Tail indices for AX + B Recursion with Triangular Matrices

Muneya Matsui¹ · Witold Świątkowski²

Received: 18 June 2019 / Revised: 10 June 2020 / Published online: 29 July 2020
© The Author(s) 2020

Abstract
We study multivariate stochastic recurrence equations (SREs) with triangular matrices. If coefficient matrices of SREs have strictly positive entries, the classical Kesten result says that the stationary solution is regularly varying and the tail indices are the same in all directions. This framework, however, is too restrictive for applications. In order to widen applicability of SREs, we consider SREs with triangular matrices and we prove that their stationary solutions are regularly varying with component-wise different tail exponents. Several applications to GARCH models are suggested.

Keywords  Stochastic recurrence equation · Kesten’s theorem · Regular variation · Multivariate GARCH(1, 1) processes · Triangular matrices

Mathematics Subject Classification (2010) Primary 60G10 · 60G70, Secondary 62M10 · 60H25

1 Introduction

1.1 Results and Motivation

A multivariate stochastic recurrence equation (SRE)

\[ W_t = A_t W_{t-1} + B_t, \quad t \in \mathbb{N} \]  

(1.1)

Muneya Matsui is partly supported by the JSPS Grant-in-Aid for Young Scientists B (16k16023) and for Scientific Research C (19K11868) and Nanzan University Pache Research Subsidy I-A-2 for the 2020 academic year. Witold Świątkowski is partly supported by the NCN Grant UMO-2014/15/B/ST1/00060.

Witold Świątkowski
witswiat@math.uni.wroc.pl

Muneya Matsui
mmuneya@nanzan-u.ac.jp

¹ Department of Business Administration, Nanzan University, 18 Yamazato-cho, Showa-ku, Nagoya 466-8673, Japan
² Institute of Mathematics, University of Wroclaw, Pl. Grunwaldzki 2/4, 50-384 Wroclaw, Poland
is studied, where $A_t$ is a random $d \times d$ matrix with nonnegative entries, $B_t$ is a random vector in $\mathbb{R}^d$ with nonnegative components and $(A_t, B_t)_{t \in \mathbb{Z}}$ are i.i.d. The sequence $(W_t)_{t \in \mathbb{N}}$ generated by iterations of SRE (1.1) is a Markov chain; however, it is not necessarily stationary. Under some mild contractivity and integrability conditions (see e.g., [5,6]), $W_t$ converges in distribution to a random variable $W$ which is the unique solution to the stochastic equation:

$$W \overset{d}{=} AW + B.$$  

(1.2)

here, $(A, B)$ denotes a generic element of the sequence $(A_t, B_t)_{t \in \mathbb{Z}}$, which is independent of $W$ and the equality is meant in distribution. If we put $W_0 \overset{d}{=} W$, then the sequence $(W_t)$ of (1.1) is stationary. Moreover, under suitable conditions a strictly stationary casual solution $(W_t)$ to (1.1) can be written by the formula:

$$W_t = B_t + \sum_{i=\infty}^{t} A_t \ldots A_i B_{i-1}$$

and $W_t \overset{d}{=} W$ for all $t \in \mathbb{Z}$.

Stochastic iteration (1.1) has been already studied for almost half a century, and it has found numerous applications to financial models (see e.g., Section 4 of [9, 12]). Various properties have been investigated. Our particular interest here is the tail behavior of the stationary solution $W$. The topic is not only interesting on its own, but it also has applications to, for example, risk management [12,22, Sec. 7.3].

The condition for the stationary solution $W$ having power decaying tails dates back to Kesten [19]. Since then, the Kesten theorem and its extensions have been used to characterize tails in various situations. An essential feature of Kesten-type results is that tail behavior is the same in all coordinates. We are going to call it Kesten property. However, this property is not necessarily shared by all interesting models—several empirical evidences support the fact (see e.g., [17,23,29] from economic data). Therefore, SREs with solutions having more flexible tails are both challenging and desirable in view of applications.

The key assumption implying Kesten property is an irreducibility condition, and it refers to the law of $A$. The simplest example when it fails is a SRE with diagonal matrices $A = diag(A_{11}, \ldots, A_{dd})$. In this case, the solution can exhibit different tail behaviors in coordinates, which is not difficult to see by the univariate Kesten result. In the present paper, we consider a particular case of triangular matrices $A$, which is much more complicated and much more applicable as well. We derive the precise tail asymptotics of solution in all coordinates. In particular, we show that it may vary in coordinates, though it is also possible that in some coordinates we have the same tail asymptotics.

More precisely, let $A$ be a nonnegative matrix matrices such that $A_{ii} > 0$, $A_{ij} = 0$, $i > j$ a.s. Suppose that $\mathbb{E}A_{ii}^{\alpha_i} = 1$ holds for $\alpha_i > 0$ in each $i$ such that $\alpha_i \neq$
We prove that when \( x \to \infty \)

\[
\mathbb{P}(W_i > x) \sim c_i x^{-\tilde{\alpha}_i}, \quad c_i > 0, \quad i = 1, \ldots, d,
\]

(1.3)

for \( \tilde{\alpha}_i > 0 \) depending on \( \alpha_1, \ldots, \alpha_d \). Here and in what follows, the notation ‘\( \sim \)’ means that the quotient of the left- and right-hand sides tends to 1 as \( x \to \infty \). For \( d = 2 \), the result (1.3) was proved in [10] with indices \( \tilde{\alpha}_1 = \min(\alpha_1, \alpha_2) \) and \( \tilde{\alpha}_2 = \alpha_2 \). The dependency of \( \tilde{\alpha}_1 \) on \( \alpha_1, \alpha_2 \) comes from the SRE:

\[
W_{1,t} = A_{11,t} W_{1,t-1} + A_{12,t} W_{2,t-1} + B_{1,t}
\]

where \( W_{1,t} \) is influenced by \( W_{2,t} \).

The same dependency holds in any dimension since \( W_i \) depends on \( W_{i+1}, \ldots, W_d \) in a similar but more complicated manner. In order to prove (1.3), we clarify this dependency with new notions and apply an induction scheme effectively.

The structure of the paper is as follows. In Sect. 2, we state preliminary assumptions and we show existence of stationary solutions to SRE (1.1). Section 3 contains the main theorem together with its proof. The proof follows by induction, and it requires several preliminary results. They are described in Sects. 4 and 5. Applications to GARCH models are suggested in Sect. 6. Section 7 contains discussion about constants of tails and open problems related to the subject of the paper.

### 1.2 Kesten’s Condition and Previous Results

We state Kesten’s result and briefly review the literature. We prepare a function

\[
h(s) = \inf_{n \in \mathbb{N}} \left( \mathbb{E} \| \Pi_n \|^s \right)^{1/n} \quad \text{with} \quad \Pi_n = A_1 \cdots A_n,
\]

where \( \| \cdot \| \) is a matrix norm. The tail behavior of \( W \) is determined by \( h \). We assume that there exists a unique \( \alpha > 0 \) such that \( h(\alpha) = 1 \). The crucial assumption of Kesten is irreducibility. It can be described in several ways, neither of which seems simple and intuitive. We are going to state a weaker property, which is much easier to understand. For any matrix \( Y \), we write \( Y > 0 \) if all entries are positive. The irreducibility assumption yields that for some \( n \in \mathbb{N} \)

\[
\mathbb{P}(\Pi_n > 0) > 0.
\]

(1.4)

Then, the tail of \( W \) is essentially heavier than that of \( \|A\| \): It decays polynomially even if \( \|A\| \) is bounded. Indeed, there exists a function \( e_\alpha \) on the unit sphere \( S^{d-1} \) such that

\[
\lim_{x \to \infty} x^\alpha \mathbb{P}(y'W > x) = e_\alpha(y), \quad y \in S^{d-1}
\]

(1.5)

and \( e_\alpha(y) > 0 \) for \( y \in S^{d-1} \cap [0, \infty)^d \), where \( y'W = \sum_{j=1}^d y_j W_j \) denotes the inner product. If \( \alpha \notin \mathbb{N} \), then (1.5) implies multivariate regular variation of \( W \), while if \( \alpha \in \mathbb{N} \), the same holds under some additional conditions (see [9, Appendix C]). Here,

\[\text{(1.4)}\]
we say that a $d$-dimensional r.v. $X$ is multivariate regularly varying with index $\alpha$ if

$$
P\left(\begin{array}{c} |X| > ux, \frac{X}{|X|} \in \cdot \\ |X| > x \end{array}\right) \overset{v}{\to} u^{-\alpha}P(\Theta \in \cdot), \quad u > 0,$$

(1.6)

where $\overset{v}{\to}$ denotes vague convergence and $\Theta$ is a random vector on the unit sphere $S^{d-1} = \{x \in \mathbb{R}^d : |x| = 1\}$. This is a common tool for describing multivariate power tails; see [3,26,27] or [9, p.279].

We proceed to the literature after Kesten. Later on, an analogous result was proved by Alsmeyer and Mentemeier [1] for invertible matrices $A$ with some irreducibility and density conditions (see also [20]). The density assumption was removed by Guivarc’h and Le Page [16] who developed the most general approach to (1.1) with signed $A$ having possibly a singular law. Moreover, their conclusion was stronger, namely they obtained existence of a measure $\mu$ on $\mathbb{R}^d$ being the weak limit of

$$
x^{\alpha}P(x^{-1}W \in \cdot) \quad \text{when } x \to \infty,$$

(1.7)

which is equivalent to (1.6). As a related work, in [8] the existence of the limit (1.7) was proved under the assumption that $A$ is a similarity.$^2$ See [9] for the details of Kesten’s result and other related results.

For all the matrices $A$ considered above, we have the same tail behavior in all directions, one of the reasons being a certain irreducibility or homogeneity of the operations generated by the support of the law of $A$. This is not the case for triangular matrices, since (1.4) does not hold then. Research in such directions would be a next natural step.$^3$

In this paper, we work out the case when the indices $\alpha_1, \ldots, \alpha_d > 0$ satisfying $E A_{ii}^{\alpha_i} = 1$ are all different. However, there remain not a few problems being settled, e.g., what happens when they are not necessarily different. This is not clear even in the case of $2 \times 2$ matrices. A particular case $A_{11} = A_{22} > 0$ and $A_{12} \in \mathbb{R}$ was studied in [11] where the result is

$$
P(W_1 > x) \sim \begin{cases} Cx^{-\alpha_1} (\log x)^{\alpha_1} & \text{if } E A_{11} A_{21}^{\alpha_1-1} \neq 0, \\ Cx^{-\alpha_1} (\log x)^{\alpha_1/2} & \text{if } E A_{11} A_{21}^{\alpha_1-1} = 0, \end{cases}$$

where $C > 0$ is a constant which may be different line by line. Our conjecture in $d (> 1)$ dimensional case is that

$$
P(W_i > x) \sim Cx^{-\tilde{\alpha}_i} \ell_i(x), \quad i = 1, \ldots, d$$

for some slowly varying functions $\ell_i$, and to get optimal $\ell_i$’s would be a real future challenge.

$^2$ $A$ is a similarity if for every $x \in \mathbb{R}^d$, $|Ax| = ||A|| \cdot |x|$.

$^3$ See also [7,8] and [9, Appendix D] for diagonal matrices.
2 Preliminaries and Stationarity

We consider $d \times d$ random matrices $A = [A_{ij}]_{i,j=1}^d$ and $d$-dimensional random vectors $B$ that satisfy the set of assumptions:

\begin{itemize}
  \item [(T-1)] $P(A \geq 0) = P(B \geq 0) = 1$,
  \item [(T-2)] $P(B_i = 0) < 1$ for $i = 1, \ldots, d$,
  \item [(T-3)] $A$ is upper triangular, i.e. $P(A_{ij} = 0) = 1$ whenever $i > j$,
  \item [(T-4)] There exist $\alpha_1, \ldots, \alpha_d$ such that $\mathbb{E}A_{ii}^{\alpha_i} = 1$ for $i = 1, \ldots, d$ and $\alpha_i \neq \alpha_j$ if $i \neq j$,
  \item [(T-5)] $\mathbb{E}A_{ij}^{\alpha_j} < \infty$ for any $i, j \leq d$,
  \item [(T-6)] $\mathbb{E}B_i^{\alpha_i} < \infty$ for $i = 1, \ldots, d$,
  \item [(T-7)] $\mathbb{E}[A_{ii}^{\alpha_i} \log |A_{ii}|] < \infty$ for $i = 1, \ldots, d$,
  \item [(T-8)] The law of $\log A_{ii}$ conditioned on $(A_{ii} > 0)$ is non-arithmetic for $i = 1, \ldots, d$.
\end{itemize}

Note that (T-1) and (T-4) imply that

$$P(A_{ii} > 0) > 0 \quad (2.1)$$

for $i = 1, \ldots, d$. For further convenience, any r.v. satisfying the inequality (2.1) will be called positive. Most conditions are similar to those needed for applying Kesten–Goldie’s result (see [9]).

Let $(A_t, B_t)_{t \in \mathbb{Z}}$ be an i.i.d. sequence with the generic element $(A, B)$. We define the products

$$\Pi_{t,s} = \begin{cases} A_t \cdots A_s, & t \geq s, \\ I_d, & t < s, \end{cases}$$

where $I_d$ denotes the identity $d \times d$ matrix. In the case $t = 0$, we are going to use a simplified notation:

$$\Pi_n := \Pi_{0,-n+1}, \quad n \in \mathbb{N},$$

so that $\Pi_0 = I_d$. For any $t \in \mathbb{Z}$ and $n \in \mathbb{N}$, the products $\Pi_{t,t-n+1}$ and $\Pi_n$ have the same distribution.

Let $\|\cdot\|$ be the operator norm of the matrix: $\|M\| = \sup_{\|x\|=1} |Mx|$, where $|\cdot|$ is the Euclidean norm of a vector. The top Lyapunov exponent associated with $A$ is defined by

$$\gamma_A = \inf_{n \geq 1} \frac{1}{n} \mathbb{E}[\log \|\Pi_n\|].$$

Notice that in the univariate case $A = A \in \mathbb{R}$ and $\gamma_A = \mathbb{E}[\log |A|]$.

If $\gamma_A < 0$, then the equation

$$W^d = AW + B \quad (2.2)$$
has a unique solution $W$ which is independent of $(A, B)$. Equivalently, the stochastic recurrence equation

$$W_t = A_t W_{t-1} + B_t$$  \hspace{1cm} (2.3)

has a unique stationary solution and $W_t \overset{d}{=} W$. Then we can write the stationary solution as a series

$$W_t = \sum_{n=0}^{\infty} \Pi_{t,t+1-n} B_{t-n}. \hspace{1cm} (2.4)$$

Indeed, it is easily checked that the process $(W_t)_{t=0}^{\infty}$ defined by (2.4) is stationary and solves (2.3), if the series on the right-hand side of (2.4) is convergent for any $t \in \mathbb{N}$. The convergence is ensured if the top Lyapunov exponent is negative (for the proof see [5]). Negativity of $\gamma_A$ follows from condition (T). We provide a proof in Appendix A.

From now on, $W = (W_1, \ldots, W_d)$ will always denote the solution to (2.2) and by $W_t = (W_{1,t}, \ldots, W_{d,t})$ we mean the stationary solution to (2.3). Since they are equal in distribution, when we consider distributional properties, we sometimes switch between the two notations, but this causes no confusion.

3 The Main Result and the Proof of the Main Part

3.1 The Main Result

We are going to determine the component-wise tail indices of the solution $W$ to stochastic equation (2.2). Since $A$ is upper triangular, the tail behavior of $W_i$ is affected only by $W_j$ for $j > i$, but not necessarily all of them: We allow some entries of $A$ above the diagonal to vanish $a.s.$ In the extreme case of a diagonal matrix, the tail of any coordinate may be determined independently of each other. On the other hand, if $A$ has no zeros above the diagonal, every coordinate is affected by all the subsequent ones. In order to describe this phenomenon precisely, we define a partial order relation on the set of coordinates. It clarifies interactions between them.

Definition 3.1 For $i, j \in \{1, \ldots, d\}$, we say that $i$ directly depends on $j$ and write $i \preceq j$ if $A_{ij}$ is positive (in the sense of (2.1)). We further write $i \prec j$ if $i \preceq j$ and $i \neq j$.

Observe that $i \preceq j$ implies $i \leq j$ since $A$ is upper triangular, while $i \preceq i$ follows from the positivity of diagonal entries. We extract each component of SRE (2.3) and may write

$$W_{i,t} = \sum_{j=1}^{d} A_{ij,t} W_{j,t-1} + B_{i,t} = \sum_{j:j \succ i} A_{ij,t} W_{j,t-1} + B_{i,t}, \hspace{1cm} (3.1)$$
where in the latter sum all coefficients $A_{ij,t}$ are positive. From this, we obtain the component-wise SRE in the form

$$W_{i,t} = A_{ii,t} W_{i,t-1} + \sum_{j:j>i} A_{ij,t} W_{j,t-1} + B_{i,t} = A_{ii,t} W_{i,t-1} + D_{i,t}, \quad (3.2)$$

where

$$D_{i,t} = \sum_{j:j>i} A_{ij,t} W_{j,t-1} + B_{i,t}. \quad (3.3)$$

Therefore, the tail of $W_{i,t}$ is determined by the comparison of the autoregressive behavior, characterized by the index $\alpha_i$, and the tail behavior of $D_{i,t}$, which depends on the indices $\alpha_j$, $j > i$. To clarify this, we define recursively new exponents $\tilde{\alpha}_i$, where $i$ decreases from $d$ to 1:

$$\tilde{\alpha}_i = \alpha_i \land \min\{\tilde{\alpha}_j : i < j\}. \quad (3.4)$$

If there is no $j$ such that $i < j$, then we set $\tilde{\alpha}_i = \alpha_i$. In particular, $\tilde{\alpha}_d = \alpha_d$. Notice that in general $\tilde{\alpha}_i \neq \min\{\alpha_j : i \leq j\}$. Depending on zeros of $A$, two relations $\tilde{\alpha}_i = \min\{\alpha_j : i \leq j\}$ and $\tilde{\alpha}_i > \min\{\alpha_j : i \leq j\}$ are possible for any $i < d$, see Example 3.4.

For further convenience, we introduce also a modified, transitive version of relation $\lessdot$.

**Definition 3.2** We say that $i$ depends on $j$ and write $i \lessdot j$ if there exists $m \in \mathbb{N}$ and a sequence $(i(k))_{0 \leq k \leq m}$ such that $i = i(0) \leq \cdots \leq i(m) = j$ and $A_{i(k)i(k+1)}$ is positive for $k = 0, \ldots, m - 1$. We write $i \prec j$ if $i \lessdot j$ and $i \neq j$.

Equivalently, the condition on the sequence $(i(k))_{0 \leq k \leq m}$ can be presented in the form $i \lessdot i(1) \lessdot \cdots \lessdot i(m) = j$. In particular, $i \prec j$ implies $i \lessdot j$. Now, we can write

$$\tilde{\alpha}_i = \min\{\alpha_j : i \lessdot j\} \quad (3.5)$$

which is equivalent to (3.4), though has a more convenient form.

Although Definitions 3.1 and 3.2 are quite similar, there is a significant difference between them. To illustrate the difference, we introduce the following notation for the entries of $\Pi_{i,s}$:

$$(\Pi_{i,s})_{ij} := \pi_{ij}(t, s).$$

If $t = 0$, we use the simplified form

$$\pi_{ij}(n) := \pi_{ij}(0, -n + 1), \quad n \in \mathbb{N}.$$ 

By $i \lessdot j$, we mean that the entry $A_{ij}$ of the matrix $A$ is positive, while $i \lessdot j$ means that for some $n \in \mathbb{N}$ the corresponding entry $\pi_{ij}(n)$ of the matrix $\Pi_n$ is positive.
The former relation gives a stronger condition. On the other hand, the latter is more convenient when products of matrices are considered, especially when transitivity plays a role. Example 3.4 gives a deeper insight into the difference between the two relations. Throughout the paper, both of them are exploited.

Now, we are ready to formulate the main theorem.

**Theorem 3.3** Suppose that condition (T) is satisfied for a random matrix $A$. Let $W$ be the solution to (2.2). Then, there exist strictly positive constants $C_1, \ldots, C_d$ such that

$$
\lim_{x \to \infty} x^{\hat{\alpha}_i} \mathbb{P}(W_i > x) = C_i, \quad i = 1, \ldots, d. \tag{3.6}
$$

The following example gives some intuition of what the statement of the theorem means in practice.

**Example 3.4** Let

$$
A = \begin{pmatrix}
A_{11} & A_{12} & 0 & 0 & A_{15} \\
0 & A_{22} & 0 & A_{24} & 0 \\
0 & 0 & A_{33} & 0 & A_{35} \\
0 & 0 & 0 & A_{44} & 0 \\
0 & 0 & 0 & 0 & A_{55}
\end{pmatrix},
$$

where $\mathbb{P}(A_{ij} > 0) > 0$ and other components are zero a.s. Suppose that $\alpha_4 < \alpha_3 < \alpha_2 < \alpha_5 < \alpha_1$ and $A$ satisfies the assumptions of Theorem 3.3. Then, $\tilde{\alpha}_1 = \alpha_4$, $\tilde{\alpha}_2 = \alpha_4$, $\tilde{\alpha}_3 = \alpha_3$, $\tilde{\alpha}_4 = \alpha_4$ and $\tilde{\alpha}_5 = \alpha_5$.

We explain the example step by step. The last coordinate $W_5$ is the solution to 1-dimensional SRE, so its tail index is $\alpha_5$. Since there is no $j$ such that $4 \prec j$, $\alpha_4$ is the tail index of $W_4$. For the third coordinate, the situation is different: We have $3 \prec 5$, so the tail of $W_3$ depends on $W_5$. But $\alpha_3 < \alpha_5$, so the influence of $W_5$ is negligible and we obtain the tail index $\tilde{\alpha}_3 = \alpha_3$. Inversely, the relations $2 \prec 4$ and $\alpha_2 > \alpha_4$ imply $\tilde{\alpha}_2 = \alpha_4$. The first coordinate clearly depends on the second and fifth, but recall that the second one also depends on the fourth. Hence, we have to compare $\alpha_1$ with $\alpha_2, \alpha_4$ and $\alpha_5$. The smallest one is $\alpha_4$; hence, $\tilde{\alpha}_1 = \alpha_4$ is the tail index. Although the dependence of $W_1$ on $W_4$ is indirect, we see it in the relation $1 \prec 4$.

### 3.2 Proof of the Main Result

Since the proof includes several preliminary results which are long and technical, they are postponed to Sects. 4 and 5. To make the argument more readable, we provide an outline of the proof here, referring to those auxiliary results. In the proof of Theorem 3.3, we fix the coordinate number $k$ and consider two cases: $\tilde{\alpha}_k = \alpha_k$ and $\tilde{\alpha}_k < \alpha_k$.

In the first case $\tilde{\alpha}_k = \alpha_k$, the proof is based on Goldie’s result [15, Theorem 2.3] and it is contained in Lemmas 4.2 and 4.3.

If $\tilde{\alpha}_k < \alpha_k$, the proof is more complicated and requires several auxiliary results. We proceed by induction. Let $j_0$ be the maximal coordinate among $j \succ k$ such that the modified tail indices $\tilde{\alpha}_j$ and $\alpha_k$ are equal. Then clearly, $\tilde{\alpha}_k = \alpha_{j_0} = \tilde{\alpha}_\ell$ for any
\( \ell \) such that \( k \leq \ell < j_0 \). We start with the maximal among such coordinates (in the standard order on \( \mathbb{N} \)) and prove (3.6) for it. Inductively we reduce \( \ell \) using results for larger tail indices and finally we reach \( \ell = k \).

We develop a component-wise series representation (5.4) for \( W_0 \). In Lemma 5.1, we prove that it indeed coincides a.s. with (2.4).

The series (5.4) is decomposed into parts step by step (Lemmas 5.4, 5.9 and 5.10). Finally, we obtain the following expression:

\[
W_{\ell, 0} = \pi_{\ell, j_0}(s)W_{j_0, -s} + R_{\ell, j_0}(s). \tag{3.7}
\]

Our goal is to prove that as \( s \to \infty \), the tail asymptotics of the first term \( \pi_{\ell, j_0}(s)W_{j_0, -s} \) approaches that of (3.6), while the second term \( R_{\ell, j_0}(s) \) becomes negligible.

The quantity \( R_{\ell, j_0}(s) \) is defined inductively by (5.31) and (5.39) as a finite collection of negligible parts, each being estimated in a different way. In the process of finding the main dominant term, we simultaneously settle the upper bounds for the negligible parts, each being estimated in a different way. In the process of finding the main dominant term, we simultaneously settle the upper bounds for the negligible parts. This is done through (5.15) in Lemmas 5.4, 5.3 and (5.29) under conditions of the main dominant term, we simultaneously settle the upper bounds for the negligible parts.

The final step is related to Breiman’s lemma applied to \( \pi_{\ell, j_0}(s)W_{j_0, -s} \). By independence between \( \pi_{\ell, j_0}(s) \) and \( W_{j_0, -s} \), we obtain the equality in asymptotics:

\[
\mathbb{P}(\pi_{\ell, j_0}(s)W_{j_0, -s} > x) \sim \mathbb{E}[\pi_{\ell, j_0}(s)^{\alpha_{j_0}}]\mathbb{P}(W_{j_0, -s} > x)
\]

for fixed \( s \). Recall that we need to let \( s \to \infty \) for \( R_{\ell, j_0}(s) \) to be negligible. The existence of \( \lim_{s \to \infty} \mathbb{E}[\pi_{\ell, j_0}(s)^{\alpha_{j_0}}] \) is assured in Lemma 5.7. Then, eventually we get \( \mathbb{P}(W_{\ell, 0} > x) \sim C \cdot \mathbb{P}(W_{j_0, -s} > x) \) for a positive constant \( C \), which can be explicitly computed. Now, (3.6) follows from stationarity and Lemma 4.3.

We move to the proof of Theorem 3.3.

**Proof of Theorem 3.3** Fix \( k \leq d \). If \( \alpha_k = \alpha_{j_0} \), then the statement of the theorem directly follows from Lemma 4.3. If \( \alpha_k < \alpha_{j_0} \), then there is a unique \( j_0 > k \) such that \( \alpha_k = \alpha_{j_0} = \alpha_{j_0} \). Another application of Lemma 4.3 proves that

\[
\lim_{x \to \infty} x^{\alpha_{j_0}}\mathbb{P}(W_{j_0} > x) = C_{j_0} \tag{3.8}
\]

for some constant \( C_{j_0} > 0 \). By stationarity, we set \( t = 0 \) without loss of generality.

The proof follows by induction with respect to number \( j \) in backward direction, namely we start with \( \ell = j_1 := \max\{j : k \leq j < j_0\} \) and reduce \( \ell \) until \( \ell = k \). Notice that \( j_1 < j_0 \), and thus \( \ell = j_1 \) satisfies conditions of Lemma 5.10. Thus, there exists \( s_0 \) such that for any \( s > s_0 \) we have

\[
W_{\ell, 0} = \pi_{\ell, j_0}(s)W_{j_0, -s} + R_{\ell, j_0}(s), \tag{3.9}
\]

where \( \pi_{\ell, j_0}(s) \) is independent of \( W_{j_0, -s} \) such that \( 0 < \mathbb{E}[\pi_{\ell, j_0}(s)^{\alpha_{j_0}}] < \infty \), and \( R_{\ell, j_0}(s) \) satisfies (5.29). We are going to estimate \( \mathbb{P}(W_{\ell, 0} > x) \) from both below and above. Since \( R_{\ell, j_0}(s) \geq 0 \) a.s.,

\[
\mathbb{P}(W_{\ell, 0} > x) \geq \mathbb{P}(\pi_{\ell, j_0}(s)W_{j_0, -s} > x)
\]
holds, and by Breiman’s lemma [9, Lemma B.5.1] for fixed \( s > s_0 \)
\[
\lim_{x \to \infty} \frac{P(\pi_{\ell, j_0}(s) W_{j_0, -s} > x)}{P(W_{j_0, -s} > x)} = E[\pi_{\ell, j_0}(s)^{\alpha_{j_0}}] =: u_\ell(s).
\]
Combining these with (3.8), we obtain the lower estimate
\[
\lim_{x \to \infty} x^{\alpha_{j_0}} P(W_{\ell, 0} > x) \geq u_\ell(s) \cdot C_{j_0}, \quad (3.10)
\]
Now, we pass to the upper estimate. Recall (5.29) in Lemma 5.10 which implies that for any \( \delta \in (0, 1) \) and \( \varepsilon > 0 \) there exists \( s_1 \) such that for \( s > s_1 \)
\[
\lim_{x \to \infty} x^{\alpha_{j_0}} P(R_{\ell, j_0}(s) > \delta x) = \delta^{-\alpha_{j_0}} \lim_{x \to \infty} (\delta x)^{\alpha_{j_0}} P(R_{\ell, j_0}(s) > \delta x) < \varepsilon.
\]
Then, for fixed \( s \geq s_1 \), we apply Lemma B.1, which is a version of Breiman’s lemma, to (3.9) and obtain
\[
\lim_{x \to \infty} x^{\alpha_{j_0}} P(W_{\ell, 0} > x) = \lim_{x \to \infty} x^{\alpha_{j_0}} P(\pi_{\ell, j_0}(s) W_{j_0, -s} + R_{\ell, j_0}(s) > x) \\
\leq \lim_{x \to \infty} x^{\alpha_{j_0}} P(\pi_{\ell, j_0}(s) W_{j_0, -s} > x (1 - \delta)) \\
+ \lim_{x \to \infty} x^{\alpha_{j_0}} P(R_{\ell, j_0}(s) > \delta x) \\
< \lim_{x \to \infty} x^{\alpha_{j_0}} P(\pi_{\ell, j_0}(s) W_{j_0, -s} > x) + \varepsilon \\
= E[\pi_{\ell, j_0}(s)^{\alpha_{j_0}}] (1 - \delta)^{-\alpha_{j_0}} \lim_{x \to \infty} x^{\alpha_{j_0}} P(W_{j_0, -s} > x) + \varepsilon \\
= (1 - \delta)^{-\alpha_{j_0}} u_\ell(s) \cdot C_{j_0} + \varepsilon, \quad (3.11)
\]
where we also use (3.8). We may let \( \varepsilon \to 0 \) and \( \delta \to 0 \) together with \( s \to \infty \) here and in (3.10). The existence and positivity of the limit \( u_\ell = \lim_{s \to \infty} u_\ell(s) \) is assured by Lemma 5.7. Thus, from (3.10) and (3.11) we have
\[
\lim_{s \to \infty} u_\ell(s) \cdot C_{j_0} = u_\ell \cdot C_{j_0}.
\]
This implies that
\[
\lim_{x \to \infty} x^{\tilde{\alpha}_\ell} P(W_{\ell, 0} > x) = u_\ell \cdot C_{j_0} =: C_\ell. \quad (3.12)
\]
Now, we go back to the induction process.
If \( j_1 = k \), then the proof is over, and if \( j_1 \neq k \), we set \( j_2 = \max\{ j : k \leq j < j_0, j \neq j_1 \} \). Then, there are two possibilities, depending on whether \( j_2 < j_1 \) or not. If \( j_2 \neq j_1 \), then the assumptions of Lemma 5.10 are satisfied with \( \ell = j_2 \) and we repeat the argument that we used for \( j_1 \). If \( j_2 < j_1 \), the assumptions of Lemma 5.9 are
fulfilled with \( \ell = j_2 \). Since the assertion of Lemma 5.9 is the same as that of Lemma 5.10, we can again repeat the argument that we used for \( j_1 \).

In general, we define \( j_{m+1} = \max(\{j : k \leq j < j_0\} \setminus \{j_1, \ldots, j_m\}) \). If \( j_{m+1} < j \) for some \( j \in \{j_1, \ldots, j_m\} \), then we use Lemma 5.9 with \( \ell = j_{m+1} \). Otherwise, we use Lemma 5.10 with the same \( \ell \). Then, by induction we prove (3.12) for every \( \ell : k \preceq \ell \prec j_0 \), particularly in the end we also obtain

\[
C_k = \lim_{x \to \infty} x^{\alpha_{j_0}} P(W_k > x).
\]

Notice that we have two limit operations, with respect to \( x \) and \( s \), and always the limit with respect to \( x \) precedes. We cannot exchange the limits, namely we have to let \( s \to \infty \) with \( s \) depending on \( x \).

### 4 Case \( \tilde{\alpha}_k = \alpha_k \)

The assumption \( \tilde{\alpha}_k = \alpha_k \) means that the tail behavior of \( W_k \) is determined by its auto-regressive property, namely the tail index is the same as that of the solution \( V \) of the stochastic equation \( V^d = A_{kk} V + B_k \). The tails of other coordinates on which \( k \) depends are of smaller order, which is rigorously shown in Lemma 4.2. In Lemma 4.3, we obtain the tail index of \( W_k \) by applying Goldie’s result. By Lemma 4.2, we observe that the perturbation induced by random elements other than those of \( k \)th coordinate \((A_{kk}, B_k)\) is negligible.

In what follows, we work on a partial sum of the stationary solution (2.4) component-wisely. To write the coordinates of the partial sum directly, the following definition is useful.

**Definition 4.1** For \( m \in \mathbb{N} \) and \( 1 \leq i, j \leq d \) let \( H_m(i, j) \) be the set of all sequences of \( m + 1 \) indices \((h(k))_{0 \leq k \leq m}\) such that \( h(0) = i, h(m) = j \) and \( h(k) \preceq h(k+1) \) for \( k = 0, \ldots, m - 1 \). For convenience, the elements of \( H_m(i, j) \) will be denoted by \( h \).

Notice that each of such sequences is non-decreasing since \( A \) is upper triangular. Moreover, \( H_m(i, j) \) is nonempty if and only if \( i \preceq j \) and \( m \) is large enough. Now, we can write

\[
\pi_{ij}(s) = \sum_{h \in H_s(i, j)} \prod_{p=0}^{s-1} A_{h(p)h(p+1),p-p}.
\]

Similar expression for \( \pi_{ij}(t, s) \) can be obtained by shifting the time indices,

\[
\pi_{ij}(t, s) = \sum_{h \in H_{t-s+1}(i, j)} \prod_{p=0}^{t-s} A_{h(p)h(p+1),t-p}, \quad t \geq s.
\]
which will be used later. Since the sum is finite, it follows from condition (T-5) that $\mathbb{E}\pi_{ij}(s)\tilde{a}_i < \infty$ for any $i, j$. Moreover, if $j \geq i$, then $\mathbb{P}(\pi_{ij}(s) > 0) > 0$ for $s$ large enough (in particular $s \geq j - i$ is sufficient). By definition, $\pi_{ij}(s)$ is independent of $W_j,-s$. Notice that when $i = \ell$, $j = j_0$ and $\ell \leq j_0$, it is the coefficient of our targeting representation (3.7).

**Lemma 4.2** For any coordinate, $i \in \{1, \ldots, d\}$, $\mathbb{E}W_i^\alpha < \infty$ if $0 < \alpha < \tilde{a}_i$.

**Proof** For fixed $i$, let us approximate $W_0$ by partial sums of the series (2.4). We will denote $S_n = \sum_{m=0}^{n} \Pi_mB_{-m}$ and $S_n = (S_{1,n}, \ldots, S_{d,n})$. We have then

$$S_{i,n} = \sum_{m=0}^{n} \sum_{j: j \geq i} \pi_{ij}(m)B_{j,-m} = \sum_{m=1}^{n} \sum_{j: j \geq i} \sum_{h \in H_m(i,j)} \left( \prod_{p=0}^{m-1} A_h(p)h(p+1),-p \right)B_{j,-m} + B_{i,0}.$$  \hspace{1cm} (4.2)

Suppose that $\alpha \leq 1$. Then, by independence of $A_h(p-1)h(p),-p$, $p = 1, \ldots, m - 1$ and $B_{j,-m}$,

$$\mathbb{E}S_{i,n}^\alpha \leq \sum_{m=1}^{n} \sum_{j: j \geq i} \sum_{h \in H_m(i,j)} \left( \prod_{p=0}^{m-1} \mathbb{E}A_h^\alpha(p)h(p+1),-p \right)\mathbb{E}B_{j,-m}^\alpha + \mathbb{E}B_{i,0}^\alpha.$$  \hspace{1cm} (2.4)

To estimate the right-hand side, we will need to estimate the number of elements of $H_m(i, j)$. To see the convergence of the series (4.2) it suffices to consider $m > 2d$. Recall that the sequences in $H_m(i, j)$ are non-decreasing; thus, for a fixed $j$ there are at most $j - i$ non-diagonal terms in each product on the right-hand side of (4.2). The non-diagonal terms in the product coincide with the time indices $t$, for which the values $h(t - 1)$ and $h(t)$ are different. If there are exactly $j - i$ such time indices, then the values are uniquely determined. There are $(\binom{m}{j-i})$ possibilities of placing the moments among $m$ terms of the sequence and $(\binom{m}{j-i}) < \binom{m}{d}$ since $m > 2d$ and $j - i < d$.

If we have $l < j - i$ non-diagonal elements in the product, then there are $(\binom{m}{l})$ possibilities of placing them among other terms and there are less than $(\binom{m}{l})$ possible sets of $l$ values. Hence, we have at most $(\binom{m}{l})^{(j-i)} < \binom{m}{d}d!$ sequences for a fixed $l$. Moreover, there are $j - i < d$ possible values of $l$, and hence there are at most $d \cdot d! \cdot \binom{m}{d}$ sequences in $H_m(i,j)$. Since $\binom{m}{d} < \frac{m^d}{d!}$, the number of sequences in $H_m(i,j)$ is further bounded by $dm^d$.

Now, recall that there is $\rho < 1$ such that $\mathbb{E}A_{j,l}^\alpha \leq \rho$ for each $j \geq i$ and that there is a uniform bound $M$ such that $\mathbb{E}A_{j,l}^\alpha < M$ and $\mathbb{E}B_{j,l}^\alpha < M$ whenever $j \geq i$, for any $l$.  \hspace{1cm} Springer
It follows that
\[
\mathbb{E} S_{i,n}^\alpha \leq C + \sum_{m=2d}^n \sum_{j:j \geq i} \sum_{h \in H_m(i,j)} M^{d+1} \rho^{m-d}
\]
\[
\leq C + \sum_{m=2d}^n \sum_{j:j \geq i} d M^{d+1} \rho^{-d} \cdot m^d \rho^m
\]
\[
\leq C + \sum_{m=2d}^n d^2 M^{d+1} \rho^{-d} \cdot m^d \rho^m
\]
\[
= C + d^2 M^{d+1} \rho^{-d} \sum_{m=2d}^n m^d \rho^m < \infty \quad (4.3)
\]
uniformly in \( n \), with \( C > 0 \), which is bounded from above. Hence, there exists the limit \( S_i = \lim_{n \to \infty} S_{i,n} \) a.s. and \( \mathbb{E} S_i^\alpha < \infty \). By (2.4), we have \( W_0 = \lim_{n \to \infty} S_n \) a.s. and we conclude that \( \mathbb{E} W_0^\alpha < \infty \).

If \( \alpha > 1 \), then by Minkowski’s inequality we obtain
\[
(\mathbb{E} S_{i,n}^\alpha)^{1/\alpha} \leq C' + \sum_{m=2d}^n \sum_{j:j \geq i} \sum_{h \in H_m(i,j)} \left( \prod_{p=0}^{m-1} \left( \mathbb{E} A_{h(p),h(p+1),-p}^\alpha \right)^{1/\alpha} \right) \left( \mathbb{E} B_{j,m}^\alpha \right)^{1/\alpha}
\]
with \( C' > 0 \). The same argument as above shows the uniform convergence. Thus, the conclusion follows.

Suppose that we have \( \tilde{\alpha}_k = \alpha_k \). This implies that \( \alpha_k < \tilde{\alpha}_j \) for each \( j \searrow k \) and hence, by Lemma 4.2, \( \mathbb{E} W_j^\alpha < \infty \). The next lemma proves the assertion of the main theorem in case \( \tilde{\alpha}_k = \alpha_k \).

**Lemma 4.3** Suppose assumptions of Theorem 3.3 are satisfied and let \( k \leq d \). Provided that \( \tilde{\alpha}_k = \alpha_k \), there exists a positive constant \( C_k \) such that
\[
\lim_{x \to \infty} x^{\tilde{\alpha}_k} \mathbb{P}(W_k > x) = C_k. \quad (4.4)
\]

**Proof** We are going to use Theorem 2.3 of Goldie [15] which asserts that if
\[
\int_0^\infty |\mathbb{P}(W_k > x) - \mathbb{P}(A_{kk} W_k > x)| x^{\alpha_k - 1} dx < \infty, \quad (4.5)
\]
then
\[
\lim_{x \to \infty} x^{\alpha_k} \mathbb{P}(W_k > x) = C_k, \quad (4.6)
\]
where
\[
C_k = \frac{1}{\mathbb{E} A_{kk}^{\alpha_k} \log A_{kk}} \int_0^\infty (\mathbb{P}(W_k > x) - \mathbb{P}(A_{kk} W_k > x)) x^{\alpha_k - 1} dx. \quad (4.7)
\]
To prove (4.5), we are going to use Lemma 9.4 from [15], which derives the equality

$$\int_0^\infty |\mathbb{P}(W_{k,0} > x) - \mathbb{P}(A_{kk,0} W_{k,-1} > x)| x^{\alpha_k-1} dx \geq \frac{1}{\alpha_k} \mathbb{E} \left| W_{k,0}^\alpha - (A_{kk,0} W_{k,-1})^\alpha \right| =: I. \quad (4.8)$$

From (3.1), we deduce that $W_{k,0} \geq A_{kk,0} W_{k,-1}$ a.s. Hence, the absolute value may be omitted on both sides of (4.8). We consider two cases depending on the value of $\alpha_k$.

**Case 1 $\alpha_k < 1$.**

Due to (3.2) and Lemma 4.2, we obtain

$$\alpha_k I \leq \mathbb{E} \left[ (W_{k,0} - A_{kk,0} W_{k,-1})^\alpha \right] = \mathbb{E} D_{k,0}^\alpha < \infty.$$  

**Case 2 $\alpha_k \geq 1$.**

For any $x \geq y \geq 0$ and $\alpha \geq 1$, the following inequality holds:

$$x^\alpha - y^\alpha = \alpha \int_y^x t^{\alpha-1} dt \leq \alpha x^{\alpha-1}(x - y).$$

Since $W_{k,0} \geq A_{kk,0} W_{k,-1} \geq 0$ a.s., we can estimate

$$W_{k,0}^\alpha - (A_{kk,0} W_{k,-1})^\alpha_k \leq \alpha_k W_{k,0}^\alpha_k - (W_{k,0} - A_{kk,0} W_{k,-1})^\alpha_k$$

$$= \alpha_k D_{k,0}^\alpha_k - (W_{k,0} - A_{kk,0} W_{k,-1})^\alpha_k$$

$$= \alpha_k D_{k,0}^\alpha_k (A_{kk,0} W_{k,-1} + D_{k,0})^\alpha_k - 1.$$

Since the formula

$$(x + y)^\alpha \leq \max \{1, 2^{\alpha-1}\} (x^\alpha + y^\alpha)$$

holds for any $x, y > 0$ and each $\alpha$, by putting $\alpha = \alpha_k - 1$ we obtain

$$I \leq \mathbb{E} \left[ D_{k,0} (A_{kk,0} W_{k,-1} + D_{k,0})^\alpha_k - 1 \right]$$

$$\leq \max \{1, 2^{\alpha_k-2}\} \left( \mathbb{E} D_{k,0}^\alpha + \mathbb{E} [D_{k,0} (A_{kk,0} W_{k,-1})^\alpha_k - 1] \right).$$

Since $\mathbb{E} D_{k,0}^\alpha < \infty$ by Lemma 4.2, it remains to prove the finiteness of the second expectation. In view of (3.2),

$$\mathbb{E} [D_{k,0} (A_{kk,0} W_{k,-1})^\alpha_k - 1] = \mathbb{E} \left[ A_{kk,0} B_{k,0} \right] \mathbb{E} W_{k,-1} + \sum_{j > k} \mathbb{E} \left[ A_{kk,0} A_{kj,0} \right] \mathbb{E} \left[ W_{k,-1} W_{j,-1} \right].$$
where we use independence of \((A_{.,0}, B_{.,0})\) and \(W_{.,-1}\). Since \(E W_{k,-1}^{\alpha_k - 1} < \infty\) by Lemma 4.2, we focus on the remaining terms. Take \(p = \frac{\alpha_k + \varepsilon}{\alpha_k + \varepsilon - 1}\) and \(q = \alpha_k + \varepsilon\) with \(0 < \varepsilon < \min(\bar{\alpha}_j : j \gg k) - \alpha_k\). Then, since

\[
p(\alpha_k - 1) = \frac{\alpha_k + \varepsilon}{\alpha_k + \varepsilon - 1}(\alpha_k - 1) = \left(1 + \frac{1}{\alpha_k + \varepsilon - 1}\right)(\alpha_k - 1) < \left(1 + \frac{1}{\alpha_k - 1}\right)(\alpha_k - 1) = \alpha_k,
\]

the Hölder’s inequality together with Lemma 4.2 yields

\[
E \left[ W_{k,-1}^{\alpha_k - 1} W_{j,-1} \right] \leq \left( E W_{k,-1}^{p(\alpha_k - 1)} \right)^{1/p} \cdot \left( E W_{j,-1}^{\alpha_k + \varepsilon} \right)^{1/q} < \infty.
\]

Similarly, \(E[A_{kk,0}^{\alpha_k - 1} B_{k,0}] < \infty\) and \(E[A_{kk,0}^{\alpha_k - 1} A_{kj,0}] < \infty\) hold and hence \(I < \infty\).

By [15, Theorem 2.3], (4.6) holds and it remains to prove that \(C_k > 0\). Since \(W_{k,0} \geq A_{kk,0} W_{k,-1} + B_{k,0}\) holds from (3.11), \(W_{k,0}^{\alpha_k} - (A_{kk,0} W_{k,-1})^{\alpha_k}\) is strictly positive on the set \(\{B_{k,0} > 0\}\) which has positive probability in view of condition (T-2). Therefore, in (4.8) we obtain \(I > 0\). Condition (T-7) implies that \(0 < E[A_{kk}^{\alpha_k} \log \bar{A}_{kk}] < \infty\). Hence, we see from (4.7) that \(C_k > 0\).

\[\square\]

5 Case \(\tilde{\alpha}_k < \alpha_k\)

The situation is now quite the opposite. The auto-regressive behavior of \(W_k\) does not play any role since \(k\) depends (in terms of Definition 3.2) on coordinates which admit dominant tails. More precisely, we prove that \(W_k\) has a regularly varying tail and its tail index is smaller than \(\alpha_k\). It is equal to the tail index \(\alpha_{j_0}\) of the unique component \(W_{j_0}\) such that \(\bar{\alpha}_k = \alpha_{j_0}(= \tilde{\alpha}_{j_0})\). The latter is due to the formula

\[
W_{\ell,0} = \pi_{\ell,j_0}(s) W_{j_0,-s} + R_{\ell,j_0}(s),
\]

and it is proved inductively for all \(\ell\) such that \(k \leq \ell < j_0\). The decomposition (5.1) is the main goal in this section. We show that \(R_{\ell,j_0}(s)\) is negligible as \(s \to \infty\), so that the tail of \(W_{\ell,0}\) comes from \(\pi_{\ell,j_0}(s) W_{j_0,-s}\). Moreover, we apply Breiman’s lemma to \(\pi_{\ell,j_0}(s) W_{j_0,-s}\). Then, we also need to describe the limit behavior of \(E \pi_{\ell,j_0}(s)^{\alpha_{j_0}}\) as \(s \to \infty\).

To reach (5.1), we first prove that \(W_{\ell,0}\) may be represented as in (5.4). Then, we divide the series (5.4) into two parts: the finite sum \(Q_F(s)\) and the tail \(Q_T(s)\) of (5.6). Next, we decompose \(Q_F(s)\) into two parts: \(Q_W(s)\) containing \(W\)-terms and \(Q_B(s)\) containing \(B\)-terms; see (5.13) and (5.14). Finally, \(Q_W(s)\) is split into \(Q'_W(s)\) and \(Q''_W(s)\), where the former contains the components with dominating tails, while the latter gathers those with lower-order tails. This decomposition suffices to settle the induction basis in Lemma 5.10, since (5.28) is satisfied if we set \(R_{\ell,j_0}(s) = Q''_W(s) + Q_B(s) + Q_T(s)\). The three parts \(Q_T(s), Q_B(s)\) and \(Q''_W(s)\) are estimated in Lemmas 5.3 and 5.4.
For the induction step in Lemma 5.9, we find \( s_1, s_2 \in \mathbb{N} \) such that \( s = s_1 + s_2 \) and extract another term \( Q_W(s_1, s_2) \) from \( Q_W(s_1) \), so that \( Q_W(s_1) = \pi_{\ell j_0}(s)W_{j_0, -s} + Q^*_W(s_1, s_2) \). Then, we set

\[
R_{\ell j_0}(s) = Q''_W(s_1) + Q_B(s_1) + Q_T(s_1) + Q^*_W(s_1, s_2).
\]

Notice that the two definitions of \( R_{\ell j_0}(s) \) above coincide, since \( Q^*_W(s, 0) = 0 \) in the framework of Lemma 5.10. The term \( Q^*_W(s_1, s_2) \) is estimated in the induction step in Lemma 5.9, where both \( s_1 \) and \( s_2 \) are required to be sufficiently large. The limit behavior of \( \mathbb{E}\pi_{\ell j_0}(s)^{a_{j_0}} \) as \( s \to \infty \) is given in Lemma 5.7.

### 5.1 Decomposition of the Stationary Solution

We need to introduce some notation. For the products of the diagonal entries of the matrix \( A \), we write

\[
\Pi^{(i)}_{t, s}(i) = \begin{cases} A_{ii, t} \cdots A_{ii, s}, & t \geq s, \\ 1, & t < s, \\ i = 1, \ldots, d; t, s \in \mathbb{Z} \end{cases}
\]

with a simplified form for \( t = 0 \)

\[
\Pi^{(i)}_{n} := \Pi^{(i)}_{0, -n+1}, \quad n \in \mathbb{N}.
\]

We also define the subset \( H'_m(i, j) \subset H_m(i, j) \) as

\[
H'_m(i, j) = \{ h \in H_m(i, j) : h(1) \neq i \}, \quad (5.2)
\]

namely \( H'_m(i, j) \) is the subset of \( H_m(i, j) \) such that the first two terms of its elements are not equal: \( i = h(0) \neq h(1) \). The latter naturally yields another definition:

\[
\pi'_{ij}(t, t - s) := \sum_{h \in H'_{m+1}(i, j)} \prod_{p=0}^{s} A_{h(p)h(p+1), t-p}; \quad \pi'_{ij}(s) := \pi'_{ij}(0, -s + 1), \quad (5.3)
\]

which looks similar to \( \pi_{ij}(t, t - s) \) and can be interpreted as an entry of the product \( A^0_t \Pi_{t-1, t-s} \), where \( A^0 \) stands for a matrix that has the same entries as \( A \) outside the diagonal and zeros on the diagonal.

Now, we are going to formulate and prove the auxiliary results of this section. We start with representing \( W_{\ell, t} \) as a series.

**Lemma 5.1** Let \( W_t = (W_{1, t}, \ldots, W_{d, t}) \) be the stationary solution of (2.3). Then

\[
W_{i, t} = \sum_{n=0}^{\infty} \Pi^{(i)}_{t, t-n+1} D_{i, t-n} \quad a.s., \quad i = 1, \ldots, d, \quad (5.4)
\]
where $D_{i,t-n}$ is as in (3.3).

**Remark 5.2** Representation (5.4) allows us to exploit the fast decay of $\Pi^{(i)}_{t,t-n+1}$ as $n \to \infty$.

**Proof** First, we show that for a fixed $\ell$ the series on the right-hand side of (5.4) converges. For this, we evaluate the moment of some order $\alpha$ from the range $0 < \alpha < \tilde{\alpha}_i$. Without loss of generality, we let $\alpha < 1$ and $t = 0$. By Fubini’s theorem,

$$\mathbb{E} \left( \sum_{n=0}^{\infty} \Pi^{(i)}_{n} D_{i,-n} \right)^{\alpha} \leq \sum_{n=0}^{\infty} \rho^n (dML + M) < \infty,$$

where $L = \max \{ \mathbb{E} W^{(j)}_{i} : j > i \}$ and $M = \mathbb{E} B_{i}^{(\alpha)} \vee \max \{ \mathbb{E} A^{(\alpha)}_{ij} : j > i \}$ are finite constants, and $\rho = \mathbb{E} A^{(\alpha)}_{ii} < 1$. Hence, the series defined in (5.4) converges.

Denote $Y_{i,t} := \sum_{n=0}^{\infty} \Pi^{(i)}_{i,t-n+1} D_{i,t-n}$. We are going to show that the vector $Y_t = (Y_1,t, \ldots, Y_d,t)$ is a stationary solution to (2.3). Since stationarity is obvious, without loss of generality set $t = 0$. For $i = 1, \ldots, d$, we have

$$Y_{i,0} = D_{i,0} + \sum_{n=1}^{\infty} \Pi^{(i)}_{n} D_{i,-n}$$

$$= D_{i,0} + A_{ii,0} \sum_{n=0}^{\infty} \Pi^{(i)}_{-1,-n} D_{i,-n-1}$$

$$= A_{ii,0} Y_{i,-1} + D_{i,0}. \quad (5.5)$$

The $d$ equations of the form (5.5) (one for each $i$) can be written together as a matrix equation:

$$Y_0 = A_0 Y_{-1} + B_0,$$

which is the special case of (2.3). It remains to prove that this implies that $Y_{i,0} = W_{i,0}$ a.s.

The series representation (2.4) allows to write $D_{i,t}$ as a measurable function of $\{(A_s, B_s) : s < t\}$. Since the sequence $(A_t, B_t)_{t \in \mathbb{Z}}$ is i.i.d. and hence ergodic, it follows from Proposition 4.3 of [21] that $(D_{i,t})_{t \in \mathbb{Z}}$ is also ergodic. Brandt [6] proved that then $Y_{i,t}$ defined in the paragraph above (5.5) is the only proper solution to (5.4) [6, Theorem 1]. Therefore, $Y_{i,t} = W_{i,t}$ a.s. \hfill \Box

Now, we divide the series (5.4) into the two parts: the finite sum $Q_F(s)$ of the $s$ first elements and the corresponding tail $Q_T(s)$:

$$W_{\ell,0} := \sum_{n=0}^{s-1} \Pi^{(\ell)}_{n} D_{\ell,-n} + \sum_{n=s}^{\infty} \Pi^{(\ell)}_{n} D_{\ell,-n} \quad (5.6)$$
For notational simplicity, we abbreviate coordinate number $\ell$ in $Q_F(s)$ and $Q_T(s)$. In the paper, the appropriate coordinate is always denoted by $\ell$. The same applies for other parts $Q_W(s)$, $Q_B(s)$, $Q'_W(s)$, $Q''_W(s)$ and $Q^*_W(s_1, s_2)$ of $W_{\ell,0}$ which will be defined in the sequel.

Both $Q_F(s)$ and $Q_T(s)$ have the same tail order for any $s$. However, if $s$ is large enough, one can prove that $Q_T(s)$ becomes negligible in the sense that the ratio of $Q_T(s)$ and $Q_F(s)$ tends to zero. The following lemma describes precisely that phenomenon. Recall that $j_0$ is the unique index with the property $\tilde{\alpha}_\ell = \alpha_{j_0}$ and so $\tilde{\alpha}_j = \alpha_{j_0}$ for all $j$ such that $\ell \preceq j \preceq j_0$.

**Lemma 5.3** Assume that for any $j$ satisfying $j_0 \preceq j \preceq \ell$,

$$\lim_{x \to \infty} x^{\alpha_{j_0}} P(W_j > x) = C_j > 0.$$  

(5.7)

Then, for every $\varepsilon > 0$ there exists $s_0 = s_0(\ell, \varepsilon)$ such that

$$\lim_{x \to \infty} x^{\alpha_{j_0}} P(Q_T(s) > x) < \varepsilon$$  

(5.8)

for any $s \geq s_0$.

**Proof** For $\rho := \mathbb{E}A^{\alpha_{j_0}}_{\ell} < 1$, we choose a constant $\gamma \in (0, 1)$ such that $\rho \gamma^{-\alpha_{j_0}} = : \delta < 1$. Notice that $\mathbb{E}(\prod_n^{\ell} A_{\ell}^{\alpha_{j_0}})$ decays to 0 exponentially fast, which is crucial in the following argument. We have

$$P(Q_T(s) > x) = \mathbb{P}\left(\sum_{n=s}^{\infty} \prod_n^{\ell} D_{\ell, -n} > x \right)$$

$$\leq \sum_{n=s}^{\infty} \mathbb{P}\left(\prod_n^{\ell} D_{\ell, -n} > x \cdot (1 - \gamma)^{-n-s} \right)$$

$$= \sum_{n=0}^{\infty} \mathbb{P}\left(\prod_n^{\ell} D_{\ell, -n-s} > x \cdot (1 - \gamma)^{-n} \right)$$

$$\leq \sum_{n=0}^{\infty} \mathbb{P}\left(\prod_n^{\ell} D'_{\ell, -n-s} > x \cdot \frac{(1 - \gamma)^{-n}}{2} \right) (: = I)$$

$$+ \sum_{n=0}^{\infty} \mathbb{P}\left(\prod_n^{\ell} D''_{\ell, -n-s} > x \cdot \frac{(1 - \gamma)^{-n}}{2} \right) (: = II),$$

where we divided $D_{\ell, -n-s} = B_{\ell, -n-s} + \sum_{j > \ell} A_{\ell, -n-s} W_{j, -n-s-1}$ into two parts:

$$D_{\ell, -n-s} = \sum_{j: \ell < j \preceq j_0} A_{\ell, -n-s} W_{j, -n-s-1} + \sum_{j: \ell < j \preceq j_0} A_{\ell, -n-s} W_{j, -n-s-1} + B_{\ell, -n-s}.$$  

(5.9)
The first part $D'_\ell, -n-s$ contains those components of $W_{-n-s-1}$ which are dominant, while $D''_\ell, -n-s$ gathers the negligible parts: the components with lower-order tails and the $B$-term. For the second sum II, we have $\mathbb{E}(D''_\ell, -n-s)^{\alpha_{j_0}} < \infty$ by Lemma 4.2 and condition (T), so that Markov’s inequality yields

$$x^{\alpha_{j_0}} II \leq \sum_{n=0}^{\infty} 2^{\alpha_{j_0}} (1 - \gamma)^{-\alpha_{j_0}} \gamma^{-n\alpha_{j_0}} \rho^{n+s} \mathbb{E}(D''_\ell, -n-s)^{\alpha_{j_0}} \leq c \rho^s \sum_{n=0}^{\infty} \delta^n < \infty.$$  

(5.10)

For the first sum I, we use conditioning in the following way:

$$x^{\alpha_{j_0}} I \leq \sum_{n=0}^{\infty} \sum_{j: j_0 < j \leq j_0} x^{\alpha_{j_0}} \mathbb{P} \left( \Pi_{n+s} A_{\ell,j,-n-s} W_{j,-n-s-1} > \frac{x \cdot (1 - \gamma) \gamma^n}{2d} \right)$$

$$= \sum_{n=0}^{\infty} \sum_{j: j_0 < j \leq j_0} \mathbb{E} \left[ x^{\alpha_{j_0}} \mathbb{P} \left( G_n W_{j,-n-s-1} > x \mid G_n \right) \right],$$

where $G_n = \Pi_{n+s} A_{\ell,j,-n-s}(1 - \gamma)^{-1} \gamma^{-n} \cdot 2d$. Notice that $G_n$ and $W_{j,-n-s-1}$ are independent and

$$\mathbb{E} G_n^{\alpha_{j_0}} \leq c_\ell \cdot \rho^{n+s} ((1 - \gamma) \gamma^n)^{-\alpha_{j_0}} (2d)^{\alpha_{j_0}},$$

where $c_\ell = \max \{ \mathbb{E} A_{\ell,j}^{\alpha_{j_0}} : \ell < j \leq j_0 \}$. Recall that by Assumption (5.7), there is a constant $c_j$ such that for every $x > 0$

$$\mathbb{P}(W_j > x) \leq c_j x^{-\alpha_{j_0}}.$$

Therefore, recalling $\rho \gamma^{-\alpha_{j_0}} =: \delta$, we further obtain

$$x^{\alpha_{j_0}} I \leq \sum_{n=0}^{\infty} \sum_{j_0 < j \leq j_0} c_j \cdot c_\ell \cdot \rho^{n+s} ((1 - \gamma) \gamma^n)^{-\alpha_{j_0}} (2d)^{\alpha_{j_0}} \leq c' \cdot \rho^s \sum_{n=0}^{\infty} \delta^n$$  

(5.11)

with $c' = d \cdot c_\ell \cdot \max \{ c_j : \ell < j \leq j_0 \} \cdot (1 - \gamma)^{-\alpha_{j_0}} (2d)^{\alpha_{j_0}}$. Now, in view of (5.10) and (5.11), since $\rho < 1$, there is $s_0$ such that

$$\lim_{x \to \infty} x^{\alpha_{j_0}} \mathbb{P}(Q_T(s) > x) < \varepsilon$$

for $s > s_0$. 

The dominating term $Q_F(s)$ of (5.6) can be further decomposed.

**Lemma 5.4** For any $\ell \leq d$ and $s \in \mathbb{N}$, the sum $Q_F(s)$ admits the decomposition

$$Q_F(s) = Q_W(s) + Q_B(s) \quad a.s.$$  

(5.12)
where

\[
Q_W(s) = \sum_{j : j > \ell} \pi_{\ell j}(s) W_{j, -s},
\]
(5.13)

\[
Q_B(s) = \sum_{n=0}^{s-1} \prod_{i}^{(\ell)} \left( B_{\ell, -n} + \sum_{j : j > \ell} \sum_{m=1}^{s-n-1} \pi_{\ell j}(-n, -n - m + 1) B_{j, -n-m} \right).
\]
(5.14)

Moreover,

\[
\mathbb{E} Q_B(s)^{\tilde{\alpha}_\ell} < \infty.
\]
(5.15)

**Remark 5.5** Each ingredient \( D_{\ell, -n} = \sum_{j : j > \ell} A_{\ell j, -n} W_{j, -n-1} + B_{\ell, -n} \) of \( Q_F(s) \) in (5.6) contains several \( W \)-terms with index \(-n - 1\) and a \( B \)-term with time index \(-n\) (see (3.3)). The idea is to change \( W_{i, -n-1} \) terms into \( W_{i, -s} \) by the iterative use of the recursion (2.3) component-wise. Meanwhile, some additional \( B \)-terms, with time indices between \(-n \) and \(-s + 1\), are produced. When all \( W \)-terms have the same time index \(-s\), we gather all ingredients containing a \( W \)-term in \( Q_W(s) \), while the remaining part \( Q_B(s) \) consists of all ingredients containing a \( B \)-term. The quantity \( Q_W(s) \) is easily treatable, and \( Q_B(s) \) is proved to be negligible in the tail (comparing to \( Q_W(s) \), for which the lower bound is shown in the proof of Theorem 3.3).

**Proof** Step 1. Decomposition of \( Q_F(s) \).

In view of (5.6) and (3.2), we have

\[
Q_F(s) = \sum_{n=0}^{s-1} \prod_{i}^{(\ell)} D_{\ell, -n} = \sum_{n=0}^{s-1} \prod_{i}^{(\ell)} \left( \sum_{i : i > \ell} A_{\ell i, -n} W_{i, -n-1} + B_{\ell, -n} \right) = \sum_{n=0}^{s-1} \xi_n.
\]
(5.16)

We will analyze each ingredient \( \xi_n \) of the last sum. We divide it into two parts, \( \eta_n \) and \( \zeta_n \), starting from \( n = s - 1 \).

\[
\xi_{s-1} = \prod_{j : j > \ell}^{(\ell)} A_{\ell j, -s+1} W_{j, -s} + \prod_{i}^{(\ell)} B_{\ell, -s+1} = \eta_{s-1} + \zeta_{s-1}.
\]
Next, for \( n = s - 2 \) applying component-wise SRE (3.1), we obtain

\[
\xi_{s-2} = \prod_{s=2}^{(\ell)} \sum_{j:j > \ell} A_{\ell j,-s+2} \left( \sum_{i:j \geq j} A_{ji,-s+1} W_{i,-s} + B_{j,-s+1} \right) + \prod_{s=2}^{(\ell)} B_{\ell,-s+2} \\
= \prod_{s=2}^{(\ell)} \sum_{j:j > \ell} A_{\ell j,-s+2} \sum_{i:j \geq j} A_{ji,-s+1} W_{i,-s} \\
+ \prod_{s=2}^{(\ell)} \left( B_{\ell,-s+2} + \sum_{j:j > \ell} A_{\ell j,-s+2} B_{j,-s+1} \right) = \eta_{s-2} + \xi_{s-2}.
\]

In this way, for \( n < s \) we define \( \eta_n \) which consists of terms including \((W_{j,-s})\) and \(\xi_n\) which contains terms \((B_{j,-i})\). Observe that in most \( \eta \)'s and \( \xi \)'s, a multiple summation appears, which is not convenient. To write them in simpler forms, we are going to use the notation (5.2). This yields

\[
\eta_{s-2} = \prod_{s=2}^{(\ell)} \sum_{i:j \geq \ell} \sum_{j:j \geq j} A_{\ell j,-s+2} A_{ji,-s+1} W_{i,-s} \\
= \prod_{s=2}^{(\ell)} \sum_{i:j \geq \ell} \sum_{h \in H_{s}^{(\ell)}, i} \left( \prod_{p=0}^{1} A_{h(p)h(p+1),-s+2-p} \right) W_{i,-s} \\
= \prod_{s=2}^{(\ell)} \sum_{i:j \geq \ell} \pi_{\ell i} (-s + 2, -s + 1) W_{i,-s}. \tag{5.17}
\]

For each \( \eta_n \), we obtain an expression of a simple form similar to (5.17). To confirm this, let us return to the decomposition of \( \xi \) and see one more step of the iteration for \( \eta \). For \( n = s - 2 \), we have

\[
\xi_{s-2} = \prod_{s=3}^{(\ell)} \left[ \sum_{j:j > \ell} A_{\ell j,-s+3} \left( \sum_{i:j \geq j} A_{ji,-s+2} \left( \sum_{u:u \geq i} A_{iu,-s+1} W_{u,-s} + B_{i,-s+1} \right) \right) \right] + B_{j,-s+2} + B_{\ell,-s+3}, \tag{5.18}
\]

so that similarly to the case \( n = s - 2 \), we obtain the expression

\[
\eta_{s-3} = \prod_{s=3}^{(\ell)} \sum_{j:j > \ell} A_{\ell j,-s+3} \sum_{i:j \geq j} A_{ji,-s+2} \sum_{u:u \geq i} A_{iu,-s+1} W_{u,-s} \\
= \prod_{s=3}^{(\ell)} \sum_{u:u \geq i, j:u \geq i, j > \ell} A_{\ell j,-s+3} A_{ji,-s+2} A_{iu,-s+1} W_{u,-s}.
\]
\[
\begin{align*}
\Pi_{s-3}^{(\ell)} & = \sum_{u:u \succ \ell} \sum_{h \in H_{\ell}(\ell,u)} \left( \prod_{p=0}^{2} A^h_{p}(p+1,-s+3-p) \right) W_{u,-s} \\
& = \sum_{u:u \succ \ell} \pi'_{\ell u} (-s + 3, -s + 1) W_{u,-s}.
\end{align*}
\]

Similarly for any \(0 \leq n \leq s - 1\), we obtain inductively

\[
\eta_n = \sum_{u:u \succ \ell} \pi'_{\ell u} (-n, -s + 1) W_{u,-s}.
\]

The expression for \(\zeta_n\) is similar but slightly different. Let us write it for \(n = s - 3\). From (5.18), we infer

\[
\begin{align*}
\zeta_{s-3} & = \sum_{j:j > \ell} \left( B_{\ell,-s+3} + \sum_{j:j > \ell} A_{\ell,j,-s+3} B_{j,-s+2} + \sum_{j:j \succ \ell} A_{j,-s+2} B_{\ell,-s+1} \right) \\
& = \sum_{j:j > \ell} \pi'_{\ell j} (-s + 3, -s + 3) B_{j,-s+2} \\
& + \sum_{i:i \succ \ell} \pi'_{i \ell} (-s + 3, -s + 2) B_{i,-s+1}
\end{align*}
\]

and the general formula for \(\zeta_n\) is:

\[
\zeta_n = \sum_{j:j \succ \ell} \pi'_{\ell j} (-n, -n - m + 1) B_{j,-n-m}.
\]

Therefore,

\[
Q_F(s) = \sum_{n=0}^{s-1} \xi_n = \sum_{n=0}^{s-1} \eta_n + \sum_{n=0}^{s-1} \zeta_n,
\]

where

\[
Q_W(s) = \sum_{n=0}^{s-1} \pi'_{\ell n} (-n, -s + 1) W_{n,-s}.
\]

Finally to obtain (5.13), we use (5.3) and the identity:

\[
\prod_{p=0}^{s-1-n} A_{h'(p)h'(p+1),-n-p} = \prod_{p=0}^{s-1} A_{h(p)h(p+1),-p}
\]
for $n < s$, $h' \in H'_{s-n}(\ell, j)$ and $h \in H_s(\ell, j)$ such that $h(0) = h(1) = \ldots = h(n) = \ell$ and $h(n+p) = h'(p)$ for $p = 0, \ldots, s-1-n$. Each element $h \in H_s(\ell, j)$ is uniquely represented in this way.

**Step 2 Estimation of $Q_B(s)$.**

Recall that $\tilde{a}_\ell \leq \alpha_j$ for $j \geq \ell$. Hence, $\mathbb{E} A_i^{\tilde{a}_\ell} < \infty$ and $\mathbb{E} B_i^{\tilde{a}_\ell} < \infty$ for any $i, j \geq \ell$. We have

$$\mathbb{E} Q_B(s)^{\tilde{a}_\ell} = \mathbb{E} \left\{ \sum_{n=0}^{s-1} \Pi_n^{(\ell)} (B_{\ell,n} + \sum_{j:j \geq \ell} \sum_{m=1}^{s-1-n} \pi_{\ell,j} (-n, -n - m + 1) B_{j,-n-m}) \right\}^{\tilde{a}_\ell}.$$  

(5.23)

Due to Minkowski’s inequality for $\tilde{a}_\ell > 1$ and Jensen’s inequality for $\tilde{a}_\ell \leq 1$, (5.23) is bounded by finite combinations of the quantities

$$\mathbb{E} \pi_{\ell,j}' (-n, -n - m + 1)^{\tilde{a}_\ell} < \infty \quad \text{and} \quad \mathbb{E} B_{j,-n-m}^{\tilde{a}_\ell} < \infty$$

for $0 \leq n \leq s-1$, $j : j \geq \ell$ and $1 \leq m \leq s-1-n$. Here, we recall that $\pi_{\ell,j}' (-n, -n - m + 1)$ is a polynomial of $A_{ij}$, $i, j \geq \ell$ (see (5.3)). Thus, $\mathbb{E} Q_B(s)^{\tilde{a}_\ell} < \infty$. \hfill $\square$

**Example 5.6** In order to grasp the intuition of decomposition (5.12), we consider the case $d = 2$ and let

$$A_t = \begin{pmatrix} A_{11,t} & A_{12,t} \\ 0 & A_{22,t} \end{pmatrix}.$$ 

Then, applying the recursions $W_{2,r} = A_{22,r} W_{2,r-1} + B_{2,r}$, $-s \leq r \leq 1$ to the quantity $Q_F(s)$ of (5.16), we obtain

$$Q_F(s) = \sum_{n=0}^{s-1} \Pi_n^{(1)} D_{1,-n}$$

$$= \sum_{n=0}^{s-1} \Pi_n^{(1)} (A_{12,-n} W_{2,-n-1} + B_{1,-n})$$

$$= \sum_{n=0}^{s-1} \Pi_n^{(1)} A_{12,-n} \Pi_{-n-1,-s+1} W_{2,-s}$$

$$+ \sum_{n=0}^{s-1} \Pi_n^{(1)} B_{1,-n} + \sum_{m=1}^{s-1-n} A_{12,-n} \Pi_{-n-1,-n+1-m} B_{2,-n-m},$$

where we recall that we use the convention $\Pi_n^{(1)} = 1$.

Let us focus on the first sum in the last expression, which is equal to $Q_W(s)$. Each term of this sum contains a product of $s+1$ factors of the form $A_{11,.}$, $A_{12,}$, or $A_{22,}$.
Each product is completely characterized by the non-decreasing sequence \((h(i))_{i=0}^{s}\) of natural numbers with \(h(0) = 1\) and \(h(s) = 2\). If \(h(i) = 1\) for \(i \leq q\) and \(h(i) = 2\) for \(i > q\), then there are \(q\) factors of the form \(A_{11}\), in front of \(A_{h(q)h(q+1),-q} = A_{12,-q}\), and \(s - 1 - q\) factors of the form \(A_{22}\), behind. All such sequences constitute \(H_s(1, 2)\) of Definition 4.1. Thus, we can write

\[
Q_W(s) = \sum_{h\in H_{s+1}(1, 2)} \prod_{p=0}^{s-1} A_{h(p)h(p+1),-p} W_2, -s. \tag*{\pi_{12}(s)}
\]

The second sum, which corresponds to \(Q_B(s)\), has another sum of the products in the \(n\)th term. All terms in these secondary sums of \(m\) are starting with \(A_{12,0}\) and then have a product of \(A_{22}\), until we reach \(B\). Each of the products is again characterized by a non-decreasing sequence \((h(i))_{i=0}^{m}\) such that \(h(0) = 1\), \(h(m) = 2\) and \(h(1) \neq 1\), because there is just one such sequence for each \(m\). Thus, we can use \(H_m'(1, 2)\) of Definition 4.1 and write

\[
Q_B(s) = \sum_{n=0}^{s-1} \prod_{n=0}^{(1)}(B_{1,-n} + \sum_{m=1}^{s-1-n} \sum_{h\in H_{m}'(1, 2)} \prod_{p=0}^{m-1} A_{h(p)h(p+1),-p} B_2, -n-m) \tag*{\pi_{12}' (n, -n-m+1)}.
\]

### 5.2 The Dominant Term

Lemmas 5.4 and 5.3 imply that the tail of \(W_{\ell,0}\) is determined by

\[
Q_W(s) = \sum_{j:j \triangleright s} \pi_{\ell,j}(s) W_{j, -s}
\]

in (5.13). In the subsequent lemmas (Lemmas 5.9 and 5.10), we will apply the recurrence to \(W_{j, -s}\), \(j < j_0\) until they reach \(W_{j_0,t}\) for some \(t < s\). Those \(W_{j, -s}\) could survive as the dominant terms, and terms \(W_{j_0, -s}\), \(j \triangleright j_0\) are proved to be negligible. In these steps, the behaviors of coefficients for all \(W_{j, -s}\) are inevitable.

The following property is crucial in subsequent steps, particularly in the main proof.

**Lemma 5.7** Let \(\alpha_i > \tilde{\alpha}_i = \alpha_{j_0}\) for some \(j_0 \triangleright i\) and \(u_i(s) := E\pi_{i,j_0}(s)^{\alpha_{j_0}}\). Then, the limit \(u_i = \lim_{s \to \infty} u_i(s)\) exists, and it is finite and strictly positive.

**Proof** First notice that \(u_i(s) > 0\) for \(s\) large enough since \(\pi_{ij}(s)\) is positive whenever \(j \triangleright i\) and \(s \geq j - i\). We are going to prove that the sequence \(u_i(s)\) is non-decreasing.
w.r.t. \( s \). Observe that

\[
\pi_{ij_0}(s + 1) = \sum_{j:j_0 \neq j \geq i} \pi_{ij}(s) A_{jj_0,s} \\
= \pi_{ij_0}(s) \cdot A_{jj_0,s}
\]

and therefore, by independence,

\[
u(s + 1) = \mathbb{E}\nu_{ij_0}(s + 1) \geq \mathbb{E}\nu_{ij_0}(s) \cdot \mathbb{E} A_{jj_0,s} = \mathbb{E}\nu_{ij_0}(s) = \nu(s).
\]

Since \( \nu(s) \) is non-decreasing in \( s \), if \( (\nu(s))_{s \in \mathbb{N}} \) is bounded uniformly in \( s \), the limit exists. For the upper bound on \( \nu(s) \), notice that each product in \( \pi_{ij}(s) \) (see (4.1)) can be divided into two parts:

\[
\prod_{p=0}^{s-1} A_{h(p)h(p+1),p} = \left( \prod_{p=0}^{s-m-1} A_{h(p)h(p+1),p} \right) \cdot \prod_{s-m+1}^{s} A_{jj_0,s},
\]

where \( m \) denotes the length of the product of \( A_{jj_0} \) terms. In the first part, all \( h(\cdot) \) but the last one \( h(s-m) = j_0 \) are strictly smaller than \( j_0 \). Since \( \mathbb{E} A_{jj_0} = 1 \), clearly

\[
\mathbb{E}\left( \prod_{p=0}^{s-1} A_{h(p)h(p+1),p} \right)^{\alpha_{j_0}} = \mathbb{E}\left( \prod_{p=0}^{s-m-1} A_{h(p)h(p+1),p} \right)^{\alpha_{j_0}}.
\]

Now, let \( \mathcal{H}'(j, j_0) \) denote the set of all sequences \( h \in \mathcal{H}(j, j_0) \) which have only one \( j_0 \) at the end. Suppose first that \( \alpha_{j_0} < 1 \). Then, by Jensen’s inequality we obtain

\[
\mathbb{E}\nu_{ij_0}(s)^{\alpha_{j_0}} = \mathbb{E}\left( \sum_{h \in \mathcal{H}'(i, j_0)} \prod_{p=0}^{s-1} A_{h(p)h(p+1),p} \right)^{\alpha_{j_0}} \\
\leq \sum_{h \in \mathcal{H}'(i, j_0)} \prod_{p=0}^{s-1} \mathbb{E} A_{h(p)h(p+1),p}^{\alpha_{j_0}} \\
= \sum_{m=0}^{s-1} \sum_{h \in \mathcal{H}'(i, j_0)} \prod_{p=0}^{s-m-1} \mathbb{E} A_{h(p)h(p+1),p}^{\alpha_{j_0}} \\
\leq \sum_{m=0}^{s-1} d(s-m)^d \cdot M^{j_0-i} \cdot \rho^{s-m-j_0} \\
\leq d \rho^{j_0-i} M^{j_0-i} \sum_{l=1}^{\infty} l^d \cdot \rho^l < \infty.
\]
Here, we put
\[
\rho = \max \{ E A_{jj}^{\alpha j_0} : j \ll j_0 \} < 1; \quad M = \max \{ E A_{uv}^{\alpha j_0} : u \ll j_0, v \ll j_0 \} < \infty.
\] (5.26)

The term \(d(s - m)^d\) in (5.24) is an upper bound on the number of elements of \(H_{s-m}(i, j)\). Another term \(M_{j_0-1}\) bounds the contribution of non-diagonal elements in the product, since any sequence \(h \in H_{s-m}(i, j_0)\) generates at most \(j_0 - i\) such elements. The last term \(\rho^{s-m-(j_0-i)}\) in (5.24) is an estimate of contribution of the diagonal elements since there are at least \(s - m - (j_0 - i)\) of them.

We obtain (5.25) from (5.24) by substituting \(l = s - m\) and extending the finite sum to the infinite series. Since the series converges and its sum does not depend on \(s\), the expectation \(E \pi_{ij}(s)^{\alpha j_0}\) is bounded from the above uniformly in \(s\).

Similarly, if \(\alpha j_0 \geq 1\), then by Minkowski’s inequality we obtain
\[
(\mathbb{E} \pi_{ij_0}(s)^{\alpha j_0})^{1/\alpha j_0} \leq \sum_{h \in H_i(i, j_0)} \prod_{p=0}^{s-1} (E A_{h(p)h(p+1), -p})^{1/\alpha j_0},
\]
which is again bounded from above, uniformly in \(s\), by the same argument. \(\Box\)

The number \(u_i\) depends only on \(i\), since the index \(j_0 = j_0(i)\) is uniquely determined for each \(i\). We are going to use \(u_i\) as an upper bound for \(\mathbb{E} \pi_{ij}(s)^{\alpha j_0}\) with any \(j\) such that \(j_0 \geq j \geq i\). This is justified by the following lemma.

**Lemma 5.8** For any \(j\) such that \(j_0 \gg j \gg i\) there is \(\hat{s}_j\) such that
\[
\mathbb{E} \pi_{ij}(s)^{\alpha j_0} < u_i \quad \text{for} \quad s > \hat{s}_j.
\] (5.27)

**Proof** The argument is similar to that in the proof of Lemma 4.2. Indeed, one finds \(\pi_{ij}(s)\) in \(S_{i,n}\) of (4.2) by setting \(s = m\). We briefly recall the argument for \(\alpha j_0 \leq 1\). The case \(\alpha j_0 > 1\) is similar. Without loss of generality, we set \(s \geq 2d - 1\). The number of sequences in \(H_i(i, j)\) is less than \(ds^d\). Taking \(\rho\) and \(M\) as in (5.26), from (4.1) and Jensen’s inequality we infer
\[ \mathbb{E} \pi_{ij}(s)^{\alpha_{j_0}} = \mathbb{E} \left( \sum_{h \in H_{i,j}} \prod_{p=0}^{s-1} A_{h(p)h(p+1),-p} \right)^{\alpha_{j_0}} \]
\[ \leq \sum_{h \in H_{i,j}} \prod_{p=0}^{s-1} \mathbb{E} A_{h(p)h(p+1)}^{\alpha_{j_0}} \]
\[ \leq d M^d s^d \rho^{s-d} \xrightarrow{s \to \infty} 0, \]
which implies (5.27).

We have already done all the preliminaries, and we are ready for the goal of this section, i.e., to establish the expression (3.7):

\[ W_{\ell,0} = \pi_{\ell,j_0}(s) W_{j_0,-s} + R_{\ell,j_0}(s) \]

and prove the negligibility of the term \( R_{\ell,j}(s) \) in the tail when \( s \to \infty \).

**Lemma 5.9** Suppose that \( \widetilde{\alpha}_\ell < \alpha_\ell \) and \( j_0 \succ \ell \) is the unique number with the property
\( \widetilde{\alpha}_\ell = \alpha_{j_0} \). Assume that
\[ \lim_{x \to \infty} x^{\alpha_{j_0}} \mathbb{P}(W_j > x) = C_j > 0 \]
whenever \( j_0 \succ j \succ \ell \). Then, for any \( \varepsilon > 0 \) there exists \( s_\ell = s_\ell(\varepsilon) \) such that if \( s > s_\ell \), \( W_{\ell,0} \) has a representation

\[ W_{\ell,0} = \pi_{\ell,j_0}(s) W_{j_0,-s} + R_{\ell,j_0}(s), \]

(5.28)

where \( R_{\ell,j_0}(s) \) satisfies
\[ \lim_{x \to \infty} x^{\alpha_{j_0}} \mathbb{P}(|R_{\ell,j_0}(s)| > x) < \varepsilon. \]

(5.29)

The proof is given by induction, and first we consider the case when indeed \( \ell \prec j_0 \), not only \( j_0 \succ \ell \). More precisely, we prove the following lemma which serves as a basic tool in each inductive step.

**Lemma 5.10** Suppose that \( \alpha_\ell > \widetilde{\alpha}_\ell = \alpha_{j_0} \) for some \( j_0 \succ \ell \) and \( \widetilde{\alpha}_j > \alpha_{j_0} \) for every \( j \succ \ell, j \neq j_0 \). Then, for any \( \varepsilon > 0 \) there exists \( s_\ell = s_\ell(\varepsilon) \) such that for any \( s > s_\ell \), \( W_{\ell,0} \) has a representation (5.28) which satisfies (5.29).

The condition in the lemma says that \( j_0 \) is the unique coordinate which has the heaviest tail among \( j : j \succ \ell \), all the other coordinates that determine \( \ell \) do not depend on \( j_0 \) and they have lighter tails.

Notice that as long as we only represent \( W_{\ell,0} \) by \( \pi_{\ell,j_0}(s) W_{j_0,-s} \) plus some r.v., we need not take a large \( s \). Indeed, \( s = 0 \) is enough to obtain \( (\pi_{\ell,j_0}(0) = A_{\ell,j_0,0}) \) if \( \ell \prec j_0 \). Thus, the number \( s_\ell \) is specific for (5.29).
**Proof** In view of (5.13), we may write

\[
Q_W(s) = \pi_{\ell,j_0}(s)W_{j_0,-s} + \sum_{j: j > \ell \quad j \neq j_0} \pi_{\ell,j}(s)W_{j,-s} + \hat{Q}_W(s)
\]

where \(\mathbb{E}\hat{Q}_W(s)^{\alpha_{j_0}} < \infty\) by the two facts;

1. \(\mathbb{E}\pi_{\ell,j}(s)^{\alpha_{j_0}} < \infty\) for any \(j\) (see the explanation (4.1)),
2. \(\mathbb{E}W_j^{\alpha_{j_0}} < \infty\) for \(j \succ \ell, j \neq j_0\), which is due to Lemma 4.2.

Thus, by Lemma 5.4 we have

\[
W_{\ell,0} = Q_W(s) + Q_B(s) + Q_T(s) = \pi_{\ell,j_0}(s)W_{j_0,-s} + \hat{Q}_W(s) + Q_B(s) + Q_T(s).
\]

where \(\hat{Q}_W(s)\) and \(Q_B(s)\) have finite moment of order \(\alpha_{j_0}\) and \(Q_T(s)\) satisfies (5.8). Now putting

\[
R_{\ell,j_0}(s) = \hat{Q}_W(s) + Q_B(s) + Q_T(s),
\]

and setting \(s_{\ell}(\varepsilon) = s_0(\ell, \varepsilon)\) as in Lemma 5.3, we obtain the result. \(\square\)

**Proof of Lemma 5.9** Here, we may allow the existence of \(j : j_0 \succ j \succ \ell\), so that there exist sequences \((j_i)_{0 \leq i \leq n}\) such that \(j_0 \succ j_1 \succ \cdots \succ j_n = \ell\). Since these sequences are strictly decreasing, their lengths are at most \(j_0 - \ell + 1\), i.e. possibly smaller than \(j_0 - \ell + 1\). Let \(n_0 = n_0(\ell)\) denote the maximal length of sequence such that \((j_i)_{1 \leq i \leq n_0}\) satisfies \(j_0 \succ j_1 \succ \cdots \succ j_{n_0} = \ell\). Then clearly \(j_0 \succ j_1 \succ \cdots \succ j_{n_0} = \ell\). In the same way, we define \(n_0(j)\) for any \(j\) in the range \(\ell < j \leq j_0\). We sometimes abbreviate to \(n_0(\ell)\). We use induction with respect to this maximal number \(n_0\) to prove (5.28) and (5.29). First, we directly prove these two properties in the cases \(n_0 = 0, 1\) which serve as the induction basis. For \(n_0 > 1\), the proof relies on the assumption that analogues of (5.28) and (5.29) hold for any \(j\) with \(n_0(j) < n_0(\ell)\). The proof is divided into four steps.

**Step 1** Scheme of the induction and the basis. Consider arbitrary coordinate \(j_0\) satisfying \(\alpha_{j_0} = \tilde{\alpha}_{j_0}\). We prove that for any \(\ell \leq j_0\) satisfying \(\tilde{\alpha}_\ell = \alpha_{j_0}\) the properties (5.28) and (5.29) hold. The proof follows by induction with respect to \(n_0(\ell)\). Namely, we assume that for any \(j \leq j_0\) satisfying \(\tilde{\alpha}_j = \alpha_{j_0}\) and \(n_0(j) < n_0(\ell)\) and for any \(\varepsilon > 0\) there is \(s_j = s_j(\varepsilon)\) such that for any \(s > s_j\) we have

\[
W_{j,0} = \pi_{j,j_0}(s)W_{j_0,-s} + R_{jj_0}(s),
\]

where \(R_{jj_0}(s)\) satisfies

\(\square\) Springer
\[
\lim_{x \to \infty} x^{\alpha_{j_0}} \mathbb{P}(|R_{j_0}(s)| > x) < \varepsilon. \tag{5.33}
\]

Then, we prove (5.28) and (5.29).

First, we settle the induction basis. Here, we consider the cases \( n_0 = 0 \) and \( n_0 = 1 \). The former case is equivalent to \( \ell = j_0 \). In that setting, Theorem 3.3 was already proved in Lemma 4.3; nevertheless, (5.28) and (5.29) need to be shown separately.

The iteration of (3.2) yields then that

\[
W_{j_0,0} = A_{j_0,j_0,0} W_{j_0,-1} + D_{j_0,0} = \ldots = \prod_{s=0}^{(j_0)} W_{j_0,-s} + \sum_{n=0}^{s-1} \Pi_{n}^{(j_0)} D_{j_0,-n}, \tag{5.34}
\]

where we recall that \( \Pi_{0}^{(j_0)} = 1 \) and \( \Pi_{1}^{(j_0)} = A_{j_0,j_0,0} \). From the definition of \( j_0 \) and Lemma 4.2, it follows that \( \mathbb{E} D_{j_0}^{\alpha_{j_0}} < \infty \). Since \( R_{j_0,j_0}(s) \) is constituted by a finite sum of ingredients which have the \( \alpha_{j_0} \)th moment, we conclude that \( \mathbb{E} R_{j_0,j_0}(s)^{\alpha_{j_0}} < \infty \) and (5.29) holds for any \( s \). Thus, we may let \( s_{\varepsilon}(\varepsilon) = 1 \) for any \( \varepsilon > 0 \). The case \( n_0 = 1 \) is precisely the setting of Lemma 5.10, in which we have already proved (5.28) and (5.29).

If \( n_0 > 1 \), then there is at least one coordinate \( j \) satisfying \( \ell < j \leq j_0 \). For any such \( j \), it follows that \( n_0(j) < n_0(\ell) \); hence, we are allowed to use the induction assumptions (5.32) and (5.33). In the next step, we prove that this range of \( j \) is essential, while for any other \( j > \ell \) the induction is not necessary.

**Step 2** Decomposition of \( Q_W(s) \) and estimation of the negligible term.

The first term in (5.28) comes from the part \( Q_W(s) \) of \( W_{\ell,0} \) in Lemma 5.4 and we further write

\[
Q_W(s) = \sum_{j: j_0 \nleq j \nleq \ell} \pi_{\ell j}(s) W_{j,-s} + \sum_{j: j_0 \nleq j \nleq \ell} \pi_{\ell j}(s) W_{j,-s}. \tag{5.35}
\]

Recall that the relation \( \nleq \) describes dependence between the components of the solution \( W_t \) after a finite number of iterations of (1.1). Therefore, the range of summation \( j_0 \nleq j \nleq \ell \) in \( Q_W'(s) \) means that \( \ell \neq j \) depends on both \( j \) and \( j_0 \) (by definition of \( j_0 \)), but \( j \) does not depend on \( j_0 \). The relation \( j \nleq \ell \) implies that \( \tilde{\alpha}_j \geq \tilde{\alpha}_\ell \), while \( j_0 \nleq j \) yields that \( \tilde{\alpha}_j \neq \alpha_{j_0} \). Recalling that \( \tilde{\alpha}_\ell = \alpha_{j_0} \), we can say that in \( Q_W'(s) \) we gather all coordinates \( j: j \nleq \ell \) such that \( \tilde{\alpha}_j = \alpha_{j_0} \) and \( Q_W'(s) \) consists of \( j: j \nleq \ell \) such that \( \tilde{\alpha}_j > \alpha_{j_0} \). Hence, by Lemma 4.2 each \( W_j \) appearing in \( Q_W'(s) \) has a tail of lower order than the tail of \( W_{j_0} \).

Notice that \( \hat{Q}_W(s) \) defined in (5.30) is the form that \( Q_W'(s) \) takes under the assumptions of Lemma 5.10. In this special case, we also have \( \hat{Q}_W(s) = \pi_{\ell,j_0}(s) W_{j_0,-s} \). We are going to study this expression in the more general setting of Lemma 5.9.

**Step 3** The induction step: decomposition of \( Q_W(s) \).
To investigate $Q'_W(s)$ in more detail, we will introduce a shifted version of $R_{ij}(s)$. First, recall that the r.v.’s $\pi_{ij}(t, s)$ and $\pi_{ij}(t - s + 1)$ have the same distribution and $\pi_{ij}(t, s)$ can be understood as a result of applying $t$ times a shift to all time indices in $\pi_{ij}(t - s - 1)$. In the same way, we define $R_{ij}(t, s)$ as a result of applying $t$ times a shift to all time indices in $R_{ij}(t - s - 1)$. In particular, we have $R_{ij}(0, -s + 1) = R_{ij}(s)$.

By the shift invariance of the stationary solution, (5.32) and (5.33) are equivalent to their time-shifted versions:

$$W_{j,t} = \pi_{jj_0}(t, t - s + 1)W_{j_0, t - s} + R_{jj_0}(t, t - s + 1), \quad t \in \mathbb{Z}$$

and

$$\lim_{x \to \infty} x^{\alpha_{j_0}} P(|R_{jj_0}(t, t - s + 1)| > x) < \varepsilon, \quad t \in \mathbb{Z},$$

respectively. Fix arbitrary numbers $s_1, s_2 \in \mathbb{N}$. Letting $t = -s_1$ in (5.36), we obtain

$$Q'_W(s_1) = \sum_{j : j_0 \geq j \geq \ell} \pi_{\ell j}(s_1) W_{j, -s_1}$$

$$= \sum_{j : j_0 \geq j \geq \ell} \pi_{\ell j}(s_1) \left\{ \pi_{jj_0}(-s_1, -s_1 - s_2 + 1)W_{j_0, -s_1 - s_2} + R_{jj_0}(-s_1, -s_1 - s_2 + 1) \right\}$$

$$= \sum_{j : j_0 \geq j \geq \ell} \pi_{\ell j}(s_1) \left\{ \pi_{jj_0}(-s_1, -s_1 - s_2 + 1)W_{j_0, -s_1 - s_2} + R_{jj_0}(-s_1, -s_1 - s_2 + 1) \right\} =: Q'_W(s_1, s_2),$$

where $\pi_{\ell j_0}(s_1 + s_2)$ consists of all combinations $\pi_{\ell j}(s_1)\pi_{jj_0}(-s_1, -s_1 - s_2 + 1)$ on $j : j_0 \geq j \geq \ell$. This is clear when we recall that $\pi_{\ell j}(s_1)$, $\pi_{jj_0}(-s_1, -s_1 - s_2 + 1)$ and $\pi_{\ell j_0}(s_1 + s_2)$ are proper entries of the matrices $\Pi_{s_1}, \Pi_{-s_1, -s_1 - s_2 + 1}$ and $\Pi_{s_1 + s_2} = \Pi_{s_1} \Pi_{-s_1, -s_1 - s_2 + 1},$ respectively.

Now, a combination of (5.12), (5.35) and (5.38) yields

$$W_{\ell,0} = Q_W(s_1) + Q_B(s_1) + Q_T(s_1)$$

$$= Q'_W(s_1) + Q''_W(s_1) + Q_B(s_1) + Q_T(s_1)$$

$$= \pi_{\ell j_0}(s_1 + s_2) W_{j_0, -s_1 - s_2} + Q''_W(s_1, s_2) + Q_B(s_1) + Q_T(s_1)$$

$$= \pi_{\ell j_0}(s_1 + s_2) W_{j_0, -s_1 - s_2} + R_{\ell j_0}(s_1 + s_2).$$

(5.39)
Our goal will be achieved by putting \( s = s_1 + s_2 \) in (5.39) and showing (5.29) for \( R_{\ell,j_0}(s) \) with \( s = s_1 + s_2 \).

**Step 4** The induction step: estimation of the negligible terms.

To obtain (5.29), we evaluate the four ingredients of \( R_{\ell,j_0}(s_1 + s_2) \) of (5.39), where the second hypothesis (5.33) of induction is used. Three of them, \( Q_W''(s_1), \ Q_B(s_1) \) and \( Q_T(s_1) \), are nonnegative; hence, it is sufficient to establish an upper bound for each of them. The fourth term, \( Q_W^*(s_1, s_2) \), may attain both positive and negative values; thus, we are going to establish an upper bound for its absolute value.

First, since the terms with \( j : j_0 \nleq j \geq \ell \) in \( Q_W''(s_1) \) of (5.35) satisfy the same condition as those of the sum \( \hat{Q}_W(s) \) of the previous lemma:

\[
\mathbb{E} Q_W''(s_1)^{\alpha_{j_0}} < \infty \tag{E.1}
\]

holds. Secondly,

\[
\mathbb{E} Q_B(s_1)^{\alpha_{j_0}} < \infty \tag{E.2}
\]

holds in view of Lemma 5.4. Moreover, by Lemma 5.3 for any \( \varepsilon' > 0 \) there is \( s_0 = s_0(\ell, \varepsilon') \) such that for \( s_1 > s_0 \)

\[
\lim_{x \to \infty} x^{\alpha_{j_0}} \mathbb{P}(Q_T(s_1) > x) < \varepsilon'. \tag{E.3}
\]

For the evaluation of \( Q_W^*(s_1, s_2) \), we will use (5.27) and therefore we have to assume that \( s_1 > \max\{\hat{s}_j; j_0 \nleq j \geq \ell\} \) where \( \hat{s}_j \) are defined in (5.27). Furthermore, we fix arbitrary \( \varepsilon'' > 0 \) and assume that \( s_2 > \max\{s_j(\varepsilon''): j_0 \nleq j \geq \ell\} \) where \( s_j(\varepsilon'') \) are defined right before (5.32). Recall that \( \pi_{\ell,j}(s_1), \ j : j_0 \nleq j \geq \ell \) is independent of \( R_{j,j_0}(-s_1, -s_1 - s_2 + 1) \) and has finite moment of order \( \alpha_{j_0} + \delta \) with some \( \delta > 0 \). We use (5.27), (5.37) and Lemma B.1 to obtain

\[
\lim_{x \to \infty} x^{\alpha_{j_0}} \mathbb{P}(\pi_{\ell,j}(s_1)R_{j,j_0}(-s_1, -s_1 - s_2 + 1) > x) \leq \mathbb{E} \pi_{\ell,j}(s_1)^{\alpha_{j_0}} \cdot \varepsilon' \leq u_{\ell} \cdot \varepsilon''.
\]

The situation is different for \( j = j_0 \). We cannot use Lemma B.1, because we do not know whether \( \mathbb{E} \pi_{\ell,j_0}(s_1)^{\alpha_{j_0} + \delta} < \infty \) for some \( \delta > 0 \). Indeed, it is possible that \( \mathbb{E} A_{j_0,j_0}^{\alpha_{j_0} + \delta} = \infty \) for all \( \delta > 0 \) and then clearly \( \pi_{\ell,j_0}(s_1) \) also does not have any moment of order greater than \( \alpha_{j_0} \). However, \( \mathbb{E} \pi_{\ell,j_0}(s_1)^{\alpha_{j_0}} < \infty \) holds for any \( s_1 \) and this is enough to obtain the desired bound. Since the term \( R_{j_0,j_0}(s_2) \) is nonnegative (see (5.34)) and it was already proved to have finite moment of order \( \alpha_{j_0} \), we obtain for any \( s_1, s_2 \)

\[
\mathbb{E} \{\pi_{\ell,j_0}(s_1)R_{j_0,j_0}(-s_1, -s_1 - s_2 + 1)\}^{\alpha_{j_0}} = \mathbb{E} \pi_{\ell,j_0}(s_1)^{\alpha_{j_0}} \cdot \mathbb{E} R_{j_0,j_0}(s_2)^{\alpha_{j_0}} < \infty
\]

and thus

\[
\lim_{x \to \infty} x^{\alpha_{j_0}} \mathbb{P}(\pi_{\ell,j_0}(s_1)R_{j_0,j_0}(-s_1, -s_1 - s_2 + 1) > x) = 0.
\]
Hence, setting \( N = \# \{ j : j_0 \geq j > \ell \} \), we obtain that

\[
\lim_{x \to \infty} x^{\alpha j_0} \mathbb{P}(Q^s_W(s_1, s_2) > x) \\
\leq \lim_{x \to \infty} x^{\alpha j_0} \mathbb{P}\left( \sum_{j: j_0 \geq j > \ell} \pi_{\ell j}(s_1) | R_{j j_0}(-s_1, -s_1 - s_2 + 1) | \right) \\
+ \pi_{\ell \ell}(s_1) \pi_{j j_0}(-s_1, -s_1 - s_2 + 1) W_{j j_0, -s_1 - s_2} > x \\
\leq \lim_{x \to \infty} x^{\alpha j_0} \left( \sum_{j: j_0 \geq j > \ell} \mathbb{P}\left( \pi_{\ell j}(s_1) | R_{j j_0}(-s_1, -s_1 - s_2 + 1) | > \frac{x}{N + 1} \right) \right) \\
+ \mathbb{P}\left( \pi_{\ell \ell}(s_1) \pi_{j j_0}(-s_1, -s_1 - s_2 + 1) W_{j j_0, -s_1 - s_2} > \frac{x}{N + 1} \right) \\
\leq \sum_{j: j_0 \geq j > \ell} \lim_{x \to \infty} x^{\alpha j_0} \mathbb{P}((N + 1) \cdot \pi_{\ell j}(s_1) | R_{j j_0}(-s_1, -s_1 - s_2 + 1) | > x) \\
+ \lim_{x \to \infty} x^{\alpha j_0} \mathbb{P}((N + 1) \cdot \pi_{\ell \ell}(s_1) \pi_{j j_0}(-s_1, -s_1 - s_2 + 1) W_{j j_0, -s_1 - s_2} > x) \\
\leq \sum_{j: j_0 \geq j > \ell} (N + 1)^{\alpha j_0} u_{\ell} \cdot \varepsilon'' + (N + 1)^{\alpha j_0} \rho^{s_1} u_{j_0} \cdot C_{j_0}.
\]

(5.40)

For the last inequality, we used Lemma B.1 and the fact that \( \pi_{\ell \ell}(s_1) = \prod_{0, -s_1 - 1}^{(\ell)} \). Since \( \rho = \mathbb{E} A_{\ell j_0}^\ell < 1 \), there is \( s' = s'(\varepsilon'') \) such that \( \rho^{s_1} u_{\ell} \cdot C_{j_0} < (N + 1)^{\alpha j_0} u_{\ell} \cdot \varepsilon'' \) for all \( s_1 > s' \). Then, recalling that the sum in the last expression contains at most \( N - 1 \) nonzero terms, the final estimate is

\[
\lim_{x \to \infty} x^{\alpha j_0} \mathbb{P}(Q^s_W(s_1, s_2) > x) < (N + 1)^{\alpha j_0 + 1} u_{\ell} \cdot \varepsilon''.
\]

(E.4)

Now, we are going to evaluate \( R_{\ell j_0}(s) \) of (5.29). The desired estimate can be obtained only if \( s \) is chosen properly.

For convenience, we briefly recall the conditions on \( s_1 \) and \( s_2 \) that were necessary to obtain the estimates (E.1–E.4). The inequalities (E.1) and (E.2) do not rely on any assumption on \( s_1 \) or \( s_2 \). The other relations are the following. Firstly, to obtain the inequality (E.3) we need to assume that \( s_1 > s_0(\ell, \varepsilon') \). Secondly, the estimates \( s_1 > \max(\hat{s}_j : j_0 \geq j > \ell) \) and \( s_2 > \max(s_j(\varepsilon'') : j_0 \geq j > \ell) \) are used to prove (5.40). Passing from (5.40) to (E.4) relies on the condition \( s_1 > s' \).

Now, let

\[
s > s_0(\ell, \varepsilon') \lor \max(\hat{s}_j : j_0 \geq j > \ell) \lor s' + \max(s_j(\varepsilon'') : j_0 \geq j > \ell) + 1,
\]

\( \odot \) Springer
where \( \cdot \lor \cdot = \max[\cdot, \cdot] \). Then, there are \( s_1 > s_0(\ell, \varepsilon') \lor \max\{s_j : j_0 \geq j > \ell\} \lor s' \) and \( s_2 > \max\{s_j(\varepsilon'') : j_0 \geq j > \ell\} \) such that \( s = s_1 + s_2 \). Then,

\[
R_{\ell, j_0}(s) = R_{\ell, j_0}(s_1 + s_2) = Q''_W(s_1) + Q''_W(s_1, s_2) + Q_B(s_1) + Q_T(s_1).
\]

The numbers \( s_1 \) and \( s_2 \) were chosen in such a way that (\textbf{E.3}) and (\textbf{E.4}) hold. The terms \( Q''_W(s_1) \) and \( Q_B(s_1) \) are negligible in the asymptotics. Therefore, it follows that

\[
\lim_{x \to \infty} x^{\alpha_{j_0}} P(R_{\ell, j_0}(s) > x) = \lim_{x \to \infty} x^{\alpha_{j_0}} P(Q''_W(s_1, s_2) + Q_T(s_1) > x) \\
\leq \lim_{x \to \infty} x^{\alpha_{j_0}} P(Q''_W(s_1, s_2) > x/2) + \lim_{x \to \infty} x^{\alpha_{j_0}} P(Q_T(s_1) > x/2) \\
\leq 2^{\alpha_{j_0}} \left( (N + 1)^{\alpha_{j_0} + 1} \mu_{\ell} \cdot \varepsilon'' + \varepsilon' \right).
\]

Since \( \varepsilon' \) and \( \varepsilon'' \) are arbitrary, we obtain (5.29).

\qed

6 Applications

Although there must be several applications, we focus on the multivariate GARCH(1, 1) processes, which is our main motivation. In particular, we consider the constant conditional correlations model by [4] and [18], which is the most fundamental multivariate GARCH process. Related results are the following. The tail of multivariate GARCH\((\rho, q)\) has been investigated in [13] but with the setting of Goldie's condition. A bivariate GARCH\((1, 1)\) series with a triangular setting has been studied in [23] and [10]. Particularly in [10], detailed analysis was presented including exact tail behaviors of both price and volatility processes. Since the detail of application is an analogue of the bivariate GARCH\((1, 1)\), we only see how the upper triangular SREs are constructed from multivariate GARCH processes.

Let \( \alpha_0 \) be a \( d \)-dimensional vector with positive elements, and let \( \alpha \) and \( \beta \) be \( d \times d \) upper triangular matrices such that nonzero elements are strictly positive. For a vector \( x = (x_1, \ldots, x_d) \), write \( x^{\gamma'} = (x_1^{\gamma'}, \ldots, x_d^{\gamma'}) \) for \( \gamma > 0 \). Then, we say that \( d \)-dimensional series \( X_t = (X_{1,t}, \ldots, X_{d,t})' \), \( t \in \mathbb{Z} \) has GARCH\((1, 1)\) structure if it satisfies

\[
X_t = \Sigma_t Z_t,
\]

where \( Z_t = (Z_{1,t}, \ldots, Z_{d,t})' \) constitute an i.i.d. \( d \)-variable random vectors and the matrix \( \Sigma_t \) is

\[
\Sigma_t = diag(\sigma_{1,t}, \ldots, \sigma_{d,t}).
\]

Moreover, the system of volatility vector \( (\sigma_{1,t}, \ldots, \sigma_{d,t})' \) is given by that of squared process \( W_t = (\sigma_{1,t}^2, \ldots, \sigma_{d,t}^2)' \). Observe that \( X_t = \Sigma_t Z_t = diag(Z_t)W_t^{1/2} \), so that
\( X_t^2 = \text{diag}(Z_t^2)W_t \). Then, \( W_t \) is given by the following auto-regressive model.

\[
W_t = \alpha_0 + \alpha X_{t-1}^2 + \beta W_{t-1} = \alpha_0 + (\text{diag}(Z_t^2) + \beta)W_{t-1}.
\]

Now, putting \( B_t := \alpha_0 \) and \( A_t := (\alpha \text{diag}(Z_t^2) + \beta) \), we obtain the SRE: \( W_t = A_tW_{t-1} + B_t \) with \( A_t \) the upper triangular with probability one. Each component of \( A_t \) is written as

\[
A_{ij,t} = \alpha_{ij} Z_{ij,t} + \beta_{ij}, \quad i \leq j \quad \text{and} \quad A_{ij,t} = 0, \quad i > j \quad \text{a.s.}
\]

Thus, we could apply our main theorem to the squared volatility process \( W_t \) and obtain the tail indices for \( W_t \). From this, we could derive tail behavior of \( X_t \) as done in [10].

Note that we have more applications in GARCH-type models. Indeed, we are considering applications in BEKK-ARCH models, of which tail behavior has been investigated with the diagonal setting (see [25]). At there, we should widen our results into the case where the corresponding SRE takes values on whole real line. The extension is possible if we assume certain restrictions and consider positive and negative extremes separately. Since the BEKK-ARCH model is another basic model in financial econometrics, the analysis with the triangular setting would provide more flexible tools for empirical analysis.

### 7 Conclusions and Further Comments

#### 7.1 Constants

In the bivariate case, we can obtain the exact form of constants for regularly varying tails (see [10]). The natural question is whether we can obtain the form of constants even in the \( d \)-dimensional case. The answer is positive. We provide an example which illustrates the method of finding these constants when \( d = 4 \). Let

\[
A = \begin{pmatrix}
A_{11} & A_{12} & A_{13} & A_{14} \\
0 & A_{22} & A_{23} & A_{24} \\
0 & 0 & A_{33} & A_{34} \\
0 & 0 & 0 & A_{44}
\end{pmatrix}
\]

and suppose that \( \alpha_3 < \alpha_4 < \alpha_2 < \alpha_1 \).

For coordinate \( k = 3, 4 \), we have

\[
C_k = \frac{\mathbb{E}[(A_{kk}W_k + B_k)^{\alpha_k} - (A_{kk}W_k)^{\alpha_k}]}{\alpha_k \mathbb{E}[A_{kk}^{\alpha_k} \log A_{kk}]},
\]

where \( W_k \) is independent of \( A_{kk} \) and \( B_k \). This is the Kesten–Goldie constant (see [15, Theorem 4.1]). Indeed, since \( W_4 \) is a solution to the univariate SRE, we immediately obtain the constant. Since the tail index of \( W_3 \) is equal to \( \tilde{\alpha}_3 = \alpha_3 \), the constant
follows by (4.6) in Lemma 4.3. For the second coordinate, we have an equation $W_2 \overset{d}{=} A_{22} W_2 + A_{23} W_3 + A_{24} W_4$. Since $\bar{\alpha}_2 = \alpha_3$ and $\bar{\alpha}_4 > \alpha_3$, the term $A_{23} W_3$ dominates all others in the asymptotics. In view of (3.12), we obtain

$$C_2 = u_2 \cdot C_3,$$

where the quantity $u_2$ is given in Lemma 5.7.

The situation seems more complicated for the first coordinate, because we have the condition $\bar{\alpha}_1 = \bar{\alpha}_2 = \alpha_3$ on the SRE: $W_1 \overset{d}{=} A_{11} W_1 + A_{12} W_2 + A_{13} W_3 + A_{14} W_4$. This means that the tail of $W_1$ comes from $W_2$ and $W_3$ both of which have dominating tails, and we could not single out the dominant term. However, by Lemmas 5.7 and 5.9 again we obtain a simple formula:

$$C_1 = u_1 \cdot C_3.$$

We can write the general recursive formula for constants in any dimension:

$$C_k = \begin{cases} 
    \frac{\mathbb{E}[(A_{kk} W_k + B_k)\alpha_k - (A_{kk} W_k)^{\alpha_k}]}{A_{kk} \mathbb{E}[A_{kk}^{\alpha_k} \log A_{kk}]} & \text{if } \tilde{\alpha}_k = \alpha_k, \\
    u_k \cdot C_{j_0} & \text{if } \tilde{\alpha}_k = \alpha_{j_0} < \alpha_k.
\end{cases}$$

Finally, we notice that these $u_k$ have only closed form including infinite sums. The exact values of $u_k$ seem to be impossible, and the only method to calculate them would be numerical approximations. The situation is similar to the Kesten–Goldie constant (see [24]).

### 7.2 Open Questions

In order to obtain the tail asymptotics of SRE such as (1.1), the Kesten’s theorem has been the key tool (see [9]). However, when the coefficients of SRE are upper triangular matrices as in our case, the assumptions of the theorem are not satisfied, so that we could not rely on the theorem. Fortunately in our setting, we can obtain the exact tail asymptotic of each coordinate, which is $\mathbb{P}(W_k > x) \sim C_k x^{-\tilde{\alpha}_k}$. However, in general setting, one does not necessarily obtain such asymptotic even in the upper triangular case.

The example is given in [11], which we briefly see. Let $A$ be an upper triangular matrix with $A_{11} = A_{22}$ having the index $\alpha > 0$. Then, depending on additional assumptions, it can be either $\mathbb{P}(W_1 > x) \sim C x^{-\alpha} (\log x)^{\alpha/2}$, or $\mathbb{P}(W_1 > x) \sim C' x^{-\alpha} (\log x)^{\alpha}$ for some constant $C, C' > 0$.

There are many natural further questions to ask. What happens to the solution if some indices $\alpha_i$ of different coordinates are equal? How could we find the tail asymptotics when the coefficient matrix is neither in the Kesten’s framework nor upper triangular? Moreover, if $A$ includes negative entries, could we derive the tail asymptotics? They are all open questions.
Acknowledgements The authors are grateful to Ewa Damek and Dariusz Buraczewski for valuable discussion on the subject of the paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Appendix A. Negativity of Top Lyapunov Exponent

We provide the proof for negativity of $\gamma_A$ in Sect. 2. It can also be deduced from more general results of Straumann [28] or Grencsér, Michaletzky and Orlovitz [14] on the top Lyapunov exponent for block triangular matrices. However, in our case there is a direct elementary approach based on equivalence of norms and Gelfand’s formula.

Since all matrix norms are equivalent, we can use the norm

$$\|A\|_1 = \sum_{i=1}^{d} \sum_{j=1}^{d} |A_{ij}|,$$

which is submultiplicative: $\|AB\|_1 \leq \|A\|_1 \|B\|_1$. Then, since $A$ has nonnegative entries, we have $E\|A\|_1 = \|EA\|_1$. Moreover, for any $\varepsilon \in (0, 1)$, $\|A\|_1^\varepsilon \leq \|A^\varepsilon\|_1$, where $A^\varepsilon$ denotes the matrix $A$ with each entry raised to the power of $\varepsilon$. We are going to apply these to the form of top Lyapunov exponent $\gamma_A$. For any $\varepsilon > 0$, by Jensen’s inequality we have

$$\gamma_A = \inf_{n \geq 1} (n\varepsilon)^{-1} E \log \|\Pi_n\|_1^\varepsilon$$

$$\leq \inf_{n \geq 1} (n\varepsilon)^{-1} \log E\|\Pi_n\|_1^\varepsilon.$$

Then, from properties of $\| \cdot \|_1$ above, we infer that

$$E\|\Pi_n\|_1^\varepsilon \leq E\|\Pi_n^\varepsilon\|_1 = E\|\Pi_n^\varepsilon\|_1 \leq E\|\Pi_n^{(\varepsilon)}\|_1,$$

(A.1)

where $\Pi_n^{(\varepsilon)} = A_0^\varepsilon \cdots A_{n-1}^\varepsilon \cdot A_{n+1}$. The last inequality follows from the superadditivity of the function $f(x) = x^\varepsilon$. Since the matrices $A_i$ are i.i.d., we have $E\Pi_n^{(\varepsilon)} = (EA^\varepsilon)^n$. Here, we take the $n$th power in terms of matrix multiplication. Hence,

$$\gamma_A \leq \inf_{n \geq 1} (n\varepsilon)^{-1} \log \|(EA^\varepsilon)^n\|_1 = \frac{1}{\varepsilon} \inf_{n \geq 1} \log \|(EA^\varepsilon)^n\|_1^{1/n}. $$

$\blacklozenge$ Springer
From Gelfand’s formula (e.g., [2, (1.3.3)]), for any matrix norm \(\| \cdot \|\) we can write
\[
\lim_{n \to \infty} \| (E A^\varepsilon)^n \|^{1/n} = \rho(E A^\varepsilon).
\]
Taking the norm \(\| \cdot \|_1\), we obtain
\[
\gamma_A \leq \frac{1}{\varepsilon} \inf_{n \geq 1} \log \| (E A^\varepsilon)^n \|_1^{1/n} \leq \frac{1}{\varepsilon} \lim_{n \to \infty} \log \| (E A^\varepsilon)^n \|_1^{1/n} = \frac{1}{\varepsilon} \log \rho(E A^\varepsilon).
\]
Hence, it suffices to show that \(\rho(E A^\varepsilon) < 1\). If \(0 < \varepsilon < \min\{\alpha_1, \ldots, \alpha_d\}\), then from condition (T-4)
\[
E A^\varepsilon_{ii} < 1, \quad i = 1, \ldots, d.
\] (A.2)
Since the spectral radius \(\rho(E A^\varepsilon)\) is the maximal eigenvalue of \(E A^\varepsilon\), stationarity is implied by (A.2).

**Remark A.1** (i) By the equivalence of matrix norms, the argument above works for any norm. In order to observe this, take a certain norm \(\| \cdot \|\) and apply the inequality \(\|A\| \leq c\|A\|_1\). Then, by (A.1) we obtain \(\|\Pi_n\|_1^{\varepsilon} \leq c^\varepsilon \|\Pi_n^{(\varepsilon)}\|_1\). Since \(\lim_{n \to \infty} c^{\varepsilon/n} = 1\) for any constant \(c > 0\), the whole argument holds.
(ii) By the submultiplicativity of \(\| \cdot \|_1\), it is immediate to see that
\[
\gamma_A \leq \frac{1}{\varepsilon} \log \|E A^\varepsilon\|_1.
\]
However, we do not have any control on the norm \(\|E A^\varepsilon\|_1\); in particular, it can be greater than 1 for any \(\varepsilon\). It is essential in our situation that \(\rho(E A^\varepsilon) \leq \|E A^\varepsilon\|_1\) and involving Gelfand’s formula is necessary to obtain the desired bound.

**Appendix B. Version of Breiman’s Lemma**

We provide a slightly modified version of the classical Breiman’s lemma (e.g., [9, Lemma B.5.1]), since it is needed in the proof for (5.29) of Lemma 5.9. In the Breiman’s lemma, we usually assume regular variation for the dominant r.v.’s of the two, which we could not apply in our situation. Instead, we require only an upper estimate of the tail. The price of weakening assumptions is also a weaker result: On behalf of the exact asymptotics of a product, we obtain just an estimate from above. The generalization is rather standard, but we include it for completeness.

**Lemma B.1** Assume that \(X\) and \(Y\) are independent r.v.’s and for some \(\alpha > 0\), the following conditions hold:
\[
\lim_{x \to \infty} x^{\alpha} P(Y > x) < M \quad \text{for a constant } M > 0; \tag{B.1}
\]
\[
E X^{\alpha + \varepsilon} < \infty \quad \text{for some } \varepsilon > 0. \tag{B.2}
\]
Then,
\[
\lim_{x \to \infty} x^\alpha P(XY > x) \leq M \cdot \mathbb{E}X^\alpha.
\]

**Proof** The idea is the same as that in the original proof of Breiman’s lemma; see e.g., [9, Lemma B.5.1]. Let \(P_X\) denote the law of \(X\). Then, for any fixed \(m > 0\) we can write
\[
P(XY > x) = \int_{(0, \infty)} P(Y > x/z) P_X(dz) = \left( \int_{(0, m]} + \int_{(m, \infty), x/z > x_0} + \int_{(m, \infty), x/z \leq x_0} \right) P(Y > x/z) P_X(dz).
\]
By (B.1), there is \(x_0\) such that \(x^\alpha P(Y > x) \leq M\) uniformly in \(x \geq x_0\). For the first integral since \((x/z) \geq x_0\) for \(x \geq mx_0\) and \(z \in (0, m]\), by Fatou’s lemma
\[
\lim_{x \to \infty} x^\alpha \int_{(0, m]} P(Y > x/z) P_X(dz) \leq \int_{(0, m]} z^\alpha \lim_{x \to \infty} (x/z)^\alpha P(Y > x/z) P_X(dz) \leq M \cdot \int_{(0, m]} z^\alpha P_X(dz) \xrightarrow{m \to \infty} M \cdot \mathbb{E}X^\alpha. \quad (B.3)
\]
Since \(x/z > x_0\), the same argument as above is applicable to the second integral:
\[
\lim_{x \to \infty} x^\alpha \int_{(m, \infty), x/z > x_0} P(Y > x/z) P_X(dz) \leq \int_{z > m} z^\alpha \lim_{x \to \infty} (x/z)^\alpha P(Y > x/z) P_X(dz) \leq M \cdot \int_{z > m} z^\alpha P_X(dz) \xrightarrow{m \to \infty} 0. \quad (B.4)
\]
The assumption (B.2) allows us to use Markov’s inequality to estimate the last integral:
\[
\int_{(m, \infty), x/z \leq x_0} P(Y > x/z) P_X(dz) \leq \int_{x/z \leq x_0} P_X(dz) = P(X > x/x_0) \leq (x/x_0)^{-(\alpha + \epsilon)} \mathbb{E}X^{\alpha + \epsilon}
\]
and therefore, the last integral is negligible as \(x \to \infty\) regardless of \(m\). Now, in view of (B.3)-(B.6), letting \(m\) to infinity, we obtain the result. \(\square\)

**References**

1. Alsmeyer, G., Mentemeier, S.: Tail behavior of stationary solutions of random difference equations: the case of regular matrices. J. Differ. Equ. Appl. **18**, 1305–1332 (2012)
2. Belitskii, G.R., Lyubich, YuI: Matrix Norms and Their Applications. Birkhäuser, Basel (1988)
3. Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular Variation. Cambridge University Press, Cambridge (UK) (1987)
4. Bollerslev, T.: Modelling the coherence in short-run nominal exchange rates: a multivariate generalised ARCH model. Rev. Econ. Stat. 72, 498–505 (1990)
5. Bougerol, P., Picard, N.: Strict stationarity of generalized autoregressive processes. Ann. Probab. 20, 1714–1730 (1992)
6. Brandt, A.: The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. Adv. Appl. Probab. 18, 211–220 (1986)
7. Buraczewski, D., Damek, E.: Regular behavior at infinity of stationary measures of stochastic recursion on NA groups. Colloq. Math. 118, 499–523 (2010)
8. Buraczewski, D., Damek, E., Guivarc’h, Y., Hulanicki, A., Urban, R.: Tail-homogeneity of stationary measures for some multidimensional stochastic recursions. Probab. Theory Relat. Fields 145, 385–420 (2009)
9. Buraczewski, D., Damek, E., Mikosch, T.: Stochastic Models with Power-Law Tails. The Equation $X = AX + B$. Springer Int. Pub., Switzerland (2016)
10. Damek, E., Matsui, M., Świa̧tkowski, W.: Componentwise different tail solutions for bivariate stochastic recurrence equations with application to GARCH(1,1) processes. Colloq. Math. 155, 227–254 (2019)
11. Damek, E., Zienkiewicz, J.: Affine stochastic equation with triangular matrices. J. Differ. Equ. Appl. 24, 520–542 (2018)
12. Embrechts, P., Klüppelberg, C., Mikosch, T.: Modelling Extremal Events for Insurance and Finance. Springer, Berlin (1997)
13. Fernández, B., Muriel, N.: Regular variation and related results for the multivariate GARCH(p, q) model with constant conditional correlations. J. Multivariate Anal. 100, 1538–1550 (2009)
14. Gerencsér, L., Michaletky, G., Orlovits, Z.: Stability of block-triangular stationary random matrices. Syst. Control Lett. 57, 620–625 (2008)
15. Goldie, C.M.: Implicit renewal theory and tails of solutions of random equations. Ann. Appl. Probab. 1, 126–166 (1991)
16. Guivarc’h, Y., Le Page, É.: Spectral gap properties for linear random walks and Pareto’s asymptotics for affine stochastic recursions. Ann. Inst. H. Poincaré Probab. Statist. 52, 503–574 (2016)
17. Horváth, R., Boril, ˘S.: GARCH models, tail indexes and error distributions: an empirical investigation. N. Am. J. Econ. Finance 37, 1–15 (2016)
18. Jeantheau, T.: Strong consistency of estimators for multivariate ARCH models. Econ. Theory 14, 70–86 (1998)
19. Kesten, H.: Random difference equations and renewal theory for products of random matrices. Acta Math. 131, 207–248 (1973)
20. Klüppelberg, C., Pergamenchtchikov, S.: The tail of the stationary distribution of a random coefficient AR(q) model. Ann. Appl. Probab. 14, 971–1005 (2004)
21. Krengel, U.: Ergodic Theorems. De Gruyter, Berlin (1985)
22. McNeil, A.J., Frey, R., Embrechts, P.: Quantitative Risk Management: Concepts, Techniques and Tools (Princeton Series in Finance). Princeton Univ. Pr, Princeton (2015)
23. Matsui, M., Mikosch, T.: The extremogram and the cross-extremogram for a bivariate GARCH(1,1) process. Adv. Appl. Probab. 48A, 217–233 (2016)
24. Mikosch, T., Samorodnitsky, G., Tafakori, L.: Fractional moments of solutions to stochastic recurrence equations. J. Appl. Probab. 50, 969–982 (2013)
25. Pedersen, R.S., Wintenberger, O.: On the tail behavior of a class of multivariate conditionally heteroskedastic processes. Extremes 21, 261–284 (2018)
26. Resnick, S.I.: Extreme Values, Regular Variation, and Point Processes. Springer, New York (1987)
27. Resnick, S.I.: Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer, New York (2007)
28. Straumann, D.: Estimation in conditionally heteroscedastic time series models. Lecture Notes in Statistics 181, Springer-Verlag, Berlin (2005)
29. Sun, P., Zhou, C.: Diagnosing the distribution of GARCH innovations. J. Empir. Finance 29, 287–303 (2014)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.