Electron. Commun. Probab. **22** (2017), no. 7, 1–12.
DOI: 10.1214/16-ECP39
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

On recurrence and transience of multivariate near-critical stochastic processes*

Götz Kersting†

Abstract

We obtain complementary recurrence/transience criteria for processes $X = (X_n)_{n\geq 0}$ with values in $\mathbb{R}_+^d$ fulfilling a non-linear equation $X_{n+1} = MX_n + g(X_n) + \xi_{n+1}$. Here M denotes a primitive matrix having Perron-Frobenius eigenvalue 1, and g denotes some function. The conditional expectation and variance of the noise $(\xi_{n+1})_{n\geq 0}$ are such that X obeys a weak form of the Markov property. The results generalize criteria for the 1-dimensional case in [5].

Keywords: Markov property; recurrence; transience; Lyapunov function; martingale.
AMS MSC 2010: 60J10.
Submitted to ECP on May 17, 2016, final version accepted on December 27, 2016.

1 Introduction and main results

For Markov chains with a higher-dimensional state space it is in general difficult to obtain criteria for recurrence or transience which cover a broader class of models. Typically this requires some specific assumptions on the type of model. In this paper we consider discrete time stochastic processes $X = (X_n)_{n\geq 0}$ taking values in the positive orthant $\mathbb{R}_+^d$ (consisting of column vectors) with $d \geq 1$, which obey non-linear equations of the form

$$X_{n+1} = MX_n + g(X_n) + \xi_{n+1} \ , \quad n \in \mathbb{N}_0 \ . \tag{1.1}$$

Here M denotes a $d \times d$ matrix with non-negative entries and $g : \mathbb{R}_+^d \to \mathbb{R}_+^d$ a measurable function. Let us successively discuss our assumptions on M, g and the random fluctuations $(\xi_{n+1})_{n\geq 0}$.

We require that M is a primitive matrix meaning that for a certain power of M all entries are (strictly) positive. Then it is known from Perron-Frobenius theory that M has left and right eigenvectors $\ell = (\ell_1, \dots, \ell_d)$ and $r = (r_1, \dots, r_d)^T$ belonging to some positive eigenvalue and possessing only positive entries. We assume that this eigenvalue is 1:

$$\ell M = \ell \ , \quad Mr = r \ .$$

Further ℓ and r are unique up to scaling factors. As is customary we choose them such that

$$\ell r = 1 \ . \tag{1.2}$$

*Partially supported by the DFG-priority program SPP 1590 "Probabilistic Structures in Evolution".
†Goethe university, Frankfurt am Main. E-mail: kersting@math.uni-frankfurt.de

For the function g we assume that

$$\|g(x)\| = o(\|x\|) \text{ as } \|x\| \to \infty \tag{1.3}$$

with some norm $\| \ \|$ on the Euclidian space $\mathbb{R}^d$.

As to the random fluctuations we demand that X is adapted to a filtration $\mathcal{F} = (\mathcal{F}_n)_{n\geq 0}$ such that

$$\mathbf{E}[\ell\xi_{n+1} \mid \mathcal{F}_n] = 0 \ , \quad \mathbf{E}[(\ell\xi_{n+1})^2 \mid \mathcal{F}_n] = \sigma^2(X_n) \text{ a.s.} \tag{1.4}$$

for some measurable function $\sigma : \mathbb{R}_+^d \to \mathbb{R}_+$ fulfilling

$$\sigma(x) = o(\|x\|) \quad \text{for } \|x\| \to \infty \ . \tag{1.5}$$

In view of applications such as branching processes we might summerize these requirements on the whole as the assumption of *near criticality*. Quite a few models fit into this framework. Here we do not dwell on them but refer to the paper [6] and to the literature cited therein. The assumption (1.4) establishes a weak form of the Markov property. We do not assume that X is a Markov chain but just formulate those assumptions which are required for the martingale considerations in our proofs. Certainly applications of our results will typically concern Markov chains.

The aim of this paper is to establish criteria which allow to decide whether $\|X_n\| \to \infty$ is an event of zero probability or not. Loosely speaking these are criteria for *recurrence* or *transience* of our models. In the univariate case $d = 1$ this question has been discussed in [5]. Ignoring some side conditions the result there was as follows: If for some $\varepsilon > 0$ and for x sufficiently large

$$xg(x) \leq \frac{1-\varepsilon}{2}\sigma^2(x) \ ,$$

then we have recurrence. If on the other hand for some $\varepsilon > 0$ and for x sufficiently large

$$xg(x) \geq \frac{1+\varepsilon}{2}\sigma^2(x) \ ,$$

then there is transience. Heuristically this can be understood as follows: In the first regime it is the noise ξ_{n+1} which dominates the drift $g(X_n)$, while in the second regime it is the other way round. We like to generalize this dichotomy to the multivariate setting.

A possible way of generalization is to suitably convert each of the two conditions to *all* $x \in \mathbb{R}_+^d$ with sufficiently large norm $\|x\|$, see Klebaner [7] and González et al [3]. A relaxation of this approach for special choices of g and σ^2 covering new examples has been obtained by Adam [1]. Yet one can do with weaker assumptions. The intuition behind this assertion is that our processes behave in a sense 1-dimensional. More precisely, if the event $\|X_n\| \to \infty$ occurs, then in view of (1.3) and (1.5) it is the term MX_n, which dominates on the right-hand side of (1.1). Thus one would expect that X_n will escape to ∞ approximately along the ray $\overline{r} = \{\nu r : \nu \geq 0\}$ spanned by the eigenvector r of M. This suggests that the two conditions above are required only in certain vicinities of this ray. (The last assertion of Theorem 2 below confirms this heuristics.)

To formalize these considerations let us introduce some notation. For any $x \in \mathbb{R}^d$ let

$$\hat{x} := r\ell x \ , \quad \check{x} := (I - r\ell)x \ , \quad \text{thus } x = \hat{x} + \check{x} \ ,$$

with the identity matrix I. Note that $\hat{x}$ is the multiple $(\ell x)r$ of the vector r and thus belongs to the ray $\overline{r}$. From (1.2) $r\ell r\ell = r\ell$ respectively $\hat{\hat{x}} = \hat{x}$ meaning that $r\ell$ is a projection matrix. Moreover $\ell\hat{x} = \ell x$ and $\ell\check{x} = 0$. The two conditions $\hat{x} \in \overline{r}$ and $\ell\hat{x} = \ell x$ determine $\hat{x} \in \mathbb{R}^d$ uniquely.

For convenience we require the additional moment condition (which could be relaxed)

$$\exists \delta > 0, c < \infty \, \forall n \in \mathbb{N}_0 \ : \ \mathbf{E}[\|\xi_{n+1}\|^p \mid \mathcal{F}_n] \leq c\sigma^p(X_n) \quad \text{with } p = 2 + \delta \ . \tag{A1}$$

Theorem 1. *Let (A1) be fulfilled and let $\varepsilon > 0$. Assume that for every $b > 0$ there exists some $a > 0$ such that for $x \in \mathbb{R}_+^d$*

$$\|x\| \geq a\ ,\ \|\check{x}\|^2 \leq b\|x\| \cdot \|g(x)\| \quad \Rightarrow \quad \ell x \cdot \ell g(x) \leq \frac{1-\varepsilon}{2}\sigma^2(x)\ . \tag{1.6}$$

Then

$$\mathbf{P}(\|X_n\| \to \infty) = 0\ .$$

In the case $d = 1$ we have $\check{x} = 0$ and $\ell x \cdot \ell g(x) = xg(x)$ such that we are back to the result from [5]. Note that due to (1.3) the above condition $\|\check{x}\|^2 \leq b\|x\| \cdot \|g(x)\|$ applies only to vectors $x \in \mathbb{R}_+^d$ with $\|\check{x}\| = o(\|x\|)$ for $\|x\| \to \infty$. Since also $\check{x} = 0$ for $x \in \overline{r}$, the condition defines a certain vicinity of the ray $\overline{r}$ (depending on g). Outside this region the relation between g and σ^2 stays arbitrary.

For our second result on divergence of $(X_n)_{n\geq 0}$ we first rule out an evident case. We assume

$$\exists u > 0\ :\ \mathbf{P}\big(\, X_n \to X_\infty \text{ with } u \leq \|X_\infty\| < \infty\big) = 0\ . \tag{A2}$$

Moreover we strengthen (1.5) to the assumption

$$\exists \kappa > 1/\delta\ :\ \sigma(x) = O(\|x\| \log^{-\kappa} \|x\|) \text{ for } \|x\| \to \infty\ , \tag{A3}$$

where δ is as in assumption (A1).

Theorem 2. *Let* (A1) *to* (A3) *be fulfilled and let $\varepsilon > 0$. Assume that for every $b > 0$ there exists some $a > 0$ such that for $x \in \mathbb{R}_+^d$*

$$\|x\| \geq a\ ,\ \|\check{x}\| \leq b\sigma(x) \quad \Rightarrow \quad \ell x \cdot \ell g(x) \geq \frac{1+\varepsilon}{2}\sigma^2(x)\ . \tag{1.7}$$

Then there is a real number $v \geq 0$ such that

$$\mathbf{P}\Big(\limsup_n \|X_n\| \leq v \text{ or } \|X_n\| \to \infty\Big) = 1\ .$$

If also $\mathbf{P}(\sup_{n\geq 0} \|X_n\| > c) > 0$ for every $c > 0$, then

$$\mathbf{P}(\|X_n\| \to \infty) > 0 \quad \text{and} \quad \mathbf{P}\Big(\frac{X_n}{\|X_n\|} \to \frac{r}{\|r\|}\ \Big|\ \|X_n\| \to \infty\Big) = 1\ .$$

Again we recover for $d = 1$ the corresponding result from [5]. Due to (A3) it is now the condition $\|\check{x}\| \leq b\sigma(x)$ giving the vicinity of the ray $\overline{r}$, where $g(x)$ and $\sigma^2(x)$ are interrelated.

Remark. Let us comment on the assumptions of Theorem 2.

1. Obviously (A2) is also a necessary requirement in Theorem 2. Typically it is easily checked in concrete examples. For Markov chains with a countable discrete state space $S \subset \mathbb{R}_+^d$ it says that away from zero there are no absorbing states. In the general case there is the following criterion: (A2) holds if $\ell g(x)$ is uniformly bounded away from zero on sets of the form $\{x \in \mathbb{R}_+^d : u \leq \ell x \leq u+1\}$ with $u > 0$ sufficiently large. For the proof of this claim adopt the arguments at the end of section 2 in [5] to the process $(\ell X_n)_{n\geq 0}$.

2. Assumption (A3) cannot be weakened substantially in our general context. This follows from example C, Section 3 in [5]. We note that (A3) is weaker than the corresponding assumption in [5] for the 1-dimensional case.

3. Remarkably, condition (1.7) cannot be relaxed in our general context. It is not enough to require (1.7) just for *some* $b > 0$ as we shall see at the end of this paper by means of a counterexample. It is tempting to conjecture that condition (1.6) cannot be weakened, too. □

So far we have not specified any choice of the norm $\| \ \|$ on $\mathbb{R}^d$. This was not necessary so far, since as is well-known all norms on a finite dimensional Euclidean space are equivalent, and one easily convinces oneself that all our conditions or statements involving norms are preserved if one passes to an equivalent norm. Thus, in examples one may work with the most convenient one, e.g. the l_1- or l_2-norm. For our proofs these norms are not appropriate. We shall utilize a norm specificially suited for our purposes. This norm is introduced in section 2. The proofs of the theorems are then presented in section 3 and 4. They use ideas from [5] and [8] and are based on the construction of Lyapunov functions of the form

$$l_{\alpha,\beta,\gamma,j}(x) = (1 + \gamma x_j/\ell x)\frac{\|\check{x}\|^2}{(\ell x)^2}(\log \ell x)^{-\beta-1} + \alpha(\log \ell x)^{-\beta}$$

with $x = (x_1, \dots, x_d)^T \in \mathbb{R}_+^d$, $1 \le j \le d$, $\alpha > 0$, $\gamma \ge 0$ and either $\beta = -1$ or $\beta > 0$. Section 5 contains the counterexample.

For notational convenience we use the symbol c for a positive constant which may change its value from line to line.

2 A useful norm

Let us briefly put together the facts on matrices which we are going to use. Recall that M is a primitive matrix with Perron-Frobenius eigenvalue 1 and corresponding left and right eigenvectors ℓ and r. Then as is well-known from Perron-Frobenius theory (see [9])

$$\max\{|\eta| : \eta \text{ is an eigenvalue of } M - r\ell\} < 1 \ .$$

This maximum is called the *spectral radius* of the matrix $M - r\ell$. It follows from matrix theory (see [4], Lemma 5.6.10) that one can construct a matrix norm $|\!|\!| \ |\!|\!|$ on the space of all $d \times d$ matrices such that

$$\rho := |\!|\!| M - r\ell |\!|\!| < 1 \ .$$

From this matrix norm we obtain (see [4], Theorem 5.7.13) a functional $\| \ \|$ on $\mathbb{R}^d$ via

$$\|x\| := |\!|\!| C_x |\!|\!| \ , \quad x \in \mathbb{R}^d \ ,$$

where C_x denotes the $d \times d$ matrix having all columns equal to x. $\| \ \|$ is a norm, since the properties of norms transfer from $|\!|\!| \ |\!|\!|$ directly to $\| \ \|$. This is the norm we are going to work with in the sequel. It has the property

$$\|Ax\| \le |\!|\!| A |\!|\!| \cdot \|x\| \tag{2.1}$$

for $x \in \mathbb{R}^d$ and any $d \times d$ matrix A. Indeed $C_{Ax} = AC_x$ and the property $|\!|\!| C_{Ax} |\!|\!| \le |\!|\!| A |\!|\!| \cdot |\!|\!| C_x |\!|\!|$ of matrix norms gives the claim. In particular

$$\|(M - r\ell)x\| \le \rho\|x\| \ . \tag{2.2}$$

Thus $M - r\ell$ induces a contraction in the norm $\| \ \|$.

By equivalence of norms we may change from $\| \ \|$ to any other norm. In particular there is a constant $\lambda < \infty$ such that

$$\|\check{x}\| \le \lambda \, \ell x \quad \text{for all } x \in \mathbb{R}_+^d \ . \tag{2.3}$$

To see this observe that from the inequality (2.1) we have $\|\check{x}\| \le \gamma\|x\|$ with $\gamma = |\!|\!| I - r\ell |\!|\!|$. Also $\|x\|' := \ell_1|x_1| + \cdots + \ell_d|x_d|$ defines a norm on $\mathbb{R}^d$, since $\ell_i > 0$ for all $i = 1, \dots, d$. Thus by equivalence of norms we arrive at (2.3).

In order to apply these results to our process $(X_n)_{n\geq 0}$ note that we have $(I-r\ell)M = M - r\ell = M(I-r\ell)$ and $\ell\check{X}_n = 0$, thus

$$\begin{aligned}\check{X}_{n+1} &= (I-r\ell)(MX_n + g(X_n) + \xi_{n+1}) \\ &= (M-r\ell)\check{X}_n + (I-r\ell)(g(X_n)+\xi_{n+1}) \ .\end{aligned}$$

From (2.2) to (2.3) it follows that

$$\|\check{X}_{n+1}\| \leq \rho\|\check{X}_n\| + c\,\ell g(X_n) + c\,\|\xi_{n+1}\| \tag{2.4}$$

for some $c<\infty$. (Here we need that $g(x)$ has only non-negative components.) Further observe that for any $\mu>0$ and $a,b\geq 0$ we have

$$(a+b)^2 \leq (1+\mu)a^2 + (1+\mu^{-1})b^2 \ . \tag{2.5}$$

Applying this estimate twice to the right-hand side of (2.4) we obtain for any $\mu>0$

$$\|\check{X}_{n+1}\|^2 \leq (1+\mu)\rho^2\|\check{X}_n\|^2 + c\,(\ell g(X_n))^2 + c\,\|\xi_{n+1}\|^2 \tag{2.6}$$

with a suitable $c<\infty$.

3 Proof of Theorem 1

First observe that if we replace X_n by $\overline{X}_n := X_n + r$ for all $n\geq 0$ then equations (1.1) and (1.4) as well as assumption (A1) still hold, if $g(x)$ and $\sigma^2(x)$ are replaced by $\overline{g}(x) := g(x-r)$ and $\overline{\sigma}^2(x) := \sigma^2(x-r)$. Note that the assumptions (1.3) and (1.5) are not affected if g and σ^2 are substituted by $\overline{g}$ and $\overline{\sigma}^2$, and the same holds true for the conditions formulated in Theorem 1 if one replaces ε by $\varepsilon/2$. Thus without loss of generality we may assume $\ell X_n \geq 1$ for all $n\geq 0$ throughout the proof. Then for any $\alpha>0$

$$L_n := \frac{\|\check{X}_n\|^2}{(\ell X_n)^2} + \alpha\log \ell X_n \ , \quad n\in\mathbb{N}_0 \ ,$$

is a sequence of non-negative random variables. We show that for large α it possesses a supermartingale property. The proof uses the following estimate, where $I(A)$ denotes the indicator variable of an event A.

Lemma 1. *For all $t>0$, $h>-t$ and $\eta>0$*

$$\log(t+h) \leq \log t + \frac{h}{t} - \frac{1}{2(1+\eta)}\frac{h^2}{t^2} I(h\leq \eta t) \ .$$

Proof. See formula (2) in [5]. □

Lemma 2. *If α is chosen large enough, then there is a number $s>0$ such that*

$$\ell X_n \geq s \quad \Rightarrow \quad \mathbf{E}[L_{n+1} \mid \mathcal{F}_n] \leq L_n \text{ a.s.}$$

Proof. Since $\ell M = \ell$ we have the equation

$$\ell X_{n+1} = \ell X_n + \ell g(X_n) + \ell\xi_{n+1} \ . \tag{3.1}$$

Thus $\ell\xi_{n+1} \geq -\mu\,\ell X_n$ implies $\ell X_{n+1} \geq (1-\mu)\,\ell X_n$. Together with (2.6) and (2.3) this entails

$$\frac{\|\check{X}_{n+1}\|^2}{(\ell X_{n+1})^2} \leq \frac{(1+\mu)\rho^2\|\check{X}_n\|^2 + c\,(\ell g(X_n))^2 + c\,\|\xi_{n+1}\|^2}{(1-\mu)^2(\ell X_n)^2} + \lambda^2 I(\ell\xi_{n+1} < -\mu\,\ell X_n) \tag{3.2}$$

for some sufficiently large $c < \infty$. Now $\rho < 1$, thus, if μ is sufficiently close to 0,

$$\frac{\|\check{X}_{n+1}\|^2}{(\ell X_{n+1})^2} \le (1-\mu)\frac{\|\check{X}_n\|^2}{(\ell X_n)^2} + c\frac{(\ell g(X_n))^2 + \|\xi_{n+1}\|^2}{(\ell X_n)^2} + \lambda^2 \frac{(\ell \xi_{n+1})^2}{\mu^2(\ell X_n)^2} \ .$$

In view of (A1), if we further enlarge c,

$$\mathbf{E}\Big[\frac{\|\check{X}_{n+1}\|^2}{(\ell X_{n+1})^2} \mid \mathcal{F}_n\Big] \le (1-\mu)\frac{\|\check{X}_n\|^2}{(\ell X_n)^2} + c\frac{(\ell g(X_n))^2 + \sigma^2(X_n)}{(\ell X_n)^2} \text{ a.s.} \tag{3.3}$$

Next from (3.1) and Lemma 1 (with $t = \ell X_n + \ell g(X_n)$ and $h = \ell \xi_{n+1}$) for $\eta > 0$

$$\log \ell X_{n+1} \le \log(\ell X_n + \ell g(X_n)) \\ + \frac{\ell \xi_{n+1}}{\ell X_n + \ell g(X_n)} - \frac{(\ell \xi_{n+1})^2}{2(1+\eta)(\ell X_n + \ell g(X_n))^2} I\big(\ell \xi_{n+1} \le \eta(\ell X_n + \ell g(X_n))\big).$$

Using the inequality $\log(1+x) \le x$ we get

$$\log \ell X_{n+1} \le \log \ell X_n + \frac{\ell g(X_n)}{\ell X_n} \\ + \frac{\ell \xi_{n+1}}{\ell X_n + \ell g(X_n)} - \frac{(\ell \xi_{n+1})^2}{2(1+\eta)(\ell X_n + \ell g(X_n))^2} + \frac{(\ell \xi_{n+1})^2}{(\ell X_n)^2} I(\ell \xi_{n+1} > \eta\, \ell X_n).$$

By means of (1.3), (1.4), (A1) and the Markov inequality and choosing η sufficiently small it follows for ℓX_n sufficiently large

$$\mathbf{E}[\log \ell X_{n+1} \mid \mathcal{F}_n] \le \log \ell X_n + \frac{\ell g(X_n)}{\ell X_n} - \frac{(1-\varepsilon/3)\sigma^2(X_n)}{2(\ell X_n)^2} + c\frac{\sigma^p(X_n)}{(\ell X_n)^p} \text{ a.s.}$$

with some $c < \infty$. Because of (1.5) there is a number $s > 0$ such that for $\ell X_n \ge s$

$$\mathbf{E}[\log \ell X_{n+1} \mid \mathcal{F}_n] \le \log \ell X_n + \frac{\ell g(X_n)}{\ell X_n} - \frac{(1-\varepsilon/2)\sigma^2(X_n)}{2(\ell X_n)^2} \text{ a.s.} \tag{3.4}$$

Now combining (3.3) and (3.4) and using (1.3) we get

$$\mathbf{E}[L_{n+1} \mid \mathcal{F}_n] \le L_n - \mu\frac{\|\check{X}_n\|^2}{(\ell X_n)^2} + (\alpha + c)\frac{\ell g(X_n)}{\ell X_n} - \Big(\frac{1-\varepsilon/2}{2}\alpha - c\Big)\frac{\sigma^2(X_n)}{(\ell X_n)^2} \text{ a.s.}$$

for $\ell X_n \ge s$ and s sufficiently large. If we let $\alpha \ge 6c/\varepsilon - c$ we arrive at

$$\mathbf{E}[L_{n+1} \mid \mathcal{F}_n] \le L_n - \mu\frac{\|\check{X}_n\|^2}{(\ell X_n)^2} + (\alpha + c)\Big(\frac{\ell g(X_n)}{\ell X_n} - \frac{1-\varepsilon}{2}\frac{\sigma^2(X_n)}{(\ell X_n)^2}\Big) \text{ a.s.}$$

for $\ell X_n \ge s$. We are now ready for the conclusion:

If $(\alpha + c)\ell g(X_n) \cdot \ell X_n \le \mu \|\check{X}_n\|^2$, then obviously $\mathbf{E}[L_{n+1} \mid \mathcal{F}_n] \le L_n$ a.s. for $\ell X_n \ge s$.

If on the other hand $\mu\|\check{X}_n\|^2 \le (\alpha + c)\, \ell g(X_n) \cdot \ell X_n$ then by equivalence of norms there is a $b < \infty$ such that $\|\check{X}_n\|^2 \le b\, \|g(X_n)\| \cdot \|X_n\|$. Now the assumption of Theorem 1 comes into play, and again $\mathbf{E}[L_{n+1} \mid \mathcal{F}_n] \le L_n$ a.s., if only ℓX_n is large enough. Thus the claim of the lemma follows. □

We complete the proof of Theorem 1 now as in [5]. Suppose that the event $\|X_n\| \to \infty$ has positive probability. Then the same holds for the event $L_n \to \infty$, and there is natural number N such that $\mathbf{P}(E) > 0$ for the event

$$E = \{\inf_{n \ge N} L_n \ge s, L_n \to \infty\} \ .$$

Define the stopping time

$$T_N := \min\{n \ge N : L_n < s\} \ .$$

In view of Lemma 2 the process $(L_{n \wedge T_N})_{n \ge N}$ is a supermartingale. It is non-negative and thus a.s. convergent. However, on the event E we have $T_N = \infty$ and $L_n \to \infty$ and consequently $L_{n \wedge T_N} \to \infty$. This contradicts the assumption $\mathbf{P}(E) > 0$, and the proof is finished.

4 Proof of Theorem 2

Here we may replace X_n by $X_n + 3r$. Therefore without loss of generality we assume $\ell X_n \geq 3$ for all $n \in \mathbb{N}_0$. Now we consider the processes $L = L^{\alpha,\beta,\gamma,j}$ given by

$$L_n = L_n^{\alpha,\beta,\gamma,j} := \frac{(1+\gamma X_{n,j}/\ell X_n)\|\check{X}_n\|^2}{(\ell X_n)^2(\log \ell X_n)^{\beta+1}} + \alpha(\log \ell X_n)^{-\beta} \ , \ n \in \mathbb{N}_0 \ ,$$

with the jth component $X_{n,j}$ of X_n, $1 \leq j \leq d$, and with $\alpha, \beta > 0$ and $\gamma \geq 0$. In view of the Jensen inequality we may without loss of generality restrict ourselves to the case $2 < p \leq 3$, in which the following estimate is valid.

Lemma 3. *Let $\beta > 0$ and $2 < p \leq 3$. Set $f(t) := (\log t)^{-\beta}$. Then there is a constant $c < \infty$ such that for all $t \geq 3$ and $h > 3 - t$*

$$f(t+h) \leq f(t) + f'(t)h + \frac{1}{2}f''(t)h^2 + \frac{c|h|^p}{(\log t)^{\beta+1}t^p} + I(h \leq -t/2) \ .$$

Proof. See formula (6) in [5]. □

Lemma 4. *Let $0 < \beta < \kappa\delta - 1$ and $\gamma \geq 0$ such that $(1+\gamma/\ell_j)\rho^2 < 1$. Then, if α is sufficiently large, there is a real number $s > 0$ such that*

$$\ell X_n \geq s \quad \Rightarrow \quad \mathbf{E}[L_{n+1}^{\alpha,\beta,\gamma,j} \mid \mathcal{F}_n] + \frac{\sigma(X_n)^p}{(\ell X_n)^p} \leq L_n^{\alpha,\beta,\gamma,j} \text{ a.s.}$$

Proof. We proceed similarly as in the proof of Lemma 2. Here instead of (3.2) we have the estimate

$$\begin{aligned}
&(1+\gamma X_{n+1,j}/\ell X_{n+1})\frac{\|\check{X}_{n+1}\|^2}{(\ell X_{n+1})^2(\log \ell X_{n+1})^{\beta+1}} \\
&\quad \leq (1+\gamma/\ell_j)\Big((1+\gamma X_{n,j}/\ell X_n)\frac{(1+\mu)\rho^2\|\check{X}_n\|^2 + c\,(\ell g(X_n))^2 + c\,\|\xi_{n+1}\|^2}{(1-\mu)^2(\ell X_n)^2(\log \ell X_n + \log(1-\mu))^{1+\beta}} \\
&\qquad + \lambda^2 I(\ell\xi_{n+1} < -\mu\,\ell X_n)\Big)
\end{aligned}$$

By assumption on γ and for $\mu > 0$ sufficiently small this implies

$$\begin{aligned}
&\mathbf{E}\Big[\frac{(1+\gamma X_{n+1,j}/\ell X_{n+1})\|\check{X}_{n+1}\|^2}{(\ell X_{n+1})^2(\log \ell X_{n+1})^{\beta+1}} \mid \mathcal{F}_n\Big] \\
&\quad \leq (1-\mu)\frac{(1+\gamma X_{n,j}/\ell X_n)\|\check{X}_n\|^2}{(\ell X_n)^2(\log \ell X_n)^{\beta+1}} + c\frac{(\ell g(X_n))^2 + \sigma^2(X_n)}{(\ell X_n)^2(\log \ell X_n)^{\beta+1}} + c\frac{\sigma^p(X_n)}{(\ell X_n)^p} \text{ a.s.}
\end{aligned} \tag{4.1}$$

with some $c < \infty$.

Next from Lemma 3 with $t = \ell X_n$ and $h = \ell g(X_n) + \ell\xi_{n+1}$, from (2.5) and (3.1) and from $\ell g(X_n) \geq 0$

$$\begin{aligned}
f(\ell X_{n+1}) &\leq f(\ell X_n) + f'(\ell X_n)(\ell g(X_n) + \ell\xi_{n+1}) \\
&\quad + \frac{1}{2}f''(\ell X_n)((1+\mu)(\ell\xi_{n+1})^2 + (1+\mu^{-1})(\ell g(X_n))^2) \\
&\quad + c\frac{(\ell g(X_n))^p + |\ell\xi_{n+1}|^p}{(\log \ell X_n)^{\beta+1}(\ell X_n)^p} + I(\ell\xi_{n+1} \leq -\ell X_n/2)
\end{aligned}$$

for a suitable $c > 0$. Since $f''(t) \sim \beta(\log t)^{-\beta-1}t^{-2}$ for $t \to \infty$,

$$\begin{aligned}
\mathbf{E}[f(\ell X_{n+1}) \mid \mathcal{F}_n] &\leq f(\ell X_n) - \beta\frac{\ell g(X_n)}{(\log \ell X_n)^{\beta+1}\ell X_n} + \frac{\beta}{2}\frac{(1+2\mu)\sigma^2(X_n) + c(\ell g(X_n))^2}{(\log \ell X_n)^{\beta+1}(\ell X_n)^2} \\
&\quad + c\frac{(\ell g(X_n))^p + \sigma^p(X_n)}{(\log \ell X_n)^{\beta+1}(\ell X_n)^p} + c\frac{\sigma^p(X_n)}{(\ell X_n)^p} \text{ a.s.}
\end{aligned}$$

for ℓX_n sufficiently large. Combining this estimate with (4.1) and rearranging terms we get

$$\begin{aligned}\mathbf{E}[L_{n+1} \mid \mathcal{F}_n] + \frac{\sigma^p(X_n)}{(\ell X_n)^p} \le L_n - \mu \frac{\|\check{X}_n\|^2}{(\ell X_n)^2(\log \ell X_n)^{\beta+1}} + ((\alpha+1)c+1)\frac{\sigma^p(X_n)}{(\ell X_n)^p} \\ - \Big(\alpha\beta - \big(c + \frac{1}{2}\alpha\beta c\big)\frac{\ell g(X_n)}{\ell X_n} + \alpha c \frac{(\ell g(X_n))^{1+\delta}}{(\ell X_n)^{1+\delta}}\Big)\frac{\ell g(X_n)}{(\log \ell X_n)^{\beta+1}\ell X_n} \\ + \Big(c + \alpha\beta\frac{1+2\mu}{2} + \alpha c \frac{\sigma(X_n)^\delta}{(\ell X_n)^\delta}\Big)\frac{\sigma^2(X_n)}{(\log \ell X_n)^{\beta+1}(\ell X_n)^2}\end{aligned}$$

Now choose α so large that $c + \alpha\beta(1+2\mu)/2 < \alpha\beta(1+3\mu)/2$. Then, after another rearrangement of terms, we obtain in view of (1.3) and (1.5) for ℓX_n sufficiently large

$$\begin{aligned}\mathbf{E}[L_{n+1} \mid \mathcal{F}_n] + \frac{\sigma^p(X_n)}{(\ell X_n)^p} \le L_n - \mu \frac{\|\check{X}_n\|^2}{(\ell X_n)^2(\log \ell X_n)^{\beta+1}} + ((\alpha+1)c+1)\frac{\sigma^p(X_n)}{(\ell X_n)^p} \\ - \frac{\alpha\beta}{(\log \ell X_n)^{\beta+1}}\Big((1-\mu)\frac{\ell g(X_n)}{\ell X_n} - \frac{1+3\mu}{2}\frac{\sigma^2(X_n)}{(\ell X_n)^2}\Big) \text{ a.s.}\end{aligned}$$

From (A3) we have for $0 < \beta < \kappa\delta - 1$

$$\frac{\sigma^p(x)}{(\ell x)^p} = O\big(\frac{\sigma^2(x)}{(\ell x)^2(\log x)^{\kappa\delta}}\big) = o\big(\frac{\sigma^2(x)}{(\ell x)^2(\log x)^{\beta+1}}\big) \quad \text{for } \|x\| \to \infty .$$

Therefore for $0 < \mu < 1$ sufficiently small

$$\begin{aligned}\mathbf{E}[L_{n+1} \mid \mathcal{F}_n] + \frac{\sigma^p(X_n)}{(\ell X_n)^p} \le L_n - \mu \frac{\|\check{X}_n\|^2}{(\ell X_n)^2(\log \ell X_n)^{\beta+1}} \\ - \frac{\alpha\beta(1-\mu)}{(\log \ell X_n)^{\beta+1}}\Big(\frac{\ell g(X_n)}{\ell X_n} - \frac{1+\varepsilon}{2}\frac{\sigma^2(X_n)}{(\ell X_n)^2}\Big) \text{ a.s.}\end{aligned}$$

if ℓX_n is large enough. We come to the conclusion:

If $\|\check{X}_n\| \ge b\sigma(X_n)$ with some sufficiently large b, then the last estimate implies the claim $\mathbf{E}[L_{n+1} \mid \mathcal{F}_n] + \sigma^p(X_n)/(\ell X_n)^p \le L_n$. If on the other hand $\|\check{X}_n\| \le b\sigma(X_n)$, then the assumption of Theorem 2 applies and again the claim follows. □

For the proof of Theorem 2 we again construct a supermartingale, this time from $L = L^{\alpha,\beta,\gamma,j}$. Observe that for some $s > 0$ and for $m, m' > 0$ and $t > s$ fulfilling

$$\alpha(\log s)^{-\beta} \ge m > m' \ge (1+\gamma/\ell_j)\lambda^2(\log t)^{-\beta-1} + \alpha(\log t)^{-\beta}$$

with $\lambda > 0$ from formula (2.3) we have

$$\begin{aligned} L_n \le m \quad &\Rightarrow \quad \ell X_n \ge s \ , \\ L_n \ge m' \quad &\Rightarrow \quad \ell X_n \le t \ . \end{aligned}$$

If we choose α, β, γ and s as demanded in Lemma 4, then $(m \wedge L_n)_{n\ge 0}$ becomes a non-negative supermartingal, which thus is a.s. convergent. Then up to a null-event there arise three possibilities. Either $L_n \to 0$, then $\ell X_n \to \infty$. Or $\liminf_n L_n \ge m$, then $\limsup_n \ell X_n \le t$. Or else L_n has a limit $0 < L_\infty < m$, then $s \le \liminf_n \ell X_n < \infty$.

In order to transfer these alternatives to the process $(X_n)_{n\ge 0}$ we choose different $\beta_1, \beta_2 > 0$ and a $\gamma > 0$ fitting the assumptions of Lemma 4. We consider the processes

$$L^0 := L^{\alpha,\beta_1,0,1} \ , \ L^1 := L^{\alpha,\beta_1,\gamma,1} \ , \ \dots \ , \ L^d := L^{\alpha,\beta_1,\gamma,d} \ , \ L^{d+1} := L^{\alpha,\beta_2,0,1}$$

and for some $s, t, m > 0$ the events

$$E := \{\ell X_n \to \infty\} \ , \ E' := \ \{\limsup_n \ell X_n \le t\} \ ,$$

$$E'' := \bigcap_{i=0}^{d+1} \{s \le \liminf_n \ell X_n < \infty, L_n^i \to L_\infty^i \text{ with } 0 < L_\infty^i < m\} \ .$$

We let α, s, t large and m small enough such that the above conclusion for $L = (L_n)_{n\ge 0}$ applies simultaneously to all processes $L^0, \dots, L^{d+1}$. Then $\mathbf{P}(E \cup E' \cup E'') = 1$.

Let us show that $\mathbf{P}(E'') = 0$ for s sufficiently large. We have

$$L_n^0 = L_n^{d+1} (\log \ell X_n)^{\beta_2 - \beta_1} \ .$$

Thus the sequence ℓX_n is convergent on E'' with $s \le \lim_n \ell X_n < \infty$. This means that the random variables $\hat{X}_n = r\ell X_n$ converge on E''. Next from the definition of L^0 it follows that the sequence $\|\check{X}_n\|$ converges on the event E'' with some limit Z. If $Z = 0$ then $\check{X}_n \to 0$, and we obtain that $X_n = \hat{X}_n + \check{X}_n$ is convergent on E''. If on the other hand $Z > 0$, then we see from the convergence of $L_n^1, \dots, L_n^d$ that the components $X_{n,1}, \dots, X_{n,d}$ all converge on E''. Again we conclude that X_n is a convergent sequence on the event E''. Let X_∞ be the limit.

Now, given $u > 0$, if we choose s sufficiently large then from $s \le \lim_n \ell X_n < \infty$ on E'' we obtain $u \le \|X_\infty\| < \infty$ by equivalence of norms. Therefore assumption (A2) may be applied and we obtain $\mathbf{P}(E'') = 0$ and consequently $\mathbf{P}(E \cup E') = 1$. By equivalence of norms this translates into the first assertion of Theorem 2.

For the second assertion we switch back to the supermartingale $m \wedge L$ with $\gamma = 0$. Let $c > t$ be such that

$$\alpha(\log c)^{-\beta} + \lambda^2 (\log c)^{-\beta-1} < \alpha(\log t)^{-\beta} \ .$$

From the assumption of this assertion and by equivalence of norms there is a natural number N such that $\mathbf{P}(\ell X_N > c) > 0$. It follows

$$\mathbf{E}[m \wedge L_N; \ell X_N > c] < \alpha(\log t)^{-\beta} \mathbf{P}(\ell X_N > c) \ .$$

From the supermartingale property of $m \wedge L$ and Fatou's Lemma

$$\mathbf{E}[\lim_n m \wedge L_n; \ell X_N > c] < \alpha(\log t)^{-\beta} \mathbf{P}(\ell X_N > c) \ .$$

If now $\mathbf{P}(E') = 1$, then $\lim_n m \wedge L_n \ge \alpha(\log t)^{-\beta}$ a.s. which contradicts the last inequality. Therefore it follows $\mathbf{P}(E) > 0$. This gives the second assertion.

For the last assertion we first show that

$$\|\xi_{n+1}\| = o(\|X_n\|) \text{ a.s. on the event } \|X_n\| \to \infty \ . \tag{4.2}$$

Define

$$L_n' := L_n + \sum_{k=0}^{n-1} \frac{\sigma^p(X_k)}{(\ell X_k)^p}$$

and for a natural number N

$$T_N := \min\{n \ge N : \ell X_n < s\} \ .$$

If again α, β, γ and s are chosen in accordance with Lemma 4 then $(L'_{n\wedge T_N})_{n\ge 0}$ is a non-negative supermartingal and thus a.s. convergent. It follows

$$\sum_{k=0}^{\infty} \frac{\sigma^p(X_k)}{(\ell X_k)^p} < \infty \text{ a.s. on the event } T_N = \infty \ .$$

Now in view of the first assertion of this theorem $\{T_N = \infty\} \uparrow \{\ell X_n \to \infty\}$ for $N \to \infty$, if only s is sufficiently large. Therefore

$$\sum_{k=0}^{\infty} \frac{\sigma^p(X_k)}{(\ell X_k)^p} < \infty \text{ a.s. on the event } \|X_n\| \to \infty \ .$$

Because of (A1) and the Markov inequality this entails for every $\eta > 0$

$$\sum_{k=0}^{\infty} \mathbf{P}(\|\xi_{k+1}\| > \eta \, \ell X_k \mid \mathcal{F}_k) < \infty \text{ a.s. on the event } \|X_n\| \to \infty \ ,$$

and the martingale version of the Borel-Cantelli Lemma (see [2], Theorem 5.3.2) implies (1.5).

Now from (2.4), (1.3) and (4.2) we obtain that

$$\|\check{X}_{n+1}\| \le \rho \|\check{X}_n\| + Y_n \quad \text{ with } Y_n = o(\|X_n\|) \text{ a.s. on } \|X_n\| \to \infty \ .$$

By induction

$$\|\check{X}_{n+1}\| \le \|\check{X}_0\| + \sum_{k=0}^{n} \rho^{n-k} Y_k \ .$$

Since $\rho < 1$ it follows

$$\|\check{X}_n\| = o(\|X_n\|) \text{ a.s. on the event } \|X_n\| \to \infty \ .$$

On the other hand $\hat{X}_n / \|\hat{X}_n\| = r/\|r\|$. This yields the last claim of Theorem 2.

5 A counterexample

We discuss an example in dimension $d = 2$, which can be easily lifted to higher dimensions. In this section we use the l_1-norm $\|x\| := |x_1| + |x_2|$ for $x = (x_1, x_2)^T$. Let

$$M = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \ , \quad r = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \ , \quad \ell = \frac{1}{2}(1,1) \ .$$

Let $\overline{g}(t)$, $\overline{\sigma}(t)$, $t \ge 0$, be two functions such that $\overline{\sigma}$ is differentiable and for $t \ge 0$

$$\lim_{t\to\infty} \overline{\sigma}'(t) = 0 \quad \text{ and } \quad \forall t > 0 : 0 < \overline{g}(t) \le \overline{\sigma}(t) \le t/2 \ , \ |\overline{\sigma}'(t)| < \frac{1}{2} \ .$$

(For definiteness make $(0,0)^T$ an absorbing state.) Define for $x \in \mathbb{R}_+^2$

$$\sigma(x) := \overline{\sigma}(\ell x) \ , \ g(x) := \begin{cases} \overline{g}(\ell x) r & \text{if } \|\check{x}\| \le \sigma(x) \\ (0,0)^T & \text{else} \ . \end{cases}$$

Let χ_n, ζ_n, $n \ge 1$, be independent, $\mathbb{R}^2$-valued random variables with

$$\mathbf{P}(\chi_n = (1,1)^T) = \mathbf{P}(\chi_n = -(1,1)^T) = \mathbf{P}(\zeta_n = (1,-1)^T) = \mathbf{P}(\zeta_n = (-1,1)^T) = \frac{1}{2} \ .$$

Define the Markov chain $X = (X_n)_{n\ge0}$ inductively by $X_0 = r$,

$$\xi_{n+1} := \sigma(X_n)\chi_{n+1} + \sigma(X_n)\zeta_{n+1} I(\|\check{X}_n\| \le \sigma(X_n))$$

and (1.1). Note that M is the orthogonal projection on the subspace spanned by r. This together with the condition $\sigma(x) = \overline{\sigma}(\ell x) \le \ell x/2$ guarantees that the process X never exits from the quadrant $\mathbb{R}_+^2$. The conditions assumptions (1.3), (1.4), (1.5) and (A1) are fulfilled, and the same is true for (A2) and (A3) under mild conditions on $\overline{g}$ and $\overline{\sigma}$. However, due to the definition of $g(x)$, the condition (1.7) will never be satisfied for $b > 1$, no matter how $\overline{g}$ and $\overline{\sigma}$ are chosen. We shall see that indeed the conclusion of Theorem 2 fails, even though (1.7) can be achieved for $b \le 1$ (but not all b). The reason is that the process X again and again leaves the region defined by the inequality $\|\check{x}\| \le \sigma(x)$.

To prove this claim notice that from our assumptions for $t > 0$

$$\overline{\sigma}(t + \overline{g}(t) \pm \overline{\sigma}(t)) < \overline{\sigma}(t) + \tfrac{1}{2}(\overline{g}(t) + \overline{\sigma}(t)) \le 2\overline{\sigma}(t)\ .$$

If now $\check{X}_n = 0$ then from the definitions

$$\ell X_{n+1} = \ell X_n + \overline{g}(\ell X_n) + \overline{\sigma}(\ell X_n)\ell\chi_{n+1} \quad \text{and} \quad \|\check{X}_{n+1}\| = \overline{\sigma}(\ell X_n)\|\zeta_{n+1}\| = 2\sigma(X_n)\ .$$

From the previous inequality it follows $\sigma(X_{n+1}) < 2\sigma(X_n)$. Thus $\sigma(X_{n+1}) < \|\check{X}_{n+1}\|$ and consequently from our definitions $\check{X}_{n+2} = 0$.

Therefore, since we started with $\check{X}_0 = 0$, we have $\check{X}_{2n} = 0$ and $\|\check{X}_{2n+1}\| > \sigma(X_{2n+1})$ for all $n \in \mathbb{N}_0$. Then $\hat{X}_{2n}$, $n \ge 0$, or (what amounts to the same thing) $\overline{X}_n := \ell X_{2n}$, $n \ge 0$, is a Markov chain. Inserting our definitions we get

$$\overline{X}_{n+1} = \overline{X}_n + \overline{g}(\overline{X}_n) + \overline{\xi}_{n+1} \quad \text{with} \quad \overline{\xi}_{n+1} := \overline{\sigma}(\overline{X}_n)\ell\chi_{2n+1} + \overline{\sigma}(\ell X_{2n+1})\ell\chi_{2n+2}\ .$$

Letting $\overline{\mathcal{F}}_n := \mathcal{F}_{2n}$

$$\mathbf{E}[\overline{\xi}_{n+1} \mid \overline{\mathcal{F}}_n] = 0\ ,\ \mathbf{E}[\overline{\xi}^2_{n+1} \mid \overline{\mathcal{F}}_n] = \tau^2(\overline{X}_n)$$

with

$$\begin{aligned}\tau^2(t) &= \overline{\sigma}^2(t) + \mathbf{E}[\overline{\sigma}^2(\ell X_1) \mid \ell X_0 = t, \check{X}_0 = 0] \\ &= \overline{\sigma}^2(t) + \frac{1}{2}\overline{\sigma}^2(t + \overline{g}(t) + \overline{\sigma}(t)) + \frac{1}{2}\overline{\sigma}^2(t + \overline{g}(t) - \overline{\sigma}(t))\end{aligned}$$

From our assumptions

$$\tau^2(t) \sim 2\overline{\sigma}^2(t) \quad \text{for } t \to \infty\ .$$

Thus we are ready to apply our theorems (with $d = 1$) to the process $\overline{X} = (\overline{X}_n)_{n\ge0}$ and see that we have recurrence if $t\overline{g}(t) \le (1-\varepsilon)\overline{\sigma}^2(t)$ for large t. Note the the factor $1/2$ dropped out on the right-hand side. Thus there are cases, where the statement is false that there is transience for $t\overline{g}(t) \ge (1+\varepsilon)\overline{\sigma}^2(t)/2$. This shows that the assertion of Theorem 2 cannot be applied to the process X.

References

[1] E. Adam (2015). Criterion for unlimited growth of critical multidimensional stochastic models. arXiv:1502.04046 [math.PR]. To appear in *Adv. Appl. Probab.*

[2] R. Durrett. Probability: Theory and examples. Cambridge University Press, 4th edition, 2010. MR-2722836

[3] M. González, R. Martínez, and M. Mota (2005). On the unlimited growth of a class of homogeneous multityp Markov chains. *Bernoulli* 11: 559–570. MR-2147775

[4] R. A. Horn and C. R. Johnson. Matrix Analysis. Cambridge University Press, 1990. MR-1084815

[5] G. Kersting (1986). On recurrence and transience of growth models. *J. Appl. Probab.* 23: 614–625.

[6] G. Kersting (2016). Recurrence and transience of near-critical multivariate growth models: criteria and examples. Branching processes and their applications, ed. I. M. del Puerto et al. *Lecture Notes in Statistics*, 219: 207–218.

[7] F. Klebaner (1991). Asymptotic behavior of near-critical multitype branching processes. *J. Appl. Probab.* 28: 512–519.

[8] J. Lamperti (1960). Criteria for recurrence or transience of stochastic processes. *J. Math. Anal. Appl.* 1: 314–330. MR-0126872

[9] E. Seneta. Non-negative matrices and Markov chains. Springer, 2nd edition, 1981.

Acknowledgments. Thanks to the anonymous referee for careful surveying the paper, which improved its presentation, and also to DFG-SPP 1590 for financial support.