J Theor Probab (2018) 31:757–788
https://doi.org/10.1007/s10959-016-0739-8

Critical Multi-type Galton–Watson Trees Conditioned to be Large

Romain Abraham[1] · **Jean-François Delmas**[2] · **Hongsong Guo**[2,3]

Received: 4 November 2015 / Revised: 12 December 2016 / Published online: 3 January 2017

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 toward a multi-type Kesten's tree. We obtain the result by generalizing Neveu's strong ratio limit theorem for aperiodic random walks on $\mathbb{Z}^d$.

Keywords Galton–Watson tree · Random tree · Local limit · Strong ratio theorem · Branching process

Mathematics Subject Classification (2010) 60J80 · 60B10

1 Introduction

In [14], Kesten shows 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

✉ Romain Abraham
romain.abraham@univ-orleans.fr

Jean-François Delmas
delmas@cermics.enpc.fr

Hongsong Guo
hsguo@mail.bnu.edu.cn

1 Laboratoire MAPMO, CNRS, UMR 7349, Fédération Denis Poisson, FR 2964, Université d'Orléans, B.P. 6759, 45067 Orléans Cedex 2, France

2 Université Paris-Est, CERMICS (ENPC), 77455 Marne la Vallée, France

3 Department of Mathematics, China University of Mining and Technology, Beijing 100083, People's Republic of China

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 subcritical GW trees, see Janson [12], Jonnson and Stefansson [13], He [9] or Abraham and Delmas [1] for results in this direction. When scaling limits of multi-type GW tree are considered, one obtains as a limit a continuous GW tree, see Miermont [17] or Gorostiza and Lopez-Mimbela [16] (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 [19] has proved that a critical d-type GW process conditioned on the total progeny to be large with a given asymptotic proportion of 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 [24] gave, under an exponential moments condition, the local convergence of a multi-type GW tree, conditioned on a linear combination of population sizes of each type to be large, toward the multi-type Kesten's tree introduced by Kurtz et al. [15]. The aim of this paper is to give minimal hypotheses to ensure the local convergence of a critical multi-type GW tree conditioned on the total progeny to be large toward the associated multi-type Kesten's tree, see Theorem 3.1. When the offspring distribution is aperiodic, the minimal hypotheses is the existence of the mean matrix which is assumed to be primitive. Furthermore, we exactly condition on the asymptotic proportion of 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 of types is not equal to the (normalized) left eigenvector associated with the Perron–Frobenius eigenvalue of the mean matrix, then under an exponential moments condition for the offspring distribution, it is possible to get a Kesten's tree as local limit, see [19]. However, without an exponential moments condition for the offspring distribution, no results are known, and results in [1] for the mono-type case suggest a condensation phenomenon (at least in the subcritical 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 the Dwass formula for multi-type GW processes given by Chaumont and Liu [5] which encodes critical or subcritical 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 [18] 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 [8] (for the sum of independent random variables), Rvaceva [22] (for the sum of d-dimensional i.i.d. random variables), or Stone [25] (for the sum of d-dimensional i.i.d. lattice or non-lattice random variables), which is given in Sect. 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 Sect. 4.1, as we believe they might be interesting by themselves.

The paper is organized as follows. We present in Sect. 2 the topology on the set 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 a multi-type GW tree with a given offspring distribution and the aperiodicity condition on the offspring distribution, see Definition 2.5. Section 3 is devoted to the main result, Theorem 3.1, and its proof. The last section collects results on the Legendre–Laplace transform in a general framework in Sect. 4.1, Gnedenko's d-dimensional local theorem in Sect. 4.2, and the strong ratio limit theorem for d-dimensional random walks in Sect. 4.3.

2 Multi-type Trees

2.1 General Notations

We denote by $\mathbb{N} = \{0, 1, 2, \ldots\}$ the set of nonnegative 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 non-empty set $A \subset \mathbb{R}^d$, we define span A as the linear subspace generated by A (i.e., 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 and B non-empty subsets of $\mathbb{R}^d$, we denote $A - B = \{x - y; x \in A, y \in B\}$.

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

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 $\mathcal{U} = \bigcup_{n\geq 0}\left((\mathbb{N}^*)^n \times [d]\right)$ 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 offspring 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 simply call them 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 $v \in \mathbf{t}\backslash\emptyset_{\mathbf{t}}$ 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 offspring 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 $\mathbf{v_t}$ the infinite spine of the tree $\mathbf{t}$. Let $\mathbb{T}_1'$ be the subset of $\mathbb{T}_1$ such that the infinite spine features each type infinitely many times:

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

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

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

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 $\mathbf{t}$. For $\mathbf{t} \in \mathbb{T}_0$, 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 toward 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 supp $(F) = \{x \in \mathbb{Z}^d, F(x) > 0\}$ the support set of F and by R_0 the smallest subgroup of $\mathbb{Z}^d$ which contains the set supp $(F) -$ supp (F).

Definition 2.3 A distribution F on $\mathbb{Z}^d$ is called aperiodic if $R_0 = \mathbb{Z}^d$.

For $x \in \mathbb{Z}^d$, let G_x be the smallest subgroup of $\mathbb{Z}^d$ that contains $-x +$ supp (F). According to the next lemma, an aperiodic distribution is called strongly aperiodic in [23, p. 42].

Lemma 2.4 *If $x \in$ supp (F), then $G_x = R_0$. The distribution F is aperiodic if and only if $G_x = \mathbb{Z}^d$ for some $x \in$ supp (F) or equivalently if and only if $G_x = \mathbb{Z}^d$ for all $x \in \mathbb{Z}^d$.*

Proof Let $x \in \mathbb{Z}^d$. Let $z \in R_0$. There exists $n, n' \in \mathbb{N}$ and $x_i, x_i', y_i, y_i' \in \text{supp }(F)$ for all $i \in \mathbb{N}^*$ such that $\sum_{i=1}^{n}(y_i - x_i) - \sum_{i=1}^{n'}(y_i' - x_i') = z$. This implies that $\sum_{i=1}^{n}(y_i - x) + \sum_{i=1}^{n'}(x_i' - x) - \sum_{i=1}^{n}(x_i - x) - \sum_{i=1}^{n'}(y_i' - x) = z$ and thus $z \in G_x$. This gives $R_0 \subset G_x$.

For $x \in \text{supp }(F)$, we get $G_x \subset R_0$ and thus $G_x = R_0$. The end of the lemma is obvious. □

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 offspring 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 the 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 to be 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 2.5 Let $p = (p^{(i)}, i \in [d])$ be an offspring distribution. We say that p is aperiodic, if the smallest subgroup of $\mathbb{Z}^d$ that contains $\bigcup_{i=1}^{d} \left(\text{supp }(p^{(i)}) - \text{supp }(p^{(i)})\right)$ is $\mathbb{Z}^d$.

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 2.6 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 $\langle a, \mathbf{1} \rangle = \langle a, a^* \rangle = 1$. Those eigenvectors correspond to the eigenvalue $\rho = 1$. 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 $k \in \mathbb{N}^d$,

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

For α a probability distribution on $[d]$, we also define the corresponding size-biased distribution $\hat{\alpha} = (\hat{\alpha}(i), i \in [d])$ by, for $i \in [d]$:

$$\hat{\alpha}(i) = \alpha(i) \frac{a_i^*}{\langle \alpha, a^* \rangle}. \tag{4}$$

Definition 2.7 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 with the root-type distribution α is defined as follows:

- Marked individuals are normal or special.
- The root of τ^* is special and its type has distribution $\hat{\alpha}$.
- 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 $\tau^* \setminus \mathbf{v}^*$ 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 (i.e., $\alpha = \delta_r$ the Dirac mass at r). From [15], we get that for $h \in \mathbb{N}$, $\mathbf{t} \in \mathbb{T}^{(h)}$ with $\mathcal{M}(\emptyset_\mathbf{t}) = r$, and $x \in \mathcal{L}_0(\mathbf{t})$ with $|x| = h$ and $\mathcal{M}(x) = i$:

$$\mathbb{P}_r(r_h(\tau^*) = \mathbf{t},\ v_h^* = x) = \frac{a_i^*}{a_r^*}\, \mathbb{P}_r(r_h(\tau) = \mathbf{t}). \tag{5}$$

Notice that if M is primitive and p is critical or subcritical, 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 2.8 *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 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 Theorem 3.1.

Lemma 2.9 *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 $\mathcal{M}(\emptyset_\mathbf{t}) = r$, $x \in \mathcal{L}_0(\mathbf{t})$ with $\mathcal{M}(x) = i \in [d]$, and $k \in \mathbb{N}^d$ such that $k \geq |\mathbf{t}|$, we have:*

$$\mathbb{P}_r(\tau \in \mathbb{T}(\mathbf{t}, x) \big|\, |\tau| = k) = \frac{a_r^*}{a_i^*} \frac{\mathbb{P}_i(|\tau| = k - |\mathbf{t}| + \mathbf{e}_i)}{\mathbb{P}_r(|\tau| = k)}\, \mathbb{P}_r(\tau^* \in \mathbb{T}(\mathbf{t}, x)). \tag{6}$$

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 (5) that:

$$\mathbb{P}_r(\tau^* \in \mathbb{T}(\mathbf{t}, x)) = \frac{a_i^*}{a_r^*}\, \mathbb{P}_r(\tau \in \mathbb{T}(\mathbf{t}, x)). \tag{7}$$

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 (7) to deduce (6). □

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 positive normalized left and right eigenvectors of the mean matrix M associated with the eigenvalue $\rho = 1$ such that $\langle a, a^* \rangle = \sum a_i = 1$. The proof of the following main theorem is given in Sect. 3.3.

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 root-type distribution* α, *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 *and* α.

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 that the population of type i, $|\tau^{(i)}|$, is large. According to Proposition 4 in [17], 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 Eq. (19), we easily get that τ conditioned on $|\tau^{(i)}|$ being large converges locally to Kesten's tree. See [24] 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 [24]. 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 [20] or from [17]. 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' is the same. Then the key Eq. (19) 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 Eq. (19).

Remark 3.4 The change in offspring distribution given in Section 1.4 of [19], when it exists, allows to extend Theorem 3.1 to subcritical multi-type GW trees. In order to consider an asymptotic proportion of types different from the one given by the (positive normalized) left eigenvector associated with the Perron–Frobenius eigenvalue, one has to change the offspring distribution, see Theorem 3 of [19]. However, this requires exponential moments for the offspring distribution.

We end this section by using Theorem 3.1 to extend results of [1] on mono-type GW tree in the following sense. Let τ be a mono-type GW tree (that is $d = 1$) with critical aperiodic offspring distribution $q = (q(\ell), \ell \in \mathbb{N})$. Let f_q denote the generating function of q and $\mathcal{Q} = \{\gamma > 0; f_q(\gamma) < +\infty\}$ its domain on $(0, +\infty)$.

Let $d \geq 2$ and assume that Card (supp q) $\geq d + 1$. Since q is critical, we have $0 \in$ supp q. Let $A_1, \ldots, A_d$ be a partition of supp q such that $0 \in A_1$ and Card $(A_1) > 1$. We set $\alpha(i) = \sum_{\ell \in A_i} q(\ell)$ for all $i \in [d]$. Notice that α is a positive probability distribution and $\alpha(1) > q(0)$. We set $|\tau| = (|\tau^{(i)}|, i \in [d])$ where $|\tau^{(i)}|$ be the number of individuals of τ whose number of offspring belongs to A_i.

For $x = (x_i, i \in [d])$, we set $\mathfrak{m}_x := \sum_{i\in[d]} x_i \inf A_i$ and for $\gamma \in \mathcal{Q}$:

$$h_x(\gamma) = \sum_{i\in[d]} x_i \frac{\gamma f'_{A_i}(\gamma)}{f_{A_i}(\gamma)} \quad \text{with} \quad f_{A_i}(\gamma) = \sum_{\ell \in A_i} \gamma^\ell q(\ell) \quad \text{for all } i \in [d].$$

Corollary 3.5 *Let q be a critical aperiodic offspring distribution. Let $\tilde{\alpha} = (\tilde{\alpha}(i), i \in [d]) \in \mathbb{R}^d$ be such that $\tilde{\alpha} > 0$ and $\langle \tilde{\alpha}, \mathbf{1} \rangle = 1$, so that $\tilde{\alpha}$ is a non-degenerate proportion. Assume that:*

$$\mathfrak{m}_{\tilde{\alpha}} < 1 \tag{8}$$

and

$$\text{there exists a (unique) } \gamma \in \mathcal{Q} \text{ such that } \; h_{\tilde{\alpha}}(\gamma) = 1. \tag{9}$$

Let $(k(n), n \in \mathbb{N}^*)$ be a sequence of $\mathbb{N}^d$ satisfying $\lim_{n\to\infty} |k(n)| = +\infty$ and $\lim_{n\to\infty} k(n)/|k(n)| = \tilde{\alpha}$. Let τ be a random mono-type GW tree with offspring distribution q, 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 $\tilde{\tau}^*$ associated with the offspring distribution $\tilde{q} = (\tilde{q}(\ell), \ell \in \mathbb{N})$ where for $i \in [d]$, $\ell \in A_i$:

$$\tilde{q}(\ell) = \frac{\tilde{\alpha}(i)}{f_{A_i}(\gamma)} \gamma^\ell q(\ell).$$

Notice that if $\tilde{\alpha} = \alpha$, then condition (8) holds as Card $(A_1) > 1$ and q is critical, and condition (9) also holds with $\gamma = 1$ as q is critical. We deduce that if $\tilde{\alpha} = \alpha$, then $\tilde{q} = q$ and $\tilde{\tau}^* = \tau^*$ is simply the Kesten's tree associated with the offspring distribution q. We now comment on the conditions (8) and (9).

Remark 3.6 One can see that condition (8) is almost optimal. This is easy to check in the binary case. Assume $q(0) + q(1) + q(2) = 1$, $q(0)q(1)q(2) > 0$, $A_1 = \{0, 1\}$ and $A_2 = \{2\}$. Since we always have $|\tau^{(1)}| > |\tau^{(2)}|$, then any asymptotic proportion has to satisfy $\tilde{\alpha}(1) \geq \tilde{\alpha}(2)$ that is $\mathfrak{m}_{\tilde{\alpha}} = 2\tilde{\alpha}(2) \leq 1$.

Remark 3.7 For $i \in [d]$, let $\mathrm{P}_{i,\gamma}$ denote the distribution of a random variable Z taking values in A_i such that $\mathrm{P}_{i,\gamma}(Z = \ell) = \mathbf{1}_{\{\ell \in A_i\}} \gamma^\ell q(\ell)/f_{A_i}(\gamma)$. In particular, we have $h_{\tilde{\alpha}}(\gamma) = \sum_{i \in [d]} \tilde{\alpha}(i) \mathrm{E}_{i,\gamma}[Z]$. An elementary computation gives that $\partial_\gamma \mathrm{E}_{i,\gamma}[Z] = \gamma^{-1} \mathrm{Var}_{\,i,\gamma}(Z)$. Using that Card $(A_1) > 1$, we get $\mathrm{Var}_{\,1,\gamma}(Z) > 0$ and thus $h'_{\tilde{\alpha}}$ is positive on $\mathcal{Q}$. We deduce that, if it exists, the root of Eq. (9) is then unique. Since $\lim_{\gamma \to 0} h_{\tilde{\alpha}}(\gamma) = \mathfrak{m}_{\tilde{\alpha}}$, we deduce that condition (8) implies $h_{\tilde{\alpha}}(0+) < 1$. A necessary and sufficient condition for the existence of a root to (9) is that $\lim_{\gamma \uparrow R} h_{\tilde{\alpha}}(\gamma) \geq 1$, with $R = \sup \mathcal{Q}$ the radius of convergence of the series f_q. A sufficient condition to get this latter condition is for example $\lim_{\gamma \uparrow R} f'_q(\gamma) = +\infty$ or even the stronger condition $R = +\infty$.

As noticed earlier, for $\tilde{\alpha} = \alpha$, as q is critical, we get that $h_\alpha(1) = 1$. So in this case no further hypothesis is needed. We also deduce that if $\tilde{\alpha}(i) \geq \alpha(i)$ for all $i \in [d]$ such that $f_{A_i}(1) > 0$, then we have $h_{\tilde{\alpha}}(1) \geq h_\alpha(1) = 1$. And thus, in this particular case also, the root of (9) exists without further assumptions on q.

Proof of Corollary 3.5 We 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 belongs to A_i. The corresponding root-type distribution is α and the corresponding offspring distribution $p = (p^{(i)}, i \in [d])$ is defined as follows: for $k = (k_1, \ldots, k_d) \in \mathbb{N}^d$,

$$p^{(i)}(k) = \mathbf{1}_{\{|k| \in A_i\}} \frac{q(|k|)}{\alpha(i)} \binom{|k|}{k} \alpha^k, \tag{10}$$

where we recall that $\alpha^k = \prod_{j \in [d]} \alpha(j)^{k_j}$ and $|k| = \sum_{i \in [d]} k_i$, and we use the following notation for the multinomial coefficient $\binom{|k|}{k} = |k|!/\prod_{i \in [d]} k_i!$. For simplicity we shall still denote the corresponding multi-type GW tree by τ. We define $\alpha^* = (\alpha^*(i), i \in [d])$ with $\alpha^*(i) = \sum_{\ell \in A_i} \ell q(\ell)/\alpha(i)$ so that $\langle \alpha, \mathbf{1} \rangle = \langle \alpha, \alpha^* \rangle = 1$. Notice that α^* is positive as $\alpha(1) > q(0)$ and $0 \in A_1$. It is easy to check that the mean matrix is given by $M = (\alpha^*)^T \alpha$. Its only nonzero eigenvalue is 1 and α and α^* are the nonnegative-associated left and right eigenvectors. The mean matrix M is primitive as all its entries are positive. We get that condition (H_1) holds. Notice (H_2) holds as we assumed q is aperiodic.

We first consider the case $\tilde{\alpha} = \alpha$. (As noticed just after Corollary 3.5, condition (8) holds and condition (9) also holds with $\gamma = 1$). We easily deduce from Theorem 3.1 that if $(k(n), n \in \mathbb{N}^*)$ is a sequence of $\mathbb{N}^d$ satisfying $\lim_{n \to \infty} k(n)/|k(n)| = \alpha$ 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 (d-type) Kesten's tree τ^* associated with the offspring distribution p and root-type distribution α.

Let $\hat{\tau}^*$ be the mono-type Kesten's tree associated with q. We shall check that τ^* is distributed as $\hat{\tau}^*$ seen as a multi-type GW tree, where an individual $u \in \hat{\tau}^*$ is of type i if the number of offspring of u belongs to A_i and that u is normal if it has a finite number of descendants (i.e., Card $(\{v \in \hat{\tau}^*;\ u \prec v\}) < +\infty$) and special otherwise.

Let $\mathbf{t} \in \mathbb{T}_0$ and $x \in \mathcal{L}_0(\mathbf{t})$. For $i \in [d]$, we set $x_i = \{\hat{x}, i\}$ and $\mathbf{t}_i = \{x_i\} \bigcup (\mathbf{t} \backslash \{x\})$, the tree which is equal to $\mathbf{t}$ except for the leaf $\hat{x}$ which is of type i instead of $\mathcal{M}(x)$. We denote by $\mathbb{T}(\mathbf{t}, \hat{x})$ the set of trees obtained by grafting trees on the leaf x of $\mathbf{t}$ with possibly changing the type of x, that is, $\mathbb{T}(\mathbf{t}, \hat{x}) = \bigcup_{i \in [d]} \mathbb{T}(\mathbf{t}_i, x_i)$. We write $\mathbb{P}_{\hat{\alpha}}(d\tau^*) = \sum_{r \in [d]} \hat{\alpha}(r) \mathbb{P}_r(d\tau^*)$. We have for $i \in [d]$:

$$\mathbb{P}_{\hat{\alpha}}(\tau^* \in \mathbb{T}(\mathbf{t}_i, x_i)) = \alpha^*(i) \sum_{r \in [d]} \frac{\hat{\alpha}(r)}{\alpha^*(r)} \mathbb{P}_r(\tau \in \mathbb{T}(\mathbf{t}_i, x_i)) = \alpha^*(i)\, \mathbb{P}(\tau \in \mathbb{T}(\mathbf{t}_i, x_i)),$$

where we used (7) for the first equality and (4) as well as $\langle \alpha, \alpha^* \rangle = 1$ for the second. Using that $\mathbb{P}(\tau \in \mathbb{T}(\mathbf{t}_i, x_i)) = \alpha(i) \mathbb{P}(\tau \in \mathbb{T}(\mathbf{t}, \hat{x}))$ and $\langle \alpha, \alpha^* \rangle = 1$, we deduce that:

$$\mathbb{P}_{\hat{\alpha}}(\tau^* \in \mathbb{T}(\mathbf{t}, \hat{x})) = \sum_{i \in [d]} \alpha^*(i)\, \mathbb{P}(\tau \in \mathbb{T}(\mathbf{t}_i, x_i)) = \mathbb{P}(\tau \in \mathbb{T}(\mathbf{t}, \hat{x})).$$

Using (7) in the mono-type case, we get $\mathbb{P}(\hat{\tau}^* \in \mathbb{T}(\mathbf{t}, \hat{x})) = \mathbb{P}(\tau \in \mathbb{T}(\mathbf{t}, \hat{x}))$ and thus $\mathbb{P}_{\hat{\alpha}}(\tau^* \in \mathbb{T}(\mathbf{t}, \hat{x})) = \mathbb{P}(\hat{\tau}^* \in \mathbb{T}(\mathbf{t}, \hat{x}))$. It is left to the reader to check that $\mathcal{F}' = \{\mathbb{T}(\mathbf{t}, \hat{x});\ \mathbf{t} \in \mathbb{T}_0, x \in \mathcal{L}_0(\mathbf{t})\}$ is a separating class on $\mathbb{T}_1'$. Hence τ^* is distributed as $\hat{\tau}^*$ and can thus be seen as the (mono-type) Kesten tree associated with the offspring distribution q.

We now shall condition on a general asymptotic proportion $\tilde{\alpha} \in \mathbb{R}^d$ satisfying condition (8) and condition (9). We assume that there exists a root to the Eq. (9), say γ. This root is unique according to Remark 3.7. The probability $\tilde{q}$ defined in Corollary 3.5 is a critical (as γ is a root of (9)) and aperiodic (as q is aperiodic) and that $\tilde{\alpha}(i) = \sum_{\ell \in A_i} \tilde{q}(\ell)$ for all $i \in [d]$. Let $\tilde{\tau}$ be a mono-type GW tree with offspring distribution $\tilde{q}$ (which can also be seen as a multi-type GW tree where the type of an individual is A_i if the number of its offspring lies in A_i). We deduce that for all $\mathbf{t} \in \mathbb{T}_0$, we have:

$$\mathbb{P}_{\tilde{\alpha}}(\tilde{\tau} = \mathbf{t}) = \prod_{u \in \mathbf{t}} \tilde{q}(k_u[\mathbf{t}]) = \gamma^{\langle |\mathbf{t}|, \mathbf{1} \rangle - 1} \Gamma^{|\mathbf{t}|}\, \mathbb{P}_\alpha(\tau = \mathbf{t}),$$

where $\Gamma = (\tilde{\alpha}(i)/f_{A_i}(\gamma), i \in [d])$. In particular, for all $k \in \mathbb{N}^d$, the random tree τ_n, which is distributed as τ conditionally on $\{|\tau| = k\}$, has the same distribution as the random tree $\tilde{\tau}_n$, which is distributed as $\tilde{\tau}$ conditionally on $\{|\tilde{\tau}| = k\}$. According to the first part, since $\tilde{q}$ is critical aperiodic, we deduce that if $(k(n), n \in \mathbb{N}^*)$ is a sequence of $\mathbb{N}^d$ satisfying $\lim_{n \to \infty} k(n)/|k(n)| = \tilde{\alpha}$ with $\lim_{n \to +\infty} |k(n)| = +\infty$, then $\tilde{\tau}_n$, and thus τ_n, converges in distribution to the mono-type Kesten's tree $\tilde{\tau}^*$ associated with $\tilde{q}$. □

3.2 Around the Dwass Formula

Let τ be a random GW tree with critical offspring distribution p. We have no assumption on p for the moment. For $i, j \in [d]$, we define the total number of individuals of type i whose parent is of type j:

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

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 $k = (k_1, \dots, k_d) \in \mathbb{N}^d$, 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 $S_{i,n}^{(j)} = \sum_{\ell=1}^{n} X_{i,\ell}^{(j)}$. 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:

$$\mathcal{S}_{ij}(k,r) = -S_{i,k_i}^{(j)} + (S_k^{(j)} + \mathbf{1}_{\{r=i\}})\mathbf{1}_{\{i=j\}}. \tag{11}$$

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

Lemma 3.8 *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 $r \in [d]$, $k = (k_1, \dots, k_d) \in (\mathbb{N}^*)^d$, $\kappa = (\kappa_{ij};\, i, j \in [d]) \in \mathbb{N}^{d\times d}$ such that

$$k = \mathbf{e}_r + \sum_{j\in[d]} \kappa_j, \tag{12}$$

we have:

$$\mathbb{P}_r(\mathcal{B} = \kappa) = \det(\Delta(k) - \kappa) \prod_{j\in[d]} \frac{\mathbb{P}(S_{j,k_j} = \kappa_j)}{k_j}, \tag{13}$$

where $\Delta(k)$ is the $d \times d$ diagonal matrix with diagonal k. Notice that additional hypotheses on the offspring distribution p were required in Theorem 1.2 from [5]. However, for fixed κ, (13) is a finite algebraic expression of p. According to [5], it holds in particular for all p such that there exists a finite constant $c \geq 2$ and $p^{(i)}(k) > 0$ if $|k| \leq c$ and $p^{(i)}(k) = 0$ if $|k| > c$ for all $i \in [d]$. This gives that (13) holds for all p.

Because of (12), we have:

$$\mathbb{P}_r(|\tau| = k,\ \mathcal{B} = \kappa) = \mathbb{P}_r(\mathcal{B} = \kappa). \tag{14}$$

Thanks to the definition of $\mathcal{S}(k, r)$, we have that $\Delta(k) - \kappa$ is equal to the transpose of $\mathcal{S}(k, r)$ on $\bigcap_{j\in[d]} \{S_{j,k_j} = \kappa_j\}$. By summing (14) and thus (13) over all the possible values of κ such that (12) 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 (15) below. (As we are interested in computing the determinant of a matrix whose all columns 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, and 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 $j \in [d]\backslash\{r\}$, 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 3.9 *For $r \in [d]$ and $k \in (\mathbb{N}^*)^d$, we have:*

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

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 $\kappa = (\kappa_{ij}; i, j \in [d]) \in \mathbb{R}^{d\times d}$

$$\det(\kappa) = (-1)^d \sum_{\mathbf{f}\in\mathbb{F}} \prod_{j\in[d]} \kappa_{j_{\mathbf{f}}, j}, \tag{15}$$

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

Thanks to Definition (11) of $\mathcal{S}(k,r)$, this implies that for $r \in [d]$ and $k \in (\mathbb{N}^*)^d$, we have:

$$\det(\mathcal{S}(k,r)) = \sum_{\mathbf{f}\in\mathbb{F}} \prod_{j\in[d]} S^{(j)}_{j_{\mathbf{f}},k_{j_{\mathbf{f}}}}, \tag{16}$$

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 $\mathbf{t}$. □

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 $\mathbf{A}^{(\ell,K)} = (A_1^K, \dots, A_\ell^K)$ 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 $m = (m_1, \dots, m_d) \in \mathbb{N}^d$ such that, for all $i \in [d]$, $m_i = 0$ if Card $(C_i(\mathbf{t})) = 0$ and $1 \le m_i \le$ Card $(C_i(\mathbf{t}))$ if 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 $k_i \ge d$ 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 $\sum_{i\in[d]} \sum_{\ell=1}^{m_i}$ Card $(A_\ell^{C_i(\mathbf{t})}) = d-1$, we deduce that:

$$0 \le G(\mathcal{A}) \le \left|\tilde{S}_{m_{\mathcal{A}}}\right|^{d-1}. \tag{17}$$

We have the following result.

Corollary 3.10 *For $r \in [d]$ and $b, k \in (\mathbb{N}^*)^d$ such that $k \ge d\mathbf{1}$, 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 Lemma 3.9 to conclude. □

3.3 Proof of Theorem 3.1

We assume that (H_1) and (H_2) hold. Let $r \in [d]$. Let $b \in \mathbb{N}^d$. We have, using Lemma 3.8 and Corollary 3.10, for every $k \geq b + \mathbf{1}$ such that $\mathbb{P}_r(|\tau| = k) > 0$:

$$\begin{aligned}
&\frac{\prod_{i\in[d]}(k_i - b_i)}{\prod_{i\in[d]} k_i}\,\frac{\mathbb{P}_r(|\tau| = k - b)}{\mathbb{P}_r(|\tau| = k)} \\
&= \frac{\mathbb{E}\left[\det(\mathcal{S}(k-b, r));\ S_{k-b} + \mathbf{e}_r = k - b\right]}{\mathbb{E}\left[\det(\mathcal{S}(k, r));\ S_k + \mathbf{e}_r = k\right]} \\
&= \frac{\sum_{\mathbf{t}\in\mathcal{T}_r}\sum_{\mathcal{A}\in\mathbb{A}_\mathbf{t}} B_{k-b}(m_\mathcal{A})\,\mathbb{E}\left[G(\mathcal{A});\ \tilde{S}_{m_\mathcal{A}} + S_{k-b-m_\mathcal{A}} = k - b - \mathbf{e}_r\right]}{\sum_{\mathbf{t}\in\mathcal{T}_r}\sum_{\mathcal{A}\in\mathbb{A}_\mathbf{t}} B_k(m_\mathcal{A})\,\mathbb{E}\left[G(\mathcal{A});\ \tilde{S}_{m_\mathcal{A}} + S_{k-m_\mathcal{A}} = k - \mathbf{e}_r\right]}.
\end{aligned} \tag{18}$$

The next lemma is an extension of the strong ratio limit theorem given in [1]. Its proof is postponed to Sect. 3.4. Recall that a is the positive normalized left eigenvector of the mean matrix M. (Notice that no moment condition is assumed for G or H).

Lemma 3.11 *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. $G \leq |H|^c$ for some $c \geq 1$.*

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:

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

Let $(k(n), n \in \mathbb{N}^*)$ be a sequence of elements 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 (17), we deduce from Lemma 3.11 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 (18) are nonnegative, and $\lim_{n\to+\infty} \prod_{i\in[d]}(k_i(n)-b_i)/k_i(n) = 1$, we deduce that:

$$\lim_{n\to+\infty} \frac{\mathbb{P}_r(|\tau| = k(n) - b)}{\mathbb{P}_r(|\tau| = k(n))} = 1. \tag{19}$$

Then, using Lemma 2.9 (with $i = r$ in (6)), we obtain that, for all $r \in [d]$, $\mathbf{t} \in \mathbb{T}_0$ and $x \in \mathcal{L}_0(\mathbf{t})$ such that $\mathcal{M}(x) = r$, $\lim_{n\to+\infty} \mathbb{P}_r(\tau_n \in \mathbb{T}(\mathbf{t}, x)) = \mathbb{P}_r(\tau^* \in \mathbb{T}(\mathbf{t}, x))$. Of course, we have for $\mathbf{t} \in \mathbb{T}_0$ and n large enough 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 toward τ^* under $\mathbb{P}_r$ for all $r \in [d]$.

Let α be a probability distribution on $[d]$. Let $\mathbf{t} \in \mathbb{T}_0$ and $x \in \mathcal{L}_0(\mathbf{t})$. Set $r = \mathcal{M}(\emptyset_{\mathbf{t}})$ and $i = \mathcal{M}(x)$. We have using (6) that:

$$\begin{aligned}\mathbb{P}_\alpha(\tau_n \in \mathbb{T}(\mathbf{t}, x)) &= \frac{\alpha(r)\,\mathbb{P}_r(\tau \in \mathbb{T}(\mathbf{t}, x),\ |\tau| = k(n))}{\sum_{j\in[d]} \alpha(j)\,\mathbb{P}_j(|\tau| = k(n))} \\ &= \frac{\alpha(r)a_r^*}{\sum_{j\in[d]} \alpha(j)a_i^*\Gamma_j(n)}\,\mathbb{P}_r(\tau^* \in \mathbb{T}(\mathbf{t}, x)),\end{aligned}$$

where

$$\Gamma_j(n) = \frac{\mathbb{P}_j(|\tau| = k(n))}{\mathbb{P}_i(|\tau| = k(n) - |\mathbf{t}| + \mathbf{e}_i)},$$

with the convention that $\Gamma_j(n) = +\infty$ if $\mathbb{P}_i(|\tau| = k(n) - |\mathbf{t}| + \mathbf{e}_i) = 0$. Let $\mathbf{t}' \in \mathbb{T}_0$ and $x' \in \mathbf{t}'$ such that $\mathcal{M}(x) = i$ and $\mathbb{P}_j(\tau^* \in \mathbb{T}(\mathbf{t}', x')) > 0$ (which is possible thanks to Lemma 2.8). Using (6) and the convergence of $(\tau_n, n \in \mathbb{N}^*)$ toward τ^* under $\mathbb{P}_j$, we deduce that $\lim_{n\to+\infty} \frac{\mathbb{P}_j(|\tau|=k(n))}{\mathbb{P}_i(|\tau|=k(n)-|\mathbf{t}'|+\mathbf{e}_i)} = a_j^*/a_i^*$. Then use (19) (with $r = i$) to deduce that $\lim_{n\to+\infty} \Gamma_j(n) = a_j^*/a_i^*$ for all $j \in [d]$. Using the definition (4) of $\hat{\alpha}$, we deduce that:

$$\lim_{n\to+\infty} \mathbb{P}_\alpha(\tau_n \in \mathbb{T}(\mathbf{t}, x)) = \frac{\alpha(r)a_r^*}{\sum_{j\in[d]} \alpha(j)a_j^*}\,\mathbb{P}_r(\tau^* \in \mathbb{T}(\mathbf{t}, x)) = \mathbb{P}_{\hat{\alpha}}(\tau^* \in \mathbb{T}(\mathbf{t}, x)),$$

where $\mathbb{P}_{\hat{\alpha}}(d\tau^*) = \sum_{r\in[d]} \hat{\alpha}(r)\mathbb{P}_r(d\tau^*)$. This proves Theorem 3.1, assuming Lemma 3.11. □

3.4 Proof of Lemma 3.11

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 positive 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 an aperiodic probability distribution.

Lemma 3.12 *Under (H_2), the distribution of Y on $\mathbb{Z}^{2d-1}$ is aperiodic.*

Proof Recall $\mathbf{v}_d = 0 \in \mathbb{N}^{d-1}$. Let H be the smallest subgroup of $\mathbb{Z}^{2d-1}$ that contains supp (F) − supp (F), with F be the probability distribution of Y. In particular, we have that H contains (supp $(p^{(i)})$ − supp $(p^{(i)})$) $\times \{\mathbf{v}_d\}$ for all $i \in [d]$ and thus their union. Since (H_2) holds, we deduce that H contains $\mathbb{Z}^d \times \{\mathbf{v}_d\}$. This implies also that $(0, \mathbf{v}_i)$ belongs to H for all $i \in [d]$, and thus $H = \mathbb{Z}^{2d-1}$. □

For $x \in \mathbb{R}^d$ and $z = (z_1, \dots, z_d) \in \mathbb{R}^d$, we set $\delta(x, z) = (x, z_1, \dots, z_{d-1})$. By definition of Y and since a is the left eigenvector of the mean matrix with eigenvalue 1, we have $\mathbb{E}[Y] = \delta(a, a)$.

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:

$$\mathbb{P}\left(W_{|k|} = \delta(s, k)\right) = D(k)\mathbb{P}(S_k = s) \quad \text{with} \quad D(k) = \frac{|k|!}{\prod_{i\in[d]} k_i!} \prod_{i\in[d]} {a_i}^{k_i}. \tag{20}$$

Recall G and H given in Lemma 3.11. We set $H' = \delta(H, 0) \in \mathbb{N}^{2d-1}$. We get for k, m, s and b in $\mathbb{N}^d$:

$$\frac{\mathbb{E}[G;\ H + S_{k-m} = s - b]}{\mathbb{E}[G;\ H + S_k = s]} = \frac{D(k)}{D(k-m)} \frac{\mathbb{E}[G;\ H' + W_{|k|-|m|} = \delta(s, k) - \delta(b, m)]}{\mathbb{E}[G;\ H' + W_{|k|} = \delta(s, k)]}. \tag{21}$$

Thanks to Lemma 3.12 and (H_2), the distribution of Y on $\mathbb{Z}^{2d-1}$ is aperiodic. Since $0 \le G \le |H|^c$, we also have $0 \le G \le |H'|^c$ 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 4.11, which is stated and proved in Sect. 4.5 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 (21) 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 3.11, assuming Lemma 4.11. □

4 Remaining Proofs

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 $\mathbb{R}^d$ with distribution F. Denote by supp (F) the closed support of F: $x \notin$ 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:

$$\varphi(\theta) = \log \left(\mathbb{E}\left[\mathrm{e}^{\langle \theta, X \rangle} \right] \right), \quad \theta \in \mathbb{R}^d. \tag{22}$$

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:

$$\psi(x) = \sup_{\theta \in \text{dom}(\varphi)} \left(\langle \theta, x \rangle - \varphi(\theta) \right), \quad x \in \mathbb{R}^d. \tag{23}$$

We recall that ψ is a lower semicontinuous (proper) convex function. Since $\varphi(0) = 0$, we deduce that ψ is nonnegative. 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). *If $\psi(x) = 0$, then we have $x \in$* ri dom(ψ).

Proof Let $x \notin$ cl ri $(F) =$ cl cv (F). According to the separation theorem, there exists $\theta \in \mathbb{R}^d$ and $\varepsilon > 0$ such that a.s. $\langle \theta, X - x \rangle \leq -\varepsilon$. This gives that for all $t > 0$, $\varphi(t\theta) - t\langle \theta, x \rangle \leq -t\varepsilon$ and thus $\psi(x) \geq \sup_{t>0} t\varepsilon = +\infty$. This implies that $\text{dom}(\psi) \subset$ cl ri (F).

Let $x \in$ ri (F). By translation invariance, we can assume that $x = 0$. We set $h(\theta) = \mathbb{E}[\max(0, \min(1, \langle \theta, X \rangle))]$. The function h is continuous and, since $0 \in$ ri (F), it is nonzero on $\mathcal{A} = \{\theta \in \text{aff } (F), |\theta| = 1\}$. Thus h has a strictly positive minimum on $\mathcal{A}$. Since $\mathbb{P}(\langle \theta, X \rangle > 0) \geq h(\theta/|\theta|)$ for $\theta \neq 0$, we deduce that $a = \inf_{\theta \in \text{aff } (F) \setminus \{0\}} \log(\mathbb{P}(\langle \theta, X \rangle > 0))$ is finite. For $\theta \in \mathbb{R}^d$, let θ_F denote its orthogonal projection on aff (F). If $\theta_F = 0$, then $\varphi(\theta) = 0$, otherwise we have $\varphi(\theta) = \varphi(\theta_F) \geq \log(\mathbb{P}(\langle \theta_F, X \rangle > 0)) \geq a$. We deduce that $\varphi \geq a$ and we get $\psi(x) = \psi(0) \leq -a$. We deduce that $x \in \text{dom}(\psi)$. This implies that ri $(F) \subset \text{dom}(\psi)$.

We deduce that ri $(F) \subset \mathrm{dom}(\psi) \subset$ cl ri (F), which gives that ri $(F) =$ ri dom(ψ).

We denote by $\partial(F) =$ cl ri $(F)\backslash$ri F the relative boundary of dom(ψ). Let $x \in \partial(F)$. Let X be a random variable with probability distribution F. According to the separation theorem, there exists $q \in \mathbb{R}^d$ such that a.s. $\langle q, X - x\rangle \leq 0$ and $\mathbb{P}(\langle q, X - x\rangle < 0) > 0$. This implies that $\varphi(q) < \langle q, x\rangle$ and thus $\psi(x) \geq \langle q, x\rangle - \varphi(q) > 0$. This gives that $\psi(x) = 0$ implies $x \in$ ri dom(ψ). □

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 Jensen's 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 (23) and that ψ is nonnegative to deduce that $\psi(\mathbb{E}[X]) = 0$. Use Lemma 4.1 to conclude. □

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

$$d\mathbb{P}_\theta(X \in dx) = \mathrm{e}^{\langle\theta, X\rangle - \varphi(\theta)}\, d\mathbb{P}(X \in dx). \tag{24}$$

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

$$m_\theta = \mathbb{E}_\theta[X] = \mathbb{E}[X\, \mathrm{e}^{\langle\theta, X\rangle - \varphi(\theta)}] = \nabla\varphi(\theta) \quad \text{and} \quad \Sigma_\theta = \mathrm{Cov}_\theta(X, X). \tag{25}$$

We set $\mathcal{I}_F = \mathrm{int}\ \mathrm{dom}(\varphi)$ 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. The elementary proof of the next lemma is left to the reader.

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

$$\sup_{\theta\in K} |\Sigma_\theta| < +\infty \quad \text{and} \quad \sup_{\theta\in K} \mathbb{E}_\theta\left[|X - m_\theta|^3\right] < +\infty. \tag{26}$$

We set $\mathcal{O}_F = \mathrm{int\ cv}\ (F)$ 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's 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 $\mathrm{dom}(\varphi) = \mathbb{R}^d$, as well as differentiable throughout $\mathbb{R}^d$. This implies that φ is smooth on $\mathbb{R}^d$ in the sense of [21] Section 26. According to Theorem 26.5 in [21], 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 is $\nabla\varphi^{-1}$. Furthermore, according to Corollary 26.4.1

in [21], we have ri $\mathrm{dom}(\psi) \subset D \subset \mathrm{dom}(\psi)$. Since D is open, we deduce that $D = \mathrm{ri}\ \mathrm{dom}(\psi) = \mathrm{int}\ \mathrm{dom}(\psi)$. Then, use Lemma 4.1 to get that $D = \mathrm{ri}\ (F) = \mathcal{O}_F$. □

Recall Definition 2.3 for an aperiodic probability distribution.

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

$$\inf_{\theta \in K} |\Sigma_\theta| > 0. \tag{27}$$

Proof Since F is aperiodic, we have aff $(F) = \mathbb{R}^d$. This implies the first part of the lemma.

Let $\theta \in \mathcal{I}_F$ be such that $|\Sigma_\theta| = 0$. Then there exists $h \in \mathbb{R}^d \backslash \{0\}$ 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. Since aff $(F) = \mathbb{R}^d$, we get $h = 0$. Since this is absurd, we deduce that $|\Sigma_\theta| > 0$ for all $\theta \in \mathcal{I}_F$. Then use the continuity of $\theta \mapsto |\Sigma_\theta|$ 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 (22), (24) and (25) and that $\mathcal{I}_F = \mathrm{int}\ \mathrm{dom}(\varphi)$. The next theorem is an extension of the one-dimensional theorem of Gnedenko [7], see also [22,25].

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:*

$$\lim_{n\to\infty} \sup_{\theta \in K} \sup_{s \in \mathbb{Z}^d} \left| n^{d/2} |\Sigma_\theta|^{1/2} \mathbb{P}_\theta(S_n = s) - (2\pi)^{-d/2}\, \mathrm{e}^{-\|z_n(\theta,s)\|^2/2} \right| = 0, \tag{28}$$

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

The end of this section is devoted to the proof of Theorem 4.6. This proof is a straightforward extension to the one-dimensional case. For completeness, we give it in detail.

Let $K \subset \mathcal{I}_F$ be compact. Thanks to Lemmas 4.3 and 4.5, we have $|\Sigma_\theta| > 0$ and $\Sigma_\theta^{-1/2}$ is well defined. We define:

$$Y = n^{-1/2} \Sigma_\theta^{-1/2}(X_1 - m_\theta) \quad \text{and} \quad f_\theta(t) = \mathbb{E}_\theta \left[\mathrm{e}^{i\langle t, Y \rangle} \right]. \tag{29}$$

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

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

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

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

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

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

We obtain:

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

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

$$\lim_{n\to\infty} C_n = \infty \quad \text{and} \quad \lim_{n\to\infty} n^{-1/(12+6d)} C_n = 0. \tag{30}$$

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:

$$c_1 := \sup_{\theta\in K,\, \|t\|=1} \langle t, \Sigma_\theta^{-1} t\rangle < \infty. \tag{31}$$

Set $J_1 = \{t \in \mathbb{R}^d;\ \| t \| \leq C_n\}$, so that $t \in J_1$ implies $\| n^{-1/2}\Sigma_\theta^{-1/2} t \|^2 \leq n^{-1} c_1 \| t \|^2 \leq n^{-1} c_1 C_n^2$. Thanks to (30), we get there exists n_1 finite, such that $J_1 \subset \mathcal{J}_\theta$ for all $n \geq n_1$ and all $\theta \in K$.

For $\varepsilon \in (0, 1)$ and $n \geq n_1$, we obtain:

$$(2\pi)^d |I_n(\theta)| \leq \int_{\mathbb{R}^d} \left| \mathbf{1}_{\mathcal{J}_\theta}(t) f_\theta(t)^n - \mathrm{e}^{-\|t\|^2/2} \right| \mathrm{d}t \leq I_{n,1}(\theta) + I_{n,2}(\theta) + I_{n,3}(\theta) + I_{n,4}, \tag{32}$$

with

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

and $I_{n,4} = \int_{J_1^c} \mathrm{e}^{-\|t\|^2/2}\, \mathrm{d}t$ as well as $J_{2,\theta} = \{t \in \mathbb{R}^d;\ \| t \| > C_n \text{ and } n^{-1/2} \| \Sigma_\theta^{-1/2} t \| < \varepsilon\}$, $J_{3,\theta} = \{t \in \mathcal{J}_\theta;\ n^{-1/2} \| \Sigma_\theta^{-1/2} t \| \geq \varepsilon\}$. The proof of the Theorem will be complete as soon as we prove the converge of the terms $I_{n,i}$ to 0 for $i \in \{1, \ldots, 4\}$ 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 (30) 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 [23, 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 $\theta \in K$:

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

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 (26) 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 (26), we have

$$a_2 := \sup_{\theta\in K} \mathbb{E}_\theta[\| X_1 - m_\theta \|^2] < \infty \quad \text{and} \quad a_3 := \sup_{\theta\in K} \mathbb{E}_\theta[\| X_1 - m_\theta \|^3] < \infty. \tag{33}$$

Using c_1 defined in (31), we can choose ε small enough such that

$$\varepsilon^2 a_2 + \varepsilon a_3 c_1 < 1. \tag{34}$$

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

$$\mathbb{E}_\theta\left[\| Y \|^2\right] = \frac{1}{n}\mathbb{E}_\theta\Big[\langle X_1 - m_\theta, \Sigma_\theta^{-1}(X_1 - m_\theta)\rangle\Big]$$
$$= \frac{1}{n}\sum_{j=1}^{d}\sum_{\ell=1}^{d}\left[\Sigma_\theta^{-1}(j,\ell)\Sigma_\theta(\ell, j)\right] = \frac{d}{n}\cdot \tag{35}$$

Using similar computations, we obtain:

$$\mathbb{E}_\theta\left[\langle t, Y\rangle^2\right] = \frac{\| t \|^2}{n}\cdot \tag{36}$$

Recall notations a_3 in (33) and c_1 in (31). For $t \in J_{2,\theta}$, we get:

$$\mathbb{E}_\theta\left[|\langle t, Y\rangle|^3\right] \le n^{-3/2} \| \Sigma_\theta^{-1/2} t \|^3 \mathbb{E}_\theta[\| X_1 - m_\theta \|^3] \le \frac{\| t \|^2}{n}\varepsilon a_3 c_1 \le \frac{\| t \|^2}{n}, \tag{37}$$

where we used $n^{-1/2} \| \Sigma_\theta^{-1/2} t \| < \varepsilon$, (33) and (31) for the second inequality and (34) for the last. Recall a_2 given in (33). From (34) and since $t \in J_{2,\theta}$, we get:

$$\mathbb{E}_\theta\left[\langle t, Y\rangle^2\right] \le \| n^{-1/2}\Sigma_\theta^{-1/2} t \|^2 \mathbb{E}_\theta[\| X_1 - m_\theta \|^2] \le \varepsilon^2 a_2 < 1. \tag{38}$$

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

$$|f_\theta(t)| = |\mathbb{E}_\theta[\mathrm{e}^{i\langle t, Y\rangle}]| = \left|1 - \frac{\mathbb{E}_\theta[|\langle t, Y\rangle|^2]}{2} - i\mathbb{E}_\theta\left[\int_0^{\langle t,Y\rangle}\int_0^v\int_0^s \mathrm{e}^{iu}\,\mathrm{d}u\mathrm{d}s\mathrm{d}v\right]\right|$$
$$\le 1 - \frac{\mathbb{E}_\theta[|\langle t, Y\rangle|^2]}{2} + \mathbb{E}_\theta\left[\int_0^{|\langle t,Y\rangle|}\int_0^v\int_0^s \mathrm{d}u\mathrm{d}s\mathrm{d}v\right]$$
$$= 1 - \frac{\mathbb{E}_\theta[|\langle t, Y\rangle|^2]}{2} + \frac{\mathbb{E}_\theta[|\langle t, Y\rangle|^3]}{6}$$
$$\le 1 - \frac{\| t \|^2}{2n} + \frac{\| t \|^2}{6n} = 1 - \frac{\| t \|^2}{3n},$$

where we used that $\mathbb{E}_\theta[Y] = 0$ for the first equality, that $\mathbb{E}_\theta[\langle t, Y\rangle^2] \le 1$ for the first inequality (see (38)) and (36) as well as (37) for the last inequality. Therefore, we get that:

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

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:

$$|f_\theta(t)^n - \mathrm{e}^{-\|t\|^2/2}| \leq n|f_\theta(t) - \mathrm{e}^{-\|t\|^2/(2n)}| \leq n|h_\theta(n,t)| + ng(n,t), \tag{39}$$

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 $t \in J_1$:

$$ng(n,t) \leq \frac{\| t \|^4}{8n} \leq n^{-1} C_n^4. \tag{40}$$

Since $\mathbb{E}_\theta[Y] = 0$ and $\mathbb{E}_\theta\left[\langle t, Y\rangle^2\right] = \| t \|^2/n$, see (36), we deduce that:

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

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

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

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:

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

where we used c_1 defined in (31) for the inequality and (36) for the last equality. Hölder's inequality gives:

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

Using a_3 defined in (33), we get:

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

Using Tchebychev's inequality and a_2 defined in (33), we get:

$$\mathbb{P}_\theta(\| X_1 - m_\theta \| \ge L_n) \le \mathbb{E}_\theta\left[\| X_1 - m_\theta \|^2\right] L_n^{-2} \le a_2 L_n^{-2}.$$

This gives:

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

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

$$n|h_\theta(n,t)| \le \frac{1}{6} n^{-1/4} C_n^3 \sqrt{c_1} + n^{-1/6} c_1 C_n^2 a_3^{2/3} a_2^{1/3}.$$

Using (39) and (40), 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:

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

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

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

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 an 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 (22) and its conjugate ψ defined in (23). We state the following strong ratio theorem, which is of interest by itself. However, in this paper we used the extension of the strong ratio theorem given in Sect. 4.5.

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:*

$$\lim_{n\to\infty} \frac{\mathbb{P}(S_{n-m} = s_n - b)}{\mathbb{P}(S_n = s_n)} = 1, \tag{41}$$

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 [18]. We first state a preliminary lemma.

Lemma 4.9 *Let F be an aperiodic probability distribution on $\mathbb{Z}^d$. Let $(s_n, n \in \mathbb{N}^*)$ be elements of $\mathbb{Z}^d$ satisfying (a) and (b) of Theorem 4.7. Then, for all $b \in \mathbb{Z}^d$ and $m \in \mathbb{Z}$, we have* $\lim_{n\to\infty} \psi(\frac{s_n+b}{n+m}) = 0$.

Proof Assume that (a) and (b) of Theorem 4.7 hold. Let x be a limit of a converging subsequence of $(s_n/n, n \in \mathbb{N}^*)$. Since ψ is lower semicontinuous and nonnegative, we deduce from (b) that $\psi(x) = 0$. Thus, the possible limits of subsequences of $((s_n + b)/(n + m),\ n + m \geq 1)$, which are also the possible limits of subsequences of $(s_n/n,\ n \in \mathbb{N}^*)$, are zeros of ψ. Then, using the second part of Lemma 4.1 and the continuity of ψ on the interior of its domain, we deduce that $\lim_{n\to\infty} \psi(\frac{s_n+b}{n+m}) = 0$. □

Since F is aperiodic, using elementary arithmetic consideration and Lemma 4.9, we see it is enough to prove (41) 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:

$$\begin{aligned}\left|\frac{\mathbb{P}(S_{n-1} = a - b)}{\mathbb{P}(S_n = a)} - 1\right| = \left|\frac{\mathbb{E}\left[\frac{N_n}{n}; S_n = a\right]}{\mathfrak{p}\mathbb{P}(S_n = a)} - 1\right| &\leq \frac{\mathbb{E}[|\frac{N_n}{n} - \mathfrak{p}|; S_n = a]}{\mathfrak{p}\mathbb{P}(S_n = a)} \\ &\leq \frac{\varepsilon}{\mathfrak{p}} + \frac{R_n(a)}{\mathfrak{p}},\end{aligned} \tag{42}$$

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 [11], since N_n is binomial with parameter $(n, \mathfrak{p})$, we get:

$$\mathbb{P}\left(\left|\frac{N_n}{n} - \mathfrak{p}\right| > \varepsilon\right) \leq 2\,\mathrm{e}^{-2n\varepsilon^2}. \tag{43}$$

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

Lemma 4.10 *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 (43) and Lemma 4.10 with $1 - \eta = \mathrm{e}^{-\varepsilon^2}$, we get:

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

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

Proof of Lemma 4.10 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 (22) is $\mathbb{R}^d$. Recall notation (24) as well as $m_\theta = \mathbb{E}_\theta[X]$ and $\Sigma_\theta = \mathrm{Cov}_\theta(X, X)$. Let K_0 be a compact subset of $\mathcal{O}_F$. According to Lemma 4.4, there exists a compact set $K \subset \mathbb{R}^d$ 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

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

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

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

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

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

Since s_n/n belongs to $\{m_\theta;\ \theta \in K\}$, we get that $\inf_{\theta\in K} u_n(\theta, s_n) = 0$. Thanks to (26) 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 (23), we get:

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

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

$$\mathbb{P}(S_n = s_n)\,\mathrm{e}^{n\psi(s_n/n)} \geq \delta n^{-d/2} > 0. \tag{44}$$

This gives Lemma 4.10 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 $M > 0$ so that $\delta_M = \mathbb{P}(|X| > M) < 1$. Let X^M be distributed as X conditionally on $\{|X| \leq M\}$. 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 (22) with F replaced by F_M and ψ_M defined by (23) with φ replaced by φ_M. Since F is aperiodic, we get that F_M is aperiodic for M large enough. We get:

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

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 $\tilde{\varphi}_M(\theta) = \log\left(\mathbb{E}\left[\mathrm{e}^{\langle\theta,X\rangle}\, \mathbf{1}_{\{|X|\leq M\}}\right]\right)$ so that $\tilde{\psi}_M(x) = \psi_M(x) - \log(1-\delta_M)$.

Notice that the sequence of continuous finite convex functions $(\tilde{\varphi}_M, M \in \mathbb{N}^*)$ is non-decreasing and converges pointwise to the convex function φ (which is not identically $+\infty$ as $\varphi(0) = 0$). By definition, the sequence of convex functions $(\tilde{\psi}_M, M \in \mathbb{N}^*)$ is non-increasing and $\tilde{\psi}_M \geq \psi$. Therefore, the sequence converges to a function say $\tilde{\psi}$ such that $\tilde{\psi} \geq \psi$. Thanks to Theorem B.3.1.4 in [10] or Theorem II.10.8 of [21], $\tilde{\psi}$ 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 [10] gives that the closure of $\tilde{\psi}$ (defined in Definition B.1.2.4 in [10]) 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 (44) 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 completes the proof. □

4.5 An Extension of Theorem 4.7

We shall need the following extension of Theorem 4.7.

Lemma 4.11 *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:*

$$\lim_{n\to\infty} \frac{\mathbb{E}[G;\ H' + W_{n-\ell} = w_n - b]}{\mathbb{E}[G;\ H' + W_n = w_n]} = 1. \tag{45}$$

Proof Since F is aperiodic and by elementary arithmetic consideration, it is enough to prove (45) 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 (42), we get:

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

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:

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

where we used $G \le |H'|^c$ a.s. and that $H' + W_n = w_n$ implies $H' \le w_n$ for the first inequality, and inequality (43) for the second. Notice that for all $\varepsilon' > 0$ we have $|w_n|^c \le \exp(\varepsilon' n)$ for n large enough.

Then use Lemma 4.10 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. □

Acknowledgements The authors would like to thank Jean-Philippe Chancelier for pointing out the references on convex analysis and his valuable advice as well as the two anonymous referees for their comments and suggestions. 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. Abraham, R., Delmas, J.-F.: Local limits of conditioned Galton–Watson trees: the condensation case. Electron. J. Probab. **19**(56), 1–29 (2014)
2. Abraham, R., Delmas, J.-F.: Local limits of conditioned Galton–Watson trees: the infinite spine case. Electron. J. Probab. **19**(2), 1–19 (2014)
3. Athreya, K.B., Ney, P.E.: Branching Processes. Springer, Berlin (1972)
4. Auslender, A., Teboulle, M.: Asymptotic Cones and Functions in Optimization and Variational Inequalities. Springer, Berlin (2006)
5. Chaumont, L., Liu, R.: Coding multitype forests: application to the law of the total population of branching forests. Trans. Am. Math. Soc. **368**, 2723–2747 (2016)
6. Delmas, J.-F., Hénard, O.: A Williams decomposition for spatially dependent superprocesses. Electron. J. Probab. **18**(37), 1–43 (2013)
7. Gnedenko, B.V.: On a local limit theorem of the theory of probability. Uspekhi Mat. Nauk **3**(3), 187–194 (1948)
8. Gnedenko, B.V., Kolmogorov, A.N.: Limit Distributions for Sums of Independent Random Variables. English translation, Addison-Wesley, Cambridge (1954)
9. He, X.: Conditioning Galton–Watson trees on large maximal out-degree. J. Theor. Probab. (2016). doi:10.1007/s10959-016-0664-x
10. Hiriart-Urruty, J.-B., Lemaréchal, C.: Fundamentals of Convex Analysis. Springer, Berlin (2001)

11. Hoeffding, W.: Probability inequalities for sums of bounded random variables. J. Am. Stat. Assoc. **58**(301), 13–30 (1963)
12. Janson, S.: Simply generated trees, conditioned Galton–Watson trees, random allocations and condensation. Probab. Surv. **9**, 103–252 (2012)
13. Jonnson, T., Stefansson, S.: Condensation in nongeneric trees. J. Stat. Phys. **142**, 277–313 (2011)
14. Kesten, H.: Subdiffusive behavior of random walk on a random cluster. Ann. de l'Inst. Henri Poincaré **22**, 425–487 (1986)
15. Kurtz, T., Lyons, R., Pemantle, R., Peres, Y.: 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., pp. 181–185. Springer (1997)
16. Luis, J.A.L.-M., Gorostiza, G.: The multitype measure branching process. Adv. Appl. Probab. **22**(1), 49–67 (1990)
17. Miermont, G.: Invariance principles for spatial multitype Galton–Watson trees. Ann. Inst. H. Poincaré Probab. Statist **44**, 1128–1161 (2007)
18. Neveu, J.: Sur le théorème ergodique de Chung-Erdős. C. R. Acad. Sci. Paris **257**, 2953–2955 (1963)
19. Pénisson, S.: Beyond Q-process: various ways of conditioning the multitype Galton–Watson process. ALEA **13**, 223–237 (2016)
20. Rizzolo, D.: 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)
21. Rockafellar, R.T.: Convex Analysis. Princeton Landmarks in Mathematics. Princeton University Press, Princeton (1997)
22. Rvaceva, E.: On domains of attraction of multi-dimensional distributions. Sel. Transl. Math. Stat. Probab. **2**, 183–205 (1961)
23. Spitzer, F.: Principles of Random Walk. Springer, Berlin (2013)
24. Stephenson, R.: Local convergence of large critical multi-type Galton-Watson trees and applications to random maps. J. Theor. Probab. (2016). doi:10.1007/s10959-016-0707-3
25. Stone, C.: On local and ratio limit theorems. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 2, no. (part II), pp. 217–224. University of California Press, Berkeley, Los Angeles (1966)