Abstract. Given $d \geq 1$, let $(A_i)_{i\geq 1}$ be a sequence of random $d \times d$ real matrices and $Q$ be a random vector in $\mathbb{R}^d$. We consider fixed points of multivariate smoothing transforms, i.e., random variables $X \in \mathbb{R}^d$ satisfying

$$X \text{ has the same law as } \sum_{i \geq 1} A_i X_i + Q,$$

where $(X_i)_{i \geq 1}$ are i.i.d. copies of $X$ and independent of $(Q, (A_i)_{i \geq 1})$. By means of contraction arguments, the existence of fixed points, which attract point masses, can be shown. Let $X$ be such a fixed point. Assuming that the action of the matrices is expanding as well with positive probability, it was shown in a number of papers that there is $\beta > 0$ with

$$\lim_{t \to \infty} t\beta P(\langle u, X \rangle > t) = K \cdot f(u),$$

where $u$ denotes an arbitrary element of the unit sphere and $f$ a positive function and $K \geq 0$. However in many cases it was not established that $K$ is indeed positive.

In this paper, under quite general assumptions, we prove that

$$\liminf_{t \to \infty} t\beta P(\langle u, X \rangle > t) > 0,$$

completing, in particular, the results of Mirek (2013) and Buraczewski et al. (2013).

1. Introduction

1.1. The (multivariate) smoothing transform. Let $d \geq 1$. Let $(Q, (A_i)_{i \geq 1})$ be a random element of $\mathbb{R}^d \times M(d \times d, \mathbb{R})^N$, that is $Q$ is a random vector and $(A_i)_{i \geq 1}$ is a sequence of random matrices. We assume that the random number $N := \max\{i : A_i \neq 0\}$ is finite a.s. If $X \in \mathbb{R}^d$ is a random variable such that

$$(1.1) \quad X \text{ has the same law as } \sum_{i=1}^N A_i X_i + Q,$$

where $(X_i)_{i \geq 1}$ are i.i.d. copies of $X$ and independent of $(Q, (A_i)_{i \geq 1})$, then we call the law $\mathbb{L}(X)$ of $X$ a fixed point of the (multivariate, if $d > 1$) smoothing transform. By a slight abuse of notation, we also call $X$ itself a fixed point.

Eq. (1.1) has drawn a lot of attention for decades. In the univariate case this equation occurs in various areas, e.g., the analysis of recursive algorithms (Rosler (1991, 2001), Neininger and Rüschendorf (2004)), branching particle systems (Durrett and Liggett (1983)), Googles PageRank algorithm (Jelenković and Olvera-Cravioto (2012a,b)). Also the multivariate situation draws a lot of attention. A classical example where (1.1) appears is the joint distribution of key comparisons and key exchanges for Quicksort (Neininger and Rüschendorf (2004)). However in this case the action of the matrices is purely contracting, and therefore all fixed points have exponential moments, which
is not in the scope of the present paper. Some recent examples are related to kinetic models, see Bassetti and Matthes (2014). Then solutions to (1.1) describe e.g. equilibrium distribution of the particle velocity in Maxwell gas.

The aim of this paper is to describe the tail behavior of fixed points, i.e. the decay rate of
\[ P(|X| > t) \quad \text{or} \quad P(\langle u, X \rangle > t), \]
as \( t \) goes to infinity, where \( u \) denotes an arbitrary element of the unit sphere \( S \).

1.2. Univariate smoothing transform. In dimension \( d = 1 \), complete results about the structure of fixed points are available under very weak assumptions, see Durrett and Liggett (1983); Liu (1998); Biggins and Kyprianou (1997); Alsmeyer et al. (2012); Alsmeyer and Meiners (2013); Buraczewski and Kolesko (2014) for the case of \( A_i \geq 0 \), and Iksanov and Meiners (2015) for the most general case of \( A_i \in \mathbb{R} \). It turns out, that the characterization depends on the function
\[ m(s) := \mathbb{E} \sum_{i=1}^{N} |A_i|^s, \]
which is log-convex, and in particular on the value \( \alpha = \inf \{ s > 0 : m(s) = 1 \} \). It is shown that there are two classes of fixed points: Fixed points are either mixtures of \( \alpha \)-stable laws and attract (only) laws with \( \alpha \)-regular varying tails, or have a finite moment of order \( \alpha + \varepsilon \) for some \( \varepsilon > 0 \) (subject to the assumption \( \mathbb{E} |Q|^{\alpha + \varepsilon} < \infty \)) and attract point masses. Since tail behavior of the first class is well understood, we focus here on the second class, which we call attracting fixed points.

Under various assumptions on \( (Q, (A_i)_{i \geq 1}) \) and \( N \), it has been shown in Guivarc’h (1990); Liu (2001); Jelenković and Olvera-Cravioto (2012b,a) that if, roughly speaking, there is \( \beta > \alpha \) with \( m(\beta) = 1 \) and \( \mathbb{E} |Q|^\beta < \infty \) (including the case \( Q \equiv 0 \)), then
\[ \lim_{t \to \infty} t^\beta P(|X| > t) = K \geq 0. \]

There is also a rich literature concerning the case when \( \alpha \) is the unique point such that \( m(\alpha) = 1 \), that is when \( m'(\alpha) = 0 \) (see Durrett and Liggett (1983); Liu (1998); Biggins and Kyprianou (2005); Buraczewski (2009); Buraczewski and Kolesko (2014)). It is obviously a very important question, whether \( K \) is indeed positive, since otherwise, \( t^\beta \) might not be the precise rate. For \( A_i \geq 0 \) and \( Q = 0 \), positivity of \( K \) is proved in Guivarc’h (1990); Liu (2001), but it remained - except for some special cases - an open question in Jelenković and Olvera-Cravioto (2012b,a), where the cases \( Q \neq 0 \) resp. \( A_i \in \mathbb{R} \) were considered. This question was answered in Alsmeyer et al. (2013) (see also Buraczewski et al. (2014)).

1.3. Main results. The contribution of this paper is to prove the positivity of \( K \) using a very general argument, that is in particular valid for the multidimensional situation \( d > 1 \). There, an analogue of the function \( m \) can be defined (details given below), and in the case where the \( A_i \), \( Q \) and \( X \) all have nonnegative entries, it has been shown in Mentemeier (2015), that again fixed points are either mixtures of multivariate \( \alpha \)-stable laws (with \( \alpha \) defined as before), or have a finite moment of order \( \alpha + \varepsilon \), if the same holds for \( Q \).

Tail behavior of such attracting fixed points has been analyzed for the case \( \alpha = 1 \) and \( Q = 0 \) in Buraczewski et al. (2014), where it has been shown that
\[ \lim_{t \to \infty} t^\beta P(\langle u, X \rangle > t) = Kr(u) \]
with a positive continuous function \( r \) on \( S_\geq := S \cap [0, \infty)^d \) for \( \beta \) being the unique value such that \( \beta > \alpha \) and \( m(\beta) = 1 \). In this case, also positivity of \( K \) has been proved.
The inhomogeneous case \( Q \neq 0 \), with \( \alpha \leq 1/2 \), has been studied in [Mirek 2013] and the existence of the limit in Eq. (1.4) is proved there, but it remained an open question, whether \( K \) is positive (at least for \( \beta < 1 \)).

The case of invertible matrices \( (A_i)_{i \geq 1} \) was studied in [Bassetti and Matthes 2014] and [Buraczewski et al. 2013], with tail behavior being studied mainly in the latter paper. There once more existence of the limit in Eq. (1.4) was proved, but not the positivity of \( K \). With this work, we will solve all these open multidimensional questions and extend known one dimensional results. Let us also underline, that up to our knowledge, there are no results in the literature concerning the multivariate situation with random \( N \). With this respect our main results are new.

In the next section we introduce the three classes of matrices which are considered in this paper and further notation relevant for the multivariate case. The main results of this paper are formulated in Section 3. The remaining sections are devoted to the proof, which mainly builds upon results obtained in [Buraczewski and Mentemeier 2014].

2. Notations

In this section, we describe, in an abbreviated form, but similar to [Buraczewski and Mentemeier 2014] three sets of assumptions for random matrices, namely condition (C) for nonnegative matrices and conditions (i-p) and (id) for invertible matrices. Each set of assumptions guarantees precise large deviation estimates extending the Furstenberg-Kesten-theorem ([Furstenberg and Kesten 1960]), i.e. the SLN for the norm of products of random matrices. These large deviation estimates will play a prominent role in our proof below. Let \( d \geq 1 \). Given a probability law \( \mu \) on the set of \( d \times d \)-matrices \( M(d \times d, \mathbb{R}) \), let \( (M_n)_{n \in \mathbb{N}} \) be a sequence of i.i.d. random matrices with law \( \mu \). Equip \( \mathbb{R}^d \) with any norm \( \| \cdot \| \), write \( \|m\| := \sup_{x \in S} |mx| \) for the operator norm of a matrix \( M \) and denote the unit sphere in \( \mathbb{R}^d \) by \( S \). We write

\[
    m \cdot x := \frac{mx}{|mx|}, \quad x \in S
\]

for the action of a matrix \( m \) on \( S \) (as soon as this is well defined). If \( S \) is invariant under the action of \( M \), we introduce a Markov random walk \( (U_n, S_n)_{n \in \mathbb{N}} \) on \( S \times \mathbb{R} \) by

\[
    U_n := M_n \cdots M_1 \cdot U_0, \quad S_n := \log |M_n \cdots M_1 U_0| = \log |M_n U_{n-1}| + S_{n-1},
\]

for some initial data \( U_0 \in S \), the value of which we note by the convention \( \mathbb{P}_u (U_0 = u) = 1 \), \( u \in S \).

Below, the following concepts will appear several times: Write \( \Gamma := [\text{supp } \mu] \) for the semigroup of matrices, generated by the support of \( \mu \). A matrix \( m \) with an algebraic simple dominant eigenvalue \( \lambda_m \), that exceeds all other eigenvalues in absolute value, will be called proximal, and we will denote by \( v^+ \) the corresponding normalized eigenvectors \( (v^+_m, -v^-_m) \), using the convention that \( \min \{i : (v^+_m)_i > 0\} < \min \{i : (v^-_m)_i > 0\} \). Note that a matrix with all entries positive is proximal by the Perron-Frobenius theorem, and that \( v^+ \) is the Perron-Frobenius eigenvector.

2.1. Invertible Matrices: Condition (i-p).

The condition (i-p) (irreducible and proximal), described below, is due to Guivarc’h, Le Page and Raugi and was studied in detail in several articles by these authors, the most comprehensive one of which is [Guivarc’h and Le Page 2012].

Let now \( \mu \) be a probability measure on the group \( GL(d, \mathbb{R}) \) of invertible \( d \times d \) matrices. Then the measure \( \mu \) is said to satisfy condition (i-p), if

1. There is no finite union \( \mathcal{W} = \bigcup_{i=1}^n W_i \) of subspaces \( 0 \neq W_i \subset \mathbb{R}^d \) which is \( \Gamma \)-invariant, i.e. \( \Gamma \mathcal{W} = \mathcal{W} \), (strong irreducibility)

2. \( \Gamma \) contains a proximal matrix. (proximality)
It may happen that there is a $\Gamma$-invariant proper closed convex cone $C$. This situation is very similar to the case of nonnegative matrices, see [Buraczewski et al. 2014]. Therefore, we will exclude it and only consider matrices satisfying $(i-p)$, and there is no $\Gamma$-invariant proper closed convex cone.

In this case, it can be shown that the Markov chain $(U_n)$ has a unique invariant probability measure, which is supported on

$$V(\Gamma) := \{v_m^\pm \in S : m \in \Gamma \text{ is proximal}\},$$

and due to the strong irreducibility, the orthogonal space of $V(\Gamma)$ is $\{0\}$. Finally, write

$$\iota(m) := \inf_{x \in S} |mx| = \|m^{-1}\|^{-1}.$$

### 2.2. Nonnegative Matrices: Condition $(C)$

Next, we introduce a condition on nonnegative matrices, i.e. all entries greater or equal to zero, which do not need to be invertible. We will use similar notation as for condition $(i-p)$, in order to highlight connections. Note, that these assumptions can be formulated more generally for matrices leaving invariant a proper closed convex cone, see [Buraczewski et al. 2014].

Denote the cone of vectors with nonnegative entries by $\mathbb{R}^d_{\geq}$ and write $S_\geq = \{x \in \mathbb{R}^d_{\geq} : |x| = 1\}$ for its intersection with unit sphere. It is invariant under the action of allowable matrices, i.e. matrices having nonnegative entries and no zero row nor column. For an allowable matrix, the quantity

$$\iota(m) := \min_{x \in S_\geq} |mx|$$

is strictly positive and is the right substitute for $\iota$ as defined for invertible matrices.

We say that a probability measure $\mu$ on nonnegative matrices satisfies condition $(C)$, if:

1. Every $m \in \text{supp} \mu$ is allowable.
2. $[\text{supp} \mu]$ contains a matrix all entries of which are strictly positive (a positive matrix).

Once again, this guarantees the existence of a unique invariant probability measure for $(U_n)$ on $S_\geq$, which is supported in

$$V(\Gamma) := \{v_m^\pm : m \in \Gamma \text{ is a positive matrix}\}.$$

Note that for nonnegative matrices, being positive is a stronger assumption than proximality, for it also asserts irreducibility: A diagonal matrix might be allowable and proximal as well, but in contrast to a positive matrix, its dominant eigenvector is not attractive on the whole set $S_\geq$. This is why no assumption on invariant subspaces is needed here. Instead, we have to impose an additional non-lattice condition for $(S_n)$, which is automatically satisfied under $(i-p)$: Define

$$S(\Gamma) := \{\log \lambda_m : m \in \Gamma \cap \text{int} (\mathcal{M}_+ )\}.$$

Then we say that $\mu$ is non-arithmetic, if the (additive) subgroup of $\mathbb{R}$ generated by $S(\Gamma)$ is dense.

### 2.3. Invertible Matrices: Condition $(id)$

The third set of assumptions, called $(id)$ for irreducible and density, appears first at the end of [Kesten 1973] and was elaborated in [Alsmeyer and Mentemeier 2012]. In fact, it can be shown to imply condition $(i-p,o)$. Due to the stronger assumption that $\mu$ is absolutely continuous, it often allows for simpler proofs, this is why we include it as an extra set of assumptions.

A probability measure $\mu$ on $GL(d,\mathbb{R})$ is said to satisfy condition $(id)$ if

1. for all open $B \subset S$ and all $x \in S$, there is $n \in \mathbb{N}$ such that $P(\Pi_n \cdot x \in B) > 0$, and
(2) there are a matrix $m_0 \in GL(d, \mathbb{R})$, $\delta, c > 0$ and $n_0 \in \mathbb{N}$ such that
\[
\mathbb{P}(\Pi_{n_0} \in dm) \geq c 1_{B_{\delta}(m_0)}(m) l(dm),
\]
where $l$ denotes the Lebesgue measure on $\mathbb{R}^{d^2} \simeq M(d \times d, \mathbb{R})$.

The classical example is $\mu$ having a density about the identity matrix.

It is shown in (Alsmeyer and Mentemeier, 2012, Lemma 5.5) that $U_n$ is a Doeblin chain under condition (id). The support of its stationary probability measure is $S$ by (Alsmeyer and Mentemeier, 2012, Proposition 4.3), therefore in the case of (id) we have $V(\Gamma) = S$.

2.4. Markov random walk and change of measure. Below, we identify $S = S_{\geq}$ in the case of nonnegative matrices and $S = S$ in the case of (i-p)- or (id)-matrices. Given a measure $\mu$ on matrices as before, set
\[
I_\mu := \{ s \geq 0 : \mathbb{E}[\|M\|^s] < \infty \}.
\]
Then, for $s \in I_\mu$, we define operators in the set $\mathcal{C}(S)$ of continuous functions on $S$ by
\[
P_s f(x) := \mathbb{E}[\|M\|^s f(M \cdot x)]
\]
(2.1)

It was proved in Kesten (1973); Buraczewski et al. (2014) for nonnegative matrices, in Guivarc’h and Le Page (2012) for invertible matrices under condition (i-p,o) and in Mentemeier (2013) under condition (id), that the spectral radii of these operators are given by the log-convex and differentiable function
\[
k(s) := \lim_{n \to \infty} (\mathbb{E}[\|M_n \ldots M_1\|^s])^{\frac{1}{s}}
\]
and that for each $s \in I_\mu$ there are an unique normalized function $r_s \in \mathcal{C}(S)$ and an unique probability measure $\nu_s \in \mathcal{P}(S)$ satisfying
\[
P_s r_s = k(s) r_s \quad \text{and} \quad P_s \nu_s = k(s) \nu_s
\]
(2.3)

Moreover, the function $r_s$ is strictly positive and $\bar{s} := \min\{s,1\}$-Hölder continuous. The supports of the measure is given by supp $\nu_s = V(\Gamma)$. Equation (2.3) yields that if $k(\gamma) = 1$, then $h(u,t) := e^{\gamma t} r_s(u)$ is an harmonic function for the Markov chain $(U_n, S_n)$. Using the idea of Doob’s $h$-transform, one can introduce new probability measures $P_{\gamma}^\infty$, and it turns out that under $P_{\gamma}^\infty$, $S_n$ has drift $k'(\gamma)$, i.e.
\[
\lim_{n \to \infty} \frac{S_n}{n} = \frac{k'(\gamma)}{k(\gamma)} \quad P_{\gamma}^\infty \text{-a.s.}
\]
(2.4)

This idea can be extended (see (Buraczewski and Mentemeier, 2014, Section 2) for details) to yield exponential shifted probability measures $P_{\gamma}^\infty$ for all $\gamma \in I_\mu$, such that the property (2.4) holds.

3. Statement of Results

We give separately the results for the one-dimensional case $d = 1$ and the multidimensional case.

We will make the following case distinction concerning the number $N$:

(N-random) $N \in \mathbb{N}$ is random with $1 < EN < \infty$, and conditioned upon $N$, $(A_i)_{i \geq 1}$ are i.i.d. with law $\mu$, and the variables $Q$ and $(N, (A_i)_{i \geq 1})$ are independent.

(N-fixed) $N \geq 2$ is fixed, $(A_1, \ldots, A_N, Q)$ having any dependence structure

The case (N-fixed), without any loss of generality, can be reduced to the situation where all the random variables $A_1, \ldots, A_N$ are identically distributed (see Buraczewski et al. (2013) for more details). Supposing identical distribution, we set
\[
\mu := \mathcal{L}(A_1^* \in \cdot),
\]
i.e. \( \mu \) is the law of the transpose of \( A_1 \). As before, let \( M, M_1, M_2, \ldots \) i.i.d. random variables, having law \( \mu \). Then, in particular, \( M \overset{d}{=} A_i^\ast, 1 \leq i \leq N \). Then the general, multivariate version of the function \( m(s) \) (see Eq. (1.2)) is given by

\[
m(s) := (\mathbb{E}N)k(s),
\]

with \( k(s) \) as defined in (2.2).

### 3.1. The univariate case: \( d = 1 \).

**Theorem 3.1.** Assume that either (\( N \)-random) or (\( N \)-fixed) is satisfied. Assume moreover that

1. there are \( 0 < \alpha < \beta \) and \( \varepsilon > 0 \) such that \( m(\alpha) = m(\beta) = 1 \), \( \mathbb{E}|Q|^{\beta + \varepsilon} < \infty \),
2. the law of \( \log M \) is nonarithmetic in the classical sense;
3. there is a nondegenerate random variable \( X \) satisfying (1.1) with \( \mathbb{E}|X|^s < \infty \) for all \( s < \beta \).

Then for this \( X \),

\[
\liminf_{t \to \infty} t^{\beta} \mathbb{P}(|X| > t) > 0.
\]

It was proved in Jelenković and Olvera-Cravioto (2012a,b), that the limit in (3.1) actually exists (but may be zero). Positivity of the limiting constant was established in Alsmeyer et al. (2013) by interpreting the limit as the residue of a meromorphic function; and later on, using completely different arguments, namely large deviation estimates in Buraczewski et al. (2014) (only for constant \( N \) and i.i.d. \( (A_i) \)) resp. comparison with a maximum recursion and symmetrization inequalities in Jelenkovic and Olvera-Cravioto (2014). Mainly as a by-product of the multidimensional case, we extend here the method introduced in Buraczewski et al. (2014) to the cases of random \( N \) or \( (A_i) \) not being necessary i.i.d. in the following way.

### 3.1.1. Multivariate case: \( d > 1 \). Here is our main result in the multidimensional situation:

**Theorem 3.2.** Assumptions:

1. Let either (\( N \)-random) be satisfied, or assume that (\( N \)-fixed) holds and \( \mu \) has a bounded support,
2. Geometrical assumptions: Assume one of the following:
   (Ga) \( A_i \) and \( Q \) are nonnegative, \( \mu \) satisfies (C) and is nonarithmetic, or
   (Gb) \( A_i \) are invertible and satisfy (i-p,o) or (id).
3. Moment assumptions: Assume all of the following
   (M1) There are \( 0 < \alpha < \beta \) and \( \varepsilon > 0 \) such that \( m(\alpha) = m(\beta) = 1 \), \( \mathbb{E}|Q|^{\beta + \varepsilon} < \infty \), \( \mathbb{E}\|A_i^\ast\|^{\beta + \varepsilon} (A_i^\ast)^{-\varepsilon} < \infty \),
   (M2) there is a nondegenerate random variable \( X \) satisfying (1.1) with \( \mathbb{E}|X|^s < \infty \) for all \( s < \beta \).

Then for this \( X \) and all \( u \in \mathcal{S} \),

\[
\liminf_{t \to \infty} t^{\beta} \mathbb{P}(\langle u, X \rangle > t) > 0.
\]

As a corollary of this results we obtain that the asymptotic behavior proved in (Buraczewski et al. 2013, Theorems 2.7, 2.9 and 2.11) and (Mirek 2013, Theorem 1.9) is exact. Tail estimates for the case of random \( N \) in the multivariate situation has not yet been considered in the literature and this is the first result in that direction.
3.2. Structure of the paper. We proceed in Section 4 by introducing the weighted branching process, which allows for the study of the fixed point equation (1.1) by iteration and for the construction of random variables, which satisfy the equation a.s. (in contrast to in law). Using that the support of these random variables is unbounded, we can estimate \( \mathbb{P}((X > t)) \) from below by a union of events of the type “one large term occurs”, this is made precise in Section 5, with the fundamental estimate being proved in Lemma 5.2. Section 6 is mainly combinatorial, there we count the number events occurring in the union, and estimate from above the probability of intersections, which we make small by an appropriate choice of parameters and thereby complete the proof of the main theorem in Section 7. An outline of the proof is given in Subsection 4.3.

Remark 3.3. We have tried hard, but were not able to avoid the case distinctions concerning \( N \). A natural way to do this would be the use of a spinal-tree-identity (many-to-one lemma), but it seems that our approach is not compatible with this technique. The main difficulty is that we consider sums over particular subtrees (as defined in (6.1)), which we were not able to reformulate in such a way that a many-to-one-lemma would be applicable.

4. Weighted branching process

In this section, we introduce the weighted branching process, i.e. a sequence of random variables which satisfy Eq. (1.1) almost surely.

4.1. Trees. Let \( \mathbb{N} = \{1, 2, \ldots\} \) be the set of positive integers and let

\[ \mathbb{U} = \bigcup_{k=0}^{\infty} \mathbb{N}^k \]

be the set of all finite sequences \( i = i_1 \ldots i_n \). By \( \emptyset \) we denote the empty sequence. For \( i = i_1 \ldots i_n \) we denote by \( |i| \) its length and by \( i|_k = i_1 \ldots i_k \) the curtailment of \( i \) up to first \( k \) terms. Given \( i \in \mathbb{U} \) and \( j \in \mathbb{N} \) we define \( ij = i_1 \ldots i_n j \) the sequence obtained by juxtaposition. In the same way we define \( ij \) for \( i, j \in \mathbb{U} \).

We introduce a partial ordering on \( \mathbb{U} \), writing \( i \leq j \) when there exists \( i_1 \in \mathbb{U} \) such that \( j = ii_1 \). If \( i, i' \in \mathbb{U} \), we write \( j = i \land i' \) for the maximal common sequence of \( i \) and \( i' \), that is, \( j \) is the longest sequence such that \( j \leq i \) and \( j \leq i' \).

We say that a subset \( T \) of \( \mathbb{U} \) is a tree if

- \( \emptyset \in T \);
- if \( i \in T \), then \( i|_k \in T \) for any \( k < |i| \);
- if \( i \in T \) and \( j \in \mathbb{N}_+ \), then \( ij \in T \) if and only if \( 1 \leq j \leq N_i \), for some integer \( N_i \geq 0 \).

Then \( \emptyset \) is the root of the tree.

In case (N-random), let \( (N_i)_{i \in \mathbb{U}} \) be a family of i.i.d. copies of \( N \), which thus determines the shape of the tree \( T \). By \( \mathcal{F}_T \) we will denote below the \( \sigma \)-algebra generated by \( (N_i)_{i \in \mathbb{U}} \). In case (N-fixed), the shape of the tree is deterministic, then \( T = \bigcup_{k=0}^{\infty} \{1, \ldots, N\}^k \).

4.2. Random variables indexed by \( \mathbb{U} \). A simple argument (see Buraczewski et al. 2013, Proposition A.1) justifies that without loss of generality, under the assumption (N-fixed) we may assume that all the random variables are identically distributed. The same argument can be used to supply the even stronger property, that for any (vector valued) function \( f \) on \( M(d \times d, \mathbb{R})^N \) and any permutation \( \sigma \) of \( \{1, \ldots, N\} \) we can assume

\[ f(A_1, \ldots, A_N) \overset{d}{=} f(A_{\sigma(1)}, \ldots, A_{\sigma(N)}). \]

From now we will assume that (4.1) is satisfied.
To each node $j \in U$ we attach an independent copy $A_j := (Q_j, (A_{ji})_{i \geq 1})$ of $A := (Q, (A_i)_{i \geq 1})$ and, given a random variable $X \in \mathbb{R}^d$, satisfying \cite{11}, an independent copy $X_j$ of $X$ as well. For simplifying notation, let $(Q_{\emptyset}, (A_{\emptyset i})_{i \geq 1}) = (Q, (A_i)_{i \geq 1})$. We refer to $A_j$ as the weight pertaining to the edge connecting $j$. Denote the total weight on the unique path connecting the edge $j$ with the edge $ji$ by

$$\Pi_{j,ji} := A_{ji}, A_{ji, i_2} \cdots A_{ji}$$

and define the empty product to be the $d \times d$ identity matrix. Due to the assumption $N < \infty$ $\mathbb{P}$-a.s., each generation of $\mathbb{T}$ has a.s. a finite size. Notice also that in view of \cite{11} the law of $\Pi_{j,ji}$ depends only on the numbers of factors and coincides with the law of $\Pi_i$.

Recall that we defined $\mu$ to be the law of $A_1$, and $M_1, M_2, \ldots$ to be a sequence of i.i.d. random variables with law $\mu$. Then $\Pi_n := M_n \cdots M_1$ has the same law as $\Pi^*_i$ for every $i \in \mathbb{T}$ with $|i| = n$ and moreover,

$$\mathbb{P}(\Pi^*_i \cdot u \in \cdot) = \mathbb{P}_u (U_n \in \cdot), \quad \mathbb{P}(\log |\Pi^*_i u| \in \cdot) = \mathbb{P}_u (S_n \in \cdot)$$

We write $[T]_j := \{i \in U : ji \in T\}$ for the subtree of $T$ rooted at $j$, and define in general the shift operator acting on functions of the family $(A_i, X_i)_{i \in U}$ by

$$[F((A_i, X_i)_{i \in U})]_j := F((A_{ji}, X_{ji})_{i \in U}).$$

With this notation, $\Pi_{j,ji} = [\Pi]_j$. The random variables

$$Y_l := \sum_{|i| < l} \Pi_i Q_i + \sum_{|i| = l} \Pi_i X_i, \quad l \geq 1,$$

(4.2)

$$Y_0 := X_0,$$

are called \emph{weighted branching process associated with $(Q, (A_i)_{i \geq 1})$ and $X$}. They satisfy

$$Y_l = \sum_{i=1}^N A_i [Y_{l-1}]_i + Q,$$

where $[Y_{l-1}]_i$ are i.i.d., with the same law as $Y_{l-1}$. Since $X_i$ are solutions to \cite{11}, then in particular, $Y_l \overset{d}{=} X$ for all $l \in \mathbb{N}$.

We define moreover

$$Z_{l,ik} := \sum_{j \neq k, j \leq N_i} A_{ij} [Y_{l-[i]-1}]_j + Q_i, \quad l > |i|,$$

(4.3)

Then for $l = |i| + 1$ we have

$$Z_{l,ik} := \sum_{j \neq k, j \leq N_i} A_{ij} X_{ij} + Q_i.$$

(4.4)

By \cite{11} the random variables $(Z_{l,ik})_{1 \leq k \leq N_i}$ are obviously identically distributed. To simplify our notation we define

$$Z_{i|i} := Z_{i|i, i}$$

Then for every $l \in \mathbb{N}$ and $i \in \mathbb{T}$ with $|i| \leq l$, we can rearrange the sum in Eq. (4.2) to obtain the a.s. identity

$$Y_l = \Pi_i [Y_{l-[i]}]_i + \sum_{k \leq |i|} \Pi_{i|k-1} Z_{l,i|k},$$

(4.5)
Observe that this implies for \( |i| \leq l \) the following identity in law.

\[
Y_l \overset{\text{d}}{=} \Pi_i X_i + \sum_{k \leq |i|} \Pi_{|i|-1} Z_{i,k}.
\]

4.3. **Outline of the proof.** This identity may give a first idea, how we are going to proceed in the proof of the main theorems: We consider sets where \( \Pi_i X_i \) is large, while the remaining sum is small. Therefore, we in turn study sets where \( \|\Pi_i\| \) is large, but smaller products are comparably small (with the comparison governed by a parameter \( C_0 \)). The probability of such sets will be estimated using large deviation results for products of random matrices, obtained in [Buraczewski and Mentemeier 2014]. Then the probability that \( X \) is large will be estimated from below by the union of sets as described above, over different \( i \). It will be convenient to not take the union over the whole tree, but rather from a sparse subtree, in order to make the events sufficiently disjoint. The relative size of the subtree will be given by a parameter \( C_1 \), which will be a free parameter of the proof.

A particular problem in the multivariate situation is to compare \( \Pi_i X_i \) with \( \|\Pi_i\| \). We deal with this question at the beginning of the next section, the better part of which is devoted to formulate precisely the heuristics we described above.

5. **First estimates**

We start this section by a lemma stating that \( X \) has unbounded support in “all” directions of \( \mathbb{R}^d \) resp. \( \mathbb{R}_{\geq} \), which we will make use of subsequently in Lemma 5.2, which gives the fundamental comparison between \( P(|X| > t) \) and the union of large deviation events.

**Lemma 5.1.** Assume that hypotheses of Theorem 3.2 are satisfied and that \( X \) is not a.s. constant. Then for all \( D > 0 \) there is \( J < \infty, \varepsilon_j > 0, \kappa > 0, 1 \leq j \leq J \), such that there are disjoint subsets \( \Omega_j \) of \( \mathbb{R}^d \) with

\[
P \left( X \in \Omega_j \text{ and } |X| > \frac{D}{\varepsilon_j} \right) \geq \kappa
\]

and moreover

\[
\mathbb{R}^d \subset \bigcup_{j=1}^J \Omega_j^*,
\]

where \( \Omega_j^* \) are the cones

\[
\Omega_j^* := \{ z \in \mathbb{R}^d : \langle z, x \rangle \geq \varepsilon_j |z| |x| \text{ for all } x \in \Omega_j \}.
\]

If \( \mu \) satisfies (C), then the same statement is valid, but with \( \Omega_j \) being subsets of \( \mathbb{R}_{\geq}^d \) and (5.1) replaced by

\[
\mathbb{R}_{\geq} \subset \bigcup_{j=1}^J \Omega_j^*.
\]

**Proof.** Let \( X_1, \ldots, X_N \) be i.i.d. copies of \( X \) (with \( N \) constant or random). Set \( B := \sum_{i=2}^N A_i X_i + Q \). Since \( X_i \) are i.i.d., nontrivial and independent of \( (A_i)_{i \geq 1} \) and \( Q \), it follows that also \( B \) must be nontrivial. Moreover, due to the moment assumptions (M1) and (M2), \( B \) has a finite moment of order \( \beta \). Then \( X \) satisfies the equation \( X \overset{\text{d}}{=} A_1 X_1 + B \), and for \( X \) satisfying such an equation, the results are shown in [Buraczewski and Mentemeier 2014, Lemma 10.2]. Note that there only the condition \( k'(\beta) > 0 \) is relevant (which follows from the convexity of \( k \)); the additional condition (stated there), that \( k(\beta) = 1 \) is not needed.

\( \square \)
Now we turn to the announced estimate from below for \( P(\langle u, X \rangle > t) \). Our estimates will be given in terms of the sets

\[
V_{i,t} := \left\{ \|\Pi_i^* u\| \geq t \text{ and } \|\Pi_{i,k}^* \| \leq e^{-\langle|i|-k\rangle \delta} C_0 t \forall k < |i| \right\},
\]

\[
W_{i,t'} := \left\{ \|\Pi_i^* u\| > t, \|\Pi_{i,k}^* \| > t, \|\Pi_{i,k}^* \| \leq C_0 t e^{-\langle|i|-k\rangle \delta} \right\},
\]

for some constants \( C_0 \) and \( \delta \) that will be defined below.

**Lemma 5.2.** For all \( u \in S, C_0 > 0 \) there is \( \kappa > 0 \) such that for all \( t > 0 \) and all a.s. subsets \( W \subset \mathbb{T} \),

\[
P(\langle u, X \rangle > t) \geq \kappa E \left[ \sum_{i \in W} P(V_{i,t}) \right] - E \left[ \sum_{i \in W} \sum_{i' \in W \setminus W|i|} \sum_{i'' |i| \neq i'} P(W_{i,t'}) \right] \in [-\infty, \infty).
\]

**Proof.** Fix \( l \in \mathbb{N} \). From Eq. (4.5) we obtain that

\[
|\langle u, Y_i \rangle| \geq \left| \langle \Pi_i^* u, [Y_i - |i|] \rangle \right| - \sum_{k \leq |i|} \|\Pi_{i,k}^* \| \left| Z_{i,k} \right| = \sum_{k \leq |i|} \|\Pi_{i,k}^* \| \left| Z_{i,k} \right| \geq D t - \sum_{k \leq |i|} \|\Pi_{i,k}^* \| \left| Z_{i,k} \right| \geq D t - \sum_{k \leq |i|} \|\Pi_{i,k}^* \| \left| Z_{i,k} \right| \geq D t - \sum_{k \leq |i|} \|\Pi_{i,k}^* \| \left| Z_{i,k} \right| \geq D t - \sum_{k \leq |i|} \|\Pi_{i,k}^* \| \left| Z_{i,k} \right| \geq D t - \sum_{k \leq |i|} \|\Pi_{i,k}^* \| \left| Z_{i,k} \right| \geq D t - \sum_{k \leq |i|} \|\Pi_{i,k}^* \| \left| Z_{i,k} \right|
\]

which is larger than \( t \) upon choosing \( D \) large enough. Here again, we need the a.s. version (4.5).

Since \( X \lesssim Y_i \) for any \( l \in \mathbb{N} \), we obtain the following estimate

\[
P\left(|\langle u, X \rangle| > t\right) \geq P\left(\bigcup_{|i| \leq l} V_{i,t} \right) \geq P\left(\bigcup_{|i| \leq l, i \in W} V_{i,t} \right)
\]

which is valid for any a.s. subset \( W \subset \mathbb{T} \). Below, we use the shorthand \( W_l = \{i \in \mathbb{T} : |i| \leq l, i \in W\} \).

Assuming that \( \sup_{i \in W} \sum_{i \in W^i} 1_{\bar{V}_{i,t'}} < \infty \) (this will be shown below), we use the inclusion-exclusion formula and separate as follows:

\[
P\left(\bigcup_{|i| \leq l, i \in W} \bar{V}_{i,t} \right) \geq E\left(\sum_{i \in W} P(\bar{V}_{i,t} | \mathcal{F}_i)\right) - E\left(\sum_{i, i' \in W : i \neq i', |i| \leq l} 1_{\bar{V}_{i,t} \cap \bar{V}_{i',t'}}\right)
\]
Then As for obtaining Eq. (4.6) from Eq. (4.5), we have that

\[ \sum_{i \in W_t} \mathbb{P} \left( \tilde{V}_{i,t} \right) \geq \mathbb{E} \left( \sum_{i \in W_t} \mathbb{P} \left( \tilde{V}_{i,t} \mid J_i \right) \right) - \mathbb{E} \left( \sum_{i \in W_t} \sum_{i \in W_t, i \neq i', |i'| \leq |i|} 1_{W_{i',t}} \right) \]

As for obtaining Eq. (4.8) from Eq. (4.7), we have that

\[ \mathbb{P} \left( \tilde{V}_{i,t} \mid J_i \right) = \mathbb{P} \left( \tilde{V}_{i,t} \right) \quad \text{P-a.s.} \]

The union in the definition of \( \tilde{V}_{i,t} \) is over disjoint sets, since the \( \Omega_j \) are disjoint. Furthermore, \( X_i \) is independent of \( V_{i,t} \) and \( \Pi_i^* \), and has the same law as \( X \), hence

\[ \mathbb{P} \left( \tilde{V}_{i,t} \right) = \sum_{j=1}^J \mathbb{P} \left( V_{i,t} \cap \{ \Pi_i^* u \in \Omega_j^* \} \right) \mathbb{P} \left( X_i \in \Omega_j \land \left| X_i \right| > D \right) \]

\[ = \mathbb{P} \left( X \in \Omega_j \land \left| X \right| > D \right) \sum_{j=1}^J \mathbb{P} \left( V_{i,t} \cap \{ \Pi_i^* u \in \Omega_j^* \} \right) \geq \kappa \mathbb{P} \left( V_{i,t} \right) \]

Thus, we have obtained the following estimate, valid for all \( l \in \mathbb{N} \):

\[ \mathbb{P} \left( \left| \left( u, X \right) \right| > t \right) \geq \kappa \mathbb{E} \left( \sum_{i \in W_t} \mathbb{P} \left( V_{i,t} \right) \right) - \mathbb{E} \left( \sum_{i \in W_t} \sum_{i \in W_t, i \neq i', |i'| \leq |i|} 1_{W_{i',t}} \right) \]

We finally have to justify that \( \sup_{t \in \mathbb{N}} \mathbb{E} \sum_{i \in W_t} 1_{\tilde{V}_{i,t}} < \infty \). But estimating similarly as in (5.4), we obtain that \( \mathbb{P} \left( \tilde{V}_{i,t} \right) \leq J \mathbb{P} \left( V_{i,t} \right) \); and the supremum is obviously bounded by

\[ J \mathbb{E} \sum_{i \in \mathbb{T}} \mathbb{P} \left( V_{i,t} \right) \leq J \mathbb{E} \sum_{i \in \mathbb{T}} \mathbb{P} \left( \left( \Pi_i^* u \right) > t \right) = JC \sum_{n=0}^\infty (EN)^n \mathbb{P}_\mu \left( S_n > \log t \right) \leq CJ \sum_{n=0}^\infty \frac{m(s+d)^n}{t^s} < \infty, \]

where we used the Markov inequality for some \( s \in (\alpha, \beta) \) in the last step.

\[ 5.1. \text{Estimates for } \mathbb{P} \left( V_{i,t} \right), \mathbb{P} \left( W_{i',t} \right). \text{ The further analysis of Eq. (5.2) splits in two parts. On the one hand, we have to estimate the probabilities appearing there, in terms of the distance between } i \text{ and } i' \text{ on the other hand, we have to do some combinatorics on the tree, in order to do the summation. In this section, we bound the probabilities. The set } W \text{ will be defined precisely in the next section. However it will be a subset of the tree } \mathbb{T} \text{ consisting of vertices } i \text{ such that } \]

\[ n_t - \sqrt{n_t} < |i| < n_t - \sqrt{n_t}/2, \]

where \( n_t = \lfloor \log t/ \rho \rfloor \) and \( \rho = m'(\beta) \). Thus the estimates provided below will be only for this particular set of indices.

\[ 5.2. \text{Probability of } V_{i,t}. \text{ In view of (4.4) the probability of } V_{i,t} \text{ does in fact only depend on } n = |i|. \text{ Let } (A_n, Z_n) \text{ be a sequence of i.i.d. copies of } \left( A_1, Z \right) \text{ for } Z = \sum_{i=2}^N A_i X_i + Q. \text{ Define, with } \Pi_i^* := A_n^* \cdots A_1^*, \]

\[ V_{n,t} := \left\{ \|\Pi_i^* u\| \geq t \text{ and } \|\Pi_k^*\| \left( |Z_{k+1}| \lor 1 \right) \leq e^{-c(n-k)\delta} C_0 t \forall k < n \right\}. \]

Then \( \mathbb{P} \left( V_{i,t} \right) = \mathbb{P} \left( V_{n,t} \right) \) as soon as \( |i| = n \).
The sets $V_{n,t}$ were already considered for $d = 1$ in \cite{Buraczewski et al. (2014)} (Theorem 2.3) and for $d \geq 2$ (under the same hypotheses as in the present paper) in \cite{Buraczewski and Mentemeier (2014)} (Lemma 10.5). We refer to these two papers for the proofs of the following lemmas.

**Lemma 5.3.** Assume that $\mathbb{E}|Z|^\gamma < \infty$ for some $\gamma > 0$, then there are constants $\delta, C_0, D_1, D_2, N_0 > 0$ such that

$$D_1 \cdot \frac{k(\beta)^n}{\sqrt{n_t e^{3n_t \hat{\sigma}}}} \leq \mathbb{P} (|\Pi_n u| > t) \leq D_2 \cdot \frac{k(\beta)^n}{\sqrt{n_t e^{3n_t \hat{\sigma}}}},$$

$$D_1 \cdot \frac{k(\beta)^n}{\sqrt{n_t e^{3n_t \hat{\sigma}}}} \leq \mathbb{P} (V_{n,t}) \leq D_2 \cdot \frac{k(\beta)^n}{\sqrt{n_t e^{3n_t \hat{\sigma}}}}.$$

for all $n_t > N_0$ and every $n_t - \sqrt{n_t} \leq n \leq n_t - \sqrt{n_t}/2$.

For the assertion of this lemma to hold, $k(\beta) = 1$ is not necessary, we only need that $k'(\beta) > 0$.

**Lemma 5.4.** There is $D_3 > 0$ such that for all $t$ large enough and all $n_t - \sqrt{n_t} \leq n \leq n_t - \sqrt{n_t}/2$ and all $s \geq e^{n_t}$,

$$\mathbb{P} (||\Pi_n|| > s) \leq \frac{D_3 k(\beta)^n}{\sqrt{n_t s^\beta}}.$$ 

5.3. **Probability of $W_{i,\hat{i},t}$.** The main result of this subsection is the following lemma.

**Lemma 5.5.** Assume that hypotheses of Theorem 3.1 or Theorem 3.2 are satisfied. Then for $|i| \geq |\hat{i}|$ we have

$$\mathbb{P} (W_{i,\hat{i},t}) \leq \frac{C_2 k(\beta)|i|}{t^\beta} \frac{k(\beta)|\hat{i}| - |i\wedge \hat{i}|}{e^{\chi(|i| - |i\wedge \hat{i}|)}},$$

which is valid for $n_t - \sqrt{n_t} \leq |i|, |\hat{i}| \leq n_t - \sqrt{n_t}/2$ as soon as $t > e^{n_0}$, with a constant $C_2$ which is independent of $t$.

**Proof.** Here, the cases (N-random) and (N-fixed) as well as $d = 1$ resp. $d > 1$ have to be treated separately at some instance. More precisely, we will consider

1. (N-random), i.e. independent weights and $d \geq 1$,
2. (N-fixed) for $d > 1$, with the additional assumption that the support $\text{supp} M = \bigcup_{i=1}^N \text{supp} A_i$ is bounded,
3. (N-fixed) for $d = 1$, assuming that $\mathbb{E}|M|^{2\beta} < \infty$.

We will start with some general calculations, and then deal with the cases separately, the most complicated one being 3.

Denote the joint law of $(A_1, A_2)$ by $\eta$. Recall that for each $i \in T, 1 \leq k < l \leq N_i$, $\mathcal{L}((A_{ik}, A_{il})) = \eta$ as well. For $i, \hat{i} \in T$, we write

$$i_0 = i \wedge \hat{i}, \quad m = |i \wedge \hat{i}| = |i_0|, \quad i_{m+1} = i_0 k, \quad \hat{i}_{m+1} = i_0 l, \quad p = |i|, \quad q = |\hat{i}|, \quad U_i := \Pi_i^* u/|\Pi_i u|$$
and recall the notation $\Pi_{ij} = A_{ij}, \cdots A_{ij}, \Pi_{ij} = A_{ij} \cdots A_{ij}$ for the product of the weights along the path between $i$ and $j$. Then

$$\mathbb{P}(W_{i,t}) = \mathbb{P}(\Pi^*_{i} > t, \Pi^*_i > t, \Pi_{iA} \leq C_0 e^{-\delta(||i-m||})$$

$$= \mathbb{P}(\Pi^*_{i} > t, \Pi^*_i > t, \Pi_{iA} \leq C_0 e^{-\delta(||i-m||})$$

$$\leq \mathbb{P}(\Pi^*_{i} > t, \Pi^*_i > t, \Pi_{iA} \leq C_0 e^{-\delta(||i-m||})$$

$$\leq \mathbb{P}(\Pi^*_{i} > t, \Pi^*_i > t, \Pi_{iA} \leq C_0 e^{-\delta(||i-m||})$$

$$= \int \mathbb{P}(\Pi^*_{i} > t, \Pi^*_i > t, \Pi_{iA} \leq C_0 e^{-\delta(||i-m||})$$

$$\leq \int \mathbb{P}(\Pi^*_{i} > t, \Pi^*_i > t, \Pi_{iA} \leq C_0 e^{-\delta(||i-m||})$$

(5.8)

where we conditioned upon $(A_{i,b}, A_{i,d})$ and used the Markov inequality in the last step. The reason that we applied it with the exponent $\alpha$ is that we will be able to replace $\mathbb{E}[\Pi_n^\alpha]$ by $k(\alpha)^n$, but the latter one equals also $k(\beta)$. For $d = 1$, we have $\mathbb{E}[\Pi_n^\beta] = k(\beta)^n$. If $d > 1$, we obtain from Corollary 4.6 for Condition (C) (Guivarch and Le Page, 2012) Lemma 2.8 for condition (i-p-o) (the proof working for (id) as well) that for all $s \in I_n$ there is a constant $c_s$, independent of $n$, such that

(5.9)

$$\mathbb{E}[\Pi_n^\alpha] \leq c_s k(s)^n.$$

From now on, we will consider the cases (1) - (3) separately.

Case (1).
In this case, $a_1$ and $a_2$ are independent, and Eq. (5.8) simplifies to

$$\mathbb{P}(W_{i,t}) \leq \mathbb{P}(\Pi_{i} > t) \mathbb{E}[\Pi_{q-m}^\alpha] C_0 \mathbb{E}[M^\alpha]$$

$$\leq \frac{D_2^\alpha k(\beta)^p e^{a\delta(p-m)} k(\beta)^q-m}{\sqrt{m} t^{3} e^{a\delta(p-m)}}$$

$$= \frac{C_2 k(\beta)^q-m}{\sqrt{m} t^{3} e^{a\delta(p-m)}}$$

by Lemma 5.3 and Eq. 5.5 (recall $k(\alpha) = (\mathbb{E}N)^{-1} = k(\beta)$), with $C_2 := D_2^\alpha c_2 C_0^\alpha$ and $\chi := a\delta$.

Case (2).
In this case, $\|a_2\|$ is bounded by some constant $c_A$, say, and Eq. 5.8 simplifies to

$$\mathbb{P}(W_{i,t}) \leq \int \mathbb{P}(\Pi_{i} > t) \mathbb{E}[\Pi_{q-m}^\alpha] C_0 c_2^{\alpha}$$

and from here, we can proceed as before to obtain the estimate 5.7 with $C_2 := D_2^\alpha c_2 C_0^\alpha e^{a\delta(p-m)}$ and $\chi = a\delta$.

Case (3).
Here, we are dealing just with the one-dimensional-case. Therefore, Eq. 5.8 simplifies to

\[
\mathbb{P}(W_{i', t}) \leq \int \mathbb{P}(|\Pi_{m+1}^i| a_1 | |\Pi_m| > t) \frac{\mathbb{E}||\Pi_{m-1}^i| C_0^a |a_2|^\alpha}{e^{\alpha \delta(p-m)}} \eta(da_1, da_2)
\]

= \int \mathbb{P}
(\frac{t}{|a_1|}) \frac{k(\beta)^{q-m-1} C_0^a |a_2|^\alpha}{e^{\alpha \delta(p-m)}} \eta(da_1, da_2).

Now we have to distinguish, whether or not Lemma 5.3 applies.

**Case (3A).** Assume \( t/|a_1| > e^{(\beta-1)\varepsilon} \), i.e. \( |a_1| < e^{-(\beta-1)\varepsilon}t \). Then

\[
\int \mathbb{P}(\frac{t}{|a_1|}) \leq \int \mathbb{D}_2(k(\beta)^{q-m-1} C_0^a |a_2|^\alpha}{e^{\alpha \delta(p-m)}} \eta(da_1, da_2)
\]

\[
\leq \int \mathbb{D}_2(k(\beta)^{p} \frac{k(\beta)^{q-m}}{\sqrt{n_t} t^\beta} e^{\alpha \delta(p-m)} \mathbb{E}[|A_1|^\beta |A_2|^\alpha] \leq \mathbb{D}_2(k(\beta)^{p} \frac{k(\beta)^{q-m}}{\sqrt{n_t} t^\beta} e^{\alpha \delta(p-m)} \mathbb{E}[|A_1|^{2\beta}]^{1/2} \mathbb{E}[|A_2|^{2\alpha}]^{1/2},
\]

where we used the Cauchy-Schwartz inequality in the last line. Since \( \alpha < \beta \), and \( \mathcal{L}(A_1) = \mathcal{L}(A_2) = \mathcal{L}(M) \), the assumption \( \mathbb{E}[M]^{2\beta} \) guarantees in particular the finiteness of \( \mathbb{E}[|A_1|^{2\beta}] \) and \( \mathbb{E}[|A_2|^{2\alpha}] \).

**Case (3B).** Now we consider \( t/|a_1| \leq e^{(\beta-1)\varepsilon} \), and thus, recalling that \( p = |i| \) is supposed to satisfy \( n_t - \sqrt{n_t} \leq |i| \leq n_t - \sqrt{n_t}/2 \),

\[
|a_1| \geq e^{-(\beta-1)\varepsilon}t \geq e^{(n_t-p)\varepsilon} \geq e^{\varepsilon \sqrt{n_t}/2}.
\]

Using the Markov inequality, we obtain

\[
\int \mathbb{P}(\frac{t}{|a_1|}) \leq \int \mathbb{E}[|A_1|^\beta 1_{|A_1| \geq \exp(\varepsilon \sqrt{n_t}/2)} |A_2|^\alpha] \leq \int \mathbb{E}[|A_1|^\beta 1_{|A_1| \geq \exp(\varepsilon \sqrt{n_t}/2)} |A_2|^\alpha]
\]

To "clean" the constant and to obtain the missing factor \( \frac{1}{\sqrt{n_t}} \), we use the Hölder inequality two times. First, with \( p_1 = 2\beta/(2\beta - \alpha) = : 2 - \varepsilon \) and \( p_2 = 2\beta/\alpha \), and subsequently with \( p'_1 = 2/(2-\varepsilon) \) and \( p'_2 = 2/\varepsilon \). Finally, we use the Markov inequality in order to obtain the following:

\[
\mathbb{E}[|A_1|^\beta 1_{|A_1| \geq \exp(\varepsilon \sqrt{n_t}/2)} |A_2|^\alpha] \leq \mathbb{E}[|A_1|^{2\beta} 1_{|A_1| \geq \exp(\varepsilon \sqrt{n_t}/2)}]^{1/(2-\varepsilon)} \mathbb{E}[|A_2|^{2\beta}]^{\alpha/(2\beta)}
\]

\[
\leq \mathbb{E}[|A_1|^{2\beta}]^{1/2} \mathbb{P}(A_1 \geq \exp(\varepsilon \sqrt{n_t}/2))^{\varepsilon/(2(2-\varepsilon))} \mathbb{E}[|A_2|^{2\beta}]^{\alpha/(2\beta)}
\]

\[
\leq \mathbb{E}[|A_1|^{2\beta}]^{1/2} \frac{\mathbb{E}[|A_1|^\alpha]}{e^{\varepsilon \sqrt{n_t}/2}} |A_2|^{\alpha/(2\beta)}
\]

For \( t \) large enough, \( \sqrt{n_t} \leq \exp(\alpha \varepsilon \sqrt{n_t}/(8 - 4\varepsilon)) \).
Lemma 6.1. For all $E \in k\{W\}$ consisting of $C\{W\}$, we have that

\[
\mathbb{P}(W_{i,t} \leq \frac{k(\beta)^{|i|}}{n_t^{\beta}} \cdot e^{-\kappa(|i| - |k\epsilon|)} \cdot \frac{C_n}{k(\beta)^2} \mathbb{E}[^{2\beta}A_1]^{1/2} \cdot (D_3 \mathbb{E}[^{2\alpha}A_2]^{1/2} + (\mathbb{E}|A_1|^{\alpha})^{n_t/\beta} \mathbb{E}[^{2\beta}A_2]^{\alpha/(2\beta)})
\]

\[
\square
\]

6. Combinatorics on the tree

This section considers $N$ random. As we mentioned above the subset $W$ of $T$ will contain only some of the nodes satisfying $n_t - \sqrt{n_t} \leq |i| \leq n_t - \sqrt{n_t}/2$ and therefore also will depend on $t$. We will consider only a sparse subset of those nodes. Namely only nodes from every $C_1\{W\}$, which moreover end with $C\{W\}$ one’s. The number $C_1 \in \mathbb{N}$ will be a parameter of the proof, to be fixed at the very end. It’s choice will be independent of $t$. Note however, that the estimate $\mathbb{P}((u, X) > t) \geq \varepsilon t^\beta$ is only valid for large enough $t$, namely $t > C_1 e^{N_\alpha}$. Since it is sufficient to show that $\liminf_{t \to \infty} t^\beta \mathbb{P}((u, X) > t) \geq \varepsilon > 0$, we will even restrict to such $t$, for which $\sqrt{n_t}/2C_1$ is an integer. This is not really necessary, but simplifies expressions.

Below, we will often use that $1 = m(\beta) = k(\beta)EN$, and therefore, $k(\beta) = (EN)^{-1}$. Keep in mind, that under (N-random), the shape of the tree is random, as will be that of its subset. Therefore, we have to take expectations with respect to the shape of the tree.

To be precise, upon fixing $t$, let $n_t = \lceil \log t/\rho \rceil$. We will consider nodes the generations of which are from the set

\[L_t = \{k \in C_1 : n_t - \sqrt{n_t} \leq kC_1 < n_t - \sqrt{n_t}/2\}.\]

Note $\#L_t = \sqrt{n_t}/C_1$ since we assume the latter to be integer. Denote by $1_C 1 = 1 \ldots 1 \in U$ the sequence consisting of $C_1$ one’s. Define

\[
\#W = \{i \in T : |i| \in L_t \text{ and } i = i|\beta| - C_1 1, C_1\}.
\]

We will calculate below several times the expected number of elements of $W$ lying on the level $k \in L_t$:

\[
\mathbb{E}[\#\{i \in W : |i| = k\}] = \mathbb{E}[\#\{i \in T : |i| = k - C_1\}] = (EN)^{k-C_1} = k(\beta)^{C_1-k}
\]

Lemma 6.1. For all $t$ large enough, (and such that $\sqrt{n_t}/2C_1 \in \mathbb{N}$)

\[
\mathbb{E}\left[\sum_{i \in W} \mathbb{P}(V_{i,t})\right] \geq \frac{D_1 k(\beta)^{C_1}}{2C_1} \cdot \frac{1}{t^\beta}.
\]

Proof. Using Lemma 5.3,

\[
\mathbb{E}\left[\sum_{i \in W} \mathbb{P}(V_{i,t})\right] \geq \frac{D_1}{t^\beta} \cdot \frac{1}{\sqrt{n_t}} \mathbb{E}\left[\sum_{i \in W} k(\beta)^{|i|}\right]
\]

\[
\geq \frac{D_1}{t^\beta} \cdot \frac{1}{\sqrt{n_t}} \sum_{i \in L_t} k(\beta)^|i| \mathbb{P}(i = i|\beta| - C_1 1)
\]

\[
= \frac{D_1 k(\beta)^{C_1}}{2C_1} \cdot \frac{1}{t^\beta}
\]

\[
\square
\]
This in particular shows the finiteness of the left-hand-side for each $t$ and this particular subset $\mathcal{W}$. Therefore, Lemma 6.2 applies to give the estimate of the first term in (5.2), and we can proceed computing the sum over the mixed terms.

6.1. Calculations involving $W_{i,i',t}$.

Lemma 6.2. There is $0 < \eta < \chi$ (independent of $t$), such that for all $t$ large enough

$$
E \left[ \sum_{i \in \mathcal{W}} \sum_{i' \in \mathcal{W}, |i'| \leq |i|, i \neq i'} P(W_{i,i',t}) \right] \leq \frac{C_2 k(\beta)^{2C_1}}{\eta^q C_1} \frac{1}{t^\beta}.
$$

Proof. Step 1: We are going to reorganize the summation over $i, i' \in \mathcal{W}$, by ordering them according to latest common ancestor, $i_0 := i \wedge i'$. Introducing as before the notation

$$
\mathcal{W}_{i_0,t} = \{ i \in \mathcal{W} : i \geq i_0 \}, \quad |i \wedge i'| = m, \quad |i| = p, \quad |i'| = q,
$$

the restrictions $i_0 = i \wedge i', i, i' \in \mathcal{W}, |i'| \leq |i|$ translate to (we omit the restriction $i \neq i'$, since anyway we want an estimate from above)

$$
i, i' \in \mathcal{W}_{i_0,t}, \quad \max\{ m + C_1, n_t - \sqrt{m_t} \} \leq p \leq \sqrt{m_t} - \sqrt{m_t}/2, \quad \sqrt{m_t} \leq q \leq \sqrt{m_t}.
$$

Similar to Eq. (5.2), we compute for $l \geq m, l \in \mathcal{T}_t$, the expected size of the $l$th generation in $T_{i_0,t}$ to be

$$
E[\# \{ i \in \mathcal{W}_{i_0,t} : |i| = l \}] = E[\# \{ i \in \mathcal{T}_t : |i| = l - C_1, i|_{m} = i_0 \}] = (\mathcal{E} \mathcal{N})^{l-C_1} = k(\beta)^{C_1+m-l}.
$$

Abbreviate the lower bound for $p$ by $p_t := \max\{ m + C_1, n_t - \sqrt{m_t} \}$, and the upper bound $n_t^* := n_t - \sqrt{m_t}/2$. Then, using (5.7)

$$
E \left[ \sum_{i \in \mathcal{W}} \sum_{i' \in \mathcal{W}, |i'| \leq |i|, i \neq i'} P(W_{i,i',t}) \right] \leq E \left[ \sum_{m \leq n_t^*} \sum_{p_t \leq n_t^*} \sum_{i \in \mathcal{T}_{i_0,t}, |i| = p} \sum_{m \leq q \leq p} \left\{ i' : i_0 = i \wedge i', |i'| = q \right\} \sum_{k(\beta)^{C_1+m-q}} \frac{C_2 k(\beta)^p}{t^\beta \sqrt{m_t}} \frac{k(\beta)^{q-m}}{e^{\mathcal{E}(p-m)}} \right]
$$

$$
\leq E \left[ \sum_{m \leq n_t^*} \sum_{p_t \leq n_t^*} \sum_{i \in \mathcal{T}_{i_0,t}, |i| = p} \sum_{m \leq q \leq p} k(\beta)^{C_1+m-q} \frac{C_2 k(\beta)^p}{t^\beta \sqrt{m_t}} \frac{k(\beta)^{q-m}}{e^{\mathcal{E}(p-m)}} \right]
$$

$$
\leq E \left[ \sum_{m \leq n_t^*} \sum_{p_t \leq n_t^*} \sum_{i \in \mathcal{T}_{i_0,t}, |i| = p} \frac{C_2 k(\beta)^{p+C_1}}{t^\beta \sqrt{m_t}} \frac{(p-m)}{e^{\mathcal{E}(p-m)}} \right]
$$

$$
\leq \sum_{m \leq n_t^*} \sum_{p_t \leq n_t^*} k(\beta)^{C_1+m-p} \frac{C_2 k(\beta)^{p+C_1}}{t^\beta \sqrt{m_t}} \frac{(p-m)}{e^{\mathcal{E}(p-m)}} \frac{C_2 k(\beta)^{p+C_1}}{t^\beta \sqrt{m_t}} \frac{(p-m)}{e^{\mathcal{E}(p-m)}}
$$

Now, we have to split the remaining summation, depending on which lower bound for $p$ we are going to use. Before we do so, we note that since $p - m \geq C_1$, we can estimate

$$
\frac{p - m}{e^{\mathcal{E}(p-m)}} \leq \frac{1}{e^{\eta(p-m)}}
$$
for some $0 < \eta < \chi$, as soon as $C_1$ is large enough.

$$
\frac{C_2}{t^\beta \sqrt{n_t}} \sum_{m \leq n_t^\beta} \sum_{p \leq n_t^\beta} k(\beta)^{m+2C_1} (p-m) e^{\eta(p-m)}
\leq \frac{C_2}{t^\beta \sqrt{n_t}} \left[ \sum_{m+1 \leq n_t} \sum_{p \leq n_t^\beta} k(\beta)^{m+2C_1} e^{\eta(p-m)} + \sum_{n_t^\beta \leq p \leq n_t} k(\beta)^{m+2C_1} e^{\eta(p-m)} \right]
\leq \frac{C_2 k(\beta)^{2C_1}}{t^\beta \sqrt{n_t}} \left[ \frac{1}{e^{\eta C_1}} \sum_{l=0}^{\infty} e^{\eta l} + \frac{\sqrt{n_t}}{2} \frac{1}{e^{\eta C_1}} \right] = \frac{C_2 k(\beta)^{2C_1}}{t^\beta e^{\eta C_1}} \left[ \frac{1}{\sqrt{n_t} (1-e-\eta)} + \frac{1}{2} \right]
$$

For $t$ large enough, the factor in the brackets becomes smaller than one, and we obtain the assertion.

$\Box$

7. PROOF OF THE MAIN THEOREM

The proof is just a consequence of the previous results. Lemma 5.2 provides lower estimates of $\mathbb{P}(\langle u, X \rangle > t)$, that is, $\mathbb{P}(\langle u, X \rangle > t)$ in terms of a subset $\mathbb{W}$ of $\mathbb{T}$ and probabilities $\mathbb{P}(V_{i,t})$, $\mathbb{P}(W_{i',t})$ for $i, i' \in \mathbb{W}$. Estimates of those probabilities were given in Lemmas 5.3 and 5.5, and the set $\mathbb{W}$ was defined as the beginning of Section 6. In view of Lemmas 6.1 and 6.2 we obtain

$$
\mathbb{P}(\langle u, X \rangle > t) \geq \kappa k(\beta)^{C_1} \left( \frac{D_1}{2C_1} - \frac{C_2}{(EN)^{C_1(e^{\eta C_1})}} \right) t^{-\beta}.
$$

Finally choose a large $C_1$ such that the last constant is positive we conclude the result.

REFERENCES

Alsmeyer, G., J. Biggins, and M. Meiners (2012). The functional equation of the smoothing transform. *Ann. Probab.* 40(5), 2069–2105.
Alsmeyer, G., E. Damek, and S. Mentemeier (2013). Precise tail index of fixed points of the two-sided smoothing transform. In *Random matrices and iterated random functions*, Volume 53 of *Springer Proc. Math. Stat.*, pp. 229–251. Springer, Heidelberg.
Alsmeyer, G. and M. Meiners (2013). Fixed points of the smoothing transform: two-sided solutions. *Probab. Theory Related Fields* 155(1-2), 165–199.
Alsmeyer, G. and S. Mentemeier (2012). Tail behaviour of stationary solutions of random difference equations: the case of regular matrices. *J. Difference Equ. Appl.* 18(8), 1305–1332.
Bassetti, F. and D. Mattes (2014). Multi-dimensional smoothing transformations: Existence, regularity and stability of fixed points. *Stochastic Processes and their Applications* 124(1), 154 – 198.
Biggins, J. D. and A. E. Kyprianou (1997). Seneta-Heyde norming in the branching random walk. *Ann. Probab.* 25(1), 337–360.
Biggins, J. D. and A. E. Kyprianou (2005). Fixed points of the smoothing transform: the boundary case. *Electron. J. Probab.* 10, no. 17, 609–631.
Buraczewski, D. (2009). On tails of fixed points of the smoothing transform in the boundary case. *Stochastic Process. Appl.* 119(11), 3955–3961.
Buraczewski, D., E. Damek, Y. Guivarc’h, and S. Mentemeier (2014). On multidimensional Mandelbrot’s cascades. *J. Difference Equ. Appl.* 20(11), 1523–1567.
Buraczewski, D., E. Damek, S. Mentemeier, and M. Mirek (2013). Heavy tailed solutions of multivariate smoothing transforms. *Stochastic Process. Appl.* 123(6), 1947–1986.
Buraczewski, D., E. Damek, and J. Zienkiewicz (2014+). Precise tail asymptotics of fixed points of the smoothing transform with general weights. to appear in Bernoulli.
Buraczewski, D. and K. Kolesko (2014). Linear stochastic equations in the critical case. J. Difference Equ. Appl. 20(2), 188–209.
Buraczewski, D. and S. Mentemeier (2014, May). Precise Large Deviation Results for Products of Random Matrices. ArXiv e-prints.
Durrett, R. and T. M. Liggett (1983). Fixed points of the smoothing transformation. Z. Wahrsch. Verw. Gebiete 64(3), 275–301.
Furstenberg, H. and H. Kesten (1960). Products of random matrices. Ann. Math. Statist. 31, 457–469.
Guivarc’h, Y. (1990). Sur une extension de la notion de loi semi-stable. Ann. Inst. H. Poincaré Probab. Statist. 26(2), 261–285.
Guivarc’h, Y. and É. Le Page (2012, April). Spectral gap properties and asymptotics of stationary measures for affine random walks. ArXiv e-prints. available online at [http://arxiv.org/abs/1204.6004](http://arxiv.org/abs/1204.6004).
Ikhsanov, A. and M. Meiners (2015). Fixed points of multivariate smoothing transforms with scalar weights. to appear in ALEA.
Jelenković, P. and M. Olvera-Cravioto (2012a). Implicit renewal theorem for trees with general weights. Stochastic Process. Appl. 122(9), 3209 – 3238.
Jelenković, P. and M. Olvera-Cravioto (2012b). Implicit renewal theory and power tails on trees. Adv. in Appl. Probab. 44(2), 528–561.
Jelenkovic, P. R. and M. Olvera-Cravioto (2014, May). Maximums on Trees. ArXiv e-prints.
Kesten, H. (1973). Random difference equations and renewal theory for products of random matrices. Acta Math. 131, 207–248.
Liu, Q. (1998). Fixed points of a generalized smoothing transformation and applications to the branching random walk. Adv. in Appl. Probab. 30(1), 85–112.
Liu, Q. (2001). Asymptotic properties and absolute continuity of laws stable by random weighted mean. Stochastic Process. Appl. 95(1), 83–107.
Mentemeier, S. (2013). On Multivariate Stochastic Fixed Point Equations: The Smoothing Transform and Random Difference Equations. Ph. D. thesis, Westfälische Wilhelms-Universität Münster.
Mentemeier, S. (2015+). The Fixed Points of the Multivariate Smoothing Transform. to appear in Probab. Theory Related Fields, 1–59.
Mirek, M. (2013). On fixed points of a generalized multidimensional affine recursion. Probab. Theory Related Fields 156(3-4), 665–705.
Neininger, R. and L. Rüschendorf (2004). A general limit theorem for recursive algorithms and combinatorial structures. Ann. Appl. Probab. 14(1), pp. 378–418.
Rösler, U. (1991). A limit theorem for “Quicksort”. RAIRO Inform. Théor. Appl. 25(1), 85–100.
Rösler, U. (2001). On the analysis of stochastic divide and conquer algorithms. Algorithmica 29(1-2), 238–261. Average-case analysis of algorithms (Princeton, NJ, 1998).

*Univiersytet Wroclawski, Instytut Matematyczny, pl. Grunwaldzki 2/4, 50-384 Wroclaw, Poland, †TU Dortmund, Fakultät für Mathematik, Vogelpothsweg 87, 44227 Dortmund, Germany

E-mail address: * dbura@math.uni.wroc.pl, † sebastian.mentemeier@tu-dortmund.de