CRITICAL MULTI-TYPE GALTON-WATSON TREES CONDITIONED TO BE LARGE

ROMAIN ABRAHAM, JEAN-FRANÇOIS DELMAS, AND HONGSONG GUO

Abstract. Under minimal condition, we prove the local convergence of a critical multi-type Galton-Watson tree conditioned on having a large total progeny by types towards a multi-type Kesten's tree. We obtain the result by generalizing Neveu's strong ratio limit theorem for aperiodic random walks on $\mathbb{Z}^d$.

1. Introduction

In [13], Kesten gives that the local limit of a critical or subcritical Galton-Watson (GW) tree conditioned on having a large height is an infinite GW tree (in fact a multi-type GW tree with one special individual per generation) with a unique infinite spine, which we shall call *Kesten's tree* in the present paper. In Abraham and Delmas [2] a sufficient and necessary condition is given for a wide class of conditionings for a critical GW tree to converge locally to Kesten's tree under minimal hypotheses on the offspring distribution. Notice that condensation may arise when considering sub-critical GW trees, see Janson [11], Jonnson and Stefansson [12], He [8] or Abraham and Delmas [1] for results in this direction. When scaling limit of multi-type GW tree are considered, one obtains as a limit a continuous GW tree, see Miermont [16] or Gorostiza and Lopez-Mimbela [15] (when the probability to give birth to different types goes down to 0). In this latter case see Delmas and Hénard [6] for the limit on the conditioned random tree to have a large height.

In the multi-type case, Pénisson [18] proved that a critical d-types GW process conditioned on the total progeny to be large with a given asymptotic proportion per types converges locally to a multi-type GW process (with a special individual per generation) under the condition that the branching process admits moments of order $d+1$. Stephenson [23] gave, under exponential moments condition, the local convergence of a multi-type GW tree conditioned on a large population for some type towards the multi-type Kesten's tree introduced by Kurtz, Lyons, Pemantle and Peres [14]. The aim of this paper is to give minimal hypotheses to insure the local convergence of a critical multi-type GW tree conditioned on the total progeny to be large towards the associated multi-type Kesten's tree, see Theorem 3.1. The minimal hypotheses are the existence of the mean matrix which is assumed to be primitive and an aperiodic condition on the offspring distribution. Furthermore, we exactly condition by the asymptotic proportion per types for the total progeny of the GW tree to be given by the (normalized) left eigenvector associated with the Perron-Frobenius eigenvalue of the mean matrix.

If the asymptotic proportion per types is not equal to the (normalized) left eigenvector associated with the Perron-Frobeniux eigenvalue of the mean matrix, then under exponential moments

Date: November 5, 2015.
2010 *Mathematics Subject Classification.* 60J80; 60B10.
Key words and phrases. Galton-Watson; random tree; local-limit; strong ratio theorem; branching process.

condition for the offspring distribution, it is possible to get a Kesten's tree as local limit, see [18]. However, without exponential moments condition for the offspring distribution no results are known, and results in [1] for the mono-type case suggests a condensation phenomenon (at least in the sub-critical case). Conditioning large multi-type (or even mono-type) continuous GW tree to have a large population in the spirit of [6] is also an open question.

The proof of Theorem 3.1 relies on two arguments. The first one is a generalization of Dwass formula for multi-type GW processes given by Chaumont and Liu [5] which encodes critical or sub-critical d-multi-type GW forests using d random walks of dimension d. The second one is the strong ratio theorem for random walks in $\mathbb{Z}^d$, see Theorem 4.7, which generalizes a result by Neveu [17] in dimension one. The proof of the strong ratio theorem relies on a uniform version of the d-dimensional local theorem of Gnedenko [7], see also Gnedenko and Kolmogorov [7] (for the sum of independent random variables), Rvaceva [21] (for the sum of d-dimensional i.i.d. random variables) or Stone [24] (for the sum of d-dimensional i.i.d. lattice or non lattice random variables), which is given in Section 4.2, and properties of the Legendre Laplace transform of a probability distribution. As we were unable to find those latter properties in the literature, we give them in a general framework in Section 4.1, as we believe they might be interesting by themselves.

The paper is organized as follows. We present in Section 2 the topology of the multi-type trees and a sufficient and necessary condition for the local convergence of random multi-type trees, see Corollary 2.2, the definition of multi-type GW tree with a given offspring distribution and the aperiodicity condition on the offspring distribution, see Definition 2.6. Section 3.1 is devoted to the main result, Theorem 3.1, and its proof. The Appendix collects results on the Legendre Laplace transform in a general framework in Section 4.1, Gnedenko's d-dimensional local theorem in Section 4.2, and the strong ratio limit theorem for d-dimensional random walks in Section 4.3.

2. Multi-type trees

2.1. **General notations.** We denote by $\mathbb{N} = \{0, 1, 2, \ldots\}$ the set of non-negative integers and by $\mathbb{N}^* = \{1, 2, \ldots\}$ the set of positive integers. For $d \in \mathbb{N}^*$, we set $[d] = \{1, \ldots, d\}$.

Let $d \geq 1$. We say $x = (x_i, i \in [d]) \in \mathbb{R}^d$ is a column vector in $\mathbb{R}^d$. We write $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^d$, $0 = (0, \ldots, 0) \in \mathbb{R}^d$ and denote by $\mathbf{e}_i$ the vector such that the i-th element is 1 and others are 0. For vectors $x = (x_i, i \in [d]) \in \mathbb{R}^d$ and $y = (y_i, i \in [d]) \in \mathbb{R}^d$, we denote by $\langle x, y\rangle$ the usual scalar product of x and y, by x^y the product $\prod_{i=1}^d x_i^{y_i}$, by $|x| = \sum_{i=1}^d |x_i|$ and $\| x \| = \sqrt{\langle x, x\rangle}$ the L^1 and L^2 norms of x, and we write $x \leq y$ (resp. $x < y$) if $x_i \leq y_i$ (resp. $x_i < y_i$) for all $i \in [d]$.

For any nonempty set $A \subset \mathbb{R}^d$, we define span A as the linear sub-space generated by A (that is span $A = \{\sum_{i=1}^n \alpha_i y_i;\, \alpha_i \in \mathbb{R}, y_i \in A, i \in [n], n \in \mathbb{N}^*\}$) and for $x \in \mathbb{R}^d$, we denote $x + A = \{x + y; y \in A\}$.

For a random variable X and an event A, we write $\mathbb{E}[X;\, A]$ for $\mathbb{E}\left[X \mathbf{1}_A\right]$.

2.2. **Notations for marked trees.** Let $d \in \mathbb{N}^*$. Denote by $[d]$ the set of types or marks, by $\widehat{\mathcal{U}} = \bigcup_{n\geq 0} (\mathbb{N}^*)^n$ the set of finite sequences of positive integers with the convention $(\mathbb{N}^*)^0 = \{\widehat{\emptyset}\}$ and by $\mathcal{U} = \bigcup_{n\geq 0} \Big((\mathbb{N}^*)^n \times [d]\Big)$ the set of finite sequences of positive integers with a type. For a marked individual $u \in \mathcal{U}$, we write $u = (\hat{u}, \mathcal{M}(u))$ with $\hat{u} \in \widehat{\mathcal{U}}$ the individual and $\mathcal{M}(u) \in [d]$ its type or mark. Let $|u| = |\hat{u}|$ be the length or height of u defined as the integer n such that $\hat{u} = (u_1, \ldots, u_n) \in (\mathbb{N}^*)^n$. If $\hat{u}$ and $\hat{v}$ are two sequences in $\widehat{\mathcal{U}}$, we denote by $\hat{u}\hat{v}$ the concatenation of the two sequences, with the convention that $\hat{u}\hat{v} = \hat{u}$ if $\hat{v} = \widehat{\emptyset}$ and $\hat{u}\hat{v} = \hat{v}$ if $\hat{u} = \widehat{\emptyset}$. For $u, v \in \mathcal{U}$, we denote by uv the concatenation of u and v such that $\widehat{uv} = \hat{u}\hat{v}$ and $\mathcal{M}(uv) = \mathcal{M}(v)$ if $|v| \geq 1$; $\mathcal{M}(uv) = \mathcal{M}(u)$ if $|v| = 0$. Let $u, v \in \mathcal{U}$. We say that v (resp. $\hat{v}$) is an ancestor of u (resp. $\hat{u}$) and write $v \preccurlyeq u$ (resp. $\hat{v} \preccurlyeq \hat{u}$) if there exists $w \in \mathcal{U}$ such that $u = vw$ (resp. $\hat{w} \in \hat{\mathcal{U}}$ such that $\hat{u} = \hat{v}\hat{w}$).

A tree $\hat{\mathbf{t}}$ is a subset of $\hat{\mathcal{U}}$ such that:

- $\widehat{\emptyset} \in \hat{\mathbf{t}}$.
- If $\hat{u} \in \hat{\mathbf{t}}$, then $\{\hat{v}; \hat{v} \preccurlyeq \hat{u}\} \subset \hat{\mathbf{t}}$.
- For every $\hat{u} \in \hat{\mathbf{t}}$, there exists $k_{\hat{u}}[\hat{\mathbf{t}}] \in \mathbb{N}$ such that, for every positive integer ℓ, $\hat{u}\ell \in \hat{\mathbf{t}}$ iff $1 \leq \ell \leq k_{\hat{u}}[\hat{\mathbf{t}}]$.

A marked tree $\mathbf{t}$ is a subset of $\mathcal{U}$ such that:

(a) The set $\hat{\mathbf{t}} = \{\hat{u}; u \in \mathbf{t}\}$ of (unmarked) individuals of $\mathbf{t}$ is a tree.
(b) There is only one type per individual: for $u, v \in \mathbf{t}$, $\hat{u} = \hat{v}$ implies $\mathcal{M}(u) = \mathcal{M}(v)$ and thus $u = v$.

Thanks to (b), the number of offsprings of the marked individual $u \in \mathbf{t}$, $k_u[\mathbf{t}]$, corresponds to $k_{\hat{u}}[\hat{\mathbf{t}}]$. In what follows we will deal only with marked trees and call them only trees.

Denote by $\emptyset_{\mathbf{t}} = (\widehat{\emptyset}, \mathcal{M}(\emptyset_{\mathbf{t}})) \in \mathcal{U}$ the root of the tree $\mathbf{t}$ and write $\emptyset$ instead of $\emptyset_{\mathbf{t}}$ when the context is clear. The parent of $v \in \mathbf{t} \setminus \emptyset_{\mathbf{t}}$ in $\mathbf{t}$, denoted by $\mathrm{Pa}_v(\mathbf{t})$ is the only $u \in \mathbf{t}$ such that $|u| = |v| - 1$ and $u \preccurlyeq v$. The set of the children of $u \in \mathbf{t}$ is

$$C_u(\mathbf{t}) = \{v \in \mathbf{t},\ \mathrm{Pa}_v(\mathbf{t}) = u\}.$$

Notice that $k_u[\mathbf{t}] = \mathrm{Card}\ (C_u(\mathbf{t}))$ for $u \in \mathbf{t}$. We set $k_u(\mathbf{t}) = (k_u^{(i)}[\mathbf{t}], i \in [d])$, where for $i \in [d]$

$$k_u^{(i)}[\mathbf{t}] = \mathrm{Card}\ (\{v \in C_u(\mathbf{t});\ \mathcal{M}(v) = i\})$$

is the number of offsprings of type i of $u \in \mathbf{t}$. We have $\sum_{i\in[d]} k_u^{(i)}[\mathbf{t}] = k_u[\mathbf{t}]$. The vertex $u \in \mathbf{t}$ is called a leaf if $k_u[\mathbf{t}] = 0$ and let $\mathcal{L}_0(\mathbf{t}) = \{u \in \mathbf{t}, k_u[\mathbf{t}] = 0\}$ be the set of leaves of $\mathbf{t}$.

We denote by $\mathbb{T}$ the set of marked trees. For $\mathbf{t} \in \mathbb{T}$, we define $|\mathbf{t}| = (|\mathbf{t}^{(i)}|, i \in [d])$ with $|\mathbf{t}^{(i)}| = \mathrm{Card}\ (\{u \in \mathbf{t}, \mathcal{M}(u) = i\})$ the number of individuals in $\mathbf{t}$ of type i. Let us denote by $\mathbb{T}_0 = \{\mathbf{t} \in \mathbb{T} : \mathrm{Card}\ (\mathbf{t}) < \infty\}$ the subset of finite trees. We say that a sequence $\mathbf{v} = (v_n, n \in \mathbb{N}) \subset \mathcal{U}$ is an infinite spine if $v_n \preccurlyeq v_{n+1}$ and $|v_n| = n$ for all $n \in \mathbb{N}$. We denote by $\mathbb{T}_1$ the subset of trees which have one and only one infinite spine. For $\mathbf{t} \in \mathbb{T}_1$, denote by $\mathbf{v}_{\mathbf{t}}$ the infinite spine of the tree $\mathbf{t}$. Let $\mathbb{T}_1'$ be the subset of $\mathbb{T}_1$ such that the infinite spine has infinitely many times all the types:

$$\mathbb{T}_1' = \{\mathbf{t} \in \mathbb{T}_1;\ \forall i \in [d],\ \mathrm{Card}\ (\{v \in \mathbf{v}_{\mathbf{t}};\ \mathcal{M}(v) = i\}) = \infty\}.$$

The height of a tree $\mathbf{t}$ is defined by $H(\mathbf{t}) = \sup\{|u|, u \in \mathbf{t}\}$. For $h \in \mathbb{N}$, we denote by $\mathbb{T}^{(h)} = \{\mathbf{t} \in \mathbb{T}; H(\mathbf{t}) \leq h\}$ the subset of marked trees with height less than or equal to h.

2.3. **Convergence determining class.** For $h \in \mathbb{N}$, the restriction function r_h from $\mathbb{T}$ to $\mathbb{T}$ is defined by $r_h(\mathbf{t}) = \{u \in \mathbf{t}, |u| \leq h\}$. We endow the set $\mathbb{T}$ with the ultra-metric distance $d(\mathbf{t}, \mathbf{t}') = 2^{-\max\{h\in\mathbb{N}, r_h(\mathbf{t}) = r_h(\mathbf{t}')\}}$. The Borel σ-field associated with the distance d is the smallest σ-field containing the singletons for which the restrictions $(r_h, h \in \mathbb{N})$ are measurable. With this distance, the restriction functions are continuous. Since $\mathbb{T}_0$ is dense in $\mathbb{T}$ and $(\mathbb{T}, d)$ is complete, we get that $(\mathbb{T}, d)$ is a Polish metric space.

Let $\mathbf{t}, \mathbf{t}' \in \mathbb{T}$ and $x \in \mathcal{L}_0(\mathbf{t})$. If the type of the root of $\mathbf{t}'$ is $\mathcal{M}(x)$, we denote by

$$\mathbf{t} \circledast (\mathbf{t}', x) = \mathbf{t} \cup \{xv, v \in \mathbf{t}'\}$$

the tree obtained by grafting the tree $\mathbf{t}'$ on the leaf x of the tree $\mathbf{t}$; otherwise, let $\mathbf{t} \otimes (\mathbf{t}', x) = \mathbf{t}$. Then we consider

$$\mathbb{T}(\mathbf{t}, x) = \{\mathbf{t} \otimes (\mathbf{t}', x), \mathbf{t}' \in \mathbb{T}\}$$

the set of trees obtained by grafting a tree on the leaf x of $\mathbf{t}$. It is easy to see that $\mathbb{T}(\mathbf{t}, x)$ is closed and also open.

Set $\mathcal{F} = \{\mathbb{T}(\mathbf{t}, x);\ \mathbf{t} \in \mathbb{T}_0, x \in \mathcal{L}_0(\mathbf{t})$ and $\mathcal{M}(\emptyset_{\mathbf{t}}) = \mathcal{M}(x)\} \cup \{\{\mathbf{t}\};\ \mathbf{t} \in \mathbb{T}_0\}$. Following the proof of Lemma 2.1 in [2], it is easy to get the following result.

Lemma 2.1. *The family $\mathcal{F}$ is a convergence determining class on $\mathbb{T}_0 \cup \mathbb{T}_1'$.*

We deduce the following corollary.

Corollary 2.2. *Let $(T_n, n \in \mathbb{N}^*)$ and T be random variables taking values in $\mathbb{T}_0 \bigcup \mathbb{T}_1'$. Then the sequence $(T_n, n \in \mathbb{N}^*)$ converges in distribution towards T if and only if we have for all $\mathbf{t} \in \mathbb{T}_0$ $\lim_{n\to+\infty} \mathbb{P}(T_n = \mathbf{t}) = \mathbb{P}(T = \mathbf{t})$ and for all $x \in \mathcal{L}_0(\mathbf{t})$ such that $\mathcal{M}(\emptyset_{\mathbf{t}}) = \mathcal{M}(x)$:*

$$\lim_{n\to+\infty} \mathbb{P}(T_n \in \mathbb{T}(\mathbf{t}, x)) = \mathbb{P}(T \in \mathbb{T}(\mathbf{t}, x)).$$

2.4. **Aperiodic distribution.** Let us consider a probability distribution $F = (F(x), x \in \mathbb{Z}^d)$ on $\mathbb{Z}^d$. In order to avoid degenerate cases, we assume that there exists $x_0 \in \mathbb{Z}^d$ such that:

$$0 < F(x_0) < 1. \tag{1}$$

Denote by $\text{supp}\,(F) = \{x \in \mathbb{Z}^d, F(x) > 0\}$ the support set of F.

Definition 2.3. *A distribution F on $\mathbb{Z}^d$ is called aperiodic if it has the property that for each $x \in \mathbb{Z}^d$, the smallest subgroup of $\mathbb{Z}^d$ which contains the set $x + \text{supp}\,(F)$ is $\mathbb{Z}^d$ itself.*

Remark 2.4. An aperiodic distribution is called strongly aperiodic in [22, p.42].

Let $x_0 \in \text{supp}\,(F)$ and denote by R_0 the smallest subgroup of $\mathbb{Z}^d$ that contains $-x_0+\text{supp}\,(F)$. The following lemma is elementary.

Lemma 2.5. *The subgroup R_0 does not dependent on x_0, and F is aperiodic if and only if $R_0 = \mathbb{Z}^d$.*

Proof. For $x \in \mathbb{Z}^d$, let G_x be the smallest subgroup of $\mathbb{Z}^d$ that contains $-x + \text{supp}\,(F)$. Let $z \in G_{x_0}$. There exists $n, n' \in \mathbb{N}$ and $y_i, y_i' \in \text{supp}\,(F)$ for all $i \in \mathbb{N}^*$ such that $\sum_{i=1}^n (y_i - x_0) - \sum_{i=1}^{n'} (y_i' - x_0) = z$. This implies that $\{n'x_0 + \sum_{i=1}^n y_i\} - \{nx_0 + \sum_{i=1}^{n'} y_i'\} = z$. We deduce that for all $x \in \mathbb{Z}^d$, $\{n'(x_0 - x) + \sum_{i=1}^n (y_i - x)\} - \{n(x_0 - x) + \sum_{i=1}^{n'} (y_i' - x)\} = z$. This gives that z belongs to G_x. In particular we deduce that $G_x = G_{x_0}$ for any $x \in \text{supp}\,(F)$. This implies that R_0 does not depend on x_0. Furthermore $R_0 = \mathbb{Z}^d$ implies that $G_x = \mathbb{Z}^d$ for all $x \in \mathbb{Z}^d$, that is F is aperiodic. Clearly if F is aperiodic then $R_0 = \mathbb{Z}^d$. □

2.5. **Multi-type offspring distribution.** We define a multi-type offspring distribution p of d types as a sequence of probability distributions: $p = (p^{(i)}, i \in [d])$, with $p^{(i)} = (p^{(i)}(k), k \in \mathbb{N}^d)$ a probability distribution on $\mathbb{N}^d$. Denote by $f = (f^{(1)}, \ldots, f^{(d)})$ the generating function of the offspring distribution p, i.e. for $i \in [d]$ and $s \in [0,1]^d$:

$$f^{(i)}(s) = \mathbb{E}[s^{X_i}], \tag{2}$$

with $X_i = (X_i^{(j)}, j \in [d])$ a random variable on $\mathbb{N}^d$ with distribution $p^{(i)}$. Denote by $m_{ij} = \partial_{s_j} f^{(i)}(\mathbf{1}) = \mathbb{E}[X_i^{(j)}] \in [0, +\infty]$ the expected number of offsprings with type j of a single individual of type i. Denote by M the mean matrix $M = (m_{ij};\ i, j \in [d])$ and set $(m_{ij}^{(n)};\ i, j \in [d]) = M^n$ for $n \in \mathbb{N}^*$. Following [3, p.184], we say that:

- p is non-singular if $f(s) \neq Ms$.
- M is finite if $m_{ij} < +\infty$ for all $i, j \in [d]$.
- M is primitive if M is finite and there exists $n \in \mathbb{N}^*$ such that for all $i, j \in [d]$, $m_{ij}^{(n)} > 0$.

By Frobenius theorem, see [3, p.185], if M is primitive, then M has a unique maximal (for the modulus in $\mathbb{C}$) eigenvalue ρ. Furthermore ρ is simple, positive ($\rho \in (0, +\infty)$), and the corresponding right and left eigenvectors can be chosen positive. If $\rho = 1$ (resp. $\rho > 1$, $\rho < 1$), we say that the offspring distribution and the associated multi-type GW tree are critical (resp. supercritical, subcritical).

Recall the definition of an aperiodic distribution given in Definition 2.3.

Definition 2.6. *Let $p = (p^{(i)}, i \in [d])$ be an offspring distribution and let $x_i \in \text{supp}\,(p^{(i)})$ for all $i \in [d]$. We say that p is aperiodic, if the smallest subgroup of $\mathbb{Z}^d$ that contains $-x_i + \text{supp}\,(p^{(i)})$ for all $i \in [d]$ is $\mathbb{Z}^d$.*

According to Lemma 2.5, we get that the definition of aperiodic offspring distribution does not depend on the choice of $x_i \in \text{supp}\,(p^{(i)})$.

For an offspring distribution p, we shall consider the following assumptions:

(H_1) **The mean matrix M of p is primitive, and p is critical and non-singular.**
(H_2) **The offspring distribution p is aperiodic.**

2.6. **Multi-type Galton-Watson tree and Kesten's tree.** We define the multi-type GW tree τ with offspring distribution p.

Definition 2.7. *Let p be an offspring distribution of d types and α a probability distribution on $[d]$. A $\mathbb{T}$-valued random variable τ is a multi-type GW tree with offspring distribution p and root type distribution α, if for all $h \in \mathbb{N}$, $\mathbf{t} \in \mathbb{T}^{(h)}$, we have:*

$$\mathbb{P}_\alpha(r_h(\tau) = \mathbf{t}) = \alpha(\mathcal{M}(\emptyset_\mathbf{t})) \prod_{u \in \mathbf{t}, |u| < h} \frac{k_u^{(1)}[\mathbf{t}]! \cdots k_u^{(d)}[\mathbf{t}]!}{k_u[\mathbf{t}]!} \, p^{(\mathcal{M}(u))}(k_u(\mathbf{t})).$$

We deduce from the definition that for $\mathbf{t} \in \mathbb{T}_0$, we have

$$\mathbb{P}_\alpha(\tau = \mathbf{t}) = \alpha(\mathcal{M}(\emptyset_\mathbf{t})) \prod_{u \in \mathbf{t}} \frac{k_u^{(1)}[\mathbf{t}]! \cdots k_u^{(d)}[\mathbf{t}]!}{k_u[\mathbf{t}]!} \, p^{(\mathcal{M}(u))}(k_u(\mathbf{t})).$$

The multi-type GW tree enjoys the branching property: an individual of type i generates children according to $p^{(i)}$ independently of any born individual, for $i \in [d]$.

Let p be an offspring distribution of d types such that (H_1) holds. Denote by a^* (resp. a) the right (resp. left) positive normalized eigenvector of M such that $\langle a, \mathbf{1} \rangle = \langle a, a^* \rangle = 1$. Notice that a is a probability distribution on $[d]$. The corresponding size-biased offspring distribution $\hat{p} = (\hat{p}^{(i)}, i \in [d])$ is defined by: for $i \in [d]$ and $k \in \mathbb{N}^d$,

$$\hat{p}^{(i)}(k) = \frac{\langle k, a^* \rangle}{a_i^*} \, p^{(i)}(k). \tag{3}$$

Definition 2.8. *Let p be an offspring distribution of d types whose mean matrix is primitive and let α be a probability distribution on $[d]$. A multi-type Kesten's tree τ^* associated with the offspring distribution p and root distribution α is defined as follows:*

- *Marked individuals are normal or special.*
- *The root of τ^* is special and its type has distribution α.*
- *A normal individual of type $i \in [d]$ produces only normal individuals according to $p^{(i)}$.*

- *A special individual of type $i \in [d]$ produces children according to $\hat{p}^{(i)}$. One of those children, chosen with probability proportional to a_j^* where j is its type, is special. The others (if any) are normal.*

Notice that the multi-type Kesten's tree is a multi-type GW tree (with $2d$ types). The individuals which are special in τ^* form an infinite spine, say $\mathbf{v}^*$, of τ^*; and the individuals of $\tau^* \backslash \mathbf{v}^*$ are normal.

Let $r \in [d]$. We shall write $\mathbb{P}_r(d\tau)$, resp. $\mathbb{P}_r(d\tau^*)$, for the distribution of τ, resp. τ^*, when the type of its root is r (that is $\alpha = \delta_r$ the Dirac mass at r). From [14], we get that for $h \in \mathbb{N}$, $\mathbf{t} \in \mathbb{T}^{(h)}$ with $\mathcal{M}(\emptyset_\mathbf{t}) = r$, and $x \in \mathcal{L}_0(t)$ with $|x| = h$ and $\mathcal{M}(x) = i$:

$$\mathbb{P}_r(r_h(\tau^*) = \mathbf{t},\, v_h^* = x) = \frac{a_i^*}{a_r^*}\, \mathbb{P}_r(r_h(\tau) = \mathbf{t}). \tag{4}$$

Notice that if M is primitive and p is critical or sub-critical, then a.s. Kesten's tree τ^* belongs to $\mathbb{T}_1$. The next lemma asserts that there are infinitely many individuals of all types on the infinite spine.

Lemma 2.9. *Let p be an offspring distribution of d types satisfying (H_1) and α a probability distribution on $[d]$. Then a.s. the multi-type Kesten tree τ^* belongs to $\mathbb{T}_1'$.*

Proof. Recall that $a^* = (a_i^*, i \in [d])$ is the normalized right eigenvalue of M such that $\langle a^*, a \rangle = 1$. By construction, the sequence $(\mathcal{M}(v_n^*), n \in \mathbb{N})$ is a Markov chain on $[d]$ and transition matrix $Q = (Q_{i,j},\, i, j \in [d])$ given by

$$Q_{i,j} = \mathbb{P}(\mathcal{M}(v_1^*) = j |\, \mathcal{M}(v_0^*) = i) = \sum_{k=(k_1,\ldots,k_d)\in\mathbb{N}^d} \frac{k_j a_j^*}{\langle k, a^* \rangle}\, \hat{p}^{(i)}(k) = \frac{a_j^*}{a_i^*}\, m_{i,j},$$

where we used (3) for the definition of $\hat{p}$ and the definition of the mean matrix M for the last equality. Since a^* is positive and M is primitive, we deduce that Q is also primitive. This implies that the Markov chain $(\mathcal{M}(v_n^*), n \in \mathbb{N})$ is recurrent on $[d]$ and hence it visits a.s. infinitely many times all the states of $[d]$. □

The next lemma will be used in the proof of the Theorem 3.1. In the next lemma, we shall consider a leaf x of a finite tree $\mathbf{t}$ with type i and the root of type r. However, we will only use the case $i = r$ in the proof of the Theorem 3.1.

Lemma 2.10. *Let p be an offspring distribution of d types satisfying (H_1) and $r \in [d]$. Let τ be a GW tree with offspring distribution p and τ^* be a Kesten's tree associated with p. For all $\mathbf{t} \in \mathbb{T}_0$ with $\mathcal{M}(\emptyset_\mathbf{t}) = r$, $x \in \mathcal{L}_0(t)$ with $\mathcal{M}(x) = i \in [d]$, and $k \in \mathbb{N}^d$ such that $k \geq |\mathbf{t}|$, we have:*

$$\mathbb{P}_r(\tau \in \mathbb{T}(\mathbf{t}, x) \big|\, |\tau| = k) = \frac{a_r^*}{a_i^*}\, \frac{\mathbb{P}_i(|\tau| = k - |\mathbf{t}| + \mathbf{e}_i)}{\mathbb{P}_r(|\tau| = k)}\, \mathbb{P}_r(\tau^* \in \mathbb{T}(\mathbf{t}, x)). \tag{5}$$

Proof. Since τ^* has a unique infinite spine $\mathbf{v}^*$ and $\mathbf{t} \in \mathbb{T}_0$, we deduce that $\tau^* \in \mathbb{T}(\mathbf{t}, x)$ implies that x belongs to $\mathbf{v}^*$ and we get in the same spirit of (4) that:

$$\mathbb{P}_r(\tau^* \in \mathbb{T}(\mathbf{t}, x)) = \frac{a_i^*}{a_r^*}\, \mathbb{P}_r(\tau \in \mathbb{T}(\mathbf{t}, x)). \tag{6}$$

We have, following the ideas of [2]:

$$\begin{aligned}\mathbb{P}_r(\tau \in \mathbb{T}(\mathbf{t}, x),\, |\tau| = k) &= \sum_{\mathbf{t}' \in \mathbb{T}_0} \mathbb{P}_r(\tau = \mathbf{t} \otimes (\mathbf{t}', x)) \mathbf{1}_{\{|\mathbf{t} \otimes (\mathbf{t}', x)| = k\}} \\ &= \sum_{\mathbf{t}' \in \mathbb{T}_0} \mathbb{P}_r(\tau \in \mathbb{T}(\mathbf{t}, x))\, \mathbb{P}_i(\tau = \mathbf{t}') \mathbf{1}_{\{|\mathbf{t} \otimes (\mathbf{t}', x)| = k\}} \\ &= \mathbb{P}_r(\tau \in \mathbb{T}(\mathbf{t}, x)) \sum_{\mathbf{t}' \in \mathbb{T}_0} \mathbb{P}_i(\tau = \mathbf{t}') \mathbf{1}_{\{|\mathbf{t}'| = k - |\mathbf{t}| + \mathbf{e}_i\}} \\ &= \mathbb{P}_r(\tau \in \mathbb{T}(\mathbf{t}, x)) \mathbb{P}_i(|\tau| = k - |\mathbf{t}| + \mathbf{e}_i),\end{aligned}$$

where we used the branching property of the multi-type GW tree for the second equality. Use (6) to deduce (5). □

3. Main results

3.1. **Conditioning on the total population size.** Recall that under (H_1), we denote by $a = (a_\ell, \ell \in [d])$ and $a^* = (a^*_\ell, \ell \in [d])$ the normalized left and right eigenvectors of the mean matrix M associated with the eigenvalue $\rho = 1$ such that $\langle a, a^* \rangle = \sum a_i = 1$.

Theorem 3.1. *Assume that (H_1) and (H_2) hold. Let $(k(n), n \in \mathbb{N}^*)$ be a sequence of $\mathbb{N}^d$ satisfying $\lim_{n\to\infty} |k(n)| = +\infty$ and $\lim_{n\to\infty} k(n)/|k(n)| = a$. Let τ be a random GW tree with critical offspring distribution p and τ_n be distributed as τ conditionally on $\{|\tau| = k(n)\}$. Then the sequence $(\tau_n, n \in \mathbb{N}^*)$ converges in distribution to the Kesten's tree τ^* associated with p.*

Remark 3.2. Let τ be a critical GW tree with offspring distribution p satisfying (H_1). We can consider τ conditionally on the event the population of type i, $|\tau^{(i)}|$, is large. According to Proposition 4 in [16], the random variable $|\tau^{(i)}|$ is distributed as the total number of vertices of a critical mono-type GW tree under $\mathcal{M}_\tau(\emptyset) = i$, or as the total number of vertices of a random number of independent mono-type critical GW trees with the same distribution under $\mathcal{M}_\tau(\emptyset) \neq i$. In particular, we deduce from [2] that, if $p^{(i)}$ is aperiodic, the key equality $\lim_{n\to+\infty} \mathbb{P}(|\tau^{(i)}| = n - b)/\mathbb{P}_r(|\tau^{(i)}| = n) = 1$ holds for any $b \in \mathbb{Z}$. And following the proof of Theorem 3.1 after Equation (15), we easily get that τ conditioned on $|\tau^{(i)}|$ being large converges locally to Kesten's tree. See [23] for a detailed proof.

Remark 3.3. The local convergence of a multi-type critical GW tree τ conditioned on the number of vertices of one fixed type being large to a Kesten's tree has been proved in [23]. It would be easy to extend Theorem 3.1, with the same minimal conditions (H_1) and (H_2) to a conditioning on an asymptotic proportion per types for d' types, with $d' < d$ by using the constructions from [19] or from [16]. The idea is to map a multi-type GW tree τ with d types onto another GW tree τ' with $d' < d$ types and offspring distribution p' so that the size of the population of types 1 to d' of τ and τ' are the same. Then the key Equation (15) is now replaced by the one for τ' which holds if the offspring distribution p' satisfies (H_1) and (H_2). Then the proof follows as in the proof of Theorem 3.1 after Equation (15).

Remark 3.4. Theorem 3.1 can be used to extend results of [1] on mono-type GW tree in the following sense. Let τ be a mono-type critical GW tree (that is $d = 1$) with offspring distribution q. Assume that $q(0) + q(1) < 1$ to avoid degenerated case and that q is aperiodic. Let $A_1, \ldots, A_d$ be pairwise disjoint subsets of $\mathbb{N}$ such that $q(A_i) > 0$ for all $i \in [d]$ and for simplicity assume $\sum_{i\in[d]} q(A_i) = 1$. (If this latter case is not satisfied, set $A_{d+1} = \mathbb{N} \setminus \bigcup_{i\in[d]} A_i$, and then use the restriction method presented in Remark 3.3.) In order to avoid degenerated cases, we shall assume that $q(A_i) > q(0)$ if $0 \in A_i$ (*i.e.* $0 \in A_i$ and there exists $\ell \in A_i$ such that $\ell > 0$ and $q(\ell) > 0$).

Then consider artificially that τ is a d dimensional multi-type GW tree, by saying that an individual $u \in \tau$ is of type i if the number of offspring of u lies in A_i. Notice that the corresponding offspring distribution $p = (p^{(i)}, i \in [d])$ is defined as follows: for $k = (k_1, \ldots, k_d) \in \mathbb{N}^d$,

$$p^{(i)}(k) = \mathbf{1}_{\{|k| \in A_i\}} \frac{q(|k|)}{q(A_i)} \frac{|k|!}{\prod_{i\in[d]} k_i!} \prod_{j\in[d]} q(A_j)^{k_j},$$

where we recall that $|k| = \sum_{i\in[d]} k_i$. Since $q(A_i) > q(0)$ if $0 \in A_i$, we deduce the mean matrix is positive and thus primitive. Its Perron-Frobenius eigenvalue is 1 since τ is a critical mono-type GW tree which implies that it is also a critical multi-type GW tree. We get the condition (H_1) holds. Notice (H_2) holds as q is aperiodic. One can easily check that the (normalized) left eigenvector associated with the Perron-Frobenius eigenvalue is $(q(A_i), i \in [d])$.

For simplicity we shall design the corresponding multi-type GW tree by τ. In particular $|\tau^{(i)}|$ denotes the number of individuals in τ with number of children in A_i, and $|\tau| = (|\tau^{(i)}|, i \in [d])$. So we easily deduce from Theorem 3.1 that if $(k(n), n \in \mathbb{N}^*)$ is a sequence of $\mathbb{N}^d$ satisfying $\lim_{n\to\infty} k_i(n)/|k(n)| = q(A_i)$ for $i \in [d]$ with $\lim_{n\to+\infty} |k(n)| = +\infty$, then τ_n, which is distributed as τ conditionally on $\{|\tau| = k(n)\}$, converges in distribution to the (mono-type) Kesten's tree τ^* associated with q.

If we condition on an asymptotic proportion different from $(q(A_i), i \in [d])$ then it is possible to use Remark 3.5 see also the Remark 3.6 for explicit computation in the binary case.

Remark 3.5. The change of offspring distribution given in Section 1.4 of [18], when it exits, allows to extend Theorem 3.1 to sub-critical multi-type GW trees. In order to consider an asymptotic proportion per types different from the one given by the (normalized) left eigenvector associated with the Perron-Frobenius eigenvalue, one has to change the offspring distribution, see Theorem 3 of [18]. However, this requires exponential moments for the offspring distribution.

Remark 3.6. As an example, we consider the local convergence of binary GW tree conditioned to have large population of nodes of out-degree in $\{0, 1\}$ or in $\{2\}$. In particular, we shall see that the distribution of the limiting Kesten's tree depends only on the asymptotic ratio of the number of nodes of out-degree in $\{0, 1\}$ or in $\{2\}$.

Let q be an offspring distribution on $\{0, 1, 2\}$ such that $q(0)+q(1)+q(2) = 1$ and $q(0)q(1)q(2) > 0$ (we don't assume that q is critical, sub-critical or super-critical). Set $A_1 = \{0, 1\}$ and $A_2 = \{2\}$. Let τ_q be a mono-type GW tree with binary offspring q. For $i \in [2]$, let $|\tau^{(i)}|$ be the number of individuals in τ with number of children in A_i. For $k = (k_1, k_2) \in (\mathbb{N}^*)^2$, let $\tau_{q,k}$ be distributed as τ_q conditionally on $\{|\tau^{(i)}| = k_i,\ i \in [2]\}$.

Let $\delta \in [1, +\infty]$. (The cases $\delta = 1$ and $\delta = +\infty$ will be degenerated.) We define an offspring distribution q_δ on $\{0, 1, 2\}$ as follows:

$$q_\delta(0) = q_\delta(2) = \frac{1}{\delta + 1} \quad \text{and} \quad q_\delta(1) = \frac{\delta - 1}{\delta + 1}\cdot$$

Notice that q_δ is critical and depends only on δ and not on the original offspring distribution q. Let τ^*_δ be the Kesten's tree associated with q_δ.

Then using Remarks 3.4 and 3.5, it is easy to check that the sequence $(\tau_{q,k(n)}, n \in \mathbb{N}^*)$, such that $\lim_{n\to+\infty} |k(n)| = +\infty$ and $\lim_{n\to+\infty} k_1(n)/k_2(n) = \delta$ (and require $k_1(n) > k_2(n)$ in the degenerated case $\delta = 1$), converges in distribution towards the Kesten's tree τ^*_δ. In particular the distribution of the limit depends on the proportion in the conditioning event and not on the initial offspring distribution.

3.2. **Proof of Theorem 3.1.** Assume that (H_1) holds. Let τ be a random GW tree with critical offspring distribution p. For $i, j \in [d]$, we define the total number of individuals of type i whose parent is of type j:

$$B_{ij} = \text{Card}\ (\{u \in \tau,\, \mathcal{M}(u) = i \text{ and } \mathcal{M}(\text{Pa}(u)) = j\})\,.$$

And we set $\mathcal{B} = (B_{ij}; i, j \in [d])$. Notice that $\sum_{j\in[d]} B_{ij} = |\tau^{(i)}|$.

Let $(X_{i,\ell}; \ell \in \mathbb{N}^*)$ for $i \in [d]$ be d independent families of independent random variables in $\mathbb{N}^d$ with $X_{i,\ell}$ having probability distribution $p^{(i)}$. For $i \in [d]$, we consider the random walk $S_{i,n} = \sum_{\ell=1}^{n} X_{i,\ell}$ for $n \in \mathbb{N}^*$ with $S_{i,0} = 0$. For $k = (k_1, \ldots, k_d) \in \mathbb{N}^d$, we set $S_k = \sum_{i\in[d]} S_{i,k_i}$. We adopt the following convention for a d-dimensional random variable X to write $X = (X^{(j)}, j \in [d])$, so that we have in particular $S^{(j)}_{i,n} = \sum_{\ell=1}^{n} X^{(j)}_{i,\ell}$. For $k \in \mathbb{N}^d$ and $r \in [d]$, we define the matrix $\mathcal{S}(k, r) = (\mathcal{S}_{ij}(k, r);\, i, j \in [d])$ of size $d \times d$ by:

$$\mathcal{S}_{ij}(k, r) = -S^{(j)}_{i,k_i} + (S^{(j)}_k + \mathbf{1}_{\{r=i\}})\mathbf{1}_{\{i=j\}}. \tag{7}$$

The following corollary is a direct consequence of the representation of Chaumont and Liu [5] for multi-type GW process, which is the generalization of Dwass formula to the multi-type case.

Corollary 3.7. *Assume that (H_1) holds. Let τ be a random GW tree with critical offspring distribution p. For $r \in [d]$ and $k \in (\mathbb{N}^*)^d$, we have:*

$$\mathbb{P}_r(|\tau| = k) = \frac{1}{\prod_{i\in[d]} k_i}\, \mathbb{E}\left[\det(\mathcal{S}(k, r));\, S_k + \mathbf{e}_r = k\right].$$

Proof. For $\kappa = (\kappa_{ij}; i, j \in [d]) \in \mathbb{N}^{d\times d}$, we denote, for $j \in [d]$, by κ_j the column vector $(\kappa_{ij}, i \in [d])$. We deduce from Theorem 1.2 in [5] that, for $r \in [d]$, $k = (k_1, \ldots, k_d) \in (\mathbb{N}^*)^d$, $\kappa = (\kappa_{ij}; i, j \in [d]) \in \mathbb{N}^{d\times d}$ such that

$$k = \mathbf{e}_r + \sum_{j\in[d]} \kappa_j, \tag{8}$$

we have:

$$\mathbb{P}_r(\mathcal{B} = \kappa) = \det(\Delta(k) - \kappa) \prod_{j\in[d]} \frac{\mathbb{P}(S_{j,k_j} = \kappa_j)}{k_j}, \tag{9}$$

where $\Delta(k)$ is the $d \times d$ diagonal matrix with diagonal k. Because of (8), we have:

$$\mathbb{P}_r(|\tau| = k,\, \mathcal{B} = \kappa) = \mathbb{P}_r(\mathcal{B} = \kappa). \tag{10}$$

Thanks to the definition of $\mathcal{S}(k, r)$, we have $\Delta(k) - \kappa = \mathcal{S}(k, r)$ on $\bigcap_{j\in[d]} \left\{S_{j,k_j} = \kappa_j\right\}$. By summing (10) and thus (9) over all the possibles values of κ such that (8) holds, we get:

$$\begin{aligned}
\mathbb{P}_r(|\tau| = k) &= \sum_{\kappa} \mathbb{P}_r(\mathcal{B} = \kappa)\mathbf{1}_{\{k=\mathbf{e}_r+\sum_{j\in[d]}\kappa_j\}}\\
&= \frac{1}{\prod_{j\in[d]} k_j} \sum_{\kappa} \det(\Delta(k) - \kappa)\, \mathbf{1}_{\{k=\mathbf{e}_r+\sum_{j\in[d]}\kappa_j\}}\mathbb{P}(\forall j \in [d],\, S_{j,k_j} = \kappa_j)\\
&= \frac{1}{\prod_{i\in[d]} k_i}\, \mathbb{E}\left[\det(\mathcal{S}(k, r));\, \mathbf{e}_r + S_k = k\right].
\end{aligned}$$

□

In order to compute the determinant $\det(\mathcal{S}(k,r))$, instead of using a development based on permutations, we shall use a development based on elementary forests, see Lemma 4.5 in [5] and Formula (11) below. (As we are interested in computing the determinant of a matrix whose all lines but one sum up to 0, we shall only consider forests reduced to one tree.)

Recall $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^d$. For $r \in [d]$, we consider $\mathcal{T}_r$ the subset of $\mathbb{T}_0$ of trees with root of type r, having exactly d individuals all of them with a distinct type:

$$\mathcal{T}_r = \{\mathbf{t} \in \mathbb{T}_0;\ |\mathbf{t}| = \mathbf{1},\ \text{and}\ \mathcal{M}(\emptyset_{\mathbf{t}}) = r\}.$$

For $\mathbf{t} \in \mathcal{T}_r$ and $j \in [d] \setminus \{r\}$, let $j_{\mathbf{t}}$ denote the type of the parent of the individual of type j: $j_{\mathbf{t}} = \mathcal{M}(\mathrm{Pa}(u_j))$, where u_j is the only element of $\mathbf{t}$ such that $\mathcal{M}(u_j) = j$. We shall use the following formula to give asymptotics on $\det(\mathcal{S}(k,r))$.

Lemma 3.8. *For $r \in [d]$ and $k \in (\mathbb{N}^*)^d$, we have:*

$$\det(\mathcal{S}(k,r)) = \sum_{\mathbf{t} \in \mathcal{T}_r} \prod_{j \in [d] \setminus \{r\}} S^{(j)}_{j_{\mathbf{t}}, k_{j_{\mathbf{t}}}}.$$

Proof. We follow the presentation of [5]. We say that a collection of trees is a forest. A forest $\mathbf{f} = (\mathbf{t}_j, j \in J)$ is called elementary if the trees are pairwise disjoint and if the forest contains exactly one individual of each type, that is $\sum_{j \in J} |\mathbf{t}_j| = \mathbf{1}$. Let $\mathbb{F}$ denote the set of elementary forests. For $\mathbf{f} \in \mathbb{F}$, set u_i the individual in $\mathbf{f}$ of type i, which belongs to a tree of $\mathbf{f}$ say $\mathbf{t}_j$, and write $i_{\mathbf{f}} = \mathcal{M}(v)$ for the type of the parent $v = \mathrm{Pa}_{u_i}(\mathbf{t}_j)$ of u_i if $|u_i| > 0$ and $i_{\mathbf{f}} = 0$ if $|u_i| = 0$.

According to Lemma 4.5 in [5], we have for $\kappa = (\kappa_{ij}; i, j \in [d]) \in \mathbb{R}^{d \times d}$

$$\det(\kappa) = (-1)^d \sum_{\mathbf{f} \in \mathbb{F}} \prod_{j \in [d]} \kappa_{j_{\mathbf{f}}, j}, \tag{11}$$

with the convention that $\kappa_{0,j} = -\sum_{i \in [d]} \kappa_{ij}$.

Thanks to Definition (7) of $\mathcal{S}(k,r)$, this implies that for $r \in [d]$ and $k \in (\mathbb{N}^*)^d$, we have:

$$\det(\mathcal{S}(k,r)) = \sum_{\mathbf{f} \in \mathbb{F}} \prod_{j \in [d]} S^{(j)}_{j_{\mathbf{f}}, k_{j_{\mathbf{f}}}}, \tag{12}$$

with the convention that if $j_{\mathbf{f}} = 0$, then $S^{(j)}_{j_{\mathbf{f}}, k_{j_{\mathbf{f}}}} = \mathbf{1}_{\{j=r\}}$. Notice that $\prod_{j \in [d]} S^{(j)}_{j_{\mathbf{f}}, k_{j_{\mathbf{f}}}} = 0$ if the forest $\mathbf{f}$ is not reduced to a single tree whose root is of type r. To conclude, use that $j_{\mathbf{f}} = j_{\mathbf{t}}$ if the forest $\mathbf{f}$ is reduced to a single tree $\mathbf{t}$. $\square$

Let $(\tilde{X}_{i,\ell};\ \ell \in \mathbb{N}^*, i \in [d])$ be a sequence of random variables independent of $(X_{i,\ell};\ \ell \in \mathbb{N}^*, i \in [d])$ with the same distribution.

For a finite subset K of $\mathbb{N}$, we shall consider partitions $\mathbf{A}^{(\ell,K)} = (A_1^K, \ldots, A_\ell^K)$ of K such that $\inf A_1^K < \cdots < \inf A_\ell^K$. For $\mathbf{t} \in \mathcal{T}_r$, $i \in [d]$, recall that u_i is the individual in $\mathbf{t}$ of type i. Denote by $C_i(\mathbf{t}) = \{j \in [d];\ j_{\mathbf{t}} = i\}$ the set of types of the children of u_i in $\mathbf{t}$. Let $\mathbb{A}_{\mathbf{t}}$ be the family of all $\mathcal{A} = (m, (\mathbf{A}^{(m_i, C_i(\mathbf{t}))}), i \in [d])$, with $m = (m_1, \ldots, m_d) \in \mathbb{N}^d$ such that $1 \leq m_i \leq \mathrm{Card}\,(C_i(\mathbf{t}))$ for all $i \in [d]$ such that $\mathrm{Card}\,(C_i(\mathbf{t})) > 0$. For convenience, we may write $m_{\mathcal{A}}$ for m. With this notation, we set:

$$\tilde{S}_{m_{\mathcal{A}}} = \sum_{i \in [d]} \sum_{\ell=1}^{m_i} \tilde{X}_{i,\ell}, \quad G(\mathcal{A}) = \prod_{i \in [d]} \prod_{\ell=1}^{m_i} \prod_{j \in A_\ell^{C_i(\mathbf{t})}} \tilde{X}^{(j)}_{i,\ell},$$

with the convention that $\sum_\emptyset = 0$ and $\prod_\emptyset = 1$, and for $k = (k_1, \dots, k_d) \in \mathbb{N}^d$ such that $k_i \geq d$ for all $i \in [d]$:

$$B_k(m_\mathcal{A}) = \prod_{i\in[d]} \frac{k_i!}{(k_i - m_i)!}.$$

Since $\tilde{X}_{i,\ell}$ for $i \in [d]$, $\ell \in \mathbb{N}^*$ takes values in $\mathbb{N}^d$ and $\sum_{i\in[d]} \sum_{\ell=1}^{m_i} \text{Card}\ (A_\ell^{C_i(\mathbf{t})}) = d-1$, we deduce that:

$$0 \leq G(\mathcal{A}) \leq \left|\tilde{S}_{m_\mathcal{A}}\right|^d. \tag{13}$$

We have the following result.

Corollary 3.9. *For $r \in [d]$ and $k \in (\mathbb{N}^*)^d$ such that $k \geq d\mathbf{1}$, we have:*

$$\mathbb{E}\left[\det(\mathcal{S}(k,r));\ S_k = b\right] = \sum_{\mathbf{t}\in\mathcal{T}_r} \sum_{\mathcal{A}\in\mathbb{A}_\mathbf{t}} B_k(m_\mathcal{A})\, \mathbb{E}\left[G(\mathcal{A}); \tilde{S}_{m_\mathcal{A}} + S_{k-m_\mathcal{A}} = b\right].$$

Proof. For $r \in [d]$, $\mathbf{t} \in \mathcal{T}_r$, and $k \in (\mathbb{N}^*)^d$, we have:

$$\prod_{j\in[d]\setminus\{r\}} S^{(j)}_{j_\mathbf{t},k_{j_\mathbf{t}}} = \prod_{i\in[d]} \prod_{j\in C_i(\mathbf{t})} \sum_{\ell=1}^{k_i} X^{(j)}_{i,\ell}.$$

Using the exchangeability of $(X_{i,\ell}; \ell \in \mathbb{N}^*)$ for all $i \in [d]$, we easily get for $b, k \in (\mathbb{N}^*)^d$ such that $k \geq d\mathbf{1}$:

$$\mathbb{E}\left[\prod_{j\in[d]\setminus\{r\}} S^{(j)}_{j_\mathbf{t},k_{j_\mathbf{t}}};\ S_k = b\right] = \sum_{\mathcal{A}\in\mathbb{A}_\mathbf{t}} B_k(m_\mathcal{A})\, \mathbb{E}\left[G(\mathcal{A}); \tilde{S}_{m_\mathcal{A}} + S_{k-m_\mathcal{A}} = b\right].$$

Then use Corollary 3.7 and Lemma 3.8 to conclude. □

The next lemma is an extension of the strong ratio limit theorem given in [1]. Its proof is postponed to Section 3.3. Recall that a is the normalized left eigenvector of the mean matrix M.

Lemma 3.10. *Assume that (H_1) and (H_2) hold. Let G and H be two random variables in $\mathbb{N}$ and $\mathbb{N}^d$ respectively, independent of $(X_{i,\ell};\ \ell \in \mathbb{N}^*, i \in [d])$ and such that $\mathbb{P}(G = 0) < 1$ and a.s. $G \leq |H|^d$.*

Set $(k(n), n \in \mathbb{N}^*)$ and $(s_n, n \in \mathbb{N}^*)$ be two sequences in $\mathbb{N}^d$ satisfying $\lim_{n\to\infty} |k(n)| = +\infty$ and $\lim_{n\to\infty} k(n)/|k(n)| = \lim_{n\to\infty} s_n/|k(n)| = a$. Then for any given m, $b \in \mathbb{N}^d$, we have:

$$\lim_{n\to\infty} \frac{\mathbb{E}[G;\ H + S_{k(n)-m} = s_n - b]}{\mathbb{E}[G;\ H + S_{k(n)} = s_n]} = 1.$$

Notice that no moment condition is assumed for G or H.

Let $b \in \mathbb{N}^d$ such that $k \geq b + \mathbf{1}$. We have using Corollary 3.9:

$$\begin{aligned} \frac{\prod_{i\in[d]}(k_i - b_i)}{\prod_{i\in[d]} k_i} \frac{\mathbb{P}_r(|\tau| = k - b)}{\mathbb{P}_r(|\tau| = k)} \\ &= \frac{\mathbb{E}\left[\det(\mathcal{S}(k-b,r));\ S_{k-b} + \mathbf{e}_r = k - b\right]}{\mathbb{E}\left[\det(\mathcal{S}(k,r));\ S_k + \mathbf{e}_r = k\right]} \\ &= \frac{\sum_{\mathbf{t}\in\mathcal{T}_r} \sum_{\mathcal{A}\in\mathbb{A}_\mathbf{t}} B_{k-b}(m_\mathcal{A})\, \mathbb{E}\left[G(\mathcal{A}); \tilde{S}_{m_\mathcal{A}} + S_{k-b-m_\mathcal{A}} = k - b - \mathbf{e}_r\right]}{\sum_{\mathbf{t}\in\mathcal{T}_r} \sum_{\mathcal{A}\in\mathbb{A}_\mathbf{t}} B_k(m_\mathcal{A})\, \mathbb{E}\left[G(\mathcal{A}); \tilde{S}_{m_\mathcal{A}} + S_{k-m_\mathcal{A}} = k - \mathbf{e}_r\right]}. \end{aligned} \tag{14}$$

Assume (H_1) and (H_2) hold. Let $(k(n), n \in \mathbb{N}^*)$ be a sequence in $\mathbb{N}^d$ such that $\lim_{n\to\infty} |k(n)| = +\infty$ and $\lim_{n\to\infty} k(n)/|k(n)| = a$. Since $\mathbb{P}(G(\mathcal{A}) = 0) < 1$ and thanks to (13), we deduce from Lemma 3.10 that:

$$\lim_{n\to+\infty} \frac{\mathbb{E}\left[G(\mathcal{A}); \tilde{S}_{m_\mathcal{A}} + S_{k(n)-b-m_\mathcal{A}} = k(n) - b - \mathbf{e}_r\right]}{\mathbb{E}\left[G(\mathcal{A}); \tilde{S}_{m_\mathcal{A}} + S_{k(n)-m_\mathcal{A}} = k(n) - \mathbf{e}_r\right]} = 1.$$

We also have:

$$\lim_{n\to+\infty} \frac{B_{k(n)-b}(m_\mathcal{A})}{B_{k(n)}(m_\mathcal{A})} = 1.$$

Since all the terms in (14) are non-negative, and $\lim_{n\to+\infty} \prod_{i\in[d]}(k_i(n)-b_i)/k_i(n) = 1$, we deduce that:

$$\lim_{n\to+\infty} \frac{\mathbb{P}_r(|\tau| = k(n) - b)}{\mathbb{P}_r(|\tau| = k(n))} = 1. \tag{15}$$

Then, using Lemmas 2.10 (with $i = r$ in (5)), we obtain that for all $r \in [d]$, $\mathbf{t} \in \mathbb{T}_0$ and $x \in \mathcal{L}_0(\mathbf{t})$ such that $\mathcal{M}(x) = r$: $\lim_{n\to+\infty} \mathbb{P}_r(\tau_n \in \mathbb{T}(t, x)) = \mathbb{P}_r(\tau^* \in \mathbb{T}(t, x))$. Of course we have for n large enough and $\mathbf{t} \in \mathbb{T}_0$ that $\mathbb{P}_r(\tau_n = \mathbf{t}) = 0 = \mathbb{P}_r(\tau^* = \mathbf{t})$. We deduce from Corollary 2.2 that $(\tau_n, n \in \mathbb{N}^*)$ converges in distribution towards τ^* under $\mathbb{P}_r$ for all $r \in [d]$. This also implies that the convergence in distribution holds under $\mathbb{P}_\alpha$ for any probability distribution α on $[d]$ of the type of the root.

3.3. Proof of Lemma 3.10.

We assume (H_1). In particular, this implies that $\mathbb{P}(X_{i,1} = 0) > 0$ for some $i \in [d]$. Without loss of generality, we can assume this holds for $i = d$: $\mathbb{P}(X_{d,1} = 0) > 0$.

Recall that a is the normalized left eigenvector of the mean matrix M such that $|a| = 1$. In particular a is a probability on $[d]$. Set $\mathbf{v}_d = 0 \in \mathbb{N}^{d-1}$ and for $i \in [d-1]$, set $\mathbf{v}_i = (v_i^{(1)}, \dots, v_i^{(d-1)}) \in \mathbb{N}^{d-1}$ such that $v_i^{(j)} = \mathbf{1}_{\{j=i\}}$ for $j \in [d-1]$. Let $Y = (U, V)$ be a random variable in $\mathbb{N}^d \times \mathbb{N}^{d-1}$ such that for $i \in [d]$, $\mathbb{P}(V = \mathbf{v}_i) = a_i$, and the distribution of U conditionally on $\{V = \mathbf{v}_i\}$ is $p^{(i)}$.

Recall Definition 2.3 of aperiodic probability distribution.

Lemma 3.11. *Under (H_2), the distribution of Y on $\mathbb{Z}^{2d-1}$ is aperiodic.*

Proof. Let F be the probability distribution of Y. Notice that $\mathbb{P}(Y = 0) \geq \mathbb{P}(X_{d,1} = 0)\mathbb{P}(V = \mathbf{v}_d) > 0$. Since $0 \in \text{supp}\,(F)$, according to Lemma 2.5, the probability distribution F is aperiodic if and only if the smallest subgroup that contains the support of F is $\mathbb{Z}^{2d-1}$ itself. Let $z \in \mathbb{Z}^d$ and $v = (v^{(1)}, \dots, v^{(d-1)}) \in \mathbb{Z}^{d-1}$. To prove that F is aperiodic, we have to find $n, n' \in \mathbb{N}^*$ and $y_1, \dots, y_n$ and $y'_1, \dots, y'_{n'}$ in supp (F) such that:

$$\sum_{\ell=1}^{n} y_\ell - \sum_{\ell=1}^{n'} y'_\ell = \begin{pmatrix} z \\ v \end{pmatrix}.$$

For $i \in [d]$, let $x_i \in \text{supp}\,(p^{(i)})$. Set $v^{(d)} = 0$. We deduce from (H_2) that for all $z' \in \mathbb{Z}^d$, there exists $k = (k_1, \dots, k_d)$ and $k = (k'_1, \dots, k'_d)$ in $\mathbb{N}^d$, $(x_{i,\ell}, \ell \in \mathbb{N}^*)$ and $(x'_{i,\ell}, \ell \in \mathbb{N}^*)$ elements of supp $(p^{(i)})$ such that:

$$\sum_{i\in[d]} \sum_{\ell=1}^{k_i} (x_{i,\ell} - x_i) - \sum_{i\in[d]} \sum_{\ell=1}^{k'_i} (x'_{i,\ell} - x_i) = z'.$$

For $a \in \mathbb{R}$, let $a_+ = \max(a, 0)$ and $a_- = \max(-a, 0)$. Taking $z' = z - \sum_{i=1}^d v^{(i)} x_i$, we get:

$$\left(\sum_{i\in[d]} \sum_{\ell=1}^{k_i} x_{i,\ell} + \sum_{i\in[d]} (k'_i + v_+^{(i)}) x_i \right) - \left(\sum_{i\in[d]} \sum_{\ell=1}^{k'_i} x'_{i,\ell} + \sum_{i\in[d]} (k_i + v_-^{(i)}) x_i \right) = z.$$

We deduce:

$$\left(\sum_{i\in[d]} \sum_{\ell=1}^{k_i} \binom{x_{i,\ell}}{\mathbf{v}_i} + \sum_{i\in[d]} (k'_i + v_+^{(i)}) \binom{x_i}{\mathbf{v}_i} \right) - \left(\sum_{i\in[d]} \sum_{\ell=1}^{k'_i} \binom{x'_{i,\ell}}{\mathbf{v}_i} + \sum_{i\in[d]} (k_i + v_-^{(i)}) \binom{x_i}{\mathbf{v}_i} \right) = \binom{z}{v}.$$

To conclude, notice that $(x_{i,\ell}, \mathbf{v}_i)$, $(x'_{i,\ell}, \mathbf{v}_i)$ as well as $(x_i, \mathbf{v}_i)$ belong to supp (F) for all $i \in [d]$. $\square$

The next lemma is an extension of Theorem 4.7.

Lemma 3.12. *Let F be a probability distribution on $\mathbb{N}^{d'}$ which is aperiodic on $\mathbb{Z}^{d'}$. Let $(Y_n, n \in \mathbb{N}^*)$ be independent random variables distributed according to F and set $W_n = \sum_{\ell=1}^n Y_\ell$ for $n \in \mathbb{N}^*$. Assume that $\mathbb{E}[|Y_1|] < +\infty$. Let G and H' be two random variables in $\mathbb{N}$ and $\mathbb{N}^{d'}$ respectively and independent of $(Y_n, n \in \mathbb{N}^*)$ such that $\mathbb{P}(G = 0) < 1$ and a.s. $G \leq |H'|^c$ for some $c \geq 1$. Let $(w_n, n \in \mathbb{N}^*)$ be a sequence of $\mathbb{N}^{d'}$ such that $\lim_{n\to+\infty} w_n/n = \mathbb{E}[Y_1]$. Then for any given $\ell \in \mathbb{N}$ and $b \in \mathbb{N}^{d'}$, we have:*

$$\lim_{n\to\infty} \frac{\mathbb{E}[G;\, H' + W_{n-\ell} = w_n - b]}{\mathbb{E}[G;\, H' + W_n = w_n]} = 1. \tag{16}$$

Proof. Since F is aperiodic and by elementary arithmetic consideration, it is enough to prove (16) for $\ell = 1$ and $b \in \mathbb{N}^{d'}$ satisfying $\mathfrak{p} = \mathbb{P}(Y_1 = b) > 0$. Let $\varepsilon > 0$. Using similar arguments as in (39), we get:

$$\left| \frac{\mathbb{E}[G;\, H' + W_{n-1} = w_n - b]}{\mathbb{E}[G;\, H' + W_n = w_n]} - 1 \right| \leq \frac{\varepsilon}{\mathfrak{p}} + \frac{R_n}{\mathfrak{p}},$$

and

$$R_n = \frac{\mathbb{E}\left[G;\, |\frac{N_n}{n} - \mathfrak{p}| > \varepsilon,\, H' + W_n = w_n\right]}{\mathbb{E}[G; H' + W_n = w_n]},$$

with $N_n = \sum_{\ell=1}^n \mathbf{1}_{\{Y_\ell = b\}}$. Choose $g \in \mathbb{N}^*$ and $h \in \mathbb{N}^{d'}$ such that $q = \mathbb{P}(G = g, H' = h) > 0$. We have:

$$R_n \leq \frac{|w_n|^c\, \mathbb{P}\left(\left|\frac{N_n}{n} - \mathfrak{p}\right| > \varepsilon\right)}{gq\mathbb{P}(W_n = w_n - h)} \leq \frac{|w_n|^c\, 2\, \mathrm{e}^{-2n\varepsilon^2}}{gq\mathbb{P}(W_n = w_n - h)},$$

where we used $G \leq |H'|^c$ a.s. and that $H' + W_n = w_n$ implies $H' \leq w_n$ for the first inequality, and inequality (40) in the Appendix for the second. Notice that for all $\varepsilon' > 0$ we have $|w_n|^c \leq \exp(\varepsilon' n)$ for n large enough.

Then use Lemma 4.9 and Remark 4.8 to conclude that if $\lim_{n\to+\infty} w_n/n = \mathbb{E}[Y_1]$, then $\lim_{n\to+\infty} R_n = 0$. Since $\varepsilon > 0$ is arbitrary, we get $\lim_{n\to+\infty} \left| \frac{\mathbb{E}[G;\, H' + W_{n-1} = w_n - b]}{\mathbb{E}[G;\, H' + W_n = w_n]} - 1 \right| = 0$, which gives the result. $\square$

For $x \in \mathbb{R}^d$ and $z = (z_1, \ldots, z_d) \in \mathbb{R}^d$, we set $\delta(x, z) = (x, z_1, \ldots, z_{d-1})$.

We consider $(Y_\ell, \ell \in \mathbb{N}^*)$ independent random variables distributed as Y. We set $W_n = \sum_{\ell=1}^n Y_\ell$. Let $s \in \mathbb{N}^d$ and $k \in (\mathbb{N}^*)^d$. We have:

$$\mathbb{P}\left(W_{|k|} = \delta(s, k)\right) = D(k)\mathbb{P}(S_k = s) \quad \text{with} \quad D(k) = \frac{|k|!}{\prod_{i\in[d]} k_i!} \prod_{i\in[d]} {a_i}^{k_i}. \tag{17}$$

Recall G and H given in Lemma 3.10. We set $H' = \delta(H, 0) \in \mathbb{N}^{2d-1}$. We get for k, m, s and b in $\mathbb{N}^d$:

$$\frac{\mathbb{E}[G;\, H + S_{k-m} = s - b]}{\mathbb{E}[G;\, H + S_k = s]} = \frac{D(k)}{D(k-m)} \, \frac{\mathbb{E}[G;\, H' + W_{|k|-|m|} = \delta(s,k) - \delta(b,m)]}{\mathbb{E}[G;\, H' + W_{|k|} = \delta(s,k)]}. \tag{18}$$

Thanks to Lemma 3.11 and (H_2), the distribution of Y on $\mathbb{Z}^{2d-1}$ is aperiodic. Since $0 \leq G \leq |H|^d$, we also have $0 \leq G \leq |H'|^d$ and $\mathbb{P}(G = 0) < 1$. Let $(k(n), n \in \mathbb{N}^*)$ and $(s_n, n \in \mathbb{N}^*)$ be two sequences in $\mathbb{N}^d$ satisfying $\lim_{n\to\infty} |k(n)| = +\infty$ and $\lim_{n\to\infty} k(n)/|k(n)| = \lim_{n\to\infty} s_n/|k(n)| = a$. Notice, this implies that $\lim_{n\to\infty} \delta(s_n, k(n))/|k(n)| = \mathbb{E}[Y_1]$. We deduce from Lemma 3.12 that:

$$\lim_{n\to+\infty} \frac{\mathbb{E}[G;\, H' + W_{|k(n)|-|m|} = \delta(s_n, k(n)) - \delta(b,m)]}{\mathbb{E}[G;\, H' + W_{|k(n)|} = \delta(s_n, k(n))]} = 1.$$

Then notice that $\lim_{n\to+\infty} D(k(n))/D(k(n) - m) = 1$ as $\lim_{n\to+\infty} k(n)/|k(n)| = a$. And use (18) to get:

$$\lim_{n\to\infty} \frac{\mathbb{E}[G;\, H + S_{k(n)-m} = s_n - b]}{\mathbb{E}[G;\, H + S_{k(n)} = s_n]} = 1.$$

This ends the proof of Lemma 3.10.

4. Appendix

4.1. **Preliminary results.** For $x \in \mathbb{R}^d$ and $\delta \geq 0$, let $\mathcal{B}(x, \delta)$ be the open ball of $\mathbb{R}^d$ centered at x with radius δ. For any non-empty subset A of $\mathbb{R}^d$, denote: cv A the convex hull of A, cl A the closure of A, int A the interior of A, aff $A = x_0 + \text{span}\,(A - x_0)$ the affine hull of A where $x_0 \in A$ and, if A is convex, ri A the relative interior of A:

$$\text{ri } A = \{x \in A;\ \text{aff } A \bigcap \mathcal{B}(x, \delta) \subset A \text{ for some } \delta > 0\}.$$

Notice that, for A convex, we have int A = ri A if and only if aff $A = \mathbb{R}^d$. For a function f on $\mathbb{R}^d$ taking its values in $\mathbb{R} \bigcup \{+\infty\}$, its domain is defined by $\text{dom}(f) = \{x \in \mathbb{R}^d : f(x) < \infty\}$.

Let F be a probability distribution on $\mathbb{R}^d$ and X be a random variable on $\mathbb{Z}^d$ with distribution F. Denote by supp (F) the closed support of F: $x \notin \text{supp}\,(F)$ if and only if $\mathbb{P}(X \in \mathcal{B}(x, \delta)) = 0$ for some $\delta > 0$. Denote also by cv (F) the convex hull of its support, aff (F) and ri (F) the affine hull and the relative interior of cv (F). We define φ the Log-Laplace of X taking values in $(-\infty, +\infty]$ as:

$$\varphi(\theta) = \log\left(\mathbb{E}\left[\mathrm{e}^{\langle \theta, X\rangle}\right]\right), \quad \theta \in \mathbb{R}^d. \tag{19}$$

The function φ is convex, $\varphi(0) = 0$ (which implies that φ is proper), and lower-semicontinuous (thanks to Fatou's lemma). Its conjugate, ψ, is defined by:

$$\psi(x) = \sup_{\theta \in \text{dom}(\varphi)} \left(\langle \theta, x\rangle - \varphi(\theta)\right), \quad x \in \mathbb{R}^d. \tag{20}$$

We recall that ψ is a lower-semicontinuous (proper) convex function. Since $\varphi(0) = 0$, we deduce that ψ is non-negative. We first give a general lemma on the domain of ψ.

Lemma 4.1. *Let F be a probability distribution on $\mathbb{R}^d$. We have* ri (F) = ri dom(ψ).

Proof. We use the asymptotic function, φ_∞, associated with φ, see Section 2.5 in [4]. Since $\varphi(0) = 0$, the asymptotic function is defined by $\varphi_\infty(\theta) = \lim_{t\to+\infty} \varphi(t\theta)/t$ for $\theta \in \mathbb{R}^d$. This gives:

$$\varphi_\infty(\theta) = \lim_{t\to+\infty} \frac{\log\left(\mathbb{E}\left[\mathrm{e}^{\langle t\theta, X\rangle}\right]\right)}{t} = \sup_{x\in\text{supp }(F)} \langle\theta, x\rangle = \sup_{x\in\text{cv }(F)} \langle\theta, x\rangle.$$

According to Proposition 2.5.8 in [4], we have $z \in \text{ri dom}(\psi)$ if and only if $\varphi_\infty(\theta) > \langle\theta, z\rangle$ for all $\theta \in \mathbb{R}^d$ except those satisfying $\varphi_\infty(-\theta) = \varphi_\infty(\theta) = 0$. We deduce that $z \in \text{ri dom}(\psi)$ if and only if $\sup_{x\in\text{cv }(F)}\langle\theta, x - z\rangle > 0$ for all $\theta \in \mathbb{R}^d$ except those satisfying $\langle\theta, x - z\rangle = 0$ for all $x \in \text{cv }(F)$.

Assume $z \in \text{ri }(F)$. Let H be the orthogonal vector sub-space of aff (F) in $\mathbb{R}^d$. Since $\text{ri }(F) \subset \text{aff }(F)$, we get that for $\theta \in H$ and all $x \in \text{cv }(F) \subset \text{aff }(F)$, we have $\langle\theta, x - z\rangle = 0$. For $\theta \in \mathbb{R}^d \setminus H$ and all $x \in \text{cv }(F)$, we have $\langle\theta, x - z\rangle = \langle\theta_F, x - z\rangle$, where θ_F is the orthogonal projection of θ on aff (F). By definition of ri (F), we have that for all $\theta' \in \text{aff }(F)$, there exists $x \in \text{ri }(F) \subset \text{cv }(F)$ such that $\langle\theta', x - z\rangle > 0$. In particular, we deduce that for $\theta \in \mathbb{R}^d \setminus H$, we have $\sup_{x\in\text{cv }(F)}\langle\theta, x - z\rangle > 0$. This implies that $\text{ri }(F) \subset \text{ri dom}(\psi)$.

Assume $z \in \text{ri dom}(\psi)$. Notice that $\text{ri }\{z\} = \{z\}$. We deduce, from Proposition 1.1.11 in [4], that $z \notin \text{ri }(F)$ is equivalent to the two convex sets $\{z\}$ and cv (F) being properly separated. Thanks to Proposition 1.1.11 in [4], this is equivalent to the existence of $\theta \in \mathbb{R}^d$, $\theta \neq 0$, such that:

$$\langle\theta, z\rangle \geq \sup_{x\in\text{cv }(F)} \langle\theta, x\rangle \quad \text{and} \quad \langle\theta, z\rangle > \inf_{x\in\text{cv }(F)} \langle\theta, x\rangle.$$

Since $z \in \text{ri dom}(\psi)$, we get $z \notin \text{ri }(F)$ is equivalent to the existence of $\theta \in \mathbb{R}^d$, $\theta \neq 0$, such that for all $x \in \text{cv }(F)$ we have $\langle\theta, z\rangle = \langle\theta, x\rangle$ and $\langle\theta, z\rangle > \inf_{x\in\text{cv }(F)}\langle\theta, x\rangle$. Since those last two assertions are incompatible, we deduce that $z \in \text{ri }(F)$ and thus $\text{ri dom}(\psi) \subset \text{ri }(F)$. This ends the proof of this lemma. □

We have the following corollary.

Corollary 4.2. *Let X be a random variable on $\mathbb{R}^d$ with probability distribution F. If X is integrable then $\mathbb{E}[X]$ belongs to* ri dom(ψ) *and $\psi(\mathbb{E}[X]) = 0$.*

Proof. Since X is integrable, it is easy to check that $\mathbb{E}[X]$ belongs to ri (F). The first part of the corollary is then a consequence of Lemma 4.1. Jensen inequality implies that $\varphi(\theta) \geq \langle\theta, \mathbb{E}[X]\rangle$. This gives $\langle\theta, \mathbb{E}[X]\rangle - \varphi(\theta) \leq 0$. Then use (20) and the fact that ψ is non-negative to deduce that $\psi(\mathbb{E}[X]) = 0$. □

For $\theta \in \text{dom}(\psi)$, we define a probability measure on $\mathbb{R}^d$ by:

$$d\mathbb{P}_\theta(X \in dx) = \mathrm{e}^{\langle\theta, X\rangle - \varphi(\theta)}\, d\mathbb{P}(X \in dx). \tag{21}$$

We denote by m_θ and Σ_θ the corresponding mean vector and covariance matrix if they exist, i.e:

$$m_\theta = \mathbb{E}_\theta[X] = \mathbb{E}[X\, \mathrm{e}^{\langle\theta, X\rangle - \varphi(\theta)}] = \nabla\varphi(\theta) \quad \text{and} \quad \Sigma_\theta = \text{Cov}_\theta(X, X). \tag{22}$$

We set $\mathcal{I}_F = \text{int dom}(\varphi)$ the interior of the domain of the log-Laplace of F. Notice that X under $\mathbb{P}_\theta$ has small exponential moment for $\theta \in \mathcal{I}_F$ and its mean and covariance matrix are thus well defined for $\theta \in \mathcal{I}_F$. For a symmetric positive semi-definite matrix Σ, we denote by $|\Sigma|$ its determinant.

Lemma 4.3. *Let F be a probability distribution on $\mathbb{R}^d$. For any compact set $K \subset \mathcal{I}_F$, we have:*

$$\sup_{\theta\in K} |\Sigma_\theta| < +\infty \quad \textit{and} \quad \sup_{\theta\in K} \mathbb{E}_\theta\left[|X - m_\theta|^3\right] < +\infty. \tag{23}$$

Proof. For all $\theta \in \mathcal{I}_F$, we get that the ball $\mathcal{B}(\theta, \varepsilon)$ centered at θ with positive radius ε belongs to $\mathcal{I}_F$ for ε small enough. This implies that X under $\mathbb{P}_\theta$ has exponential moments, in particular it belongs to L^3. By dominated convergence we also get that the application:

$$\theta \mapsto \left(m_\theta, \Sigma_\theta, \mathbb{E}_\theta\left[|X - m_\theta|^3\right]\right)$$

is continuous on $\mathcal{I}_F$. This ends the proof. □

We set $\mathcal{O}_F = \text{int cv}\ (F)$ the interior of the convex hull of the support of F.

Lemma 4.4. *Assume $\mathcal{O}_F$ is non-empty and bounded. Then the application $\theta \mapsto m_\theta$ is one-to-one from $\mathbb{R}^d$ onto $\mathcal{O}_F$ and continuous as well as its inverse. In particular, for any compact set $K \subset \mathcal{O}_F$, there exists r such that $K \subset \{m_\theta;\ |\theta| \leq r\}$.*

Proof. It is easy to check, using Hölder inequality, that if $\mathcal{O}_F$ is non-empty then φ is strongly convex on its domain. If $\mathcal{O}_F$ is bounded, then X is also bounded and the function φ is finite on $\mathbb{R}^d$, so that $\text{dom}(\varphi) = \mathbb{R}^d$, as well as differentiable throughout $\mathbb{R}^d$. This implies that φ is smooth on $\mathbb{R}^d$ in the sense of [20] section 26. According to Theorem 26.5 in [20], this implies that $\nabla\varphi$ is one-to-one from $\mathbb{R}^d$ onto the open set $D = \nabla\varphi(\mathbb{R}^d)$, continuous as well as $\nabla\varphi^{-1}$. Furthermore, according to Corollary 26.4.1 in [20], we have $\text{ri dom}(\psi) \subset D \subset \text{dom}(\psi)$. Since D is open, we deduce that $D = \text{ri dom}(\psi) = \text{int dom}(\psi)$. Then, use Lemma 4.1 to get that $D = \text{ri}\ (F) = \mathcal{O}_F$. □

Recall Definition 2.3 for aperiodic probability distribution.

Lemma 4.5. *Assume F is an aperiodic probability distribution on $\mathbb{Z}^d$. Then, we have $\mathcal{O}_F$ non-empty and for any compact set $K \subset \mathcal{I}_F$,*

$$\inf_{\theta \in K} |\Sigma_\theta| > 0. \tag{24}$$

Proof. Since F is aperiodic, we have $\text{aff}\ (F) = \mathbb{R}^d$. This readily implies that the dimension of $\text{cv}\ (F)$ is d or equivalently that $\mathcal{O}_F$ is non-empty. This proves the first part of the lemma.

Let $\theta \in \mathcal{I}_F$ be such that $|\Sigma_\theta| = 0$. Then there exists $h \in \mathbb{R}^d \setminus \{0\}$ such that $\langle h, \Sigma_\theta h\rangle = 0$. This implies that $\mathbb{P}_\theta$-a.s. $\langle h, X\rangle = c$ with $c = \langle h, m_\theta\rangle$. This equality also holds $\mathbb{P}$-a.s. as the two probability measures $\mathbb{P}$ and $\mathbb{P}_\theta$ are equivalent. In particular $\text{supp}\ (F) - x$, with $x \in \text{supp}\ (F)$, is orthogonal to h. Since F is aperiodic, the smallest subgroup in $\mathbb{Z}^d$ that contains the set $-x + \text{supp}\ (F)$ is $\mathbb{Z}^d$ itself, which contradicts $h \neq 0$. We deduce that $|\Sigma_\theta| > 0$ for all $\theta \in \mathcal{I}_F$. Then use that $\theta \mapsto |\Sigma_\theta|$ is continuous on $\mathcal{I}_F$ to get the second part of the lemma. □

4.2. **Gnedenko's d-dimensional local theorem.** Recall the definitions of φ, $\mathbb{P}_\theta$, m_θ and Σ_θ given by (19), (21) and (22) and that $\mathcal{I}_F = \text{int dom}(\varphi)$. The next theorem is an extension of the one-dimensional theorem of Gnedenko [7], see also [21, 24].

Theorem 4.6. *Let F be an aperiodic probability distribution on $\mathbb{Z}^d$ such that $\mathcal{I}_F$ is non-empty. Let $(X_\ell, \ell \in \mathbb{N}^*)$ be independent random variables with distribution F and set $S_n = \sum_{\ell=1}^n X_\ell$ for $n \in \mathbb{N}^*$. Then for any compact subset K of $\mathcal{I}_F$, we have:*

$$\lim_{n\to\infty} \sup_{\theta \in K} \sup_{s \in \mathbb{Z}^d} \left| n^{d/2} |\Sigma_\theta|^{1/2} \mathbb{P}_\theta(S_n = s) - (2\pi)^{-d/2}\, \mathrm{e}^{-\left\|z_n(\theta,s)\right\|^2/2} \right| = 0, \tag{25}$$

with $z_n(\theta, s) = n^{-1/2}\Sigma_\theta^{-1/2}(s - nm_\theta)$.

The end of this section is devoted to the proof of Theorem 4.6.

Thanks to Lemmas 4.3 and 4.5, we have $|\Sigma_\theta|>0$ and $\Sigma_\theta^{-1/2}$ is well defined. We define:

$$Y=n^{-1/2}\Sigma_\theta^{-1/2}(X_1-m_\theta)\quad\text{and}\quad f_\theta(t)=\mathbb{E}_\theta\left[\mathrm{e}^{i\langle t,\,Y\rangle}\right]. \tag{26}$$

By the inversion formula, we know that for $s\in\mathbb{Z}^d$:

$$\begin{aligned}(2\pi)^d\mathbb{P}_\theta(S_n=s)&=\int_{(-\pi,\pi)^d}\mathbb{E}_\theta\left[\mathrm{e}^{i\langle u,S_n-s\rangle}\right]du\\&=\int_{(-\pi,\pi)^d}\mathbb{E}_\theta\left[\mathrm{e}^{i\langle n^{1/2}\Sigma_\theta^{1/2}u,\,n^{-1/2}\Sigma_\theta^{-1/2}(S_n-s)\rangle}\right]du\\&=\int_{(-\pi,\pi)^d}\mathbb{E}_\theta\left[\mathrm{e}^{i\langle n^{1/2}\Sigma_\theta^{1/2}t,\,Y\rangle}\right]^n\ \mathrm{e}^{-i\langle n^{1/2}\Sigma_\theta^{1/2}u,\,z_n(\theta,s)\rangle}\,du.\end{aligned}$$

In order to simplify the notation, we shall write z for $z_n(\theta,s)$. By considering the change of variable $t=n^{1/2}\Sigma_\theta^{1/2}u$, we obtain:

$$(2\pi)^d\mathbb{P}_\theta(S_n=s)=n^{-d/2}|\Sigma_\theta|^{-1/2}\int_{\mathcal{J}_\theta}f_\theta(t)^n\ \mathrm{e}^{-i\langle t,\,z\rangle}\,dt,$$

where $\mathcal{J}_\theta=\{t\in\mathbb{R}^d:n^{-1/2}\Sigma_\theta^{-1/2}t\in(-\pi,\pi)^d\}$. We set:

$$I_n(\theta)=n^{d/2}|\Sigma_\theta|^{1/2}\mathbb{P}_\theta(S_n=s)-(2\pi)^{-d/2}\,\mathrm{e}^{-\|z\|^2/2}.$$

Notice that

$$(2\pi)^{d/2}\,\mathrm{e}^{-\|z\|^2/2}=\int_{\mathbb{R}^d}\mathrm{e}^{-\|t\|^2/2-i\langle t,z\rangle}\,dt.$$

We obtain:

$$(2\pi)^d\,I_n(\theta)=\int_{\mathbb{R}^d}\left(\mathbf{1}_{\mathcal{J}_\theta}(t)f_\theta(t)^n-\mathrm{e}^{-\|t\|^2/2}\right)\mathrm{e}^{-i\langle t,z\rangle}\,dt.$$

Let $(C_n,n\in\mathbb{N}^*)$ be a sequence of positive numbers such that:

$$\lim_{n\to\infty}C_n=\infty\quad\text{and}\quad\lim_{n\to\infty}n^{-1/(12+6d)}C_n=0. \tag{27}$$

For $\varepsilon\in(0,1)$, we obtain:

$$(2\pi)^d|I_n(\theta)|\le\int_{\mathbb{R}^d}\left|\mathbf{1}_{\mathcal{J}_\theta}(t)f_\theta(t)^n-\mathrm{e}^{-\|t\|^2/2}\right|dt\le I_{n,1}(\theta)+I_{n,2}(\theta)+I_{n,3}(\theta)+I_{n,4}, \tag{28}$$

with

$$I_{n,1}(\theta)=\int_{J_1}|f_\theta(t)^n-\mathrm{e}^{-\|t\|^2/2}|dt,\quad I_{n,2}(\theta)=\int_{J_{2,\theta}}|f_\theta(t)|^ndt,\quad I_{n,3}(\theta)=\int_{J_{3,\theta}}|f_\theta(t)|^ndt,$$

and $I_{n,4}=\int_{J_1^c}\mathrm{e}^{-\|t\|^2/2}\,dt$ as well as $J_1=\{t\in\mathbb{R}^d;\ \|\,t\,\|\le C_n\}$, $J_{2,\theta}=\{t\in\mathbb{R}^d;\ \|\,t\,\|>C_n$ and $n^{-1/2}\,\|\,\Sigma_\theta^{-1/2}t\,\|<\varepsilon\}$, $J_{3,\theta}=\{t\in\mathcal{J}_\theta;\ n^{-1/2}\,\|\,\Sigma_\theta^{-1/2}t\,\|\ge\varepsilon\}$. The proof of the Theorem will be complete as soon as we prove the converge of the terms $I_{n,i}$ to 0 for $i\in\{1,\dots,4\}$ uniformly for $\theta\in K$ (notice the terms $I_{n,i}$ do not depend on $s\in\mathbb{Z}^d$).

4.2.1. *Convergence of* $I_{n,4}$. Notice that $I_{n,4}$ does not depend on θ. And we deduce from (27) that $\lim_{n\to\infty}I_{n,4}=0$.

4.2.2. *Convergence of $I_{n,3}$.* Set $h(\theta,u)=|\mathbb{E}_\theta[\mathrm{e}^{i\langle u,X_1\rangle}]|$ for $u\in\mathbb{R}^d$ and $L=\{u\in[-2\pi+\varepsilon,2\pi-\varepsilon]^d;\ \| u \|\geq\varepsilon\}$. Since F is aperiodic, we deduce from Proposition P8 in [22, p.75], that $h(\theta,u)<1$ for $u\in L$. Since h is continuous in (θ,t) on the compact set $K\times L$, there exists $\delta<1$ such that $h(\theta,u)\leq\delta$ on $K\times L$. We get for $\theta\in K$:

$$I_{n,3}(\theta)\leq n^{d/2}|\Sigma_\theta|^{1/2}\int_{(-\pi,\pi)^d} h(\theta,u)^n\,\mathbf{1}_{\{\|u\|\geq\varepsilon\}}\,du\leq n^{d/2}|\Sigma_\theta|^{1/2}(2\pi)^d\delta^n,$$

where we used that $|f_\theta(t)|=h(\theta,u)$ with $t=n^{1/2}\Sigma_\theta^{1/2}u$ for the first inequality and that h is bounded by δ on $\{u\in(-\pi,\pi)^d;\ \| u \|\geq\varepsilon\}$. Thanks to (23) we have $\sup_{\theta\in K}|\Sigma_\theta|<\infty$ and since $\delta<1$, we get $\lim_{n\to\infty}\sup_{\theta\in K}I_{n,3}(\theta)=0$.

4.2.3. *Convergence of $I_{n,2}$.* From (23), we have

$$a_2:=\sup_{\theta\in K}\mathbb{E}_\theta[\| X_1-m_\theta \|^2]<\infty\quad\text{and}\quad a_3:=\sup_{\theta\in K}\mathbb{E}_\theta[\| X_1-m_\theta \|^3]<\infty. \tag{29}$$

We deduce, using the expression of Σ_θ^{-1} based on the cofactors, that $\theta\mapsto\Sigma_\theta^{-1}$ is continuous on $\mathcal{I}_F$. This implies that $\| \Sigma_\theta^{-1/2}t \|^2=\langle t,\Sigma_\theta^{-1}t\rangle$ is continuous in (θ,t) on $\mathcal{I}_F\times\mathbb{R}^d$. We deduce that:

$$c_1:=\sup_{\theta\in K,\,\|t\|=1}\langle t,\Sigma_\theta^{-1}t\rangle<\infty. \tag{30}$$

Hence we can choose ε small enough such that

$$\varepsilon^2a_2+\varepsilon a_3c_1<1. \tag{31}$$

Recall $Y=n^{-1/2}\Sigma_\theta^{-1/2}(X_1-m_\theta)$. By the symmetry of Σ_θ, we get that

$$\mathbb{E}_\theta\left[\| Y \|^2\right]=\frac{1}{n}\mathbb{E}_\theta\left[\langle X_1-m_\theta,\Sigma_\theta^{-1}(X_1-m_\theta)\rangle\right]=\frac{1}{n}\sum_{j=1}^d\sum_{\ell=1}^d\left[\Sigma_\theta^{-1}(j,\ell)\Sigma_\theta(\ell,j)\right]=\frac{d}{n}\cdot \tag{32}$$

Using similar computations, we obtain:

$$\mathbb{E}_\theta\left[\langle t,Y\rangle^2\right]=\frac{\| t \|^2}{n}\cdot \tag{33}$$

Recall notations a_3 in (29) and c_1 in (30). For $t\in J_{2,\theta}$, we get:

$$\mathbb{E}_\theta\left[|\langle t,Y\rangle|^3\right]\leq n^{-3/2}\| \Sigma_\theta^{-1/2}t \|^3\,\mathbb{E}_\theta[\| X_1-m_\theta \|^3]\leq\frac{\| t \|^2}{n}\,\varepsilon a_3c_1\leq\frac{\| t \|^2}{n}, \tag{34}$$

where we used $n^{-1/2}\| \Sigma_\theta^{-1/2}t \|<\varepsilon$, (29) and (30) for the second inequality and (31) for the last. Recall a_2 given in (29). From (31) and since $t\in J_{2,\theta}$, we get:

$$\mathbb{E}_\theta\left[\langle t,Y\rangle^2\right]\leq\| n^{-1/2}\Sigma_\theta^{-1/2}t \|^2\,\mathbb{E}_\theta[\| X_1-m_\theta \|^2]\leq\varepsilon^2a_2<1. \tag{35}$$

We deduce that, for all $\theta \in K$ and $t \in J_{2,\theta}$,

$$\begin{aligned}|f_\theta(t)| = |\mathbb{E}_\theta[\mathrm{e}^{i\langle t,Y\rangle}]| &= \Big|1 - \frac{\mathbb{E}_\theta[|\langle t,Y\rangle|^2]}{2} - i\mathbb{E}_\theta\left[\int_0^{\langle t,Y\rangle}\int_0^v\int_0^s \mathrm{e}^{iu}\,dudsdv\right]\Big| \\ &\leq 1 - \frac{\mathbb{E}_\theta[|\langle t,Y\rangle|^2]}{2} + \mathbb{E}_\theta\left[\int_0^{|\langle t,Y\rangle|}\int_0^v\int_0^s dudsdv\right] \\ &= 1 - \frac{\mathbb{E}_\theta[|\langle t,Y\rangle|^2]}{2} + \frac{\mathbb{E}_\theta[|\langle t,Y\rangle|^3]}{6} \\ &\leq 1 - \frac{\| t \|^2}{2n} + \frac{\| t \|^2}{6n} = 1 - \frac{\| t \|^2}{3n},\end{aligned}$$

where we used that $\mathbb{E}_\theta[Y] = 0$ for the first equality, that $\mathbb{E}_\theta[\langle t,Y\rangle^2] \leq 1$ for the first inequality (see (35)) and (33) as well as (34) for the last inequality. Therefore, we get that:

$$I_{n,2}(\theta) \leq \int_{J_{2,\theta}} |f_\theta(t)|^n dt \leq \int_{J_{2,\theta}} \left(1 - \frac{\| t \|^2}{3n}\right)^n dt \leq \int_{\|t\|>C_n} \mathrm{e}^{-\|t\|^2/3}\,dt.$$

Since $\lim_{n\to\infty} C_n = \infty$, we deduce that $\lim_{n\to\infty}\sup_{\theta\in K} I_{n,2}(\theta) = 0$.

4.2.4. *Convergence of $I_{n,1}$.* Since $|f_\theta(t)| \leq 1$, we have:

$$|f_\theta(t)^n - \mathrm{e}^{-\|t\|^2/2}| \leq n|f_\theta(t) - \mathrm{e}^{-\|t\|^2/(2n)}| \leq n|h_\theta(n,t)| + ng(n,t), \tag{36}$$

where

$$h_\theta(n,t) = f_\theta(t) - 1 + \frac{\| t \|^2}{2n} \quad \text{and} \quad g(n,t) = \left|\mathrm{e}^{-\|t\|^2/(2n)} - 1 + \frac{\| t \|^2}{2n}\right|.$$

Since $0 \leq x + \mathrm{e}^{-x} - 1 \leq x^2/2$ for $x \geq 0$, we get for $t \in J_1$:

$$ng(n,t) \leq \frac{\| t \|^4}{8n} \leq n^{-1}C_n^4. \tag{37}$$

Since $\mathbb{E}[Y] = 0$ and $\mathbb{E}\left[\langle t,Y\rangle^2\right] = \| t \|^2/n$, see (33), we deduce that:

$$h_\theta(n,t) = \mathbb{E}_\theta\left[\mathrm{e}^{i\langle t,Y\rangle} - 1 + i\langle t,Y\rangle - \langle t,Y\rangle^2\right].$$

Let $L_n = n^{\frac{1}{4}}$. We have:

$$\begin{aligned}|h_\theta(n,t)| &\leq \mathbb{E}_\theta\left[\left|\mathrm{e}^{i\langle t,Y\rangle} - 1 + i\langle t,Y\rangle + \frac{\langle t,Y\rangle^2}{2}\right|\right] \\ &= \mathbb{E}_\theta\left[\left|\mathrm{e}^{i\langle t,Y\rangle} - 1 + i\langle t,Y\rangle + \frac{\langle t,Y\rangle^2}{2}\right|; \| X_1 - m_\theta \| < L_n\right] \\ &\qquad + \mathbb{E}_\theta\left[\left|\mathrm{e}^{i\langle t,Y\rangle} - 1 + i\langle t,Y\rangle + \frac{\langle t,Y\rangle^2}{2}\right|; \| X_1 - m_\theta \| \geq L_n\right] \\ &\leq \frac{1}{6}\mathbb{E}_\theta\left[|\langle t,Y\rangle|^3; |X_1 - m_\theta| < L_n\right] + \mathbb{E}_\theta\left[\langle t,Y\rangle^2; \| X_1 - m_\theta \| \geq L_n\right],\end{aligned}$$

where we used $|\,\mathrm{e}^{i\alpha} -1-i\alpha+\frac{\alpha^2}{2}| \leq \min(|\alpha|^3/6, \alpha^2)$ for $\alpha\in\mathbb{R}$ for the second inequality. We have:

$$\begin{aligned}\mathbb{E}_\theta[|\langle t,Y\rangle|^3;\ \|\ X_1-m_\theta\ \|<L_n] &= \mathbb{E}_\theta\left[\langle t,Y\rangle^2|\langle t, n^{-1/2}\Sigma_\theta^{-1/2}(X_1-m_\theta)\rangle|;\ \|\ X_1-m_\theta\ \|<L_n\right]\\ &\leq n^{-1/2}\ \|\ t\ \|\ \sqrt{c_1}\,L_n\mathbb{E}_\theta\left[\langle t,Y\rangle^2\right]\\ &= n^{-3/2}\ \|\ t\ \|^3\ \sqrt{c_1}\,L_n,\end{aligned}$$

where we used c_1 defined in (30) for the inequality and (33) for the last equality. Hölder's inequality gives:

$$\mathbb{E}_\theta\left[\langle t,Y\rangle^2;\ \|\ X_1-m_\theta\ \|\geq L_n\right] \leq \mathbb{E}_\theta\left[|\langle t,Y\rangle|^3\right]^{2/3}\mathbb{P}_\theta(\|\ X_1-m_\theta\ \|\geq L_n)^{1/3}.$$

Using a_3 defined in (29), we get:

$$\mathbb{E}_\theta\left[|\langle t,Y\rangle|^3\right]\leq n^{-3/2}\ \|\ \Sigma_\theta^{-1/2}t\ \|^3\ \mathbb{E}_\theta\left[\|\ X_1-m_\theta\ \|^3\right]\leq n^{-3/2}c_1^{3/2}\ \|\ t\ \|^3\,a_3.$$

Using Tchebychev inequality and a_2 defined in (29), we get:

$$\mathbb{P}_\theta(\|\ X_1-m_\theta\ \|\geq L_n)\leq \mathbb{E}_\theta\left[\|\ X_1-m_\theta\ \|^2\right]L_n^{-2}\leq a_2L_n^{-2}.$$

This gives:

$$\mathbb{E}_\theta\left[\langle t,Y\rangle^2;\ \|\ X_1-m_\theta\ \|\geq L_n\right]\leq n^{-1}c_1\ \|\ t\ \|^2\,a_3^{2/3}a_2^{1/3}L_n^{-2/3}.$$

For $t\in J_1$, that is $\|\ t\ \|\leq C_n$, we get:

$$n|h_\theta(n,t)|\leq \frac{1}{6}\,n^{-1/4}C_n^3\sqrt{c_1}+n^{-1/6}c_1C_n^2a_2^{1/3}.$$

Using (36) and (37), we deduce there exists a constant c which does not depend on t, θ and n such that for $t\in J_1$, $\theta\in K$, we have:

$$|f_\theta(t)^n-\mathrm{e}^{-\|t\|^2/2}|\leq c(n^{-1/4}C_n^3+n^{-1/6}C_n^2+n^{-1}C_n^4).$$

We deduce that for $\theta\in K$:

$$I_{n,1}(\theta)=\int_{J_1}|f_\theta(t)^n-\mathrm{e}^{-\|t\|^2/2}|\leq c(n^{-1/4}C_n^3+n^{-1/6}C_n^2+n^{-1}C_n^4)2^dC_n^d.$$

Recall that $\lim_{n\to\infty} n^{-1/(12+6d)}C_n=0$. This implies $\lim_{n\to\infty}\sup_{\theta\in K} I_{n,1}(\theta)=0$.

4.3. **Strong ratio limit theorem.** Recall Definition 2.3 for aperiodic probability distribution. Consider an aperiodic distribution F on $\mathbb{Z}^d$. Let X be a random variable with distribution F. Recall the function $\varphi(\theta)=\log\mathbb{E}[\mathrm{e}^{\langle\theta,X\rangle}]$ defined in (19) and its conjugate ψ defined in (20).

Theorem 4.7. *Let F be an aperiodic probability distribution on $\mathbb{Z}^d$. Let $(X_\ell,\ell\in\mathbb{N}^*)$ be independent random variables with the same distribution F. Let $S_n=\sum_{\ell=1}^n X_\ell$ for $n\in\mathbb{N}^*$. For all $m\in\mathbb{N}$ and $b\in\mathbb{Z}^d$, we have:*

$$\lim_{n\to\infty}\frac{\mathbb{P}(S_{n-m}=s_n-b)}{\mathbb{P}(S_n=s_n)}=1, \tag{38}$$

where the sequence $(s_n,n\in\mathbb{N}^*)$ of elements of $\mathbb{Z}^d$ satisfies the following conditions:

(a) $\sup_{n\in\mathbb{N}^*}|\frac{s_n}{n}|<\infty$,
(b) $\lim_{n\to\infty}\psi(\frac{s_n}{n})=0$.

Remark 4.8. Assume that X, with distribution F, is integrable. Thanks to Corollary 4.2, $\mathbb{E}[X]$ belongs to ri dom(ψ), the relative interior of the domain of ψ and $\psi(\mathbb{E}[X]) = 0$. According to Theorem 1.2.3 in [4], the function ψ is relatively continuous on ri dom(ψ). Therefore if the sequence $(s_n, n \in \mathbb{N}^*)$ of elements of dom(ψ) satisfies $\lim_{n\to\infty} s_n/n = \mathbb{E}[X]$, then (a) and (b) of Theorem 4.7 are satisfied. Notice also that if F is aperiodic (as assumed in Theorem 4.7), then Lemmas 4.5 and 4.1 imply ri dom(ψ) is the (non-empty) interior of dom(ψ) which is also equal to $\mathcal{O}_F = \text{int cv}(F)$.

4.4. **Proof of Theorem 4.7.** We adapt the proof of Neveu [17]. Since F is aperiodic and by elementary arithmetic consideration, it is enough to prove (38) for $m = 1$ and $b \in \mathbb{Z}^d$ satisfying $\mathfrak{p} := \mathbb{P}(X_1 = b) > 0$.

We set $N_n = \text{Card}\,(\{\ell \leq n; X_\ell = b\})$. Since for $a \in \mathbb{Z}^d$ the conditional probability $\mathbb{P}(X_\ell = b|S_n = a)$ does not depend on ℓ (when $1 \leq \ell \leq n$), we get:

$$\mathbb{E}\left[\frac{N_n}{n} \,\Big|\, S_n = a\right] = \mathbb{P}(X_n = b|S_n = a) = \mathfrak{p}\frac{\mathbb{P}(S_{n-1} = a - b)}{\mathbb{P}(S_n = a)}.$$

For $\varepsilon > 0$, we have:

$$\Big|\frac{\mathbb{P}(S_{n-1} = a - b)}{\mathbb{P}(S_n = a)} - 1\Big| = \Big|\frac{\mathbb{E}\left[\frac{N_n}{n}; S_n = a\right]}{\mathfrak{p}\mathbb{P}(S_n = a)} - 1\Big| \leq \frac{\mathbb{E}[|\frac{N_n}{n} - \mathfrak{p}|; S_n = a]}{\mathfrak{p}\mathbb{P}(S_n = a)} \leq \frac{\varepsilon}{\mathfrak{p}} + \frac{R_n(a)}{\mathfrak{p}}, \tag{39}$$

with

$$R_n(a) = \frac{\mathbb{P}(|\frac{N_n}{n} - \mathfrak{p}| > \varepsilon)}{\mathbb{P}(S_n = a)}.$$

Thus, the proof will be complete as soon as we prove that for all $\varepsilon > 0$, $\lim_{n\to\infty} R_n(s_n) = 0$.

By Hoeffding's inequality, see Theorem 1 in [10], since N_n is binomial with parameter $(n, \mathfrak{p})$, we get:

$$\mathbb{P}\left(\left|\frac{N_n}{n} - \mathfrak{p}\right| > \varepsilon\right) \leq 2\,\mathrm{e}^{-2n\varepsilon^2}. \tag{40}$$

We give a lower bound of $\mathbb{P}(S_n = s_n)$ in the next lemma, whose proof is postponed at the end of this section.

Lemma 4.9. *Let F be an aperiodic probability distribution on $\mathbb{Z}^d$. Let $(X_\ell, \ell \in \mathbb{N}^*)$ be independent random variables with the same distribution F. Let $S_n = \sum_{\ell=1}^n X_\ell$ for $n \in \mathbb{N}^*$. Then for $0 < \eta < 1$, K_0 compact subset of $\mathcal{O}_F$, $(s_n, n \in \mathbb{N}^*)$ a sequence of elements of $\mathbb{Z}^d$ such that $s_n/n \in K_0$, there exists some $n_0 \geq 1$ such that for $n \geq n_0$ we have:*

$$\mathbb{P}(S_n = s_n)\,\mathrm{e}^{n\psi(s_n/n)} \geq (1 - \eta)^n.$$

Using (40) and Lemma 4.9 with $1 - \eta = \mathrm{e}^{-\varepsilon^2}$, we get:

$$R_n(s_n) = \frac{\mathbb{P}\left(|\frac{N_n}{n} - \mathfrak{p}| > \varepsilon\right)}{\mathbb{P}(S_n = s_n)} \leq 2\,\mathrm{e}^{-n\varepsilon^2 + n\psi(s_n/n)}.$$

Since $\lim_{n\to\infty} \psi(s_n/n) = 0$ by hypothesis, we get the result. □

Proof of Lemma 4.9. Since F is aperiodic, Lemma 4.5 implies that $\mathcal{O}_F$ is non-empty.

We first assume that the support of F is bounded. In particular the domain of φ defined by (19) is $\mathbb{R}^d$. Recall notation (21) as well as $m_\theta = \mathbb{E}_\theta[X]$ and $\Sigma_\theta = \text{Cov}\,_\theta(X, X)$. By hypothesis, we have K_0 is a compact subset of $\mathcal{O}_F$. According to Lemma 4.4, there exists a compact set of

$\mathbb{R}^d$ such that $K_0 \subset \{m_\theta, \theta \in K\}$. According to Theorem 4.6, we have that for all $\varepsilon > 0$, there exists n_0 such that for all $n \geq n_0$:

$$\sup_{\theta\in K} \sup_{s\in\mathbb{Z}^d} \left| n^{d/2} |\Sigma_\theta|^{1/2} \mathbb{P}_\theta(S_n = s) - (2\pi)^{-d/2} \, \mathrm{e}^{-u_n(\theta,s)} \right| < \varepsilon,$$

with

$$u_n(\theta, s) = \frac{\langle s_n - nm_\theta, \Sigma_\theta^{-1}(s_n - nm_\theta)\rangle}{2n}.$$

So we get that for all $n \geq n_0$, $\theta \in K$:

$$\begin{aligned}\mathbb{P}_\theta(S_n = s_n) &\geq (2\pi n)^{-d/2} |\Sigma_\theta|^{-1/2} \, \mathrm{e}^{-u_n(\theta)} - n^{-d/2} |\Sigma_\theta|^{-1/2} \varepsilon \\ &\geq (2\pi n)^{-d/2} \left(\sup_{q\in K} |\Sigma_q| \right)^{-1/2} \mathrm{e}^{-u_n(\theta)} - n^{-d/2} \left(\inf_{q\in K} |\Sigma_q| \right)^{-1/2} \varepsilon.\end{aligned}$$

We deduce that for all $n \geq n_0$:

$$\sup_{\theta\in K} \mathbb{P}_\theta(S_n = s_n) \geq (2\pi n)^{-d/2} \left(\sup_{q\in K} |\Sigma_q| \right)^{-1/2} \mathrm{e}^{-\inf_{\theta\in K} u_n(\theta)} - n^{-d/2} \left(\inf_{q\in K} |\Sigma_q| \right)^{-1/2} \varepsilon.$$

Since s_n/n belongs to $\{m_\theta; \theta \in K\}$, we get that $\inf_{\theta\in K} u_n(\theta) = 0$. Thanks to (23) and Lemma 4.5, we can also choose $\varepsilon > 0$ and $\delta > 0$ both small enough so that $(2\pi)^{-d/2} \left(\sup_{q\in K} |\Sigma_q|\right)^{-1/2} - \left(\inf_{q\in K} |\Sigma_q|\right)^{-1/2} \varepsilon > \delta$. Then we deduce that for all $n \geq n_0$:

$$\sup_{\theta\in\mathbb{R}^d} \mathbb{P}_\theta(S_n = s_n) \geq \sup_{\theta\in K} \mathbb{P}_\theta(S_n = s_n) \geq n^{-d/2}\delta > 0.$$

Using (20), we get:

$$\sup_{\theta\in\mathbb{R}^d} \mathbb{P}_\theta(S_n = s_n) = \sup_{\theta\in\mathbb{R}^d} \mathbb{P}(S_n = s_n) \, \mathrm{e}^{\langle\theta, s_n - n\varphi(\theta)\rangle} = \mathbb{P}(S_n = s_n) \, \mathrm{e}^{n\psi(s_n/n)}.$$

This gives, for some $\delta > 0$, for all $n \geq n_0$:

$$\mathbb{P}(S_n = s_n) \, \mathrm{e}^{n\psi(s_n/n)} \geq \delta n^{-d/2} > 0. \tag{41}$$

This gives Lemma 4.9 when the support of F is bounded.

Let F be a general aperiodic probability distribution on $\mathbb{Z}^d$, and X a random variable with distribution F. Let $M > 0$ be large enough and X^M be distributed as X conditionally on $\{|X| \leq M\}$. Write $\delta_M = \mathbb{P}(|X| > M)$. Let $(X_\ell^M, \ell \in \mathbb{N})$ be independent random variables distributed as X^M, and set $S_n^M = \sum_{\ell=1}^n X_\ell^M$. We have:

$$\mathbb{P}(S_n^M = s_n) = \frac{\mathbb{P}(S_n = s_n, \, |X_\ell| \leq M \text{ for } 1 \leq \ell \leq n)}{\mathbb{P}(|X| \leq M)^n} \leq \frac{\mathbb{P}(S_n = s_n)}{(1 - \delta_M)^n}.$$

Let F_M be the probability distribution of X^M and φ_M defined by (19) with F replaced by F_M and ψ_M defined by (20) with φ replaced by φ_M. Since F is aperiodic, we get that F_M is aperiodic for M large enough. We get:

$$\mathbb{P}(S_n = s_n) \, \mathrm{e}^{n\psi(s_n/n)} \geq \mathbb{P}(S_n^M = s_n) \, \mathrm{e}^{n\psi(s_n/n)} (1 - \delta_M)^n = \mathbb{P}(S_n^M = s_n) \, \mathrm{e}^{n\psi_M(s_n/n)} \, \mathrm{e}^{n\Delta_M(s_n/n)},$$

where we define $\Delta_M(s) = \psi(s) - \tilde{\psi}_M(s)$ and $\tilde{\psi}_M(x) = \sup_{\theta\in\mathbb{R}^d} (\langle\theta, x\rangle - \tilde{\varphi}_M(\theta))$ with $\tilde{\varphi}_M(\theta) = \log\left(\mathbb{E}\left[\mathrm{e}^{\langle\theta, X\rangle} \mathbf{1}_{\{|X|\leq M\}}\right]\right)$.

Notice that the sequence of continuous finite convex functions $(\tilde{\varphi}_M, M \in \mathbb{N}^*)$ is non-decreasing and converges point-wise to the convex function φ (which is not identically $+\infty$ as $\varphi(0) = 0$).

According to vii of Proposition E.1.3.1 in [9], we get the sequence of convex functions $(\tilde{\psi}_M, M \in \mathbb{N}^*)$ is non-increasing and $\tilde{\psi}_M \geq \psi$. Therefore the sequence converge to a function say $\tilde{\psi}$ such that $\tilde{\psi} \geq \psi$. Thanks to Theorem B.3.1.4 in [9] or Theorem II.10.8 of [20], $\tilde{\psi}$ is convex and $(\tilde{\psi}_M, M \in \mathbb{N}^*)$ converges to $\tilde{\psi}$ uniformly on any compact subset of ri dom$(\tilde{\psi})$. Theorem E.2.4.4 in [9] gives that the closure of $\tilde{\psi}$ (defined in Definition B.1.2.4 in [9]) is equal to ψ. Thanks to Proposition 1.2.5 in [4], we get that ri dom$(\tilde{\psi})$ = ri dom(ψ) and on this set we have $\tilde{\psi} = \psi$. Since ri dom(ψ) = ri $(F) = \mathcal{O}_F$, see Lemmas 4.1 and 4.5, this implies that $\lim_{M\to+\infty} \Delta_M = 0$ uniformly on any compact subset of $\mathcal{O}_F$.

Notice that $\Delta_M \leq 0$. Therefore for any $\gamma > 0$, K_0 compact subset of $\mathcal{O}_F$, there exists M_0 such that for $M \geq M_0$, $0 \geq \Delta_M \geq -\gamma$ on K_0. We deduce from (41) with S_n and ψ replaced by S_n^M and ψ_M that for some $\delta > 0$ and $\gamma > 0$, there exists $n_0 \geq 1$ such that for all $n \geq n_0$:

$$\mathbb{P}(S_n = s_n)\, \mathrm{e}^{n\psi(s_n/n)} \geq \delta n^{-d/2}\, \mathrm{e}^{-\gamma n}\,.$$

This implies that for some $\eta \in (0,1)$ and n large enough, we have $\mathbb{P}(S_n = s_n)\, \mathrm{e}^{n\psi(s_n/n)} \geq (1-\eta)^n$. □

Acknowledgements. The authors would like to thank Jean-Philippe Chancelier for pointing out the references on convex analysis and his valuable advice. H. Guo would like to express her gratitude to J.-F. Delmas for his help during her stay at CERMICS.

The research has also been supported by the ANR-14-CE25-0014 (ANR GRAAL).

References

[1] R. Abraham and J.-F. Delmas. Local limits of conditioned Galton-Watson trees: the condensation case. *Elec. J. of Probab.*, 19(56):1–29, 2014.

[2] R. Abraham and J.-F. Delmas. Local limits of conditioned Galton-Watson trees: the infinite spine case. *Elec. J. of Probab.*, 19(2):1–19, 2014.

[3] K. B. Athreya and P. E. Ney. *Branching processes.* Springer-Verlag, 1972.

[4] A. Auslender and M. Teboulle. *Asymptotic cones and functions in optimization and variational inequalities.* Springer Science & Business Media, 2006.

[5] L. Chaumont and R. Liu. Coding multitype forests: application to the law of the total population of branching forests. *Transactions of the American Mathematical Society*, To appear in 2015.

[6] J.-F. Delmas and O. Hénard. A Williams decomposition for spatially dependent superprocesses. *Elec. J. of Probab.*, 18(37):1–43, 2013.

[7] B. V. Gnedenko. On a local limit theorem of the theory of probability. *Uspekhi Mat. Nauk*, 3(3):187–194, 1948.

[8] X. He. Condition Galton-Watson trees by the maximal out-degree. arXiv:1412.1972, 2014.

[9] J.-B. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of convex analysis.* Springer Science & Business Media, 2001.

[10] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American statistical association*, 58(301):13–30, 1963.

[11] S. Janson. Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation. *Probab. Surv.*, 9:103–252, 2012.

[12] T. Jonnson and S. Stefansson. Condensation in nongeneric trees. *J. Stat. Phys.*, 142:277–313, 2011.

[13] H. Kesten. Subdiffusive behavior of random walk on a random cluster. *Ann. de l'Inst. Henri Poincaré*, 22:425–487, 1986.

[14] T. Kurtz, R. Lyons, R. Pemantle, and Y. Peres. A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes. In *Classical and modern branching processes (Minneapolis, 1994)*, volume 84 of *IMA Vol. Math. Appl.*, pages 181–185. Springer, 1997.

[15] J. A. L.-M. Luis G. Gorostiza. The multitype measure branching process. *Advances in Applied Probability*, 22(1):49–67, 1990.

[16] G. Miermont. Invariance principles for spatial multitype Galton-Watson trees. *Ann. Inst. H. Poincaré Probab. Statist*, 44:1128–1161, 2007.

[17] J. Neveu. Sur le théorème ergodique de Chung-Erdős. *C. R. Acad. Sci. Paris*, 257:2953–2955, 1963.

[18] S. Pénisson. Various ways of conditioning multitype Galton-Watson processes. *arXiv*:1412.3322, 2014.
[19] D. Rizzolo. Scaling limits of Markov branching trees and Galton–Watson trees conditioned on the number of vertices with out-degree in a given set. *Ann. de l'Inst. Henri Poincaré*, 51(2):512–532, 2015.
[20] R. T. Rockafellar. *Convex analysis*. Princeton Landmarks in Mathematics. Princeton University Press, 1997.
[21] E. Rvaceva. On domains of attraction of multi-dimensional distributions. *Select. Transl. Math. Statist. and Probability*, 2:183–205, 1961.
[22] F. Spitzer. *Principles of random walk*. Springer Science & Business Media, 2013.
[23] R. Stephenson. Local convergence of large critical multi-type Galton-Watson trees and applications to random maps. *arXiv*:1412.6911, 2014.
[24] C. Stone. On local and ratio limit theorems. *Proc. of the Fifth Berkeley sympos. on mathematical statistics and probability. Berkeley and Los Angeles: Univ. California Press*, 2(part II):217–224, 1966.

Romain Abraham, Laboratoire MAPMO, CNRS, UMR 7349, Fédération Denis Poisson, FR 2964, Université d'Orléns, B.P. 6759, 45067 Orléns cedex 2, France
E-mail address: romain.abraham@univ-orleans.fr

Jean-François Delmas, Université Paris-Est, CERMICS (ENPC), F-77455 Marne La Vallée, France
E-mail address: delmas@cermics.enpc.fr

Hongsong Guo, Université Paris-Est, CERMICS (ENPC), F-77455 Marne La Vallée, France; School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China
E-mail address: hsguo@mail.bnu.edu.cn