matrix as (deterministic) replacement matrix and  $(1, 1)$  as initial condition;
- for any  $k \in \{1, 2\}$ , an urn process  $\mathbf{U}_{(1,0)}^{(k)}$  having  $A$  in (2.1) as (random) replacement matrix and  $(1, 0)$  as initial condition;
- an urn process  $\mathbf{U}_{(0,1)}^{(2)}$  having  $A$  in (2.1) as (random) replacement matrix and  $(0, 1)$  as initial condition;
- a Bernoulli random variable  $\xi_p$  with parameter  $p$  (modeling the first replacement);

all these processes being independent of each other. Then, the process  $\mathbf{U}_{(1,0)} = (\mathbf{U}_{(1,0)}(n))_{n \geq 1}$  satisfies the distributional equality, for all  $n \geq 1$ :

$$\mathbf{U}_{(1,0)}(n) \stackrel{\mathcal{L}}{=} \mathbf{U}_{(1,0)}^{(1)}(D_1(n) - 1) + \xi_p \mathbf{U}_{(1,0)}^{(2)}(D_2(n) - 1) + (1 - \xi_p) \mathbf{U}_{(0,1)}^{(2)}(D_2(n) - 1).$$

A similar equation can be obtained for  $\mathbf{U}_{(0,1)}$ .

Since  $\mathbf{D} = (D_1, D_2)$  has the same behavior as a classical urn process, we immediately have

$$\lim_{n \rightarrow \infty} n^{-1} \mathbf{D}(n) = \mathbf{Z} \quad \text{a.s.}, \quad (2.4)$$

where  $\mathbf{Z} = (V, 1 - V)$  has Dirichlet(1, 1) distribution, with  $V \sim \mathcal{U}([0, 1])$  a uniform random variable on  $[0, 1]$  (see [39] or [28, Thm 2.1]).

Using the above considerations together with Theorem 2.1, we deduce the following result.

**Theorem 2.2.** *Assume  $a \in (\frac{1}{2}, 1)$ . Let  $W_{(1,0)}$  and  $W_{(0,1)}$  be the elementary limit distributions of a large two-color Pólya urn process with random replacement matrix  $A$  given by (2.1) and ratio  $a > 1/2$ , introduced in Theorem 2.1. Then  $W_{(1,0)}$  and  $W_{(0,1)}$  satisfy the distributional equation system*

$$\begin{cases} W_{(1,0)} \stackrel{\mathcal{L}}{=} V^a W_{(1,0)}^{(1)} + \xi_p (1 - V)^a W_{(1,0)}^{(2)} + (1 - \xi_p) (1 - V)^a W_{(0,1)}^{(2)}, \\ W_{(0,1)} \stackrel{\mathcal{L}}{=} V^a W_{(0,1)}^{(1)} + \xi_p (1 - V)^a W_{(0,1)}^{(2)} + (1 - \xi_p) (1 - V)^a W_{(1,0)}^{(2)}, \end{cases} \quad (2.5)$$

where

- $V$  is a uniformly distributed random variable on  $[0, 1]$ ;
- $\xi_p$  is a Bernoulli distributed random variable with parameter  $p$ ;
- the  $W_{(1,0)}^{(k)}$  and the  $W_{(0,1)}^{(k)}$  are respective copies of  $W_{(1,0)}$  and  $W_{(0,1)}$ ;
- all the variables are independent.

**Remark 2.3.** *The variables  $W_{(1,0)}$  and  $W_{(0,1)}$  satisfy  $W_{(1,0)} \stackrel{\text{a.s.}}{=} -W_{(0,1)}$ . Indeed, by Theorem 2.1, the urn process started from  $(1, 0)$  satisfies*

$$\lim_{n \rightarrow \infty} \frac{R_n - B_n}{n^a} = W_{(1,0)} \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{B_n - R_n}{n^a} = -W_{(1,0)} \quad \text{a.s.}$$

The two colors of the urn are exchangeable, hence the random variable  $B_n - R_n$  can be seen as  $R_n - B_n$  when started from  $(0, 1)$ , and therefore satisfies

$$\lim_{n \rightarrow \infty} \frac{B_n - R_n}{n^a} = W_{(0,1)} \quad \text{a.s.}$$

As a consequence and using (2.2), we have  $L_q \stackrel{\mathcal{L}}{=} (2\xi_q - 1)W_{(1,0)}$ , and in particular  $W_{(1,0)} \stackrel{\mathcal{L}}{=} L_1$ .

Due to the above remark, we may only focus our attention on  $W_{(1,0)}$ . In particular, the system (2.5) can be reduced to a single equation, as summarized in Theorem 1.1. Theorem 1.1 is then proved.

We now give the proof of Theorem 1.2 on existence and uniqueness of the solution with a finite second-order moment to the fixed-point Equation (1.5) when the first moment is fixed.

*Proof of Theorem 1.2.* This proof is greatly inspired by [12, Lem. 6 and Thm 7]. As pointed out in [12], it could be deduced from the general result in Neininger-Rüschendorf [38], but we choose to give instead a direct proof via the contraction method.

Let us fix  $m \in \mathbb{R}$ , and define  $\mathcal{P}_2(m)$  be the space of probability distributions on  $\mathbb{R}$  that have  $m$  as expectation and finite second moment. We consider the Wasserstein distance  $d_{\mathcal{W}}$  on  $\mathcal{P}_2(m)$ . For a brief overview on this distance, we refer the reader to [12, Sec. 4.1], and we recall that

$$d_{\mathcal{W}}(\mu, \nu) = \min_{(X, Y)} \sqrt{\mathbb{E}[(X - Y)^2]},$$

for  $(X, Y)$  a random vector with marginal distributions  $(\mu, \nu)$ . We then introduce

$$H : \mathcal{P}_2(m) \longrightarrow \mathcal{P}_2(m) \\ \mu \longmapsto \mathcal{L}\left(V^a W^{(1)} + (2\xi_p - 1)(1 - V)^a W^{(2)}\right),$$

where  $W^{(1)}, W^{(2)}$  have distribution  $\mu$ ,  $V$  is uniform on  $(0, 1)$ ,  $\xi_p$  is Bernoulli  $\mathcal{B}(p)$  and all the random variables are independent from each other.

For  $\mu \in \mathcal{P}_2(m)$ , it is immediate that  $\int xH(\mu)(dx) = m$ . We want to show that  $H$  is Lipschitz-continuous for the Wasserstein distance. Let  $\mu_1, \mu_2 \in \mathcal{P}_2(m)$ ,  $(W_1^{(1)}, W_2^{(1)})$  and  $(W_1^{(2)}, W_2^{(2)})$  two independent couples with marginal distributions  $(\mu_1, \mu_2)$ . Then

$$d_{\mathcal{W}}(H\mu_1, H\mu_2)^2 \leq \mathbb{E}\left[\left(V^a(W_1^{(1)} - W_2^{(1)}) + (2\xi_p - 1)(1 - V)^a(W_1^{(2)} - W_2^{(2)})\right)^2\right] \\ \leq \mathbb{E}[V^{2a}]\mathbb{V}\left[W_1^{(1)} - W_2^{(1)}\right] + \mathbb{E}[(1 - V)^{2a}]\mathbb{V}\left[W_1^{(2)} - W_2^{(2)}\right] \\ \leq \frac{2}{1 + 2a}\mathbb{E}\left[\left(W_1^{(1)} - W_2^{(1)}\right)^2\right].$$

Thus taking the minimum over all couples  $(W_1^{(1)}, W_2^{(1)})$  with marginal distributions  $(\mu_1, \mu_2)$ , we deduce

$$d_{\mathcal{W}}(H\mu_1, H\mu_2)^2 \leq \frac{2}{1 + 2a}d_{\mathcal{W}}(\mu_1, \mu_2)^2.$$

Hence  $H$  is  $\sqrt{\frac{2}{1+2a}}$ -Lipschitz, which implies that  $H$  is a contraction ( $a > 1/2$ ) and that it has unique fixed point in  $\mathcal{P}_2(m)$ . Theorem 1.2 is written in the specific case  $m = \frac{1}{\Gamma(1+a)}$ .  $\square$

**Remark 2.4.** Equation (1.5) turns out to be a particular case of fixed point equations studied in [31]. More specifically, in the notation of [31], the dimension should be taken to 1,  $N = 2$ ,  $T_1 = V^a \in (0, 1)$  and  $T_2 = (2\xi_p - 1)(1 - V)^a \in (-1, 1)$ , with  $V$  a uniform random variable on  $(0, 1)$  and  $\xi_p$  a Bernoulli variable with parameter  $p$ , independent of  $V$ . Since  $\mathbb{E}[|T_1|^\alpha + |T_2|^\alpha] = 2/(1 + \alpha a)$ , Assumptions (A1)–(A4) of [31] are satisfied for  $\alpha = 1/a \in (1, 2)$ .

By [31, Eq. (2.31)], we deduce that there exists  $c \in \mathbb{R}$  such that

$$L_1 \stackrel{\mathcal{L}}{=} cZ + W^a Y_{1/a},$$

where

- $Z$  is the limit value of a martingale, a.s. and in  $L^\beta$ , for  $1 < \beta < 1/a$ ;
- $W$  is a special endogenous non-negative solution to the tilted equation

$$W \stackrel{\mathcal{L}}{=} VW^{(1)} + (1 - V)W^{(2)}, \tag{2.6}$$

where  $(W^{(1)}, W^{(2)})$  are independent copies of  $W$ , and independent of  $V$ ;

- $Y_{1/a}$  is a strictly  $1/a$ -stable random variable independent of  $Z$  and  $W$ .

We notice that if  $W$  is a solution to (2.6), then

$$\mathbb{E}[W^2] = 2\mathbb{E}[V^2]\mathbb{E}[W^2] + 2\mathbb{E}[V(1 - V)]\mathbb{E}[W]^2 = \frac{2}{3}\mathbb{E}[W^2] + \frac{1}{3}\mathbb{E}[W]^2.$$

It implies  $\mathbb{E}[W^2] = \mathbb{E}[W]^2$ , and we deduce that  $W$  is a constant.

Recall that  $\alpha$ -stable distributions don't have second-order moments, unless  $\alpha = 2$ . Under this assumption, we have  $W = 0$  since  $1/2 < a < 1$ . We thus deduce that up to a constant,  $Z$  has the same distribution than  $L_1$ .

### 2.3 Existence of a density

In this section we will prove Theorem 1.3. Let  $W := W_{(1,0)}$  be the first elementary limit distribution, as defined in Theorem 2.1, and  $\varphi_W$  be its Fourier transform, defined for  $t \in \mathbb{R}$  by

$$\varphi_W(t) := \mathbb{E}[e^{itW}].$$

We recall that  $W \stackrel{\mathcal{L}}{=} L_1$  satisfies the fixed-point Equation (1.5).

**Theorem 2.5.** *Let  $W$  be the first elementary limit distribution of a large two-color Pólya urn process with random replacement matrix  $A$  and ratio  $a \in (1/2, 1)$ . Then*

- the support of  $W$  is the whole real line  $\mathbb{R}$ ;
- for any  $\rho \in (0, \frac{1}{a})$ , and any  $k \geq 1$ , there exists a constant  $C > 0$  such that for any  $t \in \mathbb{R}^*$ ,

$$|\varphi_X(t)| \leq \frac{C}{|t|^{k\rho}}. \quad (2.7)$$

Before proving the theorem, we mention a few obvious consequences.

**Corollary 2.6.** *For any initial composition  $\alpha = (1, 0)$  or  $(0, 1)$ , the random variable  $W_\alpha$  admits a positive, bounded and smooth density. Its support is the whole real line  $\mathbb{R}$ .*

*Proof of Corollary 2.6.* Recall that  $W_{(1,0)} \stackrel{a.s.}{=} -W_{(0,1)}$ . We notice that  $1 < \frac{1}{a} < 2$ , then taking  $k = 1$  in Theorem 2.5, we deduce that the characteristic function  $\varphi_W$  of  $W$  is integrable on  $\mathbb{R}$ . By Fourier inversion theorem,  $W$  has a bounded density on  $\mathbb{R}$ . Since the support of  $W$  is  $\mathbb{R}$ , its density is positive a.e. Besides, for any  $k \geq 1$ ,  $t \mapsto t^k \varphi_W(t)$  is integrable on  $\mathbb{R}$ , which implies that the density function is of class  $\mathcal{C}^k$ . The corollary is proved.  $\square$

**Remark 2.7.** *From the fixed-point equation (1.5), we deduce that the density function  $f$  of  $L_1$  satisfies the following integral equation, for all  $x \in \mathbb{R}$ ,*

$$f(x) = \int_{\mathbb{R}} \int_0^1 \frac{1}{z^a} f(x - (1-z)^a y) (pf(y) + (1-p)f(-y)) dz dy.$$

*This equation is unfortunately too complex to easily deduce properties on the distribution of  $L_1$ .*

We recall that  $L_q \stackrel{\mathcal{L}}{=} (2\xi_q - 1)W$ , where  $\xi_q$  is a Bernoulli random variable with parameter  $q$ , independent on  $W$ . Then we easily deduce that Theorem 1.3 is a direct consequence of Corollary 2.6. See Figure 4 in Section 2.5 for a few illustrations of the density.

*Proof of Theorem 2.5.* We proceed as in the proof of [12, Thm 6]; the details are slightly different but still very similar to their case. Indeed, we only add one ball at each step but we do it randomly, whereas they insisted that they needed to add at least two balls at each step since they do it deterministically. We have been inspired by [36, Lem. 3.1 and Lem. 3.2] as well.

We start by proving that  $\text{Supp}(W) = \mathbb{R}$ , with  $\text{Supp}(W)$  denoting the support of  $W$ . By Theorem 2.1, the variance of  $W$  is non-zero, so there exist  $w_1, w_2 \in \text{Supp}(W)$  with  $w_1 \neq w_2$ . Because of the fixed-point equation (1.5), for any  $t \in [0, 1]$  we have

$$w_1 t^a + w_2 (1-t)^a \in \text{Supp}(W), \quad (2.8)$$

since  $\mathbb{P}(\xi_p = 1) = p > 0$ . Moreover, it holds that  $[w_1, w_2] \subset \{w_1 t^a + w_2(1-t)^a : t \in [0, 1]\}$ . This entails that  $[w_1, w_2] \subset \text{Supp}(W)$ .

Using  $\mathbb{P}(\xi_p = -1) = 1 - p > 0$ , similar computations lead to  $[w_1, -w_2] \subset \text{Supp}(W)$ . Moreover, since  $\mathbb{E}[W] > 0$ , there exists  $w > 0$  in  $\text{Supp}(W)$ . Taking  $-w_1 = w_2 = w > 0$ , we obtain that  $[-w, w] \subset \text{Supp}(W)$ .

Next, choosing  $t = 1/2$ , Equation (2.8) ensures that as soon as  $w \in \text{Supp}(W)$ ,  $2^{1-a}t \in \text{Supp}(W)$ . Since  $a < 1$ ,  $2^{1-a} > 1$  and the images of  $[-w, w]$  by the iterates of the homothetic transformation  $w \mapsto 2^{1-a}w$  fill the whole real line, which yields that  $\text{Supp}(W) = \mathbb{R}$ .

We move to the proof of (2.7), and we split the argument in four steps.

*Step 1:* We start by showing that for any  $t \neq 0$ ,  $|\varphi_W(t)| < 1$ . Assume there exists  $t_0 \in \mathbb{R}$  such that  $|\varphi_W(t_0)| = 1$ , then  $\varphi_W(t_0) = e^{i\theta_0}$  for some  $\theta_0 \in \mathbb{R}$ . We deduce that  $\mathbb{E}[\text{Re}(1 - e^{i(t_0 W - \theta_0)})] = 0$ , and then  $e^{it_0 W} = e^{i\theta_0}$  a.s. This is possible only if  $t_0 = 0$ , since  $\text{Supp}(W) = \mathbb{R}$ .

*Step 2:* We show that

$$\limsup_{t \rightarrow \pm\infty} |\varphi_W(t)| \in \{0, 1\}. \quad (2.9)$$

Using again the fixed-point equation (1.5), we obtain by conditioning on  $V$  that for any  $t \in \mathbb{R}$ ,

$$\varphi_W(t) = p\mathbb{E}[\varphi_W(tV^a)\varphi_W(t(1-V)^a)] + (1-p)\mathbb{E}[\varphi_W(tV^a)\varphi_W(-t(1-V)^a)]. \quad (2.10)$$

In particular, since  $V$  is uniformly distributed on  $[0, 1]$ , it is a.s. non-zero. Consequently, Equation (2.10) together with Fatou's lemma ensure that

$$\begin{aligned} \limsup_{t \rightarrow \pm\infty} |\varphi_W(t)| &\leq p \limsup_{t \rightarrow \pm\infty} \mathbb{E}[|\varphi_W(tV^a)| \times |\varphi_W(t(1-V)^a)|] \\ &\quad + (1-p) \limsup_{t \rightarrow \pm\infty} \mathbb{E}[|\varphi_W(tV^a)| \times |\varphi_W(-t(1-V)^a)|] \\ &\leq p \limsup_{t \rightarrow \pm\infty} |\varphi_W(t)|^2 + (1-p) \limsup_{t \rightarrow \pm\infty} |\varphi_W(t)|. \end{aligned}$$

We deduce

$$\limsup_{t \rightarrow \pm\infty} |\varphi_W(t)| \leq \limsup_{t \rightarrow \pm\infty} |\varphi_W(t)|^2 \leq \left( \limsup_{t \rightarrow \pm\infty} |\varphi_W(t)| \right)^2,$$

by monotony of  $x \mapsto x^2$  on  $\mathbb{R}_+$ , concluding the proof of (2.9).

*Step 3:* We now show that in (2.9) the only possible value is 0. Reasoning by contradiction, we assume that  $\limsup |\varphi_W(t)| = 1$  as  $t \rightarrow \pm\infty$ . Let  $t_0 > 0$  and  $\varepsilon > 0$  be such that  $0 < |\varphi_W(t_0)| < 1 - \varepsilon$ . Choose  $t_1 = t_1(\varepsilon)$  and  $t_2 = t_2(\varepsilon)$  such that  $0 < t_1 < t_0 < t_2 < \infty$  and

$$|\varphi_W(t_1)| = |\varphi_W(t_2)| = 1 - \varepsilon \quad \text{and} \quad |\varphi_W(t)| \leq 1 - \varepsilon \quad \text{for } t \in [t_1, t_2].$$

This is possible thanks to the intermediate value theorem, since  $\varphi_W$  is continuous and we assumed that  $\limsup_{t \rightarrow \pm\infty} |\varphi_W(t)| = 1 = |\varphi_W(0)|$ . Moreover, if  $(\varepsilon_n)_{n \geq 1}$  is a sequence tending to 0 as  $n$  goes to infinity and if  $y$  is a limit point of  $(t_1(\varepsilon_n))_{n \geq 1}$ , then by definition of  $t_1$  and the continuity of  $\varphi_W$ ,  $|\varphi_W(y)| = 1$ . This entails that  $y = 0$  and that  $\lim_{\varepsilon \rightarrow 0} t_1(\varepsilon) = 0$ .

Again, Equation (2.10) implies that for all  $t > 0$ ,

$$|\varphi_W(t)| \leq p\mathbb{E}[|\varphi_W(tV^a)| \times |\varphi_W(t(1-V)^a)|] + (1-p)\mathbb{E}[|\varphi_W(tV^a)| \times |\varphi_W(-t(1-V)^a)|],$$

with  $|\varphi_W(-t(1-V)^a)| = \left| \overline{\varphi_W(t(1-V)^a)} \right| = |\varphi_W(t(1-V)^a)|$ , and, for all  $n \geq 1$  and  $t > 0$ ,

$$\begin{aligned} |\varphi_W(t)| &\leq \mathbb{E}[|\varphi_W(tV^a)| \times |\varphi_W(t(1-V)^a)|] \\ &\leq \mathbb{E}[|\varphi_W(tM)|] \leq \phi_n(t), \end{aligned} \quad (2.11)$$

where  $M = V^a$  and  $\phi_n(t) = \mathbb{E}[|\varphi_W(tM_1 \cdots M_n)|]$ , with  $M_1, \dots, M_n$  i.i.d. copies of  $M$ . (Notice that one could also take  $M = (1 - V)^a$  or  $M = \max(V^a, (1 - V)^a)$ .) On the one hand, using the definitions of  $t_1, t_2$ , we find

$$\begin{aligned} \phi_n(t) &\leq (1 - \varepsilon)\mathbb{P}(t_1 < tM_1 \cdots M_n \leq t_2) + 1 - \mathbb{P}(t_1 < tM_1 \cdots M_n \leq t_2) \\ &\leq 1 - \varepsilon\mathbb{P}(t_1 < tM_1 \cdots M_n \leq t_2) \end{aligned}$$

and by (2.11)

$$1 - \varepsilon = |\varphi_W(t_2)| \leq \mathbb{E}[\phi_n(t_2V^a)\phi_n(t_2(1 - V)^a)],$$

which leads to

$$\mathbb{P}(t_1 < t_2V^aM_1 \cdots M_n \leq t_2) + \mathbb{P}(t_1 < t_2(1 - V)^aM_1 \cdots M_n \leq t_2) \leq 1 + \varepsilon. \quad (2.12)$$

Besides, when  $\varepsilon$  goes to 0, we find that

$$\mathbb{P}(t_1 < t_2xM_1 \cdots M_n \leq t_2) \longrightarrow \mathbb{P}(xM_1 \cdots M_n \leq 1),$$

which, together with dominated convergence, ensures that as  $\varepsilon \rightarrow 0$

$$\mathbb{P}(t_1 < t_2V^aM_1 \cdots M_n \leq t_2) \longrightarrow \mathbb{P}(V^aM_1 \cdots M_n \leq 1)$$

and the same thing holds for  $(1 - V)^a$ . Thus, by (2.12) and taking  $\varepsilon \rightarrow 0$ , we deduce

$$\mathbb{P}(V^aM_1 \cdots M_n \leq 1) + \mathbb{P}((1 - V)^aM_1 \cdots M_n \leq 1) \leq 1. \quad (2.13)$$

On the other hand, we also have

$$\mathbb{P}(xM_1 \cdots M_n \leq 1) \xrightarrow[n \rightarrow \infty]{} 1$$

thanks to Markov's inequality, since  $\mathbb{E}[M^{1/a}] < 1$ . Using again dominated convergence, we find

$$\mathbb{P}(V^aM_1 \cdots M_n \leq 1) + \mathbb{P}((1 - V)^aM_1 \cdots M_n \leq 1) \xrightarrow[n \rightarrow \infty]{} 2$$

and we have a contradiction with (2.13). The third step is proved.

*Step 4:* Finally, we focus our attention on the integrability of  $\varphi_W$  and prove (2.7). Let  $\varepsilon > 0$  and  $T_\varepsilon > 0$  be such that  $|\varphi_W(t)| \leq \varepsilon$  as soon as  $|t| \geq T_\varepsilon$ . The quantity  $T_\varepsilon > 0$  exists (and is finite) due to the third step of the proof. We have from Equation (2.10) that for any  $t \in \mathbb{R}$ ,

$$\begin{aligned} |\varphi_W(t)| &\leq p\mathbb{E}[|\varphi_W(tV^a)| \times |\varphi_W(t(1 - V)^a)|] + (1 - p)\mathbb{E}[|\varphi_W(tV^a)| \times |\varphi_W(-t(1 - V)^a)|] \\ &\leq \varepsilon\mathbb{E}[|\varphi_W(tV^a)|] + \mathbb{P}((1 - V)^a|t| \leq T_\varepsilon). \end{aligned}$$

Hence, since  $\mathbb{E}[V^{-a\rho}] < \infty$  as soon as  $\rho \in (0, 1/a)$ , we find thanks to Markov's inequality that there exists a constant  $C = \frac{1}{1 - \rho a} > 0$  such that, for any  $t \neq 0$ ,

$$|\varphi_W(t)| \leq \varepsilon\mathbb{E}[|\varphi_W(tV^a)|] + C\left(\frac{T_\varepsilon}{|t|}\right)^\rho.$$

Then, by iterating the last inequality, we find that

$$|\varphi_W(t)| \leq \varepsilon^n\mathbb{E}[|\varphi_W(tV_1^a \cdots V_n^a)|] + C\left(\frac{T_\varepsilon}{|t|}\right)^{\rho n-1} \sum_{k=0}^{n-1} (\varepsilon\mathbb{E}[V^{-a\rho}])^k.$$

When  $n$  goes to infinity, this implies that

$$|\varphi_W(t)| \leq C \left( \frac{T_\varepsilon}{|t|} \right)^\rho \frac{1}{1 - \varepsilon \mathbb{E}[V^{-a\rho}]} = \frac{1}{1 - a\rho - \varepsilon} \left( \frac{T_\varepsilon}{|t|} \right)^\rho \quad (2.14)$$

if  $\varepsilon$  is chosen such that  $\varepsilon \mathbb{E}[V^{-a\rho}] < 1$ . This is enough to conclude that for any  $\rho \in (0, 1/a)$ , we have  $\varphi_X(t) \underset{\pm\infty}{=} O(t^{-\rho})$ . We implement (2.14) in Equation (2.10), and we thus obtain, for any  $\rho \in (0, 1/a)$ ,  $\varepsilon \in (0, 1 - a\rho)$ , and  $|t| > T_\varepsilon$ ,

$$\begin{aligned} |\varphi_W(t)| &\leq \varepsilon \mathbb{E}[|\varphi_W(tV^a)|] + \mathbb{E}[|\varphi_W(V^a t)| \mathbf{1}_{(1-V)^a |t| \leq T_\varepsilon}] \\ &\leq \varepsilon \mathbb{E}[|\varphi_W(tV^a)|] + \frac{1}{1 - a\rho - \varepsilon} \left( \frac{T_\varepsilon}{|t|} \right)^\rho \mathbb{E} \left[ V^{-a\rho} \mathbf{1}_{V \geq 1 - \left( \frac{T_\varepsilon}{|t|} \right)^{1/a}} \right] \\ &\leq \varepsilon \mathbb{E}[|\varphi_W(tV^a)|] + \frac{1}{1 - a\rho - \varepsilon} \left( \frac{T_\varepsilon}{|t|} \right)^\rho \mathbb{E} \left[ V^{-1} \mathbf{1}_{V \geq 1 - \left( \frac{T_\varepsilon}{|t|} \right)^\rho} \right] \\ &\leq \varepsilon \mathbb{E}[|\varphi_W(tV^a)|] - \frac{1}{1 - a\rho - \varepsilon} \left( \frac{T_\varepsilon}{|t|} \right)^\rho \ln \left( 1 - \left( \frac{T_\varepsilon}{|t|} \right)^\rho \right). \end{aligned}$$

As  $-\ln(1-u) \leq 2u$  for all  $u \in (0, 1/2)$ , there exists a constant  $C_\varepsilon$  (that may change from line to line) such that for all  $|t| > 2^{1/\rho} T_\varepsilon$ ,

$$|\varphi_W(t)| \leq \varepsilon \mathbb{E}[|\varphi_W(tV^a)|] + C_\varepsilon \left( \frac{T_\varepsilon}{|t|} \right)^{2\rho},$$

and then

$$|\varphi_W(t)| \leq C_\varepsilon \left( \frac{T_\varepsilon}{|t|} \right)^{2\rho}.$$

Iterating this idea, we obtain that for any  $k \geq 1$ ,  $|\varphi_W(t)| = O(|t|^{-k\rho})$  as  $|t| \rightarrow \infty$ .  $\square$

## 2.4 Moments of the superdiffusive distribution

Recall that  $W = W_{(1,0)}$ , where  $W_{(1,0)}$  has been defined in Theorem 2.1. The main objective of this part is to prove the following result, as well as Theorem 1.4.

**Theorem 2.8.** *The random variables  $L_q$  and  $W$  have finite moments of all order and their probability distribution are determined by their moments. Moreover, their Laplace series have an infinite radius of convergence.*

To prove these results, we will use the following Carleman's criterion (see e.g. [44]).

**Theorem 2.9** (Carleman's criterion). *Let  $X$  be a real random variable with finite moments of all order. If  $\sum_{k \geq 1} (\mathbb{E}[|X|^{2k}])^{-1/2k} = \infty$ , then the distribution of  $X$  is uniquely determined by its moments.*

It is already known that  $L_q$  has finite moments of all order, see [5, Eq. (4.12)]. By Remark 2.3, we have  $\mathbb{E}[|L_q|^k] = \mathbb{E}[|W|^k]$  for all  $k \geq 1$ . The idea is to use the fixed-point equation (1.5) satisfied by  $W$  to prove that  $W$  (and thus  $L_q$ ) satisfies the assumption of Carleman's criterion. We first prove a preliminary estimate.

**Lemma 2.10.** *There exists a constant  $C > 0$  such that for any  $k \geq 1$ ,*

$$\frac{\mathbb{E}[|W|^k]}{k!} \leq \frac{C^k}{\Gamma(ka + 1)} \left( \frac{2}{a} \right)^{k-1}. \quad (2.15)$$

*Proof.* We prove the proposition by induction. The second moment of  $W \stackrel{\mathcal{L}}{=} L_1$  is given in [32, 4] and recalled in this paper in Theorem 2.1. Let  $C = \Gamma(a+1)\mathbb{E}[L_1^2]^{1/2}$ . For  $k = 1$  and  $k = 2$ , we easily see that Equation (2.15) holds (thanks to Cauchy-Schwarz inequality for  $k = 1$ ).

Let  $k \geq 3$  and assume that (2.15) is true for all  $j \in \{1, \dots, k-1\}$ . We now study the  $k$ -th moment of  $W$ . Taking  $(W^{(1)}, W^{(2)})$  a couple of independent copies of  $W$ , by (1.5) and independence between  $V$  and  $(W^{(1)}, W^{(2)})$ , we have

$$\begin{aligned} \mathbb{E}[|W|^k] &\leq \mathbb{E}\left[\left(V^a|W^{(1)}| + (1-V)^a|W^{(2)}|\right)^k\right] \\ &\leq 2\mathbb{E}[V^{ka}]\mathbb{E}[|W|^k] + \sum_{j=1}^{k-1} \binom{k}{j} \mathbb{E}[V^{ja}(1-V)^{(k-j)a}]\mathbb{E}[|W|^j]\mathbb{E}[|W|^{k-j}]. \end{aligned}$$

As  $\mathbb{E}[V^{ka}] = \frac{1}{1+ka}$ , we deduce

$$\mathbb{E}[|W|^k] \leq \frac{ka+1}{ka-1} \sum_{j=1}^{k-1} \binom{k}{j} \mathbb{E}[V^{ja}(1-V)^{(k-j)a}]\mathbb{E}[|W|^j]\mathbb{E}[|W|^{k-j}],$$

which by induction immediately leads to

$$\begin{aligned} \frac{\mathbb{E}[|W|^k]}{k!} &\leq \frac{ka+1}{ka-1} \sum_{j=1}^{k-1} \frac{\Gamma(ja+1)\Gamma((k-j)a+1)}{\Gamma(ka+2)} \frac{\mathbb{E}[|W|^j]}{j!} \frac{\mathbb{E}[|W|^{k-j}]}{(k-j)!} \\ &\leq \frac{k-1}{ka-1} \frac{a}{2} \left(\frac{2}{a}\right)^{k-1} \frac{C^k}{\Gamma(ka+1)} \\ &\leq \left(\frac{2}{a}\right)^{k-1} \frac{C^k}{\Gamma(ka+1)}, \end{aligned}$$

since for  $k \geq 3$ ,  $a(k-1) \leq 2ka-2$ . The lemma is proved.  $\square$

*Proof of Theorem 2.8.* By Remark 2.3, if we have the result for  $W$ , we immediately obtain it for  $L_q$ . By Lemma 2.10, there exists  $C > 0$  such that for all  $k \geq 1$ ,

$$\mathbb{E}[|W|^k]^{-1/k} \geq \frac{a}{2C} \left(\frac{2}{a}\right)^{1/k} \left(\frac{k!}{\Gamma(ka+1)}\right)^{-1/k}.$$

Recalling Stirling's formula  $\Gamma(z) \underset{z \rightarrow +\infty}{\sim} \sqrt{\frac{2\pi}{z}} \left(\frac{z}{e}\right)^z$ , we deduce

$$\left(\frac{2}{a}\right)^{1/k} \left(\frac{k!}{\Gamma(ka+1)}\right)^{-1/k} \underset{k \rightarrow +\infty}{\sim} \frac{a^a e^{1-a}}{k^{1-a}}.$$

Since  $1-a < 1$ ,  $\sum_{k \geq 1} 1/k^{1-a} = \infty$ , we conclude by Carleman's criterion. In fact, the estimate of Lemma 2.10 ensures that the radius of convergence of the Laplace series of  $|W|$  is infinite.  $\square$

**Corollary 2.11.** *For any first step parameter  $q \in [0, 1]$  and memory parameter  $p \in (3/4, 1)$ , the moment-generating functions of  $L_q$  and  $|L_q|$  are well defined and finite on  $\mathbb{R}$ , meaning that for all  $t \in \mathbb{R}$ ,*

$$\mathbb{E}[e^{tL_q}] < \infty \quad \text{and} \quad \mathbb{E}[e^{t|L_q|}] < \infty.$$



*Proof.* By Tonelli's theorem and Theorem 1.4, for all  $t \in \mathbb{R}$ ,

$$\mathbb{E}\left[e^{t|L_q|}\right] = \sum_{k=0}^{\infty} \frac{\mathbb{E}[|tL_q|^k]}{k!} = \sum_{k=0}^{\infty} \frac{|t|^k \mathbb{E}[|W|^k]}{k!} < \infty.$$

Then  $\mathbb{E}\left[e^{tL_q}\right] = \sum_{k=0}^{\infty} \frac{t^k L_q^k}{k!}$  is absolutely convergent.  $\square$

Bercu noticed in [4, Rem. 3.2] that  $L_q$  is a sub-Gaussian random variable. We prove here that  $L_q^2$  has finite exponential moments, this indicates that the distribution of  $L_q$  has really thin tails.

**Corollary 2.12.** *For any first step parameter  $q \in [0, 1]$  and memory parameter  $p \in (3/4, 1)$ , the moment-generating function of  $L_q^2$  is well defined and finite on  $\mathbb{R}$ .*

*Proof.* For  $t \in \mathbb{R}$ ,

$$0 \leq e^{tL_q^2} = \sum_{k=0}^{\infty} \frac{t^k L_q^{2k}}{k!} \leq \sum_{k=0}^{\infty} \frac{|t|^k W^{2k}}{k!}.$$

Hence for all  $t \in \mathbb{R}$ ,

$$\mathbb{E}\left[e^{tL_q^2}\right] \leq \sum_{k=0}^{\infty} \frac{|t|^k \mathbb{E}[W^{2k}]}{k!}.$$

Hereafter, we have from Lemma 2.10 that

$$\sum_{k=0}^{\infty} \frac{|t|^k \mathbb{E}[W^{2k}]}{k!} \leq \frac{a}{2} \sum_{k=0}^{\infty} |t|^k \left(\frac{2C}{a}\right)^{2k} \frac{(2k)!}{k! \Gamma(2ka + 1)}.$$

Since  $a > 1/2$ , we obtain that

$$\frac{k! \Gamma(2ka + 1)}{(2k)!} \frac{(2k + 2)!}{(k + 1)! \Gamma(2ka + 2a + 1)} \sim \frac{(2k)^2}{k(2ka)^{2a}} \sim \frac{2^{2(1-a)}}{a^{2a} k^{2a-1}} \rightarrow 0$$

when  $k$  goes to  $\infty$ , and thus the radius of convergence is infinite.  $\square$

Since we know that  $L_1$  has finite moments of all order and its moment-generating function is well defined, we can now prove Theorem 1.4. However, before giving the proof, we make some comments on the result.

**Remark 2.13.** *From the recursive equation in Theorem 1.4, we easily compute by induction*

$$m_1 = 1, \quad m_2 = \frac{a}{2a-1}, \quad m_3 = \frac{a+1}{2(2a-1)} \quad \text{and} \quad m_4 = \frac{a(2a^2+2a-1)}{(4a-1)(2a-1)^2},$$

and then the successive moments of  $L_1 \stackrel{\mathcal{L}}{=} W$ , are given by

$$\begin{aligned} \mathbb{E}[L_1] &= \frac{1}{\Gamma(a+1)}, & \mathbb{E}[L_1^2] &= \frac{1}{(2a-1)\Gamma(2a)}, \\ \mathbb{E}[L_1^3] &= \frac{a+1}{a(2a-1)\Gamma(3a)}, & \mathbb{E}[L_1^4] &= \frac{6(2a^2+2a-1)}{(4a-1)(2a-1)^2\Gamma(4a)}. \end{aligned}$$

We recover the expressions obtained by Bercu in [4].

Since  $a \mapsto m_2(a) = \frac{a}{2a-1}$  is positive and decreasing on  $(1/2, 1)$ , we have the following trivial corollary of Theorem 1.4.

**Corollary 2.14.** *The following holds:*

1. For all  $k \geq 1$ ,  $\mathbb{E}[L_1^k]$  is positive.
2. For all  $k \geq 2$ ,  $a \mapsto m_k(a)$  is decreasing on  $(1/2, 1)$  with

$$\lim_{a \rightarrow 1/2} m_k(a) = \infty \quad \text{and} \quad \lim_{a \rightarrow 1} m_k(a) = 1.$$

Consequently, for all  $k \geq 1$ ,  $\lim_{a \rightarrow 1} \mathbb{E}[|L_1 - 1|^k] = 0$ .

3. Since  $L_q \stackrel{\mathcal{L}}{=} (2\xi_q - 1)L_1$  with  $\xi_q \sim \mathcal{B}(q)$  independent of  $L_1$  (see Remark 2.3), we have for all  $k \geq 1$  and  $t \in \mathbb{R}$ ,

$$\begin{aligned} \mathbb{E}[L_q^k] &= \left( q + (1-q)(-1)^k \right) \mathbb{E}[L_1^k], \\ \mathbb{E}[e^{tL_q}] &= \sum_{k \geq 0} \frac{m_k}{\Gamma(ka + 1)} \left( q + (1-q)(-1)^k \right) t^k. \end{aligned}$$

*Proof of Theorem 1.4.* Since  $L_1$  is solution of the fixed-point equation (1.5), using the same notations as in Theorem 1.1, by independence between the variables, we have for  $\mu_k := \mathbb{E}[W^k]$ ,

$$\begin{aligned} \mu_k &= \mathbb{E} \left[ \sum_{j=0}^k \binom{k}{j} (1-V)^{ja} V^{(k-j)a} (2\xi_p - 1)^j \left( L_1^{(2)} \right)^j \left( L_1^{(1)} \right)^{k-j} \right] \\ &= \sum_{j=0}^k \binom{k}{j} \mathbb{E} \left[ (1-V)^{ja} V^{(k-j)a} \right] \mathbb{E} \left[ (2\xi_p - 1)^j \right] \mu_j \mu_{k-j} \\ &= \sum_{j=0}^k \binom{k}{j} \frac{\Gamma(ja + 1) \Gamma((k-j)a + 1)}{\Gamma(ka + 2)} c_j \mu_j \mu_{k-j}, \end{aligned}$$

with  $c_j = 1$  for even  $j$  and  $c_j = a$  for odd  $j$ . Thus

$$\frac{ka - c_k}{ka + 1} \mu_k = \sum_{j=1}^{k-1} \binom{k}{j} \frac{\Gamma(ja + 1) \Gamma((k-j)a + 1)}{\Gamma(ka + 2)} c_j \mu_j \mu_{k-j}.$$

We introduce  $m_k$  such that  $\mu_k = \frac{k!}{\Gamma(ka+1)} m_k$ . We thus deduce that  $m_k$  satisfies the recursive equation:  $m_1 = 1$ , and for  $k \geq 2$ ,

$$(ka - c_k)k!m_k = \sum_{j=1}^{k-1} k!c_j m_j m_{k-j},$$

which implies (1.7). Theorem 1.4 is proved.  $\square$

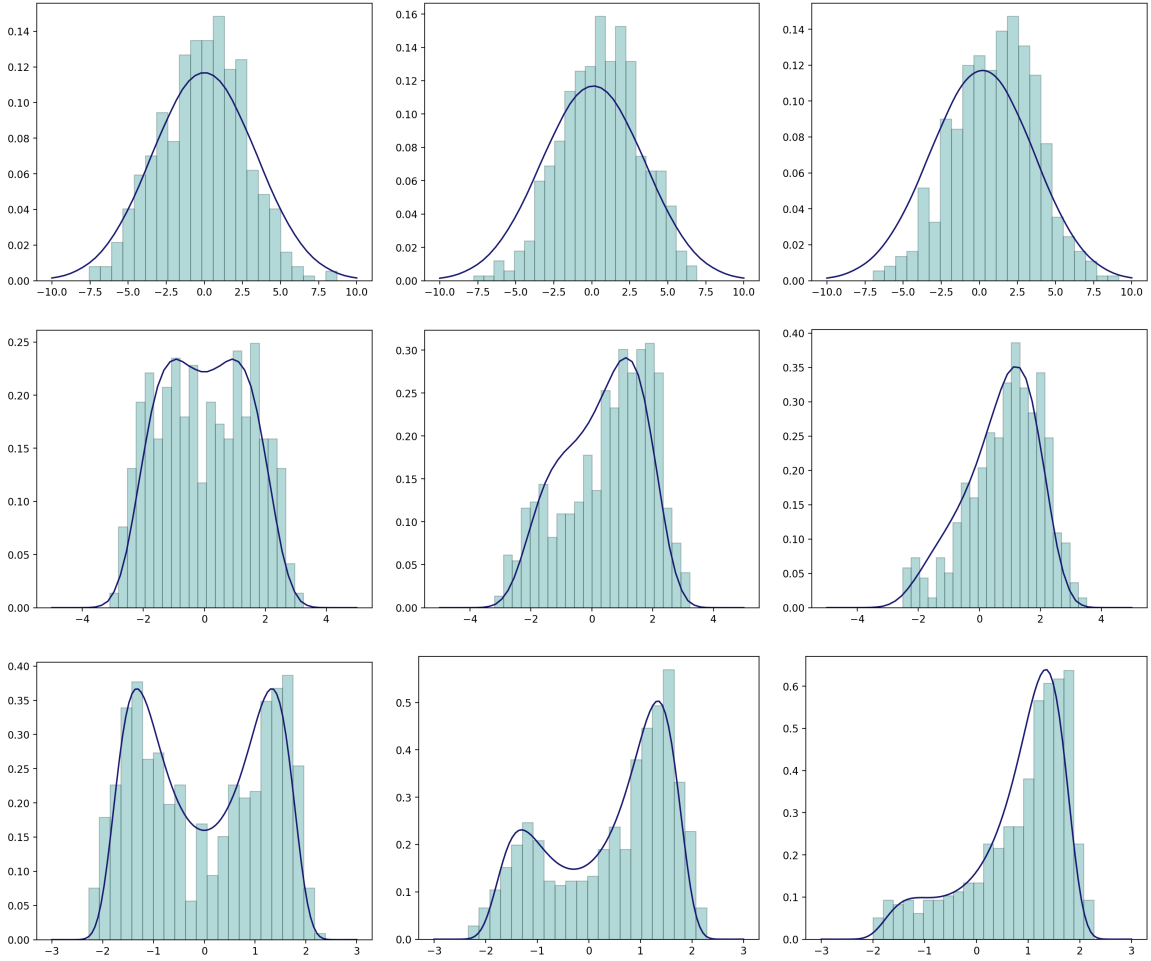


Figure 4: Histogram and method of moments approximation of the density (blue lines) of  $L_q$ . First (resp. second, third) line:  $p = 0.77$  (resp.  $p = 0.87$ ,  $p = 0.92$ ). First (resp. second, third) column:  $q = 0.5$  (resp.  $q = 0.7$ ,  $q = 0.9$ ).

## 2.5 Approximation of the density

The aim of this part is to propose a few approximations of the density of  $L_q$ , as illustrated in Figure 4.

Each display contains a histogram of the distribution of  $L_q$  as well as an approximation of the plot of the density of  $L_q$ . For each of the nine pictures, to obtain the histogram we have simulated 500 paths of ERW of length  $10^4$ .

In order to obtain the approximated curves of the density, we make use of a classical method to approximate numerically a measure knowing its moments, as presented in [29]. More specifically, given a finite number of moments (in our case, 300, quickly obtained from the recurrence (1.7)), we may introduce the so-called Hankel matrix, and then there is a canonical way to compute the associated Jacobi matrix. The eigenvalues and eigenvectors of this Jacobi matrix give an approximation of the density by a finite sum of Dirac measures. As a last step, we may consider a smooth (in our case, cubic) interpolation of this purely atomic measure, which yields the graphs of Figure 4. We kindly thank the authors of the paper [15], in particular Slim Kammoun, for providing us with an effective code.

## 2.6 A cluster decomposition of the limit variable and its relations with the computation of moments

In this part, we first state a remarkable decomposition of the limit variable  $L_q$ , which is obtained using the interpretation of ERW in terms of random recursive trees and results on cluster sizes after percolation. Although this result does not clearly appear in the literature, it follows directly from arguments in [2, 10].

**Proposition 2.15** ([2, 10]). *For  $q \in [0, 1]$  and  $a = 2p - 1 > \frac{1}{2}$ , the ERW asymptotic satisfies*

$$L_q = C_1 Z_1^{(q)} + \sum_{j=2}^{\infty} C_j Z_j \quad \text{a.s.} \quad (2.16)$$

where

- $C_1$  is a Mittag-Leffler random variable with parameter  $a$  and  $Z_1^{(q)}$  has Rademacher distribution with parameter  $q$ , independent of  $C_1$ ;
- $(Z_j)_{j \geq 2}$  are independent Rademacher random variables with parameter  $1/2$ , independent of  $Z_1^{(q)}$ ;
- $(C_j)_{j \geq 2}$  are dependent random variables such that for any  $j \geq 2$ ,  $C_j \stackrel{\mathcal{L}}{=} C_{1,j} (\beta_{\tau_j})^a$ , with
  - $C_{1,j}$  is a Mittag-Leffler random variable with parameter  $a$ ;
  - $(\beta_k)$  denote independent Beta random variables with parameters  $(1, k - 1)$ ;
  - $\tau_j - 1$  are dependent negative binomial random variables with parameters  $(j - 1, 1 - a)$ , such that  $\tau_j = \tau_{j-1} + G_j$  and  $G_j$  has geometric distribution of parameter  $1 - a$  and is independent of  $\tau_{j-1}$ ;
  - for each  $j \geq 2$ , the random variables  $C_{1,j}, \beta_{\tau_j}$  are independent.
- The random variables  $(Z_j)_{j \geq 1}$  are independent of the random variables  $(C_j)_{j \geq 1}$ .

*Proof.* The proof follows from compiling several results on the ERW and the cluster size of random recursive trees (RRT) with Bernoulli bond percolation, see [10] for the explicit construction and [8] for the definition of the ERW as a step-reinforced random walk. We give here the main arguments, but leave the details to the reader. Let  $\mathcal{T}_n$  be a RRT of size  $n$ , and denote by  $T_i$  the subtree of  $\mathcal{T}_n$  rooted at  $i$  after Bernoulli bond percolation (each edge is kept with probability  $a$ ). In particular, we have a forest after percolation. It is known from [2, Lem. 3.3], see also [10, Lem. 4 and Lem. 5], that

$$\lim_{n \rightarrow \infty} n^{-a} \#\{j \leq n : j \in T_1\} = C_1 \quad \text{a.s.},$$

where  $C_1$  has a Mittag-Leffler distribution with parameter  $a$ . More generally, for  $i \geq 1$ ,

$$\lim_{n \rightarrow \infty} n^{-a} \#\{j \leq n : j \in T_i\} = C_i > 0 \quad \text{a.s.},$$

where  $C_i$  has the same law as  $(\beta_i)^a \cdot C_{1,i}$ , where  $\beta_i$  denotes a Beta variable with parameter  $(1, i - 1)$  and  $C_{1,i}$  is a Mittag-Leffler variable with parameter  $a$  independent of  $\beta_i$ .

Hereafter, we use the connection with the ERW to write [34, 10]

$$S_n = \sum_{j=1}^{\infty} |c_n(\tau_j)| Z_j, \quad (2.17)$$

where

- $Z_1 \sim \mathcal{R}(q)$ ,  $(Z_j)_{j \geq 2}$  is a sequence of i.i.d. random variables with  $\mathcal{R}(1/2)$ -distribution;
- $c_n(j) = \{j \leq n : j \in T_i\}$  is the cluster rooted at  $j$  of the RRT associated with the ERW after percolation and  $|c_n(j)|$  is its size;
- $(\tau_j - 1)_{j \geq 1}$  are dependent negative binomial random variables with respective parameters  $(j-1, 1-a)$ , such that  $\tau_j = \tau_{j-1} + G_j$ , where  $G_j$  has geometric distribution of parameter  $1-a$  and is independent of  $\tau_{j-1}$ .

The  $(\tau_j)_{j \geq 1}$  are here to ensure that we do not count more than once the steps of the ERW (i.e., that we are looking that the roots of the trees after percolation or at the  $j$ -th tree of the forest). Gathering all of the above, we find that when  $a > 1/2$  both  $S_n/n^a$  and  $|c_n(\tau_j)|/n^a$  converge (for all  $j$ ) and the result is proved.  $\square$

An application of the previous result is that the cluster decomposition (2.16) allows us to obtain a new proof of the recursion for the moments obtained in Theorem 1.4. More precisely:

**Remark 2.16.** *Even though each  $C_j$  in (2.16) can be written in distribution as a product of a Mittag-Leffler variable of parameter  $a$  with an independent variable, and even though Equation (1.8) satisfied by the moments of  $L_q$  suggests a similar factorization, the asymptotic variable  $L_q$  cannot be written as the product of a Mittag-Leffler variable of parameter  $a$  with an independent variable having moments  $(m_k)_{k \geq 0}$ . The reason is that the sequence of  $(m_k)_{k \geq 0}$ , defined by the recursive equation (1.7), does not correspond to the sequence of moments of a random variable, since the associated Hankel matrix is not positive. The factorization of the variables  $C_j$  is just an equality in law, and the choice of  $C_{1,j}$  depends on  $j$ .*

### 3 The multidimensional Elephant random walk

#### 3.1 Definition and first properties of the process

Let  $d \geq 1$ . All vectors  $\mathbf{w} = (w_j)_{1 \leq j \leq d}$  in this section are seen as column vectors. For vectors  $\mathbf{w}, \mathbf{w}' \in \mathbb{R}^d$ ,  $\mathbf{w}^T$  stands for the transpose vector of  $\mathbf{w}$ ,  $\langle \mathbf{w}, \mathbf{w}' \rangle$  denotes the scalar product and  $\|\mathbf{w}\|$  is the Euclidean norm.

The multidimensional ERW (MERW)  $(\mathbf{S}_n)_{n \geq 0}$  on  $\mathbb{Z}^d$  is defined as follows. Let  $(\mathbf{e}_i)_{1 \leq i \leq d}$  be the canonical basis of  $\mathbb{R}^d$ . At time zero,  $\mathbf{S}_0 = 0$  and at time  $n = 1$ , the elephant moves in one of the  $2d$  directions  $(\pm \mathbf{e}_i)_{1 \leq i \leq d}$  according to a probability vector  $q = (q_1, \dots, q_{2d})$ , meaning that

$$\mathbb{P}(\mathbf{S}_1 = \mathbf{e}_k) = q_{2k-1} \quad \text{and} \quad \mathbb{P}(\mathbf{S}_1 = -\mathbf{e}_k) = q_{2k}, \quad \forall k \in \{1, \dots, d\}. \quad (3.1)$$

Afterwards, at time  $n+1 \geq 2$ , the elephant chooses uniformly at random an integer  $k$  among the previous times  $1, \dots, n$ . Then, it repeats exactly its  $k$ -th step with probability  $p$ , or it moves uniformly in one of the  $2d-1$  remaining directions, i.e., with the same probability  $\frac{1-p}{2d-1}$ , where the parameter  $p$  stands for the memory parameter of the MERW. The position of the elephant at time  $n+1$  is given by

$$\mathbf{S}_{n+1} = \mathbf{S}_n + \mathbf{X}_{n+1},$$

with  $\mathbf{X}_{n+1}$  being defined as the step of this random walk at time  $n+1$ , and satisfying

$$\mathbf{X}_{n+1} = A_{n+1} \mathbf{X}_{\mathcal{U}(n)}.$$

Here  $\mathcal{U}(n)$  is a uniform variable on  $\{1, \dots, n\}$ , and  $(A_n)_{n \geq 1}$  represents a sequence of i.i.d. random matrices, independent of  $\mathcal{U}(n)$ , given by

$$\mathbb{P}(A_n = +I_d) = p,$$

and, for  $k \in \{1, \dots, d-1\}$ ,

$$\mathbb{P}(A_n = -I_d) = \mathbb{P}(A_n = +(J_d)^k) = \mathbb{P}(A_n = -(J_d)^k) = \frac{1-p}{2d-1},$$

where  $I_d$  is the identity matrix of dimension  $d$  and the matrix  $J_d := [\mathbf{e}_2, \dots, \mathbf{e}_d, \mathbf{e}_1]$ . We observe that  $(J_d)^d = I_d$ .

Define in this section

$$a := \frac{2dp-1}{2d-1}. \quad (3.2)$$

In the superdiffusive regime (characterized, as we shall see below, by  $a \in (\frac{1}{2}, 1)$ ), Bercu and Laulin proved the convergence of the normalized MERW.

**Theorem 3.1** (Theorem 3.7 in [6]). *Let  $(\mathbf{S}_n)_{n \geq 0}$  be the  $d$ -dimensional ERW with memory parameter  $p \in [0, 1]$  and initial probability vector  $q \in [0, 1]^{2d}$  as in (3.1). When  $a > 1/2$ , there exists a non-degenerate random vector  $\mathbf{L}$  in  $\mathbb{R}^d$  such that*

$$\lim_{n \rightarrow \infty} \frac{\mathbf{S}_n}{n^a} = \mathbf{L} \quad \text{a.s.} \quad (3.3)$$

We now give some properties of the random vector  $\mathbf{L}$ . From [6, Rem. 3.5] giving the first moments of the MERW, we deduce the following lemma. Notice that we recover the result of [6, Thm 3.8] in the case  $q$  is the uniform distribution on the  $2d$  directions.

**Lemma 3.2.** *Under the same hypotheses as in Theorem 3.1, the limit  $\mathbf{L}$  in (3.3) satisfies*

$$\mathbb{E}[\mathbf{L}] = \frac{1}{\Gamma(a+1)} \begin{pmatrix} q_1 - q_2 \\ q_3 - q_4 \\ \vdots \\ q_{2d-1} - q_{2d} \end{pmatrix},$$

$$\mathbb{E}[\mathbf{L}\mathbf{L}^T] = \frac{1}{\Gamma(2a+1)} D + \frac{1-a}{da(2a-1)\Gamma(2a)} I_d,$$

where  $D$  is the diagonal matrix  $D = \text{diag}(q_1 + q_2, q_3 + q_4, \dots, q_{2d-1} + q_{2d})$ .

*Proof.* When the first step probability vector  $q$  is  $(1, 0, \dots, 0)$ , Remark 3.5 in [6] states that

$$\mathbb{E}[\mathbf{L}] = \frac{1}{\Gamma(a+1)} \mathbf{e}_1,$$

$$\mathbb{E}[\mathbf{L}\mathbf{L}^T] = \frac{1}{\Gamma(2a+1)} \left( \mathbf{e}_1 \mathbf{e}_1^T - \frac{1}{d} I_d \right) + \frac{1}{d(2a-1)\Gamma(2a)} I_d$$

$$= \frac{1}{\Gamma(2a+1)} \mathbf{e}_1 \mathbf{e}_1^T + \frac{1-a}{da(2a-1)\Gamma(2a)} I_d.$$

We have similar results when  $q$  satisfies  $q_k = 1$  and  $q_i = 0$  for  $i \neq k$ , for any  $k \in \{1, \dots, 2d\}$ . Let us now consider a general distribution  $q = (q_1, \dots, q_{2d})$  for the first step. The limit  $\mathbf{L}$  starting from the probability vector  $q$  can be written as the mixture of random vectors

$$\mathbf{L} = \sum_{k=1}^{2d} \mathbf{L}^{(k)} \mathbb{1}_{\bar{q}_{k-1} \leq U < \bar{q}_k}, \quad (3.4)$$

where  $\mathbf{L}_{(k)}$  the limit of the MERW starting from a first step in the direction  $(-1)^{k-1}\mathbf{e}_{\lfloor \frac{k-1}{2} \rfloor + 1}$ ,  $U$  an independent uniform variable on  $[0, 1]$ ,  $\bar{q}_0 = 0$  and  $\bar{q}_k = \sum_{i=1}^k q_i$ . Since

$$\begin{aligned}\mathbb{E}[\mathbf{L}_{(k)}] &= \frac{(-1)^{k-1}}{\Gamma(a+1)} \mathbf{e}_{\lfloor \frac{k-1}{2} \rfloor + 1} \\ \mathbb{E}[\mathbf{L}_{(k)} \mathbf{L}_{(k)}^T] &= \frac{1}{\Gamma(2a+1)} \mathbf{e}_{\lfloor \frac{k-1}{2} \rfloor + 1} \mathbf{e}_{\lfloor \frac{k-1}{2} \rfloor + 1}^T + \frac{1-a}{da(2a-1)\Gamma(2a)} I_d,\end{aligned}$$

the conclusion follows.  $\square$

Let  $\mathbf{Z}_1^{(q)}$  be a random variable with distribution  $q$  on the  $2d$  directions  $(\pm \mathbf{e}_i)_{1 \leq i \leq d}$  and  $(\mathbf{Z}_j)_{j \geq 2}$  be a sequence of i.i.d. random variables with uniform distribution on the directions, independent of  $\mathbf{Z}_1^{(q)}$ . As in dimension 1 (see Proposition 2.15), the superdiffusive limit  $\mathbf{L}$  of MERW can be written

$$\mathbf{L} = C_1 \mathbf{Z}_1^{(q)} + \sum_{j=2}^{\infty} C_j \mathbf{Z}_j, \quad (3.5)$$

where the variables  $(C_j)_{j \geq 1}$  are defined in Proposition 2.15 and are independent of  $\mathbf{Z}_1^{(q)}$  and  $(\mathbf{Z}_j)_{j \geq 2}$ .

The following lemma shows that the coordinates of the vector  $\mathbf{L}$  are not independent, even though its covariance matrix is diagonal by Lemma 3.2. As an important consequence for us, in order to prove that the vector  $\mathbf{L}$  has a density it is not sufficient to show that each coordinate has a density.

**Lemma 3.3.** *We assume that  $q$  is the uniform distribution on the  $2d$  directions and  $a > \frac{1}{2}$ . Then the coordinates of  $\mathbf{L} = (L_i)_{1 \leq i \leq d}$  have the same distribution and*

$$\sum_{i=1}^d L_i = \sum_{j=1}^{\infty} C_j \zeta_j \stackrel{\mathcal{L}}{=} L_{1/2},$$

where

- $(C_j)_{j \geq 1}$  are defined as in Proposition 2.15;
- $(\zeta_j)_{j \geq 1}$  are i.i.d. Rademacher  $\mathcal{R}(1/2)$  random variables, independent of the above variables;
- $L_{1/2}$  is the asymptotic variable in dimension 1 of ERW with first step probability  $1/2$ .

*Proof.* When  $q$  is the uniform distribution, for each  $j \geq 1$ , the random vectors  $\mathbf{Z}_j$  in (3.5) can be written  $\mathbf{Z}_j = \zeta_j \mathbf{M}_j$ , where  $\zeta_j$  is a Rademacher  $\mathcal{R}(1/2)$  random variable and  $\mathbf{M}_j$  is a multinomial  $\mathcal{M}(1, (\frac{1}{2d}, \dots, \frac{1}{2d}))$  random vector of  $\mathbb{R}^d$ , with  $\zeta_j$  and  $\mathbf{M}_j$  independent.

We easily deduce that each coordinates of  $\mathbf{L}$  have the same distribution. Moreover, the sum of the coordinates of each  $\mathbf{M}_j$  being equal to 1, the lemma is proved.  $\square$

We deduce from this lemma that the coordinates of  $\mathbf{L}$  are not independent. Indeed, when  $d = 2$ , for each  $j$  we have  $\mathbf{Z}_j = \zeta_j (B_j, 1 - B_j)^T$ , where  $\zeta_j$  is a Rademacher  $\mathcal{R}(1/2)$  random variable,  $B_j$  a Bernoulli  $\mathcal{B}(1/2)$  random variable, and both are independent. Consequently, for  $d = 2$  and  $q = (1/4, 1/4, 1/4, 1/4)$ ,

$$\mathbf{L} = \sum_{j=1}^{\infty} C_j \zeta_j \begin{pmatrix} B_j \\ 1 - B_j \end{pmatrix} \stackrel{\mathcal{L}}{=} \begin{pmatrix} R \\ L_{1/2} - R \end{pmatrix},$$

where  $L_{1/2}$  is the limiting random variable in dimension 1 and  $R = \sum_{j=1}^{\infty} C_j \zeta_j B_j$  is a priori not independent of  $L_{1/2}$ .

### 3.2 Urn process associated to the MERW

We now work with the space  $\mathbb{R}^{2d}$  endowed with the canonical basis  $(\mathbf{e}_i)_{1 \leq i \leq 2d}$ . Similarly as in the dimension 1 case, we associate to the MERW a Pólya-type urn process  $\mathbf{U} = (\mathbf{U}(n))_{n \geq 0}$  with  $2d$  different colors (representing the  $2d$  possible directions of the MERW), with the following replacement matrix:

$$A = \begin{cases} I_{2d} & \text{with probability } p, \\ (J_{2d})^k & \text{with probability } \frac{1-p}{2d-1} \text{ for each } k \in \{1, \dots, 2d-1\}, \end{cases} \quad (3.6)$$

where  $I_{2d}$  is the identity matrix in dimension  $2d$  and the matrix  $J_{2d} := [e_2, e_3, \dots, e_{2d}, e_1]$ , see [7]. We remark that  $(J_{2d})^{2d} = I_{2d}$  and

$$\mathbb{E}[A] = pI + \frac{1-p}{2d-1} \begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ 1 & 1 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 \\ 1 & 1 & \dots & 1 & 0 \end{pmatrix}.$$

The matrix  $\mathbb{E}[A]$  is diagonalizable and admits the eigenvalues

$$\lambda_1 = 1, \quad \lambda_2 = \lambda_3 = \dots = \lambda_{2d} = \frac{2dp-1}{2d-1},$$

with respective unit eigenvectors  $\mathbf{v}_1 = \frac{1}{2d} \sum_{i=1}^{2d} \mathbf{e}_i$ ,  $\mathbf{v}_2 = \frac{1}{2}(\mathbf{e}_1 - \mathbf{e}_2)$ ,  $\mathbf{v}_3 = \frac{1}{2}(\mathbf{e}_1 - \mathbf{e}_3)$ , up to  $\mathbf{v}_{2d} = \frac{1}{2}(\mathbf{e}_1 - \mathbf{e}_{2d})$ . The following result may be found in Theorems 3.9, 3.24 and 3.26 of [32].

**Theorem 3.4** ([32]). *Let  $a = \frac{2dp-1}{2d-1}$ . We assume  $a > 1/2$ . Assume further that  $\mathbf{U}(0) = \mathbf{e}_k$  for some  $k \in \{1, \dots, 2d\}$ , i.e., the urn process starts from a unique ball of color  $k$ . Then we have*

$$\lim_{n \rightarrow \infty} \frac{\mathbf{U}(n) - n\mathbf{v}_1}{n^a} = \sum_{i=2}^{2d} W_i(\mathbf{e}_k) \mathbf{v}_i \quad \text{a.s.}, \quad (3.7)$$

where  $W_i(\mathbf{e}_k)$  are real-valued random variables depending on  $k$ .

Moreover, the expectation of  $\mathbf{Y}_{(k)} := \sum_{i=2}^{2d} W_i(\mathbf{e}_k) \mathbf{v}_i$  is given by

$$\mathbb{E}[\mathbf{Y}_{(k)}] = \frac{1}{\Gamma(a+1)} (\mathbf{e}_k - \mathbf{v}_1). \quad (3.8)$$

We denote by  $\mathbf{U}_{(k)}$  the urn process starting from one ball of the  $k$ -th color. We adapt the idea used for an urn with two colors and divide the tree structure at time 1 into two subtrees: the first starting from a node of color  $k$ , and the second starting from a node with color  $k$  with probability  $p$ , or from another color (randomly chosen) with probability  $1-p$ . For  $i \in \{1, 2\}$ ,  $n \geq 1$ ,  $D_i(n)$  denotes the number of leaves at time  $n$  of the  $i$ -th subtree. We recall that these numbers represent the time inside each subtree. Following the ideas of Section 2.2, we deduce the lemma below.

**Lemma 3.5.** *For any  $n \geq 1$ ,*

$$\mathbf{U}_{(k)}(n) \stackrel{\mathcal{L}}{=} \mathbf{U}_{(k)}^{(1)}(D_1(n) - 1) + \sum_{j=1}^{2d} \zeta_{j,(k)} \mathbf{U}_{(j)}^{(2)}(D_2(n) - 1), \quad (3.9)$$

where  $\zeta_{(k)} = (\zeta_{j,(k)})_{1 \leq j \leq 2d}$  denotes a random vector such that  $\zeta_{(k)} = \mathbf{e}_k$  with probability  $p$  and  $\zeta_{(k)} = \mathbf{e}_{k'}$  with probability  $\frac{1-p}{2d}$ , for each  $k' \neq k$ .



Using the above almost sure convergence (2.4) and Theorem 3.4, this leads to the system of equations

$$\mathbf{Y}_{(k)} \stackrel{\mathcal{L}}{=} V^a \mathbf{Y}_{(k)}^{(1)} + (1-V)^a \sum_{j=1}^{2d} \zeta_{j,(k)} \mathbf{Y}_{(j)}^{(2)}, \quad k \in \{1, \dots, 2d\}, \quad (3.10)$$

where  $\mathbf{Y}_{(k)}$  is the limit vector of the urn process started with one ball of the  $k$ -th color. In Equation (3.10),  $V$  is a uniform random variable on the segment  $[0, 1]$ ,  $\mathbf{Y}_{(k)}^{(1)}$  and  $\mathbf{Y}_{(k)}^{(2)}$  are copies of  $\mathbf{Y}_{(k)}$ , and all random variables are independent.

### 3.3 Link between the asymptotics of the urn process and the MERW

We consider the urn process  $\mathbf{U}$  defined as above, starting from  $\mathbf{U}(0)$  chosen according to a probability vector  $q$ . We can easily check that the  $d$ -dimensional process  $(\mathbf{S}_n)_{n \geq 0}$  defined for all  $n \geq 1$  by

$$\begin{cases} S_{1,n} = U_1(n) - U_2(n) & = 2\mathbf{v}_2^T \mathbf{U}_n, \\ S_{2,n} = U_3(n) - U_4(n) & = 2(\mathbf{v}_4 - \mathbf{v}_3)^T \mathbf{U}_n, \\ \vdots & \vdots \\ S_{d,n} = U_{2d-1}(n) - U_{2d}(n) & = 2(\mathbf{v}_{2d} - \mathbf{v}_{2d-1})^T \mathbf{U}_n, \end{cases}$$

is a MERW with memory parameter  $p$  and initial distribution  $q$ .

We denote by  $\mathbf{L}_{(k)}$  the limit random vector of the MERW when the first step is in the  $k$ -th direction. When the initial distribution is given by a general probability vector  $q$ , the limit  $\mathbf{L}$  of the MERW can be written as a discrete mixture of  $(\mathbf{L}_{(1)}, \dots, \mathbf{L}_{(2d)})$ , see (3.4).

The almost sure convergences (3.3) and (3.7) imply

$$\begin{aligned} \mathbf{L}_{(k)} &= \lim_{n \rightarrow \infty} \frac{1}{n^a} \mathbf{S}_n = \lim_{n \rightarrow \infty} \frac{1}{n^a} \begin{pmatrix} U_1(n) - U_2(n) \\ U_3(n) - U_4(n) \\ \vdots \\ U_{2d-1}(n) - U_{2d}(n) \end{pmatrix} = \begin{pmatrix} Y_{1,(k)} - Y_{2,(k)} \\ Y_{3,(k)} - Y_{4,(k)} \\ \vdots \\ Y_{2d-1,(k)} - Y_{2d,(k)} \end{pmatrix} \\ &= \begin{pmatrix} W_2(\mathbf{e}_k) + \frac{1}{2} \sum_{k=3}^{2d} W_k(\mathbf{e}_k) \\ \frac{W_4(\mathbf{e}_k) - W_3(\mathbf{e}_k)}{2} \\ \vdots \\ \frac{W_{2d}(\mathbf{e}_k) - W_{2d-1}(\mathbf{e}_k)}{2} \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \cdots & \cdots & \frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \mathbf{W}(\mathbf{e}_k) \quad \text{a.s.} \end{aligned} \quad (3.11)$$

We notice that this last matrix of size  $d \times (2d - 1)$  is of rank  $d$ . Thus, if we prove that  $\mathbf{W} := (W_2, \dots, W_{2d})$  has a smooth positive density function on  $\mathbb{R}^{2d-1}$ , as what we have done in dimension 1, we will immediately deduce that the distribution of  $\mathbf{L}$  admits a smooth positive density function on  $\mathbb{R}^d$ . Consequently, we will focus our study on the random vectors  $\mathbf{W}$ .

### 3.4 Uniqueness and fixed-point equations in high dimension

We first study the system (3.10) of fixed-point equations satisfied by the limit process, and in Proposition 3.6 we prove existence and uniqueness of the associated solution in the space of random vectors with a finite second-order moment and a given expectation. Then in Proposition 3.7 we derive a second fixed-point equation, which will be the key-point for studying the distribution of the limit process.

**Existence and uniqueness of the solution to the system (3.10)**

Equation (3.10) ensures that, for all  $j, k \in \{1, \dots, 2d\}$

$$Y_{j,(k)} = V^a Y_{j,(k)}^{(1)} + (1 - V)^a \sum_{i=1}^{2d-1} A_{i,j} Y_{j,(i)}^{(2)}, \quad (3.12)$$

where  $A = (A_{i,j})$  has the same law as the random replacement matrix (3.6) of the urn process and  $Y_{j,(k)}$  is the  $j$ -th coordinate of the limiting vector  $\mathbf{Y}_{(k)}$  starting with one ball of the  $k$ -th color. Hence, for  $\mathbf{Y}^j := (Y_{j,(k)})_{1 \leq k \leq 2d}$ ,

$$\mathbf{Y}^j = V^a \mathbf{Y}^{j,(1)} + (1 - V)^a A \mathbf{Y}^{j,(2)}, \quad (3.13)$$

with  $\mathbf{Y}^{j,(1)}$  and  $\mathbf{Y}^{j,(2)}$  independent copies of  $\mathbf{Y}^j$ , and the above holds jointly for all  $j \in \{1, \dots, 2d\}$  with the same  $V$  and  $A$ .

**Proposition 3.6.** *Let  $a \in (1/2, 1)$ . Given  $\mathbf{m} = (m_2, \dots, m_{2d}) \in \mathbb{R}^{2d-1}$ , we define  $\mathcal{P}_2(\mathbf{m})$  as the space of probability measures on  $\mathbb{R}^{2d}$  with finite second-order moment and such that for all  $\mu \in \mathcal{P}_2(\mathbf{m})$ ,*

$$\int_{\mathbb{R}^{2d}} x \mu(dx) = \sum_{i=2}^{2d} m_i \mathbf{v}_i.$$

*Then there exists a unique solution in  $\mathcal{P}_2(\mathbf{m})$  to the fixed-point equation (3.13), for all  $j \in \{1, \dots, 2d\}$ .*

Thanks to the convergence Theorem 3.4, we know that there is a (not necessarily unique) solution to (3.10) when  $a > 1/2$ , and its expectation is given by (3.8).

*Proof.* As in the proof of Theorem 1.2, we consider the Wasserstein distance  $d_{\mathcal{W}}$  on  $\mathcal{P}_2(m)$ . Then we define

$$H : \mathcal{P}_2(\mathbf{m}) \longrightarrow \mathcal{P}_2(\mathbf{m}) \\ \mu \longmapsto \mathcal{L}\left(V^a \mathbf{Y}^{(1)} + (1 - V)^a A \mathbf{Y}^{(2)}\right),$$

where  $\mathbf{Y}^{(1)}, \mathbf{Y}^{(2)}$  have law  $\mu$ ,  $V$  is a uniform random variable on  $[0, 1]$ , the random matrix  $A$  is the same matrix (3.6) as in the urn process, and all the variables are independent from each other. As  $\mathbb{E}[V^a] = \frac{1}{1+a}$ , and for all  $i \in \{2, \dots, 2d\}$ ,  $\mathbb{E}[A] \mathbf{v}_i = a \mathbf{v}_i$ , we easily deduce that for all  $\mu \in \mathcal{P}_2(\mathbf{m})$ ,  $H(\mu) \in \mathcal{P}_2(\mathbf{m})$ .

We now show that  $H$  is Lipschitz-continuous for the Wasserstein metric. Let  $\mu, \nu \in \mathcal{P}_2(\mathbf{m})$ , and  $(\mathbf{Y}^{(1)}, \tilde{\mathbf{Y}}^{(1)})$  and  $(\mathbf{Y}^{(2)}, \tilde{\mathbf{Y}}^{(2)})$  two independent couples with marginal distributions  $(\mu, \nu)$ . We then compute

$$\begin{aligned} & d_{\mathcal{W}}(H\mu, H\nu)^2 \\ & \leq \mathbb{E} \left[ \left\| V^a (\mathbf{Y}^{(1)} - \tilde{\mathbf{Y}}^{(1)}) + (1 - V)^a A (\mathbf{Y}^{(2)} - \tilde{\mathbf{Y}}^{(2)}) \right\|^2 \right] \\ & \leq \mathbb{E}[V^{2a}] \mathbb{E} \left[ \left\| \mathbf{Y}^{(1)} - \tilde{\mathbf{Y}}^{(1)} \right\|^2 \right] + \mathbb{E}[(1 - V)^{2a}] \mathbb{E} \left[ \left\| A (\mathbf{Y}^{(2)} - \tilde{\mathbf{Y}}^{(2)}) \right\|^2 \right] \\ & \leq \frac{2}{1 + 2a} \mathbb{E} \left[ \left\| \mathbf{Y}^{(1)} - \tilde{\mathbf{Y}}^{(1)} \right\|^2 \right], \end{aligned}$$

because  $\|A\mathbf{Y}\| = \|\mathbf{Y}\|$  for all  $\mathbf{Y} \in \mathbb{R}^{2d}$ . Consequently, taking the infimum over all couples  $(\mathbf{Y}^{(1)}, \tilde{\mathbf{Y}}^{(1)})$  with marginal distributions  $(\mu, \nu)$ , we deduce

$$d_{\mathcal{W}}(H\mu, H\nu) \leq \sqrt{\frac{2}{1+2a}} d_{\mathcal{W}}(\mu, \nu)$$

and  $H$  is a contraction. The conclusion follows.  $\square$

Equation (3.13) is not a fully convenient (fixed-point) equation to study the distributions of  $\mathbf{W}$  and  $\mathbf{L}$ , as we cannot easily relate  $\mathbf{Y}^j$  to the  $\mathbf{L}_{(k)}$ 's. Besides, for any  $k \geq 1$ , by definition of  $\mathbf{Y}_{(k)}$  in Theorem 3.4 and by definition of the eigenvectors  $(\mathbf{v}_i)_{1 \leq i \leq 2d}$ , we easily observe that

$$2\mathbf{Y}_{(k)} = \left( \sum_{i=2}^{2d} W_i(\mathbf{e}_k) \right) \mathbf{e}_1 - \sum_{i=2}^{2d} W_i(\mathbf{e}_k) \mathbf{e}_i. \quad (3.14)$$

The support of  $\mathbf{Y}_{(k)}$  is thus included in the hyperplane  $\{x \in \mathbb{R}^{2d} : \sum_{i=1}^{2d} x_i = 0\}$ , and  $\mathbf{Y}_{(k)}$  cannot have a density with respect to the Lebesgue measure on  $\mathbb{R}^{2d}$ . However, the fixed-point equation (3.13) will be useful to study the moments of  $\mathbf{Y}$  and  $\mathbf{L}$  in Section 3.8.

In the next paragraph, we introduce another fixed-point equation, which will be the key-point to show that  $\mathbf{W}$  and  $\mathbf{L}$  have a positive density with respect to the Lebesgue measure.

### A second fixed-point equation in high dimension

We now focus on the study on  $\mathbf{U}_{(1)} = (\mathbf{U}_{(1)}(n))_{n \geq 0}$ , i.e., the urn process starting from a unique ball of color 1:  $\mathbf{U}_{(1)}(0) = \mathbf{e}_1$ . The other cases are similar by symmetry of the model. Indeed, we notice that the urn process starting from a unique ball of any color  $k \in \{1, \dots, 2d\}$  is equal to

$$\left( (J_{2d})^{k-1} \mathbf{U}_{(1)}(n) \right)_{n \geq 0}.$$

More generally, for  $i, k \in \{1, \dots, 2d\}$ , we have  $\mathbf{U}_{(i)}(n) = J_{2d}^{i-k} \mathbf{U}_{(k)}(n)$ . For clarity of the proofs, we will omit the subscript in the sequel, and we will denote  $\mathbf{U} := \mathbf{U}_{(1)}$ ,  $\mathbf{Y} := \mathbf{Y}_{(1)}$ , and for  $i \in \{2, \dots, 2d\}$ ,  $W_i := W_i(\mathbf{e}_1)$ .

We notice that Equation (3.10) leads to

$$\mathbf{Y} \stackrel{\mathcal{L}}{=} V^a \mathbf{Y}^{(1)} + (1-V)^a \sum_{j=1}^{2d} \zeta_j J^{j-1} \mathbf{Y}^{(2)}, \quad (3.15)$$

where  $\mathbf{Y}$  is the limit vector of the urn process started with one ball of color 1 (see Theorem 3.4), and  $\zeta := \zeta_{(1)}$  is defined in Lemma 3.5.

We now give a useful fixed-point equation in dimension  $2d-1$ . Define  $K$  as the following variant of a companion matrix of dimension  $(2d-1) \times (2d-1)$ :

$$K = \begin{pmatrix} -1 & -1 & \cdots & \cdots & -1 \\ 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}. \quad (3.16)$$

Observe that  $K$  is a cyclic matrix:  $K^{2d} = I_{2d-1}$ .

**Proposition 3.7.** *The limit vector  $\mathbf{W} = (W_2, \dots, W_{2d}) \in \mathbb{R}^{2d-1}$ , defined in Theorem 3.4 for  $\mathbf{U}(0) = \mathbf{e}_1$ , satisfies the fixed-point equation*

$$\mathbf{W} \stackrel{\mathcal{L}}{=} V^a \mathbf{W}^{(1)} + (1 - V)^a B \mathbf{W}^{(2)}, \quad (3.17)$$

where

- $V$  denotes a uniform random variable on  $[0, 1]$ ;
- $\mathbf{W}^{(1)}$  and  $\mathbf{W}^{(2)}$  are copies of  $\mathbf{W}$ ;
- $B$  is a random matrix equal to  $B = K^k$  with probability  $p_k$  for each  $k \in \{0, \dots, 2d - 1\}$ , where  $p_0 = p$  and  $p_1 = \dots = p_{2d-1} = \frac{1-p}{2d-1}$ ;
- all the variables  $V$ ,  $\mathbf{W}^{(1)}$ ,  $\mathbf{W}^{(2)}$  and  $B$  are assumed independent.

*Proof.* By (3.14), we note that  $\mathbf{W} = (W_2, \dots, W_{2d}) = -2(Y_2, \dots, Y_{2d})$ . Then, we deduce from (3.15) that  $\mathbf{W}$  satisfies the following fixed-point equation

$$\mathbf{W} \stackrel{\mathcal{L}}{=} V^a \mathbf{W}^{(1)} + (1 - V)^a \widetilde{\mathbf{W}}^{(2)},$$

where  $\widetilde{\mathbf{W}}^{(2)}$  is the random vector

$$\begin{aligned} \widetilde{\mathbf{W}}^{(2)} &= \begin{pmatrix} (\zeta_1 - \zeta_2)W_2^{(2)} & + & (\zeta_3 - \zeta_2)W_{2d}^{(2)} & + & \dots & + & (\zeta_{2d} - \zeta_2)W_3 \\ (\zeta_1 - \zeta_3)W_3^{(2)} & + & (\zeta_2 - \zeta_3)W_2^{(2)} & + & \dots & + & (\zeta_{2d} - \zeta_3)W_4^{(2)} \\ \vdots & & \vdots & & & & \vdots \\ (\zeta_1 - \zeta_{2d})W_{2d}^{(2)} & + & (\zeta_2 - \zeta_{2d})W_{2d-1}^{(2)} & + & \dots & + & (\zeta_{2d-1} - \zeta_{2d})W_2^{(2)} \end{pmatrix} \\ &= \begin{pmatrix} \zeta_1 - \zeta_2 & \zeta_{2d} - \zeta_2 & \dots & \zeta_4 - \zeta_2 & \zeta_3 - \zeta_2 \\ \zeta_2 - \zeta_3 & \zeta_1 - \zeta_3 & \dots & \zeta_5 - \zeta_3 & \zeta_4 - \zeta_3 \\ & \vdots & & \vdots & \\ \zeta_{2d-1} - \zeta_{2d} & \zeta_{2d-2} - \zeta_{2d} & \dots & \zeta_2 - \zeta_{2d} & \zeta_1 - \zeta_{2d} \end{pmatrix} \mathbf{W}^{(2)}. \end{aligned}$$

The conclusion follows by computing the distribution of latest matrix from the distribution of  $\zeta = \zeta_{(1)}$ , defined in Lemma 3.5.  $\square$

### 3.5 Existence of a density for $\mathbf{W}$

In this section, we assume that

$$\text{Supp}(\mathbf{W}) = \mathbb{R}^{2d-1}. \quad (3.18)$$

The support will be carefully studied in Section 3.6. We follow the same scheme of proof as in dimension 1. Consequently, our aim is to study the integrability of the characteristic function of  $\mathbf{W}$ , defined on  $\mathbb{R}^{2d-1}$  by

$$\varphi_{\mathbf{W}}(t) := \mathbb{E} \left[ e^{i \langle t, \mathbf{W} \rangle} \right].$$

**Theorem 3.8.** *Assuming (3.18), let  $\mathbf{W}$  be a non-zero solution of (3.17) with a general probability vector  $(p_k)_{0 \leq k \leq 2d-1}$  satisfying  $p_k > 0$ , for all  $k \in \{0, \dots, 2d - 1\}$ . Then*

- for any  $\rho \in (0, 1/a)$ , and any  $k \geq 1$ , there exists  $C > 0$  such that for any  $t \in \mathbb{R}^{2d-1}$ ,

$$|\varphi_{\mathbf{W}}(t)| \leq \frac{C}{\|t\|^{k\rho}}; \quad (3.19)$$

- the distribution of  $\mathbf{W}$  has a bounded smooth density function with respect to the Lebesgue measure on  $\mathbb{R}^{2d-1}$ .

To prove Theorem 3.8, we adapt the arguments of the proof of Theorem 2.5 to a higher dimension. The proof is divided in three different lemmas.

**Lemma 3.9.** *Let  $\mathbf{W}$  be a solution of the fixed-point equation (3.17) satisfying the assumptions of Theorem 3.8. Its characteristic function satisfies  $|\varphi_{\mathbf{W}}(t)| < 1$  for all non-zero  $t \in \mathbb{R}^{2d-1}$ .*

*Proof.* Using our assumption (3.18) on the support of  $\mathbf{W}$ , we remark that for  $t_0 \neq 0$ ,  $\text{Supp}(\langle t_0, \mathbf{W} \rangle) = \mathbb{R}$ . Consequently, the proof is exactly the same as in dimension 1 (see Step 1 in the proof of Theorem 2.5).  $\square$

**Lemma 3.10.** *Let  $\mathbf{W}$  be a solution of the fixed-point equation (3.17) satisfying the assumptions of Theorem 3.8. Then  $\lim_{\|t\| \rightarrow \infty} \varphi_{\mathbf{W}}(t) = 0$ .*

*Proof.* Let  $t \in \mathbb{R}^{2d-1}$ . By (3.17) and by conditioning on  $V$ , we remark that

$$\varphi_{\mathbf{W}}(t) = p_0 \mathbb{E}[\varphi_{\mathbf{W}}(V^a t) \varphi_{\mathbf{W}}((1-V)^a t)] + \sum_{k=1}^{2d-1} p_k \mathbb{E}[\varphi_{\mathbf{W}}(V^a t) \varphi_{K^k \mathbf{W}}((1-V)^a t)]. \quad (3.20)$$

Consequently,

$$|\varphi_{\mathbf{W}}(t)| \leq p_0 \mathbb{E}[|\varphi_{\mathbf{W}}(tV^a)| |\varphi_{\mathbf{W}}(t(1-V)^a)|] + (1-p_0) \mathbb{E}[|\varphi_{\mathbf{W}}(tV^a)|].$$

We conclude as in dimension 1 (Step 2 in the proof of Theorem 2.5), that

$$\limsup_{\|t\| \rightarrow \infty} |\varphi_{\mathbf{W}}(t)| \in \{0, 1\}. \quad (3.21)$$

We now prove the result by contradiction, assuming that the above limit (3.21) is 1. The difficulty is that  $K$  is not an isometric matrix ( $\|K\mathbf{W}\| \neq \|\mathbf{W}\|$ ), however it satisfies  $K^{2d} = I_{2d-1}$ .

Adapting Step 3 in the proof of Theorem 2.5 in dimension 1, and the idea used in the proof of [11, Thm 7.1], we introduce the function  $\psi_{\mathbf{W}}$  defined for  $r \geq 0$  by

$$\psi_{\mathbf{W}}(r) := \max_{\|t\|=r} \max_{k \in \{0, \dots, 2d-1\}} |\varphi_{K^k \mathbf{W}}(t)| = \max_{\|t\|=r} \max_{k \in \{0, \dots, 2d-1\}} \left| \varphi_{\mathbf{W}} \left( \left( K^k \right)^T t \right) \right|.$$

By continuity of the function  $|\varphi_{\mathbf{W}}|$ , compactness of the subset  $\{t \in \mathbb{R}^{2d-1} : \|t\| = r\}$ , invertibility of  $K$  and Lemma 3.9, we deduce that  $\psi_{\mathbf{W}}(0) = 1$  and  $\psi_{\mathbf{W}}(r) < 1$  for all  $r > 0$ .

By assumption of the proof by contradiction,  $\limsup_{r \rightarrow \infty} \psi_{\mathbf{W}}(r) = 1$ . The function  $\psi_{\mathbf{W}}$  is continuous with  $\psi_{\mathbf{W}}(0) = 1$ , and for all  $r > 0$ ,  $\psi_{\mathbf{W}}(r) < 1$ . Let  $r_0 > 0$  be fixed and  $\varepsilon > 0$  such that  $\psi_{\mathbf{W}}(r_0) < 1 - \varepsilon$ . By the intermediate value theorem, there exists  $(r_1, r_2) = (r_1(\varepsilon), r_2(\varepsilon))$  such that  $0 < r_1 < r_0 < r_2$  and

$$\psi_{\mathbf{W}}(r_1) = \psi_{\mathbf{W}}(r_2) = 1 - \varepsilon \quad \text{and} \quad \psi_{\mathbf{W}}(r) \leq 1 - \varepsilon, \quad \forall r \in (r_1, r_2).$$

Moreover, taking a sequence  $(\varepsilon_n)_{n \geq 0}$  converging to 0, we deduce by Lemma 3.9 that 0 is the unique accumulation point of  $r_1(\varepsilon_n)$ . Then  $\lim_{\varepsilon \rightarrow 0} r_1(\varepsilon) = 0$ .

Since  $V \in (0, 1)$  a.s., we deduce from (3.17) and  $K^{2d} = I_{2d-1}$  that for  $\|t\| = r$ ,  $k \in \{0, \dots, 2d-1\}$ ,

$$|\varphi_{K^k \mathbf{W}}(t)| \leq \mathbb{E}[\psi_{\mathbf{W}}(rV^a) \psi_{\mathbf{W}}(r(1-V)^a)].$$

Then

$$\psi_{\mathbf{W}}(r) \leq \mathbb{E}[\psi_{\mathbf{W}}(rV^a) \psi_{\mathbf{W}}(r(1-V)^a)]. \quad (3.22)$$

Introducing  $\phi_n(r) = \psi_{\mathbf{W}}(rV_1^a \cdots V_n^a)$  with  $(V_1, \dots, V_n)$  i.i.d. copies of  $V$ , we obtain by induction

$$\psi_{\mathbf{W}}(r) \leq \mathbb{E}[\psi_{\mathbf{W}}(rV^a)] \leq \phi_n(r).$$

The remainder of the proof is exactly as in Step 3 of the proof of Theorem 2.5. We then conclude by contradiction that the limit (3.21) should be 0.  $\square$

**Lemma 3.11.** *Let  $\mathbf{W}$  be a solution of the fixed-point equation (3.17) satisfying the assumptions of Theorem 3.8. Then for all  $k \geq 1$ , there exists a constant  $C > 0$  such that for all non-zero  $t \in \mathbb{R}^{2d-1}$ , (3.19) holds.*

*Proof.* Let  $\varepsilon \in (0, 1)$  and  $T_\varepsilon > 0$  such that  $\max_{k \in \{0, \dots, 2d-1\}} \left| \varphi_{\mathbf{W}} \left( (K^k)^T t \right) \right| \leq \varepsilon$  as soon as  $\|t\| \geq T_\varepsilon$ . The existence (finiteness) of  $T_\varepsilon$  is guaranteed by Lemma 3.10. We have from Equation (3.20) that for any non-zero  $t$ ,

$$\begin{aligned} |\varphi_{\mathbf{W}}(t)| &\leq \varepsilon \mathbb{E}[|\varphi_{\mathbf{W}}(tV^a)|] + \mathbb{P}((1-V)^a \|t\| \leq T_\varepsilon) \\ &\leq \varepsilon \mathbb{E}[|\varphi_{\mathbf{W}}(tV^a)|] + \frac{1}{1-\rho a} \left( \frac{T_\varepsilon}{\|t\|} \right)^\rho, \end{aligned}$$

for any  $\rho \in (0, 1/a)$  and  $\|t\| > T_\varepsilon$ . The conclusion follows using the same arguments as in the proof of Theorem 2.5 (Step 4).  $\square$

*Proof of Theorem 3.8.* The first point is proved in Lemma 3.11, and choosing  $\rho \in (0, 1/a)$  and  $k \geq 1$  such that  $k\rho > 2d - 1$ , we deduce that  $\varphi_{\mathbf{W}}$  is integrable on  $\mathbb{R}^{2d-1}$ . By Fourier inversion theorem, we deduce that the random vector  $\mathbf{W}$  has a bounded density with respect to the Lebesgue measure on  $\mathbb{R}^{2d-1}$ . Then, since  $k \geq 1$  may be chosen arbitrarily large, we obtain that the density is smooth on  $\mathbb{R}^{2d-1}$ .  $\square$

### 3.6 Support of the urn limit

The aim of this section is to prove that the support of the process  $\mathbf{W}$ , defined in Theorem 3.4, is  $\mathbb{R}^{2d-1}$ . The study of the support is not difficult for  $d = 2$ , but as we will see it is more challenging for a general dimension  $d$ . Consequently, we first show the result in dimension 2. Then we explain how we can study this support for any dimension  $d \geq 1$ , but give a complete proof only in the case  $d = 3$ .

**Support of the urn limit when  $d = 2$**  We are interested in studying non-null solutions of the fixed-point equation (3.17), with  $d = 2$  and

$$K = \begin{pmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad (\text{see (3.16)})$$

for a general probability vector  $(p_k)_{0 \leq k \leq 3}$ , and then in our specific case (see Proposition 3.7)

$$(p_0, p_1, p_2, p_3) = (p, \frac{1-p}{3}, \frac{1-p}{3}, \frac{1-p}{3}), \quad \text{with } p = \frac{3a+1}{4}. \quad (3.23)$$

**Proposition 3.12.** *Let  $a \in (1/2, 1)$  and  $\mathbf{W}$  be a non-null solution of the fixed-point equation (3.17) with a general probability vector  $(p_k)_{0 \leq k \leq 3}$  with positive coordinates.*

1. *The support of  $\mathbf{W}$  is either  $\text{Span}\{(1, -1, 1)\}$ , or  $\text{Span}\{(1, 0, -1), (0, 1, 0)\}$ , or  $\mathbb{R}^3$ .*
2. *In our specific case (3.23), when  $\mathbf{W}$  is the limit (3.7) of the urn process starting from  $\mathbf{U}(0) = \mathbf{e}_1$ , then  $\text{Supp}(\mathbf{W}) = \mathbb{R}^3$  and  $\mathbf{W}$  has a smooth bounded density on  $\mathbb{R}^3$ .*

*Proof.* Let  $\mathbf{w} \in \text{Supp}(\mathbf{W})$  with  $\mathbf{w} \neq 0$ . Then by Equation (3.17), since  $\mathbb{P}(B = I_3) > 0$ ,  $t^a \mathbf{w} + (1-t)^a \mathbf{w} \in \text{Supp}(\mathbf{W})$  for any  $t \in [0, 1]$ . We deduce that  $\text{Supp}(\mathbf{W})$  is not a singleton ( $a \neq 1$ ). Since  $\mathbb{P}(B = K^k) > 0$  for  $k \in \{1, 2, 3\}$ , we also have  $K\mathbf{w}, K^2\mathbf{w}, K^3\mathbf{w} \in \text{Supp}(\mathbf{W})$ . We remark that if  $\mathbf{w} = (x, y, z)$ , then the determinant

$$\det(\mathbf{w}, K\mathbf{w}, K^2\mathbf{w}) = (x+z)((x+y)^2 + (y+z)^2). \quad (3.24)$$

We now examine three cases. If first  $\mathbf{w} \in \text{Span}\{(1, -1, 1)\}$ , then  $K\mathbf{w} = -\mathbf{w}$ . Using similar arguments as in dimension 1, we deduce that

$$\text{Span}\{\mathbf{w}\} = \text{Span}\{(1, -1, 1)\} \subset \text{Supp}(\mathbf{W}).$$

We also note that  $K(\text{Span}\{(1, -1, 1)\}) = \text{Span}\{(1, -1, 1)\}$ , which implies that  $\text{Supp}(\mathbf{W}) = \text{Span}\{(1, -1, 1)\}$  is possible.

If now  $\mathbf{w} \in \text{Span}\{(1, 0, -1), (0, 1, 0)\}$ , then  $K^2\mathbf{w} = -\mathbf{w}$ . We deduce

$$\text{Span}\{\mathbf{w}, K\mathbf{w}\} = \text{Span}\{(1, 0, -1), (0, 1, 0)\} \subset \text{Supp}(\mathbf{W}).$$

We note that  $K(\text{Span}\{(1, 0, -1), (0, 1, 0)\}) = \text{Span}\{(1, 0, -1), (0, 1, 0)\}$ , and then  $\text{Supp}(\mathbf{W}) = \text{Span}\{(1, 0, -1), (0, 1, 0)\}$  is also possible.

If finally  $\mathbf{w} \notin \text{Span}\{(1, -1, 1)\}$  and  $\mathbf{w} \notin \text{Span}\{(1, 0, -1), (0, 1, 0)\}$ , then  $\mathbf{w}, K\mathbf{w}, K^2\mathbf{w}$  are linearly independent by (3.24). By (3.17), for every  $\mathbf{v}, \mathbf{v}' \in \{\mathbf{w}, K\mathbf{w}, K^2\mathbf{w}\}$  and all  $t \in [0, 1]$ ,  $t^a \mathbf{v} + (1-t)^a \mathbf{v}' \in \text{Supp}(\mathbf{W})$ . Then, as in dimension 1, we deduce  $[\mathbf{v}, \mathbf{v}'] \subset \text{Supp}(\mathbf{W})$  and  $2^{1-a} \mathbf{v} \in \text{Supp}(\mathbf{W})$  (taking  $t = 1/2$ ). We obtain that

$$\text{Span}_+\{\mathbf{w}, K\mathbf{w}, K^2\mathbf{w}\} := \left\{ \sum_{k=0}^2 \lambda_k K^k \mathbf{w} : \lambda_k \in \mathbb{R}_+ \right\} \subset \text{Supp}(\mathbf{W}).$$

Moreover, we observe that  $\mathbf{w} + K\mathbf{w}$  and  $K\mathbf{w} + K^2\mathbf{w}$  belong to  $\text{Span}\{(1, 0, -1), (0, 1, 0)\}$ , then  $K^2\mathbf{w} + K^3\mathbf{w} = -(\mathbf{w} + K\mathbf{w})$  and  $K^3\mathbf{w} + \mathbf{w} = -(K\mathbf{w} + K^2\mathbf{w})$ . We deduce that  $-\mathbf{w}, -K\mathbf{w}, -K^2\mathbf{w} \in \text{Supp}(\mathbf{W})$ , and consequently

$$\text{Span}\{\mathbf{w}, K\mathbf{w}, K^2\mathbf{w}\} = \mathbb{R}^3 = \text{Supp}(\mathbf{W}).$$

The first result of the proposition is proved.

Let us now consider the specific case of the asymptotic urn process starting from  $\mathbf{U}(0) = \mathbf{e}_1$ . By Theorem 3.4, we have  $\mathbb{E}[\mathbf{W}] = \frac{-1}{4\Gamma(a+1)}(1, 1, 1)^T$ , which means that all coordinates of  $\mathbf{W}$  have the same expectation. We easily deduce that, under this condition, the only possible support for  $\mathbf{W}$  is  $\text{Supp}(\mathbf{W}) = \mathbb{R}^3$ . We conclude by Theorem 3.8.  $\square$

**Support of  $\mathbf{W}$  for a general dimension  $d \geq 1$**  As in the case  $d = 2$ , the support of a general non-null solution  $\mathbf{W}$  of the fixed-point equation (3.17) is not unique (we exclude the case  $\mathbf{W} = 0$  a.s.).

Let  $\mathbf{w} \in \text{Supp}(\mathbf{W})$  with  $\mathbf{w} \neq 0$ . If  $\mathbf{w}$  is such that its determinant

$$\det(\mathbf{w}, K\mathbf{w}, \dots, K^{2d-2}\mathbf{w}) \quad (3.25)$$

is non-zero, then using the same proof as in dimension 2 (see Proposition 3.12), we deduce that  $\text{Supp}(\mathbf{W}) = \mathbb{R}^{2d-1}$ . We thus want to identify all vectors  $\mathbf{w} \in \mathbb{R}^{2d-1}$  which have the property that the determinant (3.25) is (non-)zero. To that purpose, it is convenient to introduce the Krylov space  $\text{Span}\{\mathbf{w}, K\mathbf{w}, \dots, K^{2d-2}\mathbf{w}\}$  generated by  $\mathbf{w}$  (see e.g. [14]). Let us mention that for all  $\ell \in \{0, \dots, 2d-1\}$ ,

$$\text{Span}\{(K^j \mathbf{w})_{0 \leq j \leq 2d-1, j \neq \ell}\} = \text{Span}\{\mathbf{w}, K\mathbf{w}, \dots, K^{2d-2}\mathbf{w}\},$$

since  $K^{2d} = I_{2d-1}$ .

**Proposition 3.13.** For  $\mathbf{w} = (w_1, \dots, w_{2d-1}) \in \mathbb{R}^{2d-1}$ , the determinant (3.25) is zero if and only if at least one of the following linear equations is satisfied

$$\sum_{k=1}^{2d-1} w_k (e^{i \frac{kj\pi}{d}} - 1) = 0,$$

with  $j \in \{1, \dots, 2d-1\}$ . The codimension of  $\text{Span}\{\mathbf{w}, K\mathbf{w}, \dots, K^{2d-2}\mathbf{w}\}$  is exactly the number of the above equations satisfied.

*Proof.* Let  $(\mathbf{f}_i)_{1 \leq i \leq 2d-1}$  be the canonical basis of  $\mathbb{R}^{2d-1}$ . We first prove that for  $\mathbf{w} = \mathbf{f}_1$ , the determinant (3.25) is non-zero. To that aim, we study the matrix  $K$  in (3.16). We have  $\mathbf{w} = \mathbf{f}_1$  and for  $\ell \in \{1, \dots, 2d-2\}$ ,  $K^\ell \mathbf{f}_1 = -\mathbf{f}_\ell + \mathbf{f}_{\ell+1}$ . As a consequence

$$\text{Span}\{\mathbf{f}_1, K\mathbf{f}_1, \dots, K^{2d-2}\mathbf{f}_1\} = \mathbb{R}^{2d-1},$$

which proves our claim about (3.25) being non-zero for  $\mathbf{w} = \mathbf{f}_1$ .

As a second step, we give a formula for the determinant (3.25). The matrix  $K$  satisfies  $K^{2d} = I_{2d-1}$ . It is easy to see that 1 is not an eigenvalue of  $K$ . The eigenvalues in  $\mathbb{C}$  of  $K$  are  $\eta_k = \eta^k$  with  $\eta = e^{i\pi/d}$ . Solving  $K\mathbf{z} = \eta^k \mathbf{z}$ , we obtain the respective eigenvectors, which are given by

$$\mathbf{z}_k = (\eta^{-k}, \eta^{-2k}, \dots, \eta^{-(2d-1)k})^T \in \mathbb{C}^{2d-1},$$

for  $k \in \{1, \dots, 2d-1\}$ . When taking  $k = d$ , we observe that  $-1$  is an eigenvalue with associated eigenvector  $\mathbf{z}_d = (-1, 1, -1, \dots, 1, -1)$ . It is the unique real eigenvalue of  $K$ . We introduce the matrix  $P = (\mathbf{z}_1, \dots, \mathbf{z}_{2d-1})$ , and for  $\mathbf{w} \in \mathbb{R}^{2d-1}$ , we define

$$\mathbf{u} := P^{-1}\mathbf{w} = (u_1, \dots, u_{2d-1}). \quad (3.26)$$

As  $K = P \text{diag}(\eta, \eta^2, \dots, \eta^{2d-1}) P^{-1}$ , we deduce that

$$\det(\mathbf{w}, K\mathbf{w}, \dots, K^{2d-2}\mathbf{w}) = \det(P) \det \left( \left( \eta^{(i-1)(k-1)} \right)_{1 \leq i, k \leq 2d-1} \right) \prod_{j=1}^{2d-1} u_j.$$

Our previous computation with  $\mathbf{w} = \mathbf{f}_1$  shows that the second determinant in the right-hand side of the above identity is non-zero (this could also be seen using a direct computation). Accordingly, the determinant (3.25) is zero if and only if there exists  $j$  such that  $u_j = 0$ . Moreover, we have

$$\dim(\text{Span}\{\mathbf{w}, K\mathbf{w}, \dots, K^{2d-2}\mathbf{w}\}) = \text{card}\{j : u_j \neq 0\}.$$

We now express the matrix  $P^{-1}$ , as it gives an explicit relation between  $\mathbf{u}$  and  $\mathbf{w}$ , see (3.26). More precisely, we show that

$$P^{-1} = \frac{1}{2d} \begin{pmatrix} \bar{\mathbf{z}}_1^T - 1 \\ \vdots \\ \bar{\mathbf{z}}_{2d-1}^T - 1 \end{pmatrix}. \quad (3.27)$$

By definition of  $P$ , we have

$$\mathbf{w} = P\mathbf{u} = \sum_{j=1}^{2d-1} u_j \mathbf{z}_j. \quad (3.28)$$



Moreover, for all  $k, \ell \in \{1, \dots, 2d-1\}$ ,

$$\langle \mathbf{z}_k, \mathbf{z}_\ell \rangle = \sum_{j=1}^{2d-1} \bar{\mathbf{z}}_{j,k} \mathbf{z}_{j,\ell} = \sum_{j=1}^{2d-1} \eta^{j(k-\ell)} = \eta^{k-\ell} \frac{1 - \eta^{(2d-1)(k-\ell)}}{1 - \eta^{k-\ell}} = \begin{cases} -1 & \text{if } k \neq \ell, \\ 2d-1 & \text{if } k = \ell, \end{cases}$$

because  $\eta^{2d(k-\ell)} = 1$ . We deduce from (3.28) that

$$\langle \mathbf{z}_k, \mathbf{w} \rangle = \sum_{j=1}^{2d-1} u_j \langle \mathbf{z}_k, \mathbf{z}_j \rangle = - \sum_{j=1}^{2d-1} u_j + 2du_k \quad (3.29)$$

and

$$\left\langle \sum_{k=1}^{2d-1} \mathbf{z}_k, \mathbf{w} \right\rangle = \sum_{k=1}^{2d-1} u_k,$$

with  $\sum_{k=1}^{2d-1} \mathbf{z}_k = (-1, 1, \dots, -1)^T$ . Then Equation (3.29) implies  $2du_k = \langle \mathbf{z}_k, \mathbf{w} \rangle - \sum_{j=1}^{2d-1} w_j = (\bar{\mathbf{z}}_k - 1)^T \mathbf{w}$  and we obtain (3.27).

To conclude, we recall that for  $j < d$ ,  $\eta^j = \bar{\eta}^{2d-j}$ , thus  $\mathbf{z}_j = \bar{\mathbf{z}}_{2d-j}$ . Moreover, since  $\mathbf{z}_d = \sum_{k=1}^{2d-1} (-1)^k e_k$ , we have  $\bar{\mathbf{z}}_d - 1 = -2 \sum_{j=1}^d \mathbf{f}_{2j-1}$ . By definition (3.26) of  $\mathbf{u}$ , we derive

- for  $j \in \{1, \dots, d-1\}$ ,  $u_j = 0 \iff u_{2d-j} = 0 \iff \sum_{k=1}^{2d-1} w_k (e^{i \frac{kj\pi}{d}} - 1) = 0$ ;
- $u_d = 0 \iff \sum_{j=1}^d w_{2j-1} = 0$ .

The proposition is proved.  $\square$

As a first consequence of Proposition 3.13, we recover the result proved in Proposition 3.12 on the support of the variable  $\mathbf{W}$  in dimension 2.

**Corollary 3.14.** *In dimension 2, discarding the case when  $\mathbf{W}$  is zero, the only possibilities for  $\text{Supp}(\mathbf{W})$  are  $\text{Span}\{\mathbf{f}_1 - \mathbf{f}_2 + \mathbf{f}_3\}$ ,  $\text{Span}\{\mathbf{f}_2, \mathbf{f}_1 - \mathbf{f}_3\}$  and  $\mathbb{R}^3$ .*

*Proof.* For  $d = 2$ , one has  $\eta = i$  and with (3.27) we obtain

$$P^{-1} = \frac{1}{4} \begin{pmatrix} -1+i & -2 & -1-i \\ -2 & 0 & -2 \\ -1-i & -2 & -1+i \end{pmatrix}$$

Then using Proposition 3.13 (or directly (3.26)), we have

- $u_1 = 0 \iff u_3 = 0 \iff w_1 = -w_2 = w_3$ ;
- $u_2 = 0 \iff w_1 = -w_3$ .  $\square$

More interestingly, Proposition 3.13 also allows us to describe the possible supports of  $\mathbf{W}$  in dimension three.

**Corollary 3.15.** *In dimension  $d = 3$ , discarding the case when  $\mathbf{W}$  is zero,  $\text{Supp}(\mathbf{W})$  is one of the following vector spaces:*

- (i)  $\text{Supp}(\mathbf{W}) = \text{Span}\{\mathbf{f}_1 - \mathbf{f}_2 + \mathbf{f}_3 - \mathbf{f}_4 + \mathbf{f}_5\}$ ;
- (ii)  $\text{Supp}(\mathbf{W}) = \text{Span}\{\mathbf{f}_2 - \mathbf{f}_3 + \mathbf{f}_5, -\mathbf{f}_1 + \mathbf{f}_3 - \mathbf{f}_4\}$ ;

(iii)  $\text{Supp}(\mathbf{W}) = \text{Span}\{\mathbf{f}_2 + \mathbf{f}_3 - \mathbf{f}_5, -\mathbf{f}_1 + \mathbf{f}_3 + \mathbf{f}_4\};$

(iv)  $\text{Supp}(\mathbf{W}) = \text{Span}\{\mathbf{f}_3 - 2\mathbf{f}_4 + 2\mathbf{f}_5, -\mathbf{f}_1 + \mathbf{f}_4 - 2\mathbf{f}_5, 2\mathbf{f}_1 - \mathbf{f}_2 + \mathbf{f}_5\};$

(v)  $\text{Supp}(\mathbf{W}) = \text{Span}\{\mathbf{f}_3, -\mathbf{f}_1 + \mathbf{f}_4, -\mathbf{f}_2 + \mathbf{f}_5\};$

(vi)  $\text{Supp}(\mathbf{W}) = \text{Span}\{\mathbf{f}_2, -\mathbf{f}_1 + \mathbf{f}_3, -\mathbf{f}_2 + \mathbf{f}_4, -\mathbf{f}_3 + \mathbf{f}_5\};$

(vii)  $\text{Supp}(\mathbf{W}) = \mathbb{R}^5.$

*Proof.* If  $d = 3$  then  $\eta = e^{i\pi/3}$  and with (3.27)

$$P^{-1} = \frac{1}{6} \begin{pmatrix} e^{i\pi/3} - 1 & e^{i2\pi/3} - 1 & -2 & e^{-i2\pi/3} - 1 & e^{-i\pi/3} - 1 \\ e^{i2\pi/3} - 1 & e^{-2i\pi/3} - 1 & 0 & e^{i2\pi/3} - 1 & e^{-i2\pi/3} - 1 \\ -2 & 0 & -2 & 0 & -2 \\ e^{-i2\pi/3} - 1 & e^{i2\pi/3} - 1 & 0 & e^{-i2\pi/3} - 1 & e^{i2\pi/3} - 1 \\ e^{-i\pi/3} - 1 & e^{-i2\pi/3} - 1 & -2 & e^{2\pi/3} - 1 & e^{i\pi/3} - 1 \end{pmatrix}.$$

We notice that

- $u_1 = 0 \iff u_5 = 0 \iff$

$$\begin{aligned} w_1 + 3w_2 + 4w_3 + 3w_4 + w_5 = 0 \quad \text{and} \quad w_1 + w_2 - w_4 - w_5 = 0 \\ \iff w_1 + w_2 - w_4 - w_5 = 0 \quad \text{and} \quad w_2 + 2w_3 + 2w_4 + w_5 = 0; \end{aligned}$$

- $u_2 = 0 \iff u_4 = 0 \iff$

$$\begin{aligned} w_1 + w_2 + w_4 + w_5 = 0 \quad \text{and} \quad w_1 - w_2 + w_4 - w_5 = 0 \\ \iff w_1 + w_4 = 0 = w_2 + w_5; \end{aligned}$$

- $u_3 = 0 \iff w_1 + w_3 + w_5 = 0.$

Accordingly, we obtain that the Krylov space generated by  $\mathbf{w}$  may be of

- *dimension 1*, when only  $u_1 = u_2 = 0$ . Then  $\mathbf{w} \in \text{Span}\{\mathbf{f}_1 - \mathbf{f}_2 + \mathbf{f}_3 - \mathbf{f}_4 + \mathbf{f}_5\}$ . We also notice that  $K(\mathbf{f}_1 - \mathbf{f}_2 + \mathbf{f}_3 - \mathbf{f}_4 + \mathbf{f}_5) = -(\mathbf{f}_1 - \mathbf{f}_2 + \mathbf{f}_3 - \mathbf{f}_4 + \mathbf{f}_5)$ . We obtain (i).
- *dimension 2*, when only  $u_1 = u_3 = 0$  or  $u_2 = u_3 = 0$ . We observe that

$$\{\mathbf{w} \in \mathbb{R}^{2d-1} : u_1 = 0 = u_3\} = \text{Span}\{\mathbf{f}_2 - \mathbf{f}_3 + \mathbf{f}_5, -\mathbf{f}_1 + \mathbf{f}_3 - \mathbf{f}_4\} \quad (\text{case (ii)})$$

and

$$\{\mathbf{w} \in \mathbb{R}^{2d-1} : u_2 = 0 = u_3\} = \text{Span}\{\mathbf{f}_2 + \mathbf{f}_3 - \mathbf{f}_5, -\mathbf{f}_1 + \mathbf{f}_3 + \mathbf{f}_4\} \quad (\text{case (iii)}).$$

Moreover, we have

$$K(\mathbf{f}_2 - \mathbf{f}_3 + \mathbf{f}_5) = -\mathbf{f}_1 + \mathbf{f}_3 - \mathbf{f}_4, \quad K^2(\mathbf{f}_2 - \mathbf{f}_3 + \mathbf{f}_5) = -(\mathbf{f}_2 - \mathbf{f}_3 + \mathbf{f}_5) - (-\mathbf{f}_1 + \mathbf{f}_3 - \mathbf{f}_4),$$

and

$$K(\mathbf{f}_2 + \mathbf{f}_3 - \mathbf{f}_5) = -\mathbf{f}_1 + \mathbf{f}_3 + \mathbf{f}_4, \quad K^2(\mathbf{f}_2 + \mathbf{f}_3 - \mathbf{f}_5) = -\mathbf{f}_1 + \mathbf{f}_3 + \mathbf{f}_4 - (\mathbf{f}_2 + \mathbf{f}_3 - \mathbf{f}_5).$$

- *dimension 3*, when only  $u_1 = 0$  or  $u_2 = 0$ . We have

$$\begin{aligned}\{\mathbf{w} \in \mathbb{R}^{2d-1} : u_1 = 0\} &= \text{Span}\{\mathbf{f}_3 - 2\mathbf{f}_4 + 2\mathbf{f}_5, -\mathbf{f}_1 + \mathbf{f}_4 - 2\mathbf{f}_5, 2\mathbf{f}_1 - \mathbf{f}_2 + \mathbf{f}_5\} \quad (\text{case (iv)}), \\ \{\mathbf{w} \in \mathbb{R}^{2d-1} : u_2 = 0\} &= \text{Span}\{\mathbf{f}_3, -\mathbf{f}_1 + \mathbf{f}_4, -\mathbf{f}_2 + \mathbf{f}_5\} \quad (\text{case (v)}).\end{aligned}$$

On one hand, we have

$$\begin{aligned}K(\mathbf{f}_3 - 2\mathbf{f}_4 + 2\mathbf{f}_5) &= -\mathbf{f}_1 + \mathbf{f}_4 - 2\mathbf{f}_5, \\ K^2(\mathbf{f}_3 - 2\mathbf{f}_4 + 2\mathbf{f}_5) &= 2\mathbf{f}_1 - \mathbf{f}_2 + \mathbf{f}_5 \\ K^3(\mathbf{f}_3 - 2\mathbf{f}_4 + 2\mathbf{f}_5) &= 2\mathbf{f}_1 + 2\mathbf{f}_2 - \mathbf{f}_3 \\ &= -(\mathbf{f}_3 - 2\mathbf{f}_4 + 2\mathbf{f}_5) - 2(-\mathbf{f}_1 + \mathbf{f}_4 - 2\mathbf{f}_5) - 2(2\mathbf{f}_1 - \mathbf{f}_2 + \mathbf{f}_5),\end{aligned}$$

and in the other hand,

$$K\mathbf{f}_3 = -\mathbf{f}_1 + \mathbf{f}_4, \quad K^2\mathbf{f}_3 = -\mathbf{f}_2 + \mathbf{f}_5, \quad K^3\mathbf{f}_3 = -\mathbf{f}_3.$$

- *dimension 4*, when only  $u_3 = 0$ . We have

$$\{\mathbf{w} \in \mathbb{R}^{2d-1} : u_3 = 0\} = \text{Span}\{\mathbf{f}_2, -\mathbf{f}_1 + \mathbf{f}_3, -\mathbf{f}_2 + \mathbf{f}_4, -\mathbf{f}_3 + \mathbf{f}_5\}$$

and  $K\mathbf{f}_2 = -\mathbf{f}_1 + \mathbf{f}_3$ ,  $K^2\mathbf{f}_2 = -\mathbf{f}_2 + \mathbf{f}_4$ ,  $K^3\mathbf{f}_2 = -\mathbf{f}_3 + \mathbf{f}_5$  and  $K^4\mathbf{f}_2 = -\mathbf{f}_4 = -\mathbf{f}_2 - K^2\mathbf{f}_2$ . This corresponds to case (vi).

- *dimension 5*, when  $u_j \neq 0$  for all  $j$ , and  $\text{Supp}(K) = \mathbb{R}^5$  (case (vii)). □

We now focus on  $d = 3$ .

**Proposition 3.16.** *In dimension 3, by (3.2), we have  $a = \frac{6p-1}{5}$ . Assume that  $\frac{1}{2} < a < 1$ . The process  $\mathbf{W}$  of the urn limit for  $d = 3$  starting from a ball of color 1, defined in Theorem 3.4, has a smooth bounded density on  $\mathbb{R}^5$ , and its support is  $\mathbb{R}^5$ .*

*Proof.* By Theorem 3.8, we only have to prove that  $\text{Supp}(\mathbf{W}) = \mathbb{R}^5$ .

Let  $\mathbf{Y}_{(1)} \in \mathbb{R}^6$  be the asymptotics of the urn process starting from a unique ball of color 1, introduced in Theorem 3.4. We recall that  $\mathbf{Y}_{(1)} = \sum_{i=2}^6 W_i \mathbf{v}_i$  with  $\mathbf{v}_i = \frac{1}{2}(\mathbf{e}_1 - \mathbf{e}_i)$  for  $i \geq 2$ . Then, by Equation (3.8),

$$\mathbb{E}[\mathbf{Y}_{(1)}] = \frac{1}{2} \mathbb{E} \left[ \begin{pmatrix} W_2 + \dots + W_6 \\ -W_2 \\ \vdots \\ -W_6 \end{pmatrix} \right] = \frac{-1}{6\Gamma(a+1)} \begin{pmatrix} -5 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Let us consider all possible supports of  $\mathbf{W}$  described in Corollary 3.15. Due to the fact that  $\mathbb{E}[W_2] = \dots = \mathbb{E}[W_6]$ , all cases happen to be impossible, except Case (vii). Indeed,

- (i)  $\text{Supp}(\mathbf{W}) = \text{Span}\{\mathbf{f}_1 - \mathbf{f}_2 + \mathbf{f}_3 - \mathbf{f}_4 + \mathbf{f}_5\}$  implies  $W_2 = -W_3$ , while we should have  $\mathbb{E}[W_2] = \mathbb{E}[W_3]$ ;
- (ii)  $\text{Supp}(\mathbf{W}) = \text{Span}\{\mathbf{f}_2 - \mathbf{f}_3 + \mathbf{f}_5, -\mathbf{f}_1 + \mathbf{f}_3 - \mathbf{f}_4\}$  implies  $W_4 = -W_2 - W_3$ , while we should have  $\mathbb{E}[W_4] = \mathbb{E}[W_2] = \mathbb{E}[W_3]$ ;
- (iii)  $\text{Supp}(\mathbf{W}) = \text{Span}\{\mathbf{f}_2 + \mathbf{f}_3 - \mathbf{f}_5, -\mathbf{f}_1 + \mathbf{f}_3 + \mathbf{f}_4\}$  implies  $W_2 = -W_5$ , but we should have  $\mathbb{E}[W_2] = \mathbb{E}[W_5]$ ;

- (iv)  $\text{Supp}(\mathbf{W}) = \text{Span}\{\mathbf{f}_3 - 2\mathbf{f}_4 + 2\mathbf{f}_5, -\mathbf{f}_1 + \mathbf{f}_4 - 2\mathbf{f}_5, 2\mathbf{f}_1 - \mathbf{f}_2 + \mathbf{f}_5\}$  implies  $W_5 = -W_2 - 2W_3 - 2W_4$  and we should have  $\mathbb{E}[W_5] = \mathbb{E}[W_4] = \mathbb{E}[W_2] = \mathbb{E}[W_3]$ ;
- (v)  $\text{Supp}(\mathbf{W}) = \text{Span}\{\mathbf{f}_3, -\mathbf{f}_1 + \mathbf{f}_4, -\mathbf{f}_2 + \mathbf{f}_5\}$  implies  $W_2 = -W_5$  and we should have  $\mathbb{E}[W_2] = \mathbb{E}[W_5]$ ;
- (vi)  $\text{Supp}(\mathbf{W}) = \text{Span}\{\mathbf{f}_2, -\mathbf{f}_1 + \mathbf{f}_3, -\mathbf{f}_2 + \mathbf{f}_4, -\mathbf{f}_3 + \mathbf{f}_5\}$  implies  $W_4 = -W_2 - W_6$  and we should have  $\mathbb{E}[W_2] = \mathbb{E}[W_4] = \mathbb{E}[W_6]$ .

Consequently,  $\text{Supp}(\mathbf{W}) = \mathbb{R}^5$ , which means that Theorem 3.8 is satisfied and the result is proved.  $\square$

We see in the above proof of Proposition 3.16 that we do not need to describe exactly all the possibilities for  $\text{Supp}(\mathbf{W})$  to obtain the result, but we only need a well-adapted relationship between the coordinates of  $\mathbf{W}$  when the support is not  $\mathbb{R}^{2d-1}$ . Even if it is effective, the method presented in this article does not seem optimal, because we are looking for too much information on the general solutions of the fixed-point equation (3.17) to deduce the support of  $\mathbf{W}$  in the specific case of the asymptotic urn process.

### 3.7 The distribution of superdiffusive limit of the MERW

We now focus on the asymptotic limit  $\mathbf{L} \in \mathbb{R}^d$  of the superdiffusive MERW to prove Theorem 1.5. We only explain the ideas when the first step of MERW is a.s. in direction  $\mathbf{e}_1$ . The general case will follow immediately, because  $\mathbf{L}$  is a discrete mixture of the variables  $\mathbf{L}_{(k)}$ ,  $k \in \{1, \dots, 2d\}$ , see (3.4).

We recall the relation (3.11) between  $\mathbf{L}$  and the urn asymptotic  $\mathbf{W}$ , defined in Theorem 3.4. As mentioned earlier (see page 24), if  $\text{Supp}(\mathbf{W}) = \mathbb{R}^{2d-1}$ , we immediately deduce that  $\mathbf{L}$  admits a smooth positive density with respect to the Lebesgue measure in  $\mathbb{R}^d$ . As we do not have a general result on the support of the solution of the fixed-point equation (3.17), for any dimension  $d \geq 2$ , we cannot have a general result for  $\mathbf{L}$ . However, by Propositions 3.12 and 3.16 in dimension 2 and 3, Theorem 1.5 is easily deduced from Theorem 3.8, and Corollaries 3.14 and 3.15.

### 3.8 Moments of the asymptotic distributions in high dimension

We now study the moments of the limit urn process  $\mathbf{Y}$  and the limit  $\mathbf{L}$  of the MERW. To that aim, we use the fixed-point equation (3.13) and not (3.17), because in (3.17) the norm of the matrix  $B$  is not smaller than 1, which do not allow us to have good estimates of the moments of  $\mathbf{W}$ .

When the first step probability vector  $q$  of the MERW is the uniform distribution on the  $2d$ -directions, it is proved in [6, Thm 3.8] that  $\mathbb{E}[\mathbf{L}] = 0$  and

$$\mathbb{E}[\mathbf{L}\mathbf{L}^T] = \frac{1}{d(2a-1)\Gamma(2a)} I_d.$$

**Theorem 3.17.** *For  $d \geq 1$ , let  $\mathbf{Y}$  be the asymptotic urn process defined in Theorem 3.4, and  $\mathbf{L}$  be the asymptotics of the MERW with memory parameter  $p$  and first step distribution  $q$ , defined in Theorem 3.1. Then  $\mathbf{Y}$  and  $\mathbf{L}$  have finite moments of all order, and their probability distribution are determined by their moments. Moreover, the moment-generating function of  $\mathbf{Y}$ ,  $\mathbf{L}$ ,  $\|\mathbf{Y}\|$  and  $\|\mathbf{L}\|^2$  are respectively defined on  $\mathbb{R}^{2d}$ ,  $\mathbb{R}^d$  and  $\mathbb{R}$ : for all  $t \in \mathbb{R}^{2d}$ ,  $r \in \mathbb{R}^d$  and  $s \in \mathbb{R}$ ,*

$$\mathbb{E}\left[e^{\langle t, \mathbf{Y} \rangle}\right] < \infty, \quad \mathbb{E}\left[e^{s\|\mathbf{Y}\|^2}\right] < \infty, \quad \mathbb{E}\left[e^{\langle r, \mathbf{L} \rangle}\right] < \infty \quad \text{and} \quad \mathbb{E}\left[e^{s\|\mathbf{L}\|^2}\right] < \infty.$$

To prove this theorem, we use the following extended Carleman's criterion (see [23, Thm 2.3]) on the distribution of  $\mathbf{Y}$ .

**Theorem 3.18** (Extended Carleman's criterion). *Let  $\mathbf{X} = (X_1, \dots, X_D)$  be a random vector with finite moments of all order. If, for any  $i \in \{1, \dots, D\}$ ,  $\sum_{\ell \geq 1} (\mathbb{E}[|X_i|^{2\ell}])^{-1/2\ell} = \infty$ , the distribution of  $\mathbf{X}$  is uniquely determined by its moments.*

*Proof of Theorem 3.17.* We study the process  $\mathbf{Y}$ ; the result for  $\mathbf{L}$  will easily follow from (3.11). We first consider the random vector  $\mathbf{Y}^j = (Y_{j,(k)})_{1 \leq k \leq 2d}$  introduced in Section 3.4, where  $Y_{j,(k)}$  is the  $j$ -th coordinate of the limiting vector  $\mathbf{Y}_{(k)}$  starting with one ball of the  $k$ -th color. As noticed in that section, this vector is solution to the fixed-point equation (3.13), recalled below

$$\mathbf{Y}^j = V^a \mathbf{Y}^{j,(1)} + (1 - V)^a A \mathbf{Y}^{j,(2)},$$

where  $A = (A_{ij})$  has the same law as the random replacement matrix (3.6) of the urn process.

It is known [32, Thm 3.9 (i)] that the asymptotic urn process has finite second-order moment. We also easily note that for all  $\mathbf{y} \in \mathbb{R}^{2d}$ ,  $\|A\mathbf{y}\| = \|\mathbf{y}\|$ . Then, for

$$C_j = \Gamma(a + 1) \mathbb{E} \left[ \|\mathbf{Y}^j\|^2 \right]^{1/2},$$

using exactly the same proof as for Lemma 2.10, we prove by induction that for all  $j \in \{1, \dots, 2d\}$  and all  $\ell \geq 1$ ,

$$\frac{\mathbb{E} \left[ \|\mathbf{Y}^j\|^\ell \right]}{\ell!} \leq \frac{C_j^\ell}{\Gamma(\ell a + 1)} \left( \frac{2}{a} \right)^{\ell-1}. \quad (3.30)$$

Since for all  $k$ ,  $\|Y_{(k)}\| \leq \sum_{j=1}^{2d} \|\mathbf{Y}^j\|$ , taking  $C = \max_{1 \leq j \leq 2d} C_j$ , the moments of  $\mathbf{Y}$  satisfy, for all  $\ell \geq 1$ ,

$$\frac{\mathbb{E} \left[ \|\mathbf{Y}\|^\ell \right]}{\ell!} \leq \frac{2dC^\ell}{\Gamma(\ell a + 1)} \left( \frac{2}{a} \right)^{\ell-1}.$$

As in Theorem 2.8, we deduce that  $\mathbf{Y}$  and  $\|\mathbf{Y}\|$  satisfy Carleman's criterions (Theorems 2.9 and 3.18). We also easily obtain the existence of exponential moments for  $\mathbf{Y}$  and  $\|\mathbf{Y}\|^2$ , using the same proofs as for Corollaries 2.11 and 2.12.  $\square$

## Acknowledgements

This project has benefited from discussions with several colleagues, who we would like to thank for their enthusiasm and their generosity. We especially thank Gerold Alsmeyer, Bernard Bercu, Jean Bertoin, Brigitte Chauvin, Tony Guttmann, Slim Kammoun, Daniela Portillo Del Valle and Nicolas Pouyanne.

## References

- [1] G. Alsmeyer, J. D. Biggins, and M. Meiners. The functional equation of the smoothing transform. *Ann. Probab.*, 40(5):2069–2105, 2012.
- [2] E. Baur and J. Bertoin. The fragmentation process of an infinite recursive tree and Ornstein-Uhlenbeck type processes. *Electron. J. Probab.*, 20:no. 98, 20, 2015.
- [3] E. Baur and J. Bertoin. Elephant random walks and their connection to Pólya-type urns. *Phys. Rev. E*, 94:052134, Nov 2016.
- [4] B. Bercu. A martingale approach for the elephant random walk. *J. Phys. A*, 51(1):015201, 16, 2018.

- [5] B. Bercu, M.-L. Chabanol, and J.-J. Ruch. Hypergeometric identities arising from the elephant random walk. *J. Math. Anal. Appl.*, 480(1):123360, 12, 2019.
- [6] B. Bercu and L. Laulin. On the multi-dimensional elephant random walk. *J. Stat. Phys.*, 175(6):1146–1163, 2019.
- [7] M. Bertenghi. Functional limit theorems for the multi-dimensional elephant random walk. *Stochastic Models*, 38(1):37–50, 2022.
- [8] J. Bertoin. Universality of Noise Reinforced Brownian Motions. In *In and Out of Equilibrium 3: Celebrating Vladas Sidoravicius*, volume 77, pages 147–161. Springer International Publishing, 2021.
- [9] J. Bertoin. Counting the zeros of an elephant random walk. *Trans. Amer. Math. Soc.*, (375):5539–5560, 2022.
- [10] S. Businger. The shark random swim (Lévy flight with memory). *J. Stat. Phys.*, 172(3):701–717, 2018.
- [11] 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.
- [12] B. Chauvin, C. Mailler, and N. Pouyanne. Smoothing equations for large Pólya urns. *J. Theoret. Probab.*, 28(3):923–957, 2015.
- [13] B. Chauvin, N. Pouyanne, and R. Sahnoun. Limit distributions for large Pólya urns. *Ann. Appl. Probab.*, 21(1):1–32, 2011.
- [14] G.-S. Cheon and H. Kim. A new aspect of Hankel matrices via Krylov matrix. *Linear Algebra Appl.*, 438(1):361–373, 2013.
- [15] R. Chhaibi, F. Gamboa, S. Kammoun, and M. Velasco. Estimation of large covariance matrices via free deconvolution: computational and statistical aspects. *arXiv:2305.05646*, 2023.
- [16] C. F. Coletti, R. Gava, and G. M. Schütz. Central limit theorem and related results for the elephant random walk. *J. Math. Phys.*, 58(5):053303, 8, 2017.
- [17] C. F. Coletti, R. Gava, and G. M. Schütz. A strong invariance principle for the elephant random walk. *J. Stat. Mech. Theory Exp.*, (12):123207, 8, 2017.
- [18] C. F. Coletti and I. Papageorgiou. Asymptotic analysis of the elephant random walk. *J. Stat. Mech. Theory Exp.*, (1):Paper No. 013205, 12, 2021.
- [19] J. C. Cressoni, M. A. A. da Silva, and G. M. Viswanathan. Amnestically induced persistence in random walks. *Phys. Rev. Lett.*, 98(7):070603, 4, 2007.
- [20] J. C. Cressoni, G. M. Viswanathan, and M. A. A. da Silva. Exact solution of an anisotropic 2D random walk model with strong memory correlations. *J. Phys. A*, 46(50):505002, 13, 2013.
- [21] N. Curien and L. Laulin. Recurrence of the plane elephant random walk. *arXiv:2310.10314*, 2023.
- [22] M. A. A. Da Silva, J. C. Cressoni, G. M. Schütz, G. M. Viswanathan, and S. Trimper. Non-gaussian propagator for elephant random walks. *Phys. Rev. E*, 88:022115, 2013.
- [23] M. de Jeu. Determinate multidimensional measures, the extended Carleman theorem and quasi-analytic weights. *Ann. Probab.*, 31(3):1205–1227, 2003.
- [24] R. Durrett and T. M. Liggett. Fixed points of the smoothing transformation. *Z. Wahrsch. Verw. Gebiete*, 64(3):275–301, 1983.
- [25] X. Fan, H. Hu, and M. Xiaohui. Cramér moderate deviations for the elephant random walk. *J. Stat. Mech. Theory Exp.*, (2):023402, 2021.
- [26] 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*, volume AG of *Discrete Math. Theor. Comput. Sci. Proc.*, pages 59–118. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2006.

- [27] P. Flajolet, J. Gabarró, and H. Pekari. Analytic urns. *Ann. Probab.*, 33(3):1200–1233, 2005.
- [28] D. A. Freedman. Bernard Friedman’s urn. *Ann. Math. Stat.*, 36:956–970, 1965.
- [29] W. Gautschi. *Orthogonal polynomials: computation and approximation*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2004. Oxford Science Publications.
- [30] M. González-Navarrete. Multidimensional walks with random tendency. *J. Stat. Phys.*, 181(4):1138–1148, 2020.
- [31] A. Iksanov and M. Meiners. Fixed points of multivariate smoothing transforms with scalar weights. *ALEA Lat. Am. J. Probab. Math. Stat.*, 12(1):69–114, 2015.
- [32] S. Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Process. Appl.*, 110(2):177–245, 2004.
- [33] N. Kubota and M. Takei. Gaussian fluctuation for superdiffusive elephant random walks. *J. Stat. Phys.*, 177(6):1157–1171, 2019.
- [34] R. Kürsten. Random recursive trees and the elephant random walk. *Phys. Rev. E*, 93(3):032111, 11, 2016.
- [35] W. V. Li. Lim inf results for the Wiener process and its increments under the  $L_2$ -norm. *Probab. Theory Related Fields*, 92(1):69–90, 1992.
- [36] Q. Liu. Asymptotic properties and absolute continuity of laws stable by random weighted mean. *Stochastic Process. Appl.*, 95(1):83–107, 2001.
- [37] M. Meiners and S. Mentemeier. Solutions to complex smoothing equations. *Probab. Theory Related Fields*, 168(1-2):199–268, 2017.
- [38] R. Neininger and L. Rüschendorf. A general limit theorem for recursive algorithms and combinatorial structures. *Ann. Appl. Probab.*, 14(1):378–418, 2004.
- [39] G. Pólya. Sur quelques points de la théorie des probabilités. *Annales de l’Institut Henri Poincaré*, 1(2):117–161, 1930.
- [40] S. Qin. Recurrence and transience of multidimensional elephant random walks. *arXiv:2309.09795*, 2023.
- [41] R. Roy, M. Takei, and H. Tanemura. The elephant random walk in the triangular array setting. *arXiv:2403.02881*, 2024.
- [42] R. Roy, M. Takei, and H. Tanemura. Phase transitions for a unidirectional elephant random walk with a power law memory. *arXiv:2402.02048*, 2024.
- [43] G. M. Schütz and S. Trimper. Elephants can always remember: Exact long-range memory effects in a non-markovian random walk. *Phys. Rev. E*, 70:045101, Oct 2004.
- [44] B. Simon. The classical moment problem as a self-adjoint finite difference operator. *Adv. Math.*, 137(1):82–203, 1998.
- [45] V. H. Vázquez Guevara. On the almost sure central limit theorem for the elephant random walk. *J. Phys. A*, 52(1):475201, 2019.