Persistence of One-Dimensional AR(1)-Sequences

Günter Hinrichs[1] · Martin Kolb[2] · Vitali Wachtel[1]

Abstract
For a class of one-dimensional autoregressive sequences (X_n), we consider the tail behaviour of the stopping time $T_0 = \min\{n \geq 1 : X_n \leq 0\}$. We discuss existing general analytical approaches to this and related problems and propose a new one, which is based on a renewal-type decomposition for the moment generating function of T_0 and on the analytical Fredholm alternative. Using this method, we show that $\mathbb{P}_x(T_0 = n) \sim V(x)R_0^n$ for some $0 < R_0 < 1$ and a positive R_0-harmonic function V. Further, we prove that our conditions on the tail behaviour of the innovations are sharp in the sense that fatter tails produce non-exponential decay factors.

Keywords Persistence · Quasistationarity · Autoregressive sequence

Mathematics Subject Classification (2010) Primary 60J05 · Secondary 60G40

1 Introduction and Setting

The analysis of first hitting times of subsets of the state space by a Markov chain $(X_n)_{n\geq 0}$ is a subject with a long history, but still many recent contributions. For many applications, it is important to gain precise control of the tail behaviour of hitting times. One of the aims of this work is to demonstrate the usefulness of the combination of analytic and probabilistic techniques using the example of autoregressive processes.

✉ Vitali Wachtel
vitali.wachtel@mathematik.uni-augsburg.de

Günter Hinrichs
Guenter.Hinrichs@math.uni-augsburg.de

Martin Kolb
kolb@math.uni-paderborn.de

1 Institut für Mathematik, Universität Augsburg, 86135 Augsburg, Germany

2 Institut für Mathematik, Universität Paderborn, 33098 Paderborn, Germany

Let ξ_k be independent, identically distributed random variables, and let $a \in (0, 1)$ be a fixed constant. An AR(1)-sequence is defined by

$$X_n = aX_{n-1} + \xi_n, \quad n \geq 1, \tag{1}$$

where the starting point X_0 of this process may be either deterministic or distributed according to any probabilistic measure ν. If $X_0 = x$, then we write $\mathbb{P}_x$ for the distribution of the process, and if X_0 is distributed according to ν, then we shall write $\mathbb{P}_\nu$ for the distribution of the process. The sequence $(X_n)_{n \in \mathbb{N}_0}$ defines a Markov chain with state space $\mathbb{R}$, whose properties have been analysed in a great number of papers, and we only refer to some of the more recent contributions, such as [2–4,7,14,17,18]. Closest to our present contribution and the main stimulus for the present paper is the recent work [2], where persistence probabilities for the process (1) and its multidimensional versions have been studied. The focus of [2] has been on deriving the existence and basic properties such as positivity and monotonicity of the persistence exponents

$$\lambda_a := \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(X_0 > 0, X_1 > 0, \ldots, X_n > 0\big)$$

for rather general Markov chains (including multidimensional cases) and its calculation for some very specific chains. For our purposes, let us define

$$T_0 := \min\{k \geq 1 : \ X_k \leq 0\},$$

i.e. the first time at which the process becomes negative. Similar to [2], we are going to study the tail behaviour of this stopping time, but, in contrast to [2], we are aiming at precise instead of rough asymptotic results. This means we aim to find the precise, without logarithmic scaling, asymptotics of

$$n \mapsto \mathbb{P}_x(T_0 > n)$$

as $n \to \infty$. Under natural and in some sense minimal conditions, we aim to show that the tail of the stopping time T_0 has an exactly exponential decay. We will focus on the one-dimensional situation.

Denoting by $P(x, dy)$ the transition probability of the Markov chain $(X_n)_{n\geq 0}$, we observe that, for every $x > 0$,

$$\begin{aligned}
&\mathbb{P}_x\big(X_1 > 0, X_2 > 0, \ldots, X_n > 0\big) \\
&\quad = \int_{(0,\infty)} P(x, dy_1) \int_{(0,\infty)} P(y_1, dy_2) \ldots \int_{(0,\infty)} P(y_{n-1}, (0, \infty))
\end{aligned}$$

and therefore the probability $\mathbb{P}_x\big(X_1 > 0, X_2 > 0, \ldots, X_n > 0\big)$ can be interpreted as the nth power of the total mass of the substochastic transition kernel given by

$$P_+(x, A) := \mathbf{1}_{(0,\infty)}(x) P(x, A \cap (0, \infty)).$$

From the Gelfand formula for the spectral radius, it is tempting to connect λ_a to the spectral radius of some operator induced by P_+. From an operator theoretic perspective, the problem consists in the fact that, due to the unboundedness of the state space $(0, \infty)$, the substochastic kernel P_+ is usually not a (quasi)compact operator on the standard Banach spaces of continuous or pth power integrable functions. One way out is to find better adapted Banach spaces, which is often possible. A different strategy which we are going to present as well consists in analysing the behaviour of the Laplace transform

$$\lambda \mapsto \mathbb{E}_x\left[\mathrm{e}^{\lambda T_0}\right]$$

near the critical line and in this respect is classical. In fact, similar arguments appear in the investigation of other large time problems in the theory of stochastic processes such as [13,23]. In order to deduce the required properties, we will show that $\mathbb{E}_x\left[\mathrm{e}^{\lambda T_0}\right]$ satisfies a suitable renewal equation and study some operator theoretic properties of the corresponding transition operator. This leads to a meromorphic representation for the function $\mathbb{E}_x\left[\mathrm{e}^{z T_0}\right]$. The final step consists in showing that all singularities near the critical line $\{z : \Re z = \lambda_a\}$ are simple poles and in the subsequent application of the Wiener–Ikehara theorem. We want to emphasize that related results have been recently derived in [6] using completely different methods and we will comment on connections below.

All our results will be valid for any stopping time

$$T_r := \min\{k \geq 1 : \ X_k \leq r\}, \quad r \in \mathbb{R}.$$

This is immediate from the observation that the sequence $X_n^{(r)} := X_n - r$, $n \geq 0$ satisfies (1) with innovations $\xi_n^{(r)} := \xi_n - (1-a)r$, $n \geq 1$.

We want to stress at this point that even though we exclusively deal with the one-dimensional situation in this work, the approaches we present are more generally applicable, also in multidimensional situations and to processes of different type. As it seems that our analytic approaches are not well known in the probabilistic literature, they are at least rarely used in standard literature, we hope that the present contribution also serves as template on powerful analytic methods for persistence and quasistationary problems.

Several authors, see [7,14,17], have obtained exact expressions for the Laplace transform in the case when the distribution of the innovations is related to the exponential distribution. Unfortunately, the expressions are very complicated, and it is not clear how to invert them or how to use them for deriving the tail asymptotics for T_0. Moreover, it is not clear whether one can obtain an explicit expression for λ_a. If, for example, the innovations ξ_n have density $\mathrm{e}^{-\mu|x|}/(2\mu)$ then, as it has been shown in [14],

$$\mathbb{E}_0 s^{T_0} = \frac{s(as, a^2)_\infty}{(as, a^2)_\infty + (s, a^2)_\infty},$$

where $(u, q)_\infty = \prod_{k=0}^{\infty}(1 - uq^k)$. Therefore, e^{λ_a} is the minimal positive solution to the equation $(as, a^2)_\infty + (s, a^2)_\infty = 0$. It is obvious that this solution lies between 1 and a^{-1}, but an explicit expression is not accessible. In order to analyse the tail behaviour of T_0, we have to take into account all singularities of this function on the circle of radius e^{λ_a}, but this information is also rather hard to extract from the exact expression. Expressions in [7,17] are even more complicated. Summarizing, all known explicit expressions for the Laplace transform of T_0 do not seem to provide any useful information on the asymptotic properties of T_0.

The structure of the paper is the following. In Sect. 2, we prove under rather general assumptions the existence and positivity of the decay rate

$$\lambda_a := -\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}_x\big(T_0 > n\big).$$

In Sect. 3, we consider the situation where the distribution of the innovations has bounded support. In this case, one can use results in [5] in order to show that the tails have a precise exponential decay. For unbounded innovations, this approach is no longer possible.

In order to be able to cover also unbounded innovations, we introduce in Sect. 4 a different functional analytic approach relying on the concept of a quasicompact operator and the associated spectral theory.

Section 5 presents another way to deal with unbounded innovations, which is based on a renewal argument in combination with some basic operator theoretic arguments applied to the renewal operator.

Section 6 deals with the situation of regularly varying case, where the foregoing theory is not applicable.

In Sect. 7, the essential ingredients of our techniques and their applicability to different problems are discussed.

2 Rough Asymptotics for Persistence Probabilities

In this section, we establish a general result concerning the exponential decay of the tails of the hitting time T_0. Results of this type will be an essential ingredient for the further investigation of more detailed properties of the first hitting time T_0. More precisely, we shall derive asymptotics for the logarithmically scaled tail of T_0. In contrast to [2], we study the one-dimensional situation, but there we are able to work under much weaker hypotheses.

Let us first recall some basic properties of the Markov process $(X_n)_{n\in\mathbb{N}_0}$. If $\mathbb{E}\log(1 + |\xi_1|) < \infty$, then X_n converges weakly to the distribution of the series

$$X_\infty := \sum_{k=1}^{\infty} a^{k-1}\xi_k.$$

This is immediate from the following expression for X_n:

$$X_n = a^n X_0 + a^{n-1}\xi_1 + a^{n-2}\xi_2 + \cdots + \xi_n. \tag{2}$$

Let π denote the distribution of X_∞. This distribution is stationary for the Markov chain X_n, that is,

$$\mathbb{P}_\pi(X_n \in dx) = \pi(dx), \quad n \geq 1. \tag{3}$$

Theorem 1 *Assume that the innovations* $(\xi_n)_{n\in\mathbb{N}}$ *satisfy*

$$\mathbb{E}\log(1+|\xi_1|) < \infty,\ \mathbb{E}(\xi_1^+)^\delta < \infty \textit{ for some } \delta > 0 \quad \textit{and} \quad \mathbb{P}(\xi_1 > 0)\mathbb{P}(\xi_1 < 0) > 0\,.$$

Then, for every $a \in (0, 1)$,

$$-\lim_{n\to\infty}\frac{1}{n}\log\mathbb{P}_x\big(T_0 > n\big) = \lambda_a \in (0,\infty), \quad x \in (0,\infty). \tag{4}$$

Furthermore, if the distribution of the innovations satisfies

$$\lim_{x\to\infty}\frac{\log\mathbb{P}(\xi_1 > x)}{\log x} = 0, \tag{5}$$

then

$$-\lim_{n\to\infty}\frac{1}{n}\log\mathbb{P}_x\big(T_0 > n\big) = 0, \quad x \in (0,\infty). \tag{6}$$

If $\mathbb{P}(\xi_1 < 0) = 0$, then $T_0 = \infty$ almost surely. Furthermore, the assumption $\mathbb{P}(\xi_1 > 0) > 0$ is imposed to avoid a trivial situation when the chain X_n is monotone decreasing before hitting negative numbers. As it has been mentioned at the beginning of this section, the existence of the logarithmic moment yields the ergodicity of X_n. Finally, the finiteness of $\mathbb{E}(\xi_1^+)^\delta$ is needed for the positivity of λ_a only.

Existence and finiteness of the limit λ_a follow also from Theorem 2.3 in [2] in a multidimensional situation at least under the condition that an exponential moment exists. We use a weaker moment assumption $\mathbb{E}(\xi_1^+)^\delta < \infty$. Since the finiteness of $\mathbb{E}(\xi_1^+)^\delta$ implies that

$$\limsup_{x\to\infty}\frac{\log\mathbb{P}(\xi_1 > x)}{\log x} \leq -\delta,$$

we conclude that the moment assumption $\mathbb{E}(\xi_1^+)^\delta < \infty$ is optimal for the positivity of λ_a.

Relation (6) is a simple consequence of relation (2.4) in [2]. It should be noted that the proof of (2.4) does not use the assumption that the innovations are normal distributed and, consequently, we may use (2.4) in the proof of (6).

Proof of Theorem 1 It follows from (2) that, for $0 \le x \le y$,

$$\mathbb{P}_x(T_0 > n) = \mathbb{P}\left(\min_{k\le n}\left(xa^k + \sum_{j=1}^{k} a^{k-j}\xi_j\right) > 0\right)$$
$$\le \mathbb{P}\left(\min_{k\le n}\left(ya^k + \sum_{j=1}^{k} a^{k-j}\xi_j\right) > 0\right) = \mathbb{P}_y(T_0 > n). \tag{7}$$

Besides T_0, we consider a slightly modified stopping time

$$\widetilde{T}_0 := \min\{k \ge 0 : X_k \le 0\}.$$

The monotonicity property (7) implies that, for every $x > 0$,

$$\mathbb{P}_\pi(\widetilde{T}_0 > n) = \int_0^\infty \pi(dy)\mathbb{P}_y(T_0 > n) \ge \int_x^\infty \pi(dy)\mathbb{P}_y(T_0 > n)$$
$$\ge \pi[x,\infty)\mathbb{P}_x(T_0 > n).$$

In other words,

$$\mathbb{P}_x(T_0 > n) \le \frac{1}{\pi[x,\infty)}\mathbb{P}_\pi(\widetilde{T}_0 > n). \tag{8}$$

Next, multiplying (2) by a^{-n}, we have

$$a^{-n}X_n = X_0 + \sum_{j=1}^{n} a^{-j}\xi_j =: X_0 + S_n$$

and

$$\{T_0 > n\} = \left\{X_0 + \min_{k\le n} S_k > 0\right\}.$$

Since X_0+S_k are sums of independent random variables, we may apply FKG inequality for product spaces, which gives the estimate

$$\mathbb{P}(X_n > y | T_0 > n) = \mathbb{P}\left(X_0 + S_n > a^{-n}y \middle| X_0 + \min_{k\le n} S_k > 0\right)$$
$$\ge \mathbb{P}\left(X_0 + S_n > a^{-n}y\right) = \mathbb{P}(X_n > y), \tag{9}$$

which holds for every distribution of X_0. Applying this inequality to the stationary process, we obtain

$$\begin{aligned}\mathbb{P}_\pi(\widetilde{T}_0 > n+m) &= \int_0^\infty \mathbb{P}_\pi(X_n \in dy, \widetilde{T}_0 > n)\mathbb{P}_y(T_0 > m) \\ &= \mathbb{P}_\pi(\widetilde{T}_0 > n)\int_0^\infty \mathbb{P}_\pi(X_n \in dy | \widetilde{T}_0 > n)\mathbb{P}_y(T_0 > m) \\ &\ge \mathbb{P}_\pi(\widetilde{T}_0 > n)\int_0^\infty \mathbb{P}_\pi(X_n \in dy)\mathbb{P}_y(T_0 > m) \\ &= \mathbb{P}_\pi(\widetilde{T}_0 > n)\mathbb{P}_\pi(\widetilde{T}_0 > m).\end{aligned}$$

Then, by the Fekete lemma,

$$-\lim_{n\to\infty}\frac{1}{n}\log \mathbb{P}_\pi(\widetilde{T}_0 > n) =: \lambda_a(\pi) \in [0,\infty). \tag{10}$$

By the same argument, for every fixed x we have

$$\begin{aligned}\mathbb{P}_x(T_0 > n+m) &\ge \mathbb{P}_x(T_0 > n)\int_0^\infty \mathbb{P}_x(X_n \in dy)\mathbb{P}_y(T_0 > m) \\ &\ge \mathbb{P}_x(T_0 > n)\int_x^\infty \mathbb{P}_x(X_n \in dy)\mathbb{P}_y(T_0 > m).\end{aligned}$$

Using now the monotonicity property (7), we get

$$\mathbb{P}_x(T_0 > n+m) \ge \mathbb{P}_x(T_0 > n)\mathbb{P}_x(X_n \ge x)\mathbb{P}_x(T_0 > m). \tag{11}$$

If x is such that $\pi[x,\infty) > 0$, then $\mathbb{P}_x(X_n \ge x) \to \pi[x,\infty)$ and we may apply again the Fekete lemma:

$$-\lim_{n\to\infty}\frac{1}{n}\log \mathbb{P}_x(T_0 > n) =: \lambda_a(x) \in [0,\infty). \tag{12}$$

By (7), $\lambda_a(x)$ is decreasing in x. Moreover, from (8), (10) and (12), we infer that

$$\lambda_a(x) \ge \lambda_a(\pi) \quad \text{for all } x \text{ such that } \pi[x,\infty) > 0.$$

If $\pi[x,\infty) > 0$, then there exist $\varepsilon > 0$ and m_0 such that $\mathbb{P}_x(X_{m_0} > x+\varepsilon) > 0$. Then, using the Markov property at time m_0 and the monotonicity of $P_y(T_0 > n)$, we conclude that $\lambda_a(x) = \lambda_a(y)$ for all $y \in [x, x+\varepsilon]$. As a result, we have

$$\lambda_a(x) = \lambda_a(\pi) =: \lambda_a \quad \text{for all } x \text{ such that } \pi[x,\infty) > 0.$$

According to Theorem 1 in [18], the assumption $\mathbf{E}(\xi_1^+)^\delta < \infty$ for some $\delta > 0$ implies that $\lambda_a > 0$. Thus, we have the same exponential rate for all starting points in the support of the measure π.

If $\mathbb{P}(\xi_1 > x) > 0$ for all $x > 0$, then we have the logarithmic asymptotic behaviour for all positive starting points.

Consider now the case when innovations are bounded from above. Let R denote the essential supremum of ξ_1, that is, $\mathbb{P}(\xi_1 \le R) = 1$ and $\mathbb{P}(\xi_1 > R - \varepsilon) > 0$ for every $\varepsilon > 0$. It is easy to see that $\pi([x, \infty)) > 0$ for every $x \in (0, R/(1-a))$ and $\pi([x, \infty)) = 0$ for each $x > R/(1-a)$. If the starting point x is greater than $R_* := R/(1-a)$, then the sequence X_n decreases before it hits $(0, R_*)$ and there exists $m = m(x, \varepsilon)$ such that $X_m \le R_* + \varepsilon$. Further, for every starting point in $[R_*, R_* + \varepsilon)$ the hitting time of $(-\infty, R_*)$ is stochastically bounded by a geometric random variable with parameter $1 - p_\varepsilon$, where $p_\varepsilon := \mathbb{P}(\xi_1 > R - \varepsilon) > 0$. Since we know the exponent for all starting points in $(0, R_*)$, we conclude that

$$-\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}_x(T_0 > n) = \min\{\lambda_a, \log(1/p_\varepsilon)\}, \quad x \ge R_*.$$

Letting now $\varepsilon \to 0$, we obtain

$$-\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}_x(T_0 > n) = \min\{\lambda_a, \log(1/p)\}, \quad x \ge R_*,$$

where $p = \mathbb{P}(\xi_1 = R)$. Now, noting that

$$\mathbb{P}_x(T_0 > n) \ge \mathbb{P}(\xi_1 = \xi_2 = \cdots = \xi_n = R) = p^n, \quad x > 0,$$

we infer that $\lambda_a \le \log(1/p)$. Consequently,

$$-\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}_x(T_0 > n) = \lambda_a, \quad x \ge R_*.$$

This completes the proof of (4).

It is clear that the distribution of ξ_1 is a uniform minorant for distributions of X_1, that is,

$$\mathbb{P}_x(X_1 > y) \ge \mathbb{P}(\xi_1 > y), \quad \text{for all } x, y > 0.$$

Therefore,

$$\mathbb{P}_x(T_0 > n) \ge \mathbb{P}_\mu(T_0 > n - 1),$$

where μ denotes the distribution of ξ_1. (6) follows now from (2.4) in [2]. □

We conclude this section with the following open problem: For which starting distributions $\mu(dx) = \mathbb{P}(X_0 \in dx)$ the statement of Theorem 1 remains valid? A general observation that the persistence exponent may depend on the initial distribution of an AR(1)-sequence is made in Proposition 2.2 in [2]. The reader can find conditions on the initial distribution in [2], in [6] and also in Sect. 4.2 ensuring the validity of Theorem 1. Until now, a complete characterization does not seem to exist. At this point, we want to emphasize that this phenomenon is well known in the theory of quasistationary distributions and is strongly related to the fact that quasistationary distributions are not unique in general (see, e.g. [8]). Conditions ensuring uniqueness of quasistationary distributions are given in [5].

3 Innovations with Bounded to the Right Support: Approach Via Quasistationarity

Exponential decay of $\mathbb{P}(T_0 > n)$ is often related to a quasistationary behaviour of (X_n) conditioned on the event $\{T_0 > n\}$. Recall that the quasistationarity implies that

$$\mathbb{P}(T_0 > n+1 | T_0 > n) \to \mathrm{e}^{-\lambda_a}.$$

So, the logarithmic asymptotics in Theorem 1 should also follow from the quasistationarity of (X_n). Furthermore, it is quite natural to expect that the knowledge on the rate of convergence towards a quasistationary distribution will imply preciser statements on the tail behaviour of T_0. This has been recently confirmed in [5], where necessary and sufficient conditions for an exponential speed of convergence in the total variation norm to a quasistationary distribution have been provided. There, it has also been shown that such fast convergence yields a purely exponential decay of $\mathbb{P}(T_0 > n)$.

Let us formulate conditions from [5] in terms of the AR(1)-sequence $(X_n)_{n\in\mathbb{N}_0}$. *First condition* there exist a probability measure ν and constants $n_0 \geq 1$, $c_1 > 0$ such that

$$\mathbb{P}_x(X_{n_0} \in \cdot | T_0 > n_0) \geq c_1 \nu(\cdot) \quad \text{for all } x > 0. \tag{13}$$

Second condition there exists a constant c_2 such that

$$\mathbb{P}_\nu(T_0 > n) \geq c_2 \mathbb{P}_x(T_0 > n) \quad \text{for all } n \geq 1 \text{ and } x > 0. \tag{14}$$

It is rather obvious that (14) cannot be valid for all $x > 0$. This observation implies that the results in [5] are not applicable to AR(1)-sequences with unbounded to the right innovations. So, we shall assume that innovations ξ_k are bounded. Let R denote, as in the previous section, the essential supremum of ξ_1. Then, the invariant measure lives on the set $(-\infty, R_*]$, where $R_* = R/(1-a)$. If the starting point lies in $(0, R_*]$, then the chain X_n does not exceed R_* at all times. Consequently, we have to find restrictions on the distribution of innovations which will ensure the validity of (13) and (14) for $x \leq R_*$ only.

We begin by showing that (13) holds for a quite wide class of innovations.

Lemma 2 *Assume that the distribution of innovations satisfies*

$$\mathbb{P}(\xi_1 \leq -aR_*) + \mathbb{P}(\xi_1 = R) < 1. \tag{15}$$

Then, there exist $\delta > 0$ and a constant c such that, for every $x \in [R_* - \delta, R_*]$,

$$\mathbb{P}_{R_*}(T_0 > n) \leq c\mathbb{P}_x(T_0 > n), \quad n \geq 1. \tag{16}$$

In particular, the condition (14) *is valid for any* ν *with* $\nu[R_* - \delta, R_*] > 0$.

Remark 3 If the assumption (15) does not hold, i.e.

$$\mathbb{P}(\xi_1 \le -aR_*) + \mathbb{P}(\xi_1 = R) = 1,$$

then one can easily see that, for all $x \in (0, R_*]$ and $n \ge 1$,

$$\mathbb{P}_x(T_0 > n) = (\mathbb{P}(\xi_1 = R))^n .$$

Therefore, (15) does not restrict the generality. $\diamond$

Proof of Lemma 2 Clearly, (15) yields the existence of $\gamma > 0$ such that

$$\mathbb{P}(\xi_1 > -aR_* + \gamma) > \mathbb{P}(\xi_1 = R).$$

Further, there exists $m = m(\gamma)$ such that, uniformly in starting points $x \in (0, R_*]$,

$$\mathbb{P}_x(X_{m-1} > R_* - \gamma/a, T_0 > m-1) \ge (\mathbb{P}(\xi_1 = R))^{m-1} .$$

Consequently,

$$\begin{aligned}\mathbb{P}_x(T_0 > m) &\ge \mathbb{P}_x(X_{m-1} > R_* - \gamma/a, T_0 > m-1)\mathbb{P}(\xi_m > -aR_* + \gamma)\\ &> (\mathbb{P}(\xi_1 = R) + \varepsilon)^m\end{aligned}$$

for all $x \in (0, R_*]$ and some $\varepsilon > 0$. Since

$$\mathbb{P}_x(T_0 > nm) \ge \left(\min_{x>0} \mathbb{P}_x(T_0 > m)\right)^{\lfloor \frac{n}{m} \rfloor} = (\mathbb{P}(\xi_1 = R) + \varepsilon)^{\lfloor \frac{n}{m} \rfloor m} ,$$

we infer that

$$\mathbb{P}(\xi_1 = R) < \mathrm{e}^{-\lambda_a}.$$

Therefore, there exists $\delta > 0$ such that

$$\mathbb{P}(\xi_1 > R - \delta) < \mathrm{e}^{-\lambda_a},$$

which is equivalent to

$$\varepsilon(\delta) := \mathbb{P}_{R_*}(X_1 > R_* - \delta) < \mathrm{e}^{-\lambda_a}. \tag{17}$$

Taking into account the monotonicity property (7), we get

$$\begin{aligned}\mathbb{P}_{R_*}(T_0 > n) \le &\mathbb{P}_{R_*}(X_1 \le R_* - \delta)\mathbb{P}_{R_*-\delta}(T_0 > n-1)\\ &+ \mathbb{P}_{R_*}(X_1 > R_* - \delta)\mathbb{P}_{R_*}(T_0 > n-1)\\ =&(1 - \varepsilon(\delta))\mathbb{P}_{R_*-\delta}(T_0 > n-1) + \varepsilon(\delta)\mathbb{P}_{R_*}(T_0 > n-1).\end{aligned}$$

Iterating this estimate, we obtain

$$\mathbb{P}_{R_*}(T_0 > n) \le \frac{1-\varepsilon(\delta)}{\varepsilon(\delta)} \sum_{k=1}^{n} \varepsilon^k(\delta)\mathbb{P}_{R_*-\delta}(T_0 > n-k). \tag{18}$$

It follows from (11) that

$$\mathbb{P}_{R_*-\delta}(T_0 > n-k) \le \frac{\mathbb{P}_{R_*-\delta}(T_0 > n)}{\mathbb{P}_{R_*-\delta}(T_0 > k)\mathbb{P}_{R_*-\delta}(X_k > R_* - \delta)}.$$

Plugging this into (18), we have

$$\mathbb{P}_{R_*}(T_0 > n) \le \frac{1-\varepsilon(\delta)}{\varepsilon(\delta)} \frac{\mathbb{P}_{R_*-\delta}(T_0 > n)}{\inf_k \mathbb{P}_{R_*-\delta}(X_k > R_* - \delta)} \sum_{k=1}^{\infty} \frac{\varepsilon^k(\delta)}{\mathbb{P}_{R_*-\delta}(T_0 > k)}.$$

The summability of the series on the right-hand side follows from Theorem 1 and from estimate (17). Furthermore, the convergence of X_n towards the stationary distribution π implies that $\inf_k \mathbb{P}_x(X_k > x)$ is positive. Thus, there exists a constant c such that

$$\mathbb{P}_{R_*}(T_0 > n) \le c\mathbb{P}_{R_*-\delta}(T_0 > n), \quad n \ge 1.$$

The monotonicity of $\mathbb{P}_x(T_0 > n)$ completes the proof of the first claim.

Using the monotonicity property once again and applying (16), we obtain

$$\begin{aligned}\mathbb{P}_\nu(T_0 > n) &\ge \frac{\nu[R_* - \delta, R_*]}{c}\mathbb{P}_{R_*}(T_0 > n) \\ &\ge \frac{\nu[R_* - \delta, R_*]}{c}\mathbb{P}_x(T_0 > n), \quad x \in (0, R_*].\end{aligned}$$

This completes the proof of the lemma. □

We now turn to the condition (13).

Lemma 4 *Assume that the distribution of ξ_1 has an absolutely continuous component with the density function $\varphi(x)$ satisfying*

$$\varphi(y) \ge \varkappa > 0 \quad \textit{for all } y \in [R - y_0, R], \ \ y_0 > 0. \tag{19}$$

Then, for every measurable $A \subseteq [R_* - y_0, R_* - ay_0]$,

$$\liminf_{n\to\infty} \inf_{x\in[0,R_*]} \mathbb{P}_x(X_n \in A | T_0 > n) \ge \varkappa\pi[R_* - y_0, R_*)\mathrm{Leb}(A)$$

*(*Leb *denoting the Lebesgue measure).*

Proof By the Markov property,

$$\mathbb{P}_x(X_{n+1} \in A | T_0 > n+1) = \frac{\mathbb{P}_x(X_{n+1} \in A, T_0 > n+1)}{\mathbb{P}_x(T_0 > n+1)}$$
$$\geq \frac{\int_A \left(\int_0^{R_*} \varphi(z - ay) \mathbb{P}_x(X_n \in dy, T_0 > n) \right) \mathrm{d}z}{\mathbb{P}_x(T_0 > n)}.$$

Applying (19), we get

$$\mathbb{P}_x(X_{n+1} \in A | T_0 > n+1) \geq \varkappa \int_A \mathbb{P}_x \left(X_n \in \left[\frac{z - R}{a}, \frac{z - R + y_0}{a} \right] \Big| T_0 > n \right) \mathrm{d}z.$$

For every $z \geq R_* - y_0$, we have $\frac{z-R+y_0}{a} \geq R_*$. Therefore,

$$\mathbb{P}_x(X_{n+1} \in A | T_0 > n+1) \geq \varkappa \int_A \mathbb{P}_x \left(X_n \geq \frac{z - R}{a} \Big| T_0 > n \right) \mathrm{d}z.$$

Furthermore, for every $z \leq R_* - ay_0$ one has $\frac{z-R}{a} \leq R_* - y_0$. Thus, using now (9) and recalling that X_n is increasing in the starting point, we conclude that

$$\inf_{x \in [0, R_*)} \mathbb{P}_x(X_{n+1} \in A | T_0 > n+1) \geq \varkappa \mathbb{P}_0(X_n \geq R_* - y_0) \mathrm{Leb}(A).$$

Letting here $n \to \infty$, we get the desired estimate. □

Combining these two lemmata with Proposition 1.2 in [5], we get

Theorem 5 *Assume that the innovations ξ_i are a.s. bounded and that their distribution possesses an absolutely continuous component satisfying* (19)*. Then, there exists a positive function $V(x)$ such that, for each $x \in (0, R_*]$,*

$$\mathbb{P}_x(T_0 > n) \sim V(x) \mathrm{e}^{-\lambda_a n} \quad \textit{as } n \to \infty.$$

4 Functional Analytic Approaches

In this section, we combine probabilistic insights with some basic functional analytic observations in order to derive the precise tail behaviour of T_0. We want to stress that even though we call this approach functional analytic, we will only make use of rather fundamental properties of compact operators combined with assertions of Perron–Frobenius type. The functional analytic ingredients can be found in standard references such as [1,9,16,21].

4.1 Quasicompactness Approach for Bounded Innovations.

The initial idea from the introduction can be most straightforwardly carried through for bounded innovations ξ_i. We assume that they have a density φ which is strictly positive on all of its support $[-A, B]$ $(A, B > 0)$ and consider $P_+ f(x) := \mathbb{E} f(ax+\xi_1)$ (where $f(y) := 0$ for $y < 0$) as an operator on $C\left(\left[0, \frac{B}{1-a}\right]\right)$ with the supremum norm. In this case, we are going to show that P_+ is compact with a simple largest eigenvalue $\mathrm{e}^{-\lambda_a}$ strictly between 0 and 1 and can then conclude

$$\mathbb{P}_x(T_0 > n) = P_+^n \mathbf{1}_{[0,\infty)}(x) = V(x)\mathrm{e}^{-\lambda_a n} + O(\mathrm{e}^{-(\lambda_a+\varepsilon)n}) \tag{20}$$

for some non-negative V and $\varepsilon > 0$. Apart from condition (19), which is not needed in this approach, the result is contained in Theorem 5. Our main purpose here is to finally lead over to those cases of unbounded innovations in which conditions from [5] are not valid.

For any continuous f and $x, y \in \left[0, \frac{B}{1-a}\right]$, $|P_+ f(x)| \leq \|f\|$ and

$$|P_+ f(x) - P_+ f(y)| = \left| \int [\varphi(z-ax) - \varphi(z-ay)] f(z) \mathrm{d}z \right| \tag{21}$$

$$\leq \|f\| \int |\varphi(z-ax) - \varphi(z-ay)| \mathrm{d}z\,. \tag{22}$$

This goes to zero for $y \to x$, i.e. P_+ maps bounded families to equicontinuous ones and, therefore, is compact. (X_n) clearly reaches zero from any starting point in $\left[0, \frac{B}{1-a}\right]$ if $\xi_1, \ldots, \xi_{\left\lceil \frac{B}{1-a} \frac{2}{A} \right\rceil} < -\frac{A}{2}$, which happens with positive probability, so

$$\left\| P_+^{\left\lceil \frac{B}{1-a} \frac{2}{A} \right\rceil} \right\| = \left\| P_+^{\left\lceil \frac{B}{1-a} \frac{2}{A} \right\rceil} 1_{[0,\infty)} \right\| = \sup_x \mathbb{P}_x \left(T_0 > \left\lceil \frac{B}{1-a} \frac{2}{A} \right\rceil \right) < 1\,.$$

Consequently, all eigenvalues of P_+ must have modulus less than 1. We now invoke the following generalization of the Perron–Frobenius theorem (see Theorems 6 and 7 in [20]):

Theorem A (see [20]). *Let K be a proper closed cone in a Banach space B which is fundamental and assume that B is a lattice with respect to the ordering induced by K. Let $T : B \to B$ be quasicompact operator, which is positive with respect to K, i.e. $TK \subset K$. Further assume that for each $B \ni f > 0$, $B^* \ni f^* > 0$ there exists an integer $n(f, f^*) \geq 0$ such that $f^*(T^n f) > 0$ for $n > n(f, f^*)$. Then*

(a) *the spectral radius $r(T) \in \sigma(T)$ has algebraic multiplicity 1 and is the only element in $\sigma(T)$ with absolute value equal to $r(T)$;*
(b) *the eigenspace corresponding to the eigenvalue $r(T)$ is one-dimensional and is spanned by a strictly positive element u;*
(c) *there exists a strictly positive element u^* such that $T^* u^* = r(T) u^*$.*

"Quasicompact" should now be thought of as "compact", the general version will be needed and explained in the proof of Theorem 7.

In our case, $B = C\left(\left[0, \frac{B}{1-a}\right]\right)$ and for K we take the cone of non-negative functions. It is closed, proper (i.e. $K \cap -K = \{0\}$) and fundamental. (This means that K spans a dense subset of B, which is clear since $K - K = B$.) P_+, taking the role of T, is positive. The density φ was assumed to be strictly positive, so positivity holds true even in the stronger sense that P_+ maps non-negative functions which are somewhere strictly positive to functions which are everywhere strictly positive. For our choice of the space, B its dual space B^* consists of all functionals $f \mapsto \int f \mathrm{d}\mu$ with finite signed Borel measures μ, positive functionals correspond to positive measures. Therefore, this strong positivity implies that $f^*(Tf) > 0$ for all $f > 0$ and all positive $f^* \in B^*$.

Theorem A is therefore applicable. It allows for the general conclusion that, for some $\varepsilon > 0$ and any $f > 0$,

$$T^n f = r(T)^n\, u^*(f)\, u + \mathcal{O}\big((r(T) - \varepsilon)^n\big) \tag{23}$$

in the space B, which, applied to our setting with $f = \mathbf{1}_{\left[0, \frac{B}{1-a}\right]}$, yields (20).

In fact, the computation (21) also works for bounded measurable f and shows that P_+ is compact on the corresponding space $B\left(\left[0, \frac{B}{1-a}\right]\right)$. It should be noted that the corresponding dual space consists of signed finitely additive and absolutely continuous measures. Consequently, (23) is also applicable for indicator functions $f = \mathbf{1}_A$ with measurable $A \subset \left[0, \frac{B}{1-a}\right]$ and yields

$$\mathbb{P}_x(X_n \in A, T_0 > n) = \nu(A) V(x) \mathrm{e}^{-\lambda_a n} + \mathcal{O}(\mathrm{e}^{-(\lambda_a + \varepsilon)n})$$

with same factor $\nu(A)$ which is strictly positive unless A has measure zero and, together with (20),

$$\mathbb{P}_x(X_n \in A \mid T_0 > n) = \nu(A) + \mathcal{O}(\mathrm{e}^{-\varepsilon n}).$$

In other words, we have an exponentially fast convergence to a quasistationary distribution ν.

4.2 Quasicompactness Approach for Unbounded Innovations

If the innovations ξ_i are absolutely continuous, but are unbounded, P_+ still maps bounded families to equicontinuous ones and has the same positivity properties, but is in general not compact. A way out is to choose a more suitable Banach space. This can be, for example, explicitly carried through in the case of innovations with standard normal distributions. In this case, X is a discretized Ornstein–Uhlenbeck process: The latter is given, e.g. by the stochastic differential equation

$$dZ_t = \theta Z_t\, dt + \sigma\, dB_t, \quad Z_0 = x,$$

where $\theta, \sigma > 0$ are constants and $(B_t)_{t\geq 0}$ denotes a Brownian motion. The random variable Z_t is normal distributed with

$$\mathbb{E}[Z_t] = x\mathrm{e}^{-\theta t} \quad \text{and} \quad \mathrm{Var}[Z_t] = \frac{\sigma^2}{2\theta}(1 - \mathrm{e}^{-2\theta t}).$$

For the chain X_n with standard normal distributed innovations, we have

$$\mathbb{P}_x\big(X_1 \in B\big) = \int_B \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{(ax-y)^2}{2}}\,\mathrm{d}y.$$

Taking

$$\theta = \log(a^{-1}) \quad \text{and} \quad \sigma^2 = \frac{2\log(a^{-1})}{1-a^2},$$

we see that Z_1 and X_1 are identically distributed.

Z has $\mathcal{N}(0, \frac{\sigma^2}{2\theta})$ as the stationary distribution and $Pf(x) := \mathbb{E}f(ax + \xi_1)$ is a self-adjoint compact operator with norm 1 on $B := L^2\left(\mathbb{R}, \mathcal{N}\left(0, \frac{\sigma^2}{2\theta}\right)\right)$. This is a well-known result and can be seen, e.g. from its diagonal representation in the Hermite polynomials, which form a complete set of eigenfunctions. Consequently, also

$$P_+ := 1_{[0,\infty)} P 1_{[0,\infty)}$$

is compact and self-adjoint. Its norm is strictly below 1. (Otherwise, one could find a normalized sequence (f_n) with $\|P_+ f_n\| \geq 1 - \frac{1}{n}$ for all $n \in \mathbb{N}$. A subsequence (f_{n_k}) would have a weak limit f with $\|f\| \leq 1$ and, by compactness, $P_+ f_{n_k} \to P_+ f$ in norm, in particular $\|P_+ f\| = 1$. Then, $P 1_{[0,\infty)} f$ would have norm 1, but the spectral representation of P shows that $g \equiv 1$ is the only function with $\|g\| \leq 1$ and $\|Pg\| = 1$. Therefore, one can apply Theorem A in the same way as for bounded innovations. A result of a similar type has been shown in [3], but the conclusion drawn has been somewhat weaker.

With somewhat more effort, one could also work on the space of continuous functions with the norm $\|f\| := \sup_x |f(x)\sqrt{\pi(x)}|$, where π is the density of the $\mathcal{N}(0, \frac{\sigma^2}{2\theta})$ distribution, so the crucial step was to introduce a suitable weight function, whereas L^2 instead of C brought only computational benefits in this particular case.

Let us now consider general AR(1)-processes with possibly unbounded innovations. Assume that for some $M > 0$, some $\varepsilon > 0$ and $\tilde{\lambda} = \lambda_a + \varepsilon$ we have

$$\Lambda(x) := \mathbb{E}_x\left[\mathrm{e}^{\tilde{\lambda} T_M}\right] < \infty, \tag{24}$$

where

$$T_M := \min\{k \geq 1 : X_k \leq M\}.$$

This property will be later shown to hold, whenever the innovations have moments of all orders, i.e. we do not require existence of an exponential moment. We introduce the Banach space $B(\mathbb{R}_0^+)$ of measurable functions on $[0,\infty)$ equipped with the norm

$$\|f\|_\Lambda := \|\Lambda^{-1} f\|_\infty < \infty\,.$$

Λ^{-1} is decreasing in $x \in \mathbb{R}_0^+$ and clearly takes values in $(0, 1]$. In contrast to other weight functions with these properties, it is computationally well tractable by the aid of the Markov property.

Actually, the approach we now present is an adaption of a standard approach to the ergodicity of Markov chains via quasicompactness (see, e.g. Chapter 6 in [19] as well as [12]). In the literature on quasistationary distributions, an approach of this type has been presented in [11]. Our main goal is to indicate via the example of autoregressive processes that the results of [6] can be to a large extent reproved via the concept of quasicompact operators. We hope that our outline of standard analytic ideas in a probabilistic context illuminates further the interrelationship between probabilistic and analytic concepts related to persistence exponents/quasistationarity on the one hand and spectral theoretic ideas on the other.

Proposition 6 *The transition operator P*

$$Pf(x) := \mathbb{E}_x\big[f(X_1), T_0 > 1\big]$$

defines a bounded operator on $(B(\mathbb{R}_0^+), \|\cdot\|_\Lambda)$. Moreover, the spectral radius $r(P)$ is lower-bounded by $\mathrm{e}^{-\lambda_a}$.

Proof For the boundedness, we observe that for $f \in B(\mathbb{R}_+)$ with $\|f\|_\Lambda \le 1$

$$\|Pf\|_\Lambda \le \sup_{x>0}\big|\Lambda(x)^{-1}(P\Lambda)(x)\big|\,.$$

By the Markov property,

$$\mathbb{E}_x\left[\mathrm{e}^{\tilde{\lambda}T_M} \mid X_1\right] = \mathrm{e}^{\tilde{\lambda}}\mathbf{1}_{0<X_1\le M} + \mathrm{e}^{\tilde{\lambda}}\Lambda(X_1)\mathbf{1}_{X_1>M}\,. \tag{25}$$

Therefore, for every $x \ge 0$ we have

$$\begin{aligned}(P\Lambda)(x) =& \mathbb{E}_x[\Lambda(X_1), 0 < X_1 \le M] + \mathbb{E}_x[\Lambda(X_1), X_1 > M]\\ \le& \Lambda(M) + \mathrm{e}^{-\tilde{\lambda}}\mathbb{E}_x\left[\mathrm{e}^{\tilde{\lambda}T_M}, X_1 > M\right]\\ \le& \Lambda(M) + \mathrm{e}^{-\tilde{\lambda}}\Lambda(x)\,,\end{aligned}$$

proving that P is bounded. In order to prove the assertion concerning the spectral radius, we observe that, for $x > 0$,

$$\|P^n\| \ge \frac{1}{\Lambda(x)}(P^n\mathbf{1})(x)$$

and that therefore

$$r(P) := \lim_{n\to\infty} \|P^n\|^{1/n} \geq e^{-\lambda_a}$$

by the Gelfand formula. □

Theorem 7 *Assume that condition (24) is satisfied and that the innovations are distributed according to a density φ. Then, the operator P is quasicompact on $(B(\mathbb{R}_0^+), \|\cdot\|_\Lambda)$. If, in addition, $\varphi > 0$ a.e., the spectral radius $r(P)$ is an isolated eigenvalue with algebraic multiplicity 1 and all other spectral values have absolute value strictly smaller than $r(P)$.*

Proof Let us decompose P into the sum of the following operators:

$$U_1 := P\mathbf{1}_{[0,M]}\,, \quad U_2 := \mathbf{1}_{[0,M]}P\mathbf{1}_{(M,\infty)}\,, \quad U_3 := \mathbf{1}_{(M,\infty)}P\mathbf{1}_{(M,\infty)}\,.$$

The operators U_1 and U_2 are easily seen to be compact. For this, we first note that

$$\{U_1 f;\, \|f\|_\Lambda \leq 1\} = \{\mathbb{E}_x[f(X_1)\Lambda(X_1), 0 < X_1 \leq M];\, \|f\| \leq 1\}$$

is bounded and equicontinuous. This follows immediately from the fact that $\Lambda(z) \leq \Lambda(M)$ for $0 \leq z \leq M$ and

$$\mathbb{E}_x[f(X_1)\Lambda(X_1), 0 < X_1 \leq M] = \int_0^M \Lambda(z)f(z)\varphi(z-ax)\mathrm{d}z\,.$$

The boundedness also implies

$$\lim_{x\to\infty} \sup_{f:\|f\|_\Lambda\leq 1} \Lambda^{-1}(x)\mathbb{E}_x[f(X_1), 0 < X_1 \leq M)] = 0\,,$$

so $\{U_1 f;\, \|f\|_\Lambda \leq 1\}$ is precompact and U_1 is compact.

By (25), for $x \leq M$,

$$U_2(\Lambda f)(x) = \mathbb{E}_x\left[f(X_1)\Lambda(X_1), X_1 > M\right] = e^{-\tilde{\lambda}}\mathbb{E}_x\left[f(X_1)e^{\tilde{\lambda}T_M}, X_1 > M\right],$$

so for all f with $\|f\| \leq 1$ we have

$$|U_2(\Lambda f)(x)| \leq e^{-\tilde{\lambda}}\Lambda(M)\,,$$

i.e. $\{U_2 f;\, \|f\|_\Lambda \leq 1\}$ is bounded w.r.t. $\|\cdot\|_\Lambda$, and

$$|U_2(\Lambda f)(y) - U_2(\Lambda f)(x)| \leq e^{-\tilde{\lambda}} \sum_{k=2}^{N} |I_{f,k}(y) - I_{f,k}(x)| + 2e^{-\tilde{\lambda}} \sup_{0\leq x\leq M} |R_{1,N}(x)|$$

with

$$
\begin{aligned}
I_{f,k}(x) &:= \mathrm{e}^{\tilde{\lambda}k}\mathbb{E}_x\left[f(X_1)\mathbf{1}_{X_1>M,\,T_M=k}\right] \\
&= \mathrm{e}^{\tilde{\lambda}k}\int_M^\infty \dots \int_M^\infty \int_{-\infty}^M f(x_1)\Phi_{x,k}(x_1,\dots,x_k)\mathrm{d}(x_k,\dots,x_1)\,,
\end{aligned}
$$

where

$$
\Phi_{x,k}(x_1,\dots,x_k) := \varphi(x_1 - ax)\varphi(x_2 - ax_1)\dots\varphi(x_k - ax_{k-1})
$$

stands for the common density of $X_1,\dots,X_k$ given $X_0 = x$, and

$$
R_{f,N}(x) := \mathbb{E}_x\left[f(X_1)\mathrm{e}^{\tilde{\lambda}T_M}, T_M > N\right].
$$

Clearly,

$$
\sup_{0\le x\le M} R_{1,N}(x) = R_{1,N}(M) \to 0 \quad \text{as } N \to \infty.
$$

Moreover,

$$
\Phi_{y,k}(x_1,\dots,x_k) = \Phi_{x,k}(x_1 - a(y-x),\dots,x_k - a^k(y-x))\,,
$$

so

$$
|I_{f,k}(y) - I_{f,k}(x)| \le \mathrm{e}^{\tilde{\lambda}k}\int \dots \int |\Phi_{y,k} - \Phi_{x,k}|\mathrm{d}(x_1,\dots,x_k) \xrightarrow{y\to x} 0\,.
$$

(This is clear for continuous Φ with compact support and follows easily for general Φ if one approximates them in L^1 by such ones.) Consequently, $\{\,(U_2 f)|_{[0,M]}\,;\,\|f\|_\Lambda \le 1\}$ is equicontinuous, U_2 is compact, too, and

$$
P = L + U_3,
$$

where L is a compact operator.

We now estimate the operator norm of U_3: For $x > M$, (25) yields

$$
(U_3\Lambda)(x) = \mathbb{E}_x\left[\Lambda(X_1), X_1 > M\right] \le \mathrm{e}^{-\tilde{\lambda}}\Lambda(x)
$$

and therefore we have for the operator norm of U_3

$$
\|U_3\| \le \mathrm{e}^{-\tilde{\lambda}},
$$

which, again by the Gelfand formula, tells us that

$$
r(U_3) = \lim_{n\to\infty}\|U_3^n\|^{1/n} \le \mathrm{e}^{-\tilde{\lambda}} < \mathrm{e}^{-\lambda_a} \le< r(P)\,,
$$

the latter by Proposition 6.

According to the definition in [20], which calls an operator P quasicompact if $P^n = L + U$ for some $n \in \mathbb{N}$, compact operator L and bounded operator U with $\rho(U) < \rho(P)^n$, P is quasicompact. Applying Theorem A as in Sect. 4.1 allows to deduce the remaining statements of the theorem. □

Remark 8 Observe that the main ingredient in the above proof is only the finiteness of $\mathbb{E}_x\big[\mathrm{e}^{\tilde{\lambda} T_M}\big]$ for some $M > 0$, which together with some 'local' compact perturbation argument ensures that there is a gap separating the largest eigenvalue from the remaining parts of the spectrum. Therefore, we expect that the method will work in other settings, too.

Corollary 9 *Assume that condition (24) is satisfied and let us assume that the innovations have a strictly positive continuous density. Then, there exists $\delta > 0$ such that for every $x > 0$ and every measurable set $A \subset (0, \infty)$*

$$\mathbb{P}_x\big(X_n \in A; T_0 > n\big) = V(x)\mathrm{e}^{-\lambda_a n} + O\Big((\mathrm{e}^{-\lambda_a} - \delta)^n\Big).$$

This result is partly contained in the recent work [6] by Champagnat and Villemonais. If $\mathbb{E}\mathrm{e}^{\theta(\xi_1)\log \xi_1} < \infty$ for some function $\theta(x)$ such that $\theta(x) \uparrow \infty$ and $\ell(x) := \theta(x)\log x$ is concave, then

$$\frac{\mathbb{E}\mathrm{e}^{\ell(ax+\xi_1)}}{\mathrm{e}^{\ell(x)}} \le \frac{\mathrm{e}^{\ell(ax)}\mathbb{E}\mathrm{e}^{\ell(\xi_1)}}{\mathrm{e}^{\ell(x)}} \le \mathrm{e}^{-\theta(ax)\log(a)}\mathbb{E}\mathrm{e}^{\ell(\xi_1)} \to 0 \quad \text{as } x \to \infty$$

and, consequently, in this case their Proposition 7.2 can be applied to AR(1)-sequences and gives the same type of convergence towards a quasistationary distribution. It is obvious that $\theta(x) = \log^b(x)$ with $b > 0$ satisfies the conditions mentioned above. We shall see later that (24) holds for innovations having all power moments. Therefore, the condition $\mathbb{E}\mathrm{e}^{\log^{1+b}\xi_1} < \infty$ is slightly stronger than the existence of all power moments.

5 Alternative Approach for Unbounded Innovations

In this section, we present another approach to investigate the tails of the hitting times, which, in contrast to Perron–Frobenius-type methods, has the potential to deal with situations where the transition operator is not quasicompact and to identify additional polynomial decay factors. Apart from birth/death processes and one-dimensional diffusions quasistationary convergence in cases with no spectral gap has not yet been established. Moreover, only basic properties of compact operators play a role and therefore the functional analytic machinery will be more straightforward.

As an alternative to the search for a weight function, we now start with the following observation: The larger X_n is, the more it is diminished by the prefactor a in the recursion $X_{n+1} = aX_n + \xi_{n+1}$. In contrast, adding ξ_{n+1} has always the same absolute effect, no matter how large X_n is. Therefore, in some sense, it is easier for the process to reach average positive values from extremely large values than to reach zero from average values, so, also in the case of unbounded innovations, the main part in estimating $\mathbb{P}_x(T_0 > n)$ should still be to analyse what happens for not too large x.

The transition operator will in general not be (quasi)compact on the usual space $C(\mathbb{R}_0^+)$, but, as shown on the next pages, a modified functional analytic approach, in which one basically works on the continuous functions on some interval and keeps under control what happens outside, is possible. In contrast, the "coming down from infinity" which we just described has not the uniform character that would be needed to apply the results in [5].

Here is the main result of this section:

Theorem 10 *Assume that the distribution of innovations has a density $\varphi(x)$ which is positive a.e. on $\mathbb{R}$. Assume also that $\mathbb{E}(\xi_1^+)^t < \infty$ for all $t > 0$ and $\mathbb{E}(\xi_1^-)^\delta < \infty$ for some $\delta > 0$. Then, there exist $\gamma > 0$ and a positive function V such that*

$$\mathbb{P}_x(T_0 = n) = \mathrm{e}^{-\lambda_a(n+1)} V(x) + O\left(\mathrm{e}^{-(\lambda_a+\gamma)n}\right). \tag{26}$$

The function V is e^{λ_a}-harmonic for the transition kernel P_+, that is,

$$V(x) = \mathrm{e}^{\lambda_a} \int_0^\infty P_+(x, dy)V(y) = \mathrm{e}^{\lambda_a}\mathbb{E}[V(X_1); T_0 > 1], \quad x \geq 0.$$

Again, this result describes not only the exact asymptotic behaviour of $\mathbb{P}_x(T_0 = n)$ but states also that the remainder term decays exponentially faster than the main term. It is worth mentioning that the existence of all power moments required in Theorem 10 is the minimal moment condition. More precisely, we shall show in Proposition 19 that if the tail of ξ_1 is regularly varying, then it may happen that $\mathrm{e}^{\lambda_a n}\mathbb{P}(T_0 > n) \to 0$.

The starting point of the approach, which we are going to use in this section, is based on the following renewal-type decomposition for the moment generating function of T_0. First define

$$\sigma_r := \inf\{n \geq 1 : X_n > r\}, \quad r > 0.$$

Fix $\lambda < \lambda_a$. Then, for $x \leq r$ we have

$$\begin{aligned}
\mathbb{E}_x\left[\mathrm{e}^{\lambda T_0}\right] =& \mathbb{E}_x\left[\mathrm{e}^{\lambda T_0}; T_0 < \sigma_r\right] + \mathbb{E}_x\left[\mathrm{e}^{\lambda T_0}; T_0 > \sigma_r\right] \\
=& \mathbb{E}_x\left[\mathrm{e}^{\lambda T_0}; T_0 < \sigma_r\right] + \mathbb{E}_x\left[\mathrm{e}^{\lambda \sigma_r}\mathbf{1}_{\{T_0>\sigma_r\}}\mathbb{E}_{X_{\sigma_r}}\left[\mathrm{e}^{\lambda T_0}\right]\right] \\
=& \mathbb{E}_x\left[\mathrm{e}^{\lambda T_0}; T_0 < \sigma_r\right] + \mathbb{E}_x\left[\mathrm{e}^{\lambda \sigma_r}\mathbf{1}_{\{T_0>\sigma_r\}}\mathbb{E}_{X_{\sigma_r}}\left[\mathrm{e}^{\lambda T_0}; T_r = T_0\right]\right] \\
&+ \mathbb{E}_x\left[\mathrm{e}^{\lambda \sigma_r}\mathbf{1}_{\{T_0>\sigma_r\}}\mathbb{E}_{X_{\sigma_r}}\left[\mathrm{e}^{\lambda T_0}; T_r < T_0\right]\right].
\end{aligned}$$

Using now the Markov property at time T_r, we obtain the equation

$$\mathbb{E}_x\left[\mathrm{e}^{\lambda T_0}\right] = F_\lambda(x) + \int_0^r K_\lambda(x, dy)\mathbb{E}_y\left[\mathrm{e}^{\lambda T_0}\right], \tag{27}$$

where

$$F_\lambda(x) = \mathbb{E}_x \left[\mathrm{e}^{\lambda T_0}; T_0 < \sigma_r \right] + \mathbb{E}_x \left[\mathrm{e}^{\lambda \sigma_r} \mathbf{1}_{\{T_0 > \sigma_r\}} \mathbb{E}_{X_{\sigma_r}} \left[\mathrm{e}^{\lambda T_0}; T_r = T_0 \right] \right] \tag{28}$$

and

$$K_\lambda(x, dy) = \mathbb{E}_x \left[\mathrm{e}^{\lambda \sigma_r} \mathbf{1}_{\{T_0 > \sigma_r\}} \mathbb{E}_{X_{\sigma_r}} \left[\mathrm{e}^{\lambda T_r} \mathbf{1}_{\{T_r < T_0\}} \mathbf{1}_{\mathrm{d}y}(X_{T_r}) \right] \right]. \tag{29}$$

To analyse the renewal equation (27), we first have to derive some properties of the functions F_λ and the operators K_λ. More precisely, we first show that there exists $r > 0$ such that $F_\lambda(x)$ and $K_\lambda(x, dy)$ can be extended analytically for $\Re\lambda < \lambda_a + \varepsilon$.

5.1 Estimates for Stopping Times $T_0 \wedge \sigma_r$ and T_r

The main purpose of this paragraph is to show that, under the conditions of Theorem 10, $T_0 \wedge \sigma_r$ and T_r have lighter tails than T_0. This fact will play a crucial role in the study of properties of F_λ and K_λ.

Lemma 11 *Assume that $\mathbb{E}|\xi_1|^\delta$ is finite for some $\delta > 0$. Then, for every r such that $\pi[r, \infty) > 0$ there exists $\varepsilon_r > 0$ such that*

$$\sup_{x \in (0,r)} \mathbb{P}_x(T_0 \wedge \sigma_r > n) \le C_r \mathrm{e}^{-(\lambda_a + \varepsilon_r)n}, \quad n \ge 0. \tag{30}$$

Proof Clearly,

$$\mathbb{P}_x(T_0 \wedge \sigma_r > n) = \mathbb{P}_x \left(\max_{k \le n} X_k < r, T_0 > n \right), \quad n \ge 1.$$

Fix some $n_0 \ge 1$ and consider the sequence $\mathbb{P}_x \left(\max_{k \le \ell n_0} X_k < r, T_0 > \ell n_0 \right)$ in $\ell \in \mathbb{N}$. By the Markov property,

$$\begin{aligned} &\mathbb{P}_x \left(\max_{k \le \ell n_0} X_k < r, T_0 > \ell n_0 \right) \\ &\quad = \int_0^r \mathbb{P}_x \left(X_{(\ell-1)n_0} \in dy; \max_{k \le (\ell-1)n_0} X_k < r, T_0 > (\ell-1)n_0 \right) \\ &\quad \times \mathbb{P}_y \left(\max_{k \le n_0} X_k < r, T_0 > n_0 \right). \end{aligned}$$

It is easy to see that functions $\mathbf{1}_{\{\max_{k \le n_0} X_k \ge r\}}$ and $\mathbf{1}_{\{T_0 > n_0\}}$ are increasing functions in every innovation ξ_k, $k \le n_0$. Thus, by the FKG inequality for product spaces,

$$\mathbb{P}_y \left(\max_{k \le n_0} X_k \ge r, T_0 > n_0 \right) \ge \mathbb{P}_y \left(\max_{k \le n_0} X_k \ge r \right) \mathbb{P}_y (T_0 > n_0).$$

In other words,

$$\mathbb{P}_y\left(\max_{k\le n_0} X_k < r, T_0 > n_0\right) \le \mathbb{P}_y\left(\max_{k\le n_0} X_k < r\right)\mathbb{P}_y\left(T_0 > n_0\right)$$
$$\le \mathbb{P}_0\left(\max_{k\le n_0} X_k < r\right)\mathbb{P}_r\left(T_0 > n_0\right).$$

Consequently,

$$\mathbb{P}_x\left(\max_{k\le \ell n_0} X_k < r, T_0 > \ell n_0\right)$$
$$\le \mathbb{P}_0\left(\max_{k\le n_0} X_k < r\right)\mathbb{P}_r(T_0 > n_0)\mathbb{P}_x\left(\max_{k\le (\ell-1)n_0} X_k < r, T_0 > (\ell-1)n_0\right)$$
$$\le \ldots \le \left(\mathbb{P}_0\left(\max_{k\le n_0} X_k < r\right)\mathbb{P}_r(T_0 > n_0)\right)^{\ell}, \quad x \in (0, r). \tag{31}$$

By Theorem 1, $\mathbb{P}_r(T_0 > n_0) = \mathrm{e}^{-\lambda_a n_0 + o(n_0)}$ as $n_0 \to \infty$. Since $r - X_n$ is an AR(1)-sequence, we may apply Theorem 1 to this sequence:

$$\mathbb{P}_0\left(\max_{k\le n_0} X_k < r\right) = \mathrm{e}^{-\widetilde{\lambda}_a n_0 + o(n_0)}, \quad n_0 \to \infty$$

for some $\widetilde{\lambda}_a > 0$. Therefore, there exists n_0 such that

$$\frac{1}{n_0}\log \mathbb{P}_0\left(\max_{k\le n_0} X_k < r\right)\mathbb{P}_r(T_0 > n_0) < -\lambda_a - \frac{\widetilde{\lambda}_a}{2}.$$

Combining this estimate with (31), we obtain (30). □

We next show that a similar estimate holds for T_r.

Lemma 12 *Assume that $\mathbb{E}(\xi_1^+)^t < \infty$ for all $t > 0$. Then, for all $A \in (1, 1/a)$ and all $\lambda > 0$ there exists $r_0 = r_0(A, \lambda)$ such that, for all $r \ge r_0$,*

$$\mathbb{E}\left[\mathrm{e}^{\lambda T_r}\right] \le 2\mathrm{e}^{\lambda}\mathbb{E}\left(\frac{X_0}{r}\right)^{\lambda/\log A}$$

for any distribution of X_0 with support in (r, ∞).

Proof Define

$$y_j = rA^j, \quad j \ge 0.$$

Let us consider an 'aggregated' chain $(Y_n)_{n\ge 0}$ defined by the transition kernel

$$\mathbb{P}(Y_1 = j | Y_0 = k) = \mathbb{P}_{y_k}(X_1 \in (y_{j-1}, y_j]), \; k, j \ge 1$$

and

$$\mathbb{P}(Y_1 = 0|Y_0 = 0) = 1, \quad \mathbb{P}(Y_1 = 0|Y_0 = k) = \mathbb{P}_{y_k}(X_1 \le y_0)\,,\ k \ge 1.$$

Similarly, we define the initial distribution:

$$\mathbb{P}(Y_0 = j) = \mathbb{P}(X_0 \in (y_{j-1}, y_j])\,,\ j \ge 1\,.$$

Define also the stopping times

$$\tau := \inf\{n : Y_n = 0\} \quad \text{and} \quad \theta := \inf\{n : Y_n \ge Y_{n-1}\}.$$

We first estimate the distribution of Y_θ. Noting that $Y_\theta = j \ge 1$ implies that $Y_{\theta-1} \le j$, we have

$$\max_{k\ge1} \mathbb{P}(Y_\theta = j|Y_0 = k) = \max_{k\le j} \mathbb{P}(Y_\theta = j|Y_0 = k).$$

Furthermore, using the fact that $\theta \le j$ for $Y_0 \le j$ and $Y_\theta = j$, we infer that

$$\max_{k\ge1} \mathbb{P}(Y_\theta = j|Y_0 = k) \le j \max_{k\le j} \mathbb{P}(Y_1 = j|Y_0 = k).$$

By the definition of Y_n,

$$\begin{aligned}\mathbb{P}(Y_1 = j|Y_0 = k) &= \mathbb{P}_{y_k}(X_1 \in (y_{j-1}, y_j]) \\ &\le \mathbb{P}_{y_k}(X_1 > y_{j-1}) \le \mathbb{P}(\xi_1 > (1 - aA)u_{j-1}), \quad k \le j.\end{aligned}$$

As a result,

$$\max_{k\ge1} \mathbb{P}(Y_\theta = j|Y_0 = k) \le j\mathbb{P}(\xi_1 > (1 - aA)u_{j-1}) =: q_j(r), \quad j \ge 1. \tag{32}$$

Set $q_0(r) := 1 - \sum_{j=1}^{\infty} q_j(r)$. It is easy to see that $\lim_{r\to\infty} \sum_{j=1}^{\infty} q_j(r) = 0$. Therefore, $\{q_j(r)\}$ is a probability distribution for all r large enough.

Noting that the chain (Y_n) is stochastically monotone and using (32), we have, for arbitrary initial Y_0,

$$\tau_0 \le Y_0 + \mathbf{1}_{\{Y_\theta\ge1\}}\tau_0^{(q)} \quad \text{in distribution,}$$

where $\tau_0^{(q)}$ is independent of $Y_0, \dots, Y_\theta$ and is distributed as τ_0 corresponding to the initial distribution $q_j(r)/(1-q_0(r)),\ j \ge 1$. Combining this inequality with the bound $\mathbb{P}(Y_\theta \ge 1) \le 1 - q_0(r)$, we obtain

$$\mathbb{E}e^{\lambda\tau_0} \le \mathbb{E}e^{\lambda Y_0}\left(1 + (1 - q_0(r))\mathbb{E}e^{\lambda\tau_0^{(q)}}\right). \tag{33}$$

If Y_0 is distributed according to $q_j(r)/(1-q_0(r))$, then we conclude from (33) that

$$\mathbb{E}e^{\lambda\tau_0^{(q)}} \le \frac{\mathbb{E}e^{\lambda Y_0}}{1-(1-q_0(r))\mathbb{E}e^{\lambda Y_0}}. \tag{34}$$

It follows from the Chebyshev inequality that

$$q_j(r) \le j\frac{\mathbb{E}(\xi_1^+)^t}{(1-aA)^t u_{j-1}^t} = \frac{\mathbb{E}(\xi_1^+)^t A^t}{(1-aA)^t r^t} jA^{-jt}, \quad j \ge 1.$$

Consequently,

$$\begin{aligned}(1-q_0(r))\mathbb{E}e^{\lambda Y_0} &= \sum_{j=1}^{\infty} q_j(r)e^{\lambda j} \\ &\le \frac{\mathbb{E}(\xi_1^+)^t e^{\lambda}}{(1-aA)^t r^t}\sum_{j=1}^{\infty} j(e^{\lambda}A^{-t})^{j-1} = \frac{\mathbb{E}(\xi_1^+)^t e^{\lambda}}{(1-aA)^t r^t}\left(1-e^{\lambda}A^{-t}\right)^{-2}.\end{aligned}$$

For all t such that $A^t \ge 2e^{\lambda}$, we have

$$(1-q_0(r))\mathbb{E}e^{\lambda Y_0} \le 4\frac{\mathbb{E}(\xi_1^+)^t e^{\lambda}}{(1-aA)^t r^t}.$$

As a result,

$$(1-q_0(r))\mathbb{E}e^{\lambda Y_0} \le \frac{1}{2}$$

for all r large enough. Combining this estimate with (34), we obtain

$$(1-q_0(r))\mathbb{E}e^{\lambda\tau_0^{(q)}} \le 1.$$

Plugging this into (33) leads to

$$\mathbb{E}e^{\lambda\tau_0} \le 2\mathbb{E}e^{\lambda Y_0}.$$

The stochastic monotonicity of X_n implies that $T_r \le \tau_0$ in distribution. Combining this with the fact that $Y_0 \le \frac{\log(X_0/r)}{\log A}+1$, we obtain the desired inequality for the chain X_n. □

Corollary 13 *Under the assumptions of Lemma 12, there exist $r > 0$ and $\varepsilon > 0$ such that, for all $\lambda \le \lambda_a + 2\varepsilon$,*

$$\sup_{x\in[0,r]} \mathbb{E}_x\left[e^{\lambda\sigma_r}\mathbf{1}\{T_0 > \sigma_r\}\mathbb{E}_{X_{\sigma_r}}[e^{\lambda T_r}]\right] < \infty.$$

Proof By the Markov property at time σ_r and by Lemma 12,

$$\begin{aligned}
&\mathbb{E}_x\left[\mathrm{e}^{\lambda\sigma_r}\mathbf{1}\{T_0>\sigma_r\}\mathbb{E}_{X_{\sigma_r}}[\mathrm{e}^{\lambda T_r}]\right]\\
&\quad=\sum_{k=1}^{\infty}\mathrm{e}^{\lambda k}\int_r^{\infty}\mathbb{P}_x(T_o>\sigma_r=k,\,X_{\sigma_r}\in dy)\mathbb{E}_y[\mathrm{e}^{\lambda T_r}]\\
&\quad\le\frac{2\mathrm{e}^{\lambda}}{r^{\lambda/\log A}}\sum_{k=1}^{\infty}\mathrm{e}^{\lambda k}\int_r^{\infty}\mathbb{P}_x(T_o>\sigma_r=k,\,X_{\sigma_r}\in dy)y^{\lambda/\log A}. \qquad (35)
\end{aligned}$$

Noting that

$$\mathbb{P}_x(T_0>\sigma_r=k,\,X_{\sigma_r}>z)\le\mathbb{P}_x(T_0>\sigma_r>k-1)\mathbb{P}(\xi_1>(1-a)z),\quad z>r, \qquad (36)$$

we obtain

$$\mathbb{E}_x\left[\mathrm{e}^{\lambda\sigma_r}\mathbf{1}\{T_0>\sigma_r\}\mathbb{E}_{X_{\sigma_r}}[\mathrm{e}^{\lambda T_r}]\right]\le C(\lambda,A,r)\sum_{k=1}^{\infty}\mathrm{e}^{\lambda k}\mathbb{P}_x(T_0>\sigma_r>k-1).$$

The desired bound follows now from Lemma 11. □

5.2 Properties of F_λ and K_λ

We start this subsection by stating properties of the function F_λ that are important for our approach.

Lemma 14 *For every complex z with real part $\Re(z)\le\lambda_a+\varepsilon$, the function F_z defines a continuous function on $[0,r]$, which is strictly positive if additionally $z\in\mathbb{R}$.*

Proof We first show that F_z is well defined for all z with $\Re(z)<\lambda_a+\varepsilon$. Indeed, it follows from Lemma 11 that

$$\left|\mathbb{E}_x\left[\mathrm{e}^{zT_0};\,T_0<\sigma_r\right]\right|\le\mathbb{E}_x\left[\mathrm{e}^{\Re(z)T_0\wedge\sigma_r}\right]\le C$$

uniformly in $x\in[0,r]$ and in z with $\Re(z)<\lambda_a+\varepsilon$ for every $\varepsilon<\varepsilon_r$. Applying Corollary 13, we also conclude that

$$\begin{aligned}
&\left|\mathbb{E}_x\left[\mathrm{e}^{z\sigma_r}\mathbf{1}_{\{T_0>\sigma_r\}}\mathbb{E}_{X_{\sigma_r}}\left[\mathrm{e}^{zT_0};\,T_r=T_0\right]\right]\right|\\
&\quad\le\mathbb{E}_x\left[\mathrm{e}^{\Re(z)\sigma_r}\mathbf{1}_{\{T_0>\sigma_r\}}\mathbb{E}_{X_{\sigma_r}}\left[\mathrm{e}^{\Re(z)T_r}\right]\right]\le C
\end{aligned}$$

uniformly in $x\in[0,r]$ and in z with $\Re(z)<\lambda_a+\varepsilon$. Therefore, $F_z(x)$ is bounded for all $x\in[0,r]$ and all z with $\Re(z)<\lambda_a+\varepsilon$.

It is also clear that, for all $x \in [0, r]$,

$$F_\lambda(x) \geq \mathbb{E}_x\left[e^{\lambda(T_0\wedge\sigma_r)}\right] \geq 1, \quad \lambda \in [0, \lambda_a + \varepsilon).$$

Thus, it remains to show the continuity of F_z. Fix some $N \geq 1$. Since the innovations have an absolutely continuous distribution, the probabilities $\mathbb{P}_x(\sigma_r > T_0 = k)$ are continuous in x. As a result, $\sum_{k=1}^{N} e^{zk}\mathbb{P}_x(\sigma_r > T_0 = k)$ is continuous in x. Noting that, according to Lemma 11,

$$\max_{x\in[0,r]}\left|\mathbb{E}_x\left[e^{zT_0}\mathbf{1}_{\{\sigma_r>T_0>N\}}\right]\right| \to 0 \quad \text{as } N \to \infty,$$

we infer that $\mathbb{E}_x\left[e^{zT_0}; T_0 < \sigma_r\right]$ is continuous on $[0, r]$.

Using the continuity of the distribution of innovations once again, we conclude that

$$\mathbb{E}_x\left[e^{z\sigma_r}\mathbf{1}_{\{\sigma_r<T_0\wedge N\}}\mathbb{E}_{X_{\sigma_r}}[e^{zT_r}\mathbf{1}_{\{T_r=T_0<N\}}]\right]$$

is continuous in $x \in [0, r]$. Therefore, it remains to show that, as $N \to \infty$,

$$\sup_{x\in[0,r]}\left|\mathbb{E}_x\left[e^{z\sigma_r}\mathbf{1}_{\{T_0>\sigma_r\geq N\}}\mathbb{E}_{X_{\sigma_r}}[e^{zT_r}]\right]\right| \to 0 \tag{37}$$

and

$$\left|\mathbb{E}_x\left[e^{z\sigma_r}\mathbb{E}_{X_{\sigma_r}}[e^{zT_r}\mathbf{1}_{\{T_r\geq N\}}]\right]\right| \to 0. \tag{38}$$

Similar to the derivation of (35),

$$\begin{aligned}
&\left|\mathbb{E}_x\left[e^{z\sigma_r}\mathbf{1}_{\{T_0>\sigma_r\geq N\}}\mathbb{E}_{X_{\sigma_r}}[e^{zT_r}]\right]\right| \\
&\quad\leq \mathbb{E}_x\left[e^{\Re(z)\sigma_r}\mathbf{1}_{\{T_0>\sigma_r\geq N\}}\mathbb{E}_{X_{\sigma_r}}[e^{\Re(z)T_r}]\right] \\
&\quad\leq \frac{2e^{\Re(z)}}{r^{\Re(z)/\log A}}\sum_{k=N}^{\infty} e^{\Re(z)k}\int_r^\infty \mathbb{P}_x(T_o > \sigma_r = k, X_{\sigma_r} \in dy)y^{\Re(z)/\log A}.
\end{aligned}$$

Using now (36), one gets

$$\left|\mathbb{E}_x\left[e^{z\sigma_r}\mathbf{1}_{\{T_0>\sigma_r\geq N\}}\mathbb{E}_{X_{\sigma_r}}[e^{zT_r}]\right]\right| \leq C(\Re(z), A, r)\sum_{k=N}^{\infty} e^{\Re(z)k}\mathbb{P}(T_0 > \sigma_r > k-1).$$

From this bound and Lemma 11, we infer that (37) is valid for all z with $\Re(z) < \lambda_a + 2\varepsilon$.

Applying the Cauchy-Schwarz inequality, we obtain

$$\left|\mathbb{E}_y\left[e^{zT_r}\mathbf{1}_{\{T_r\geq N\}}\right]\right| \leq \left(\mathbb{E}_y\left[e^{2\Re(z)T_r}\right]\right)^{1/2}\mathbb{P}_y^{1/2}(T_r \geq N).$$

From this bound and Lemma 12, we get

$$\left|\mathbb{E}_x\left[\mathrm{e}^{z\sigma_r}\mathbb{E}_{X_{\sigma_r}}\left[\mathrm{e}^{zT_r}\mathbf{1}_{\{T_r\geq N\}}\right]\right]\right| \leq C\mathbb{E}_x\left[\mathrm{e}^{\Re(z)\sigma_r}X_{\sigma_r}^{\Re(z)/\log A}\mathbb{P}^{1/2}_{X_{\sigma_r}}(T_r \geq N)\right]$$

Since $\lim_{N\to\infty}\mathbb{P}^{1/2}_{X_{\sigma_r}}(T_r \geq N)$ almost surely and the family $\mathrm{e}^{\Re(z)\sigma_r}X_{\sigma_r}^{\Re(z)/\log A}$, $\Re(z) \leq \lambda_a + \varepsilon$ is uniformly integrable, we conclude that (38) is valid. □

We now establish an essential further property of the family of kernels K_λ, which exactly is the reason for introducing them.

Lemma 15 *Consider $K_\lambda f := \int_0^r K_\lambda(\cdot, dy)f(y)$, where r is as in Lemma 12. Then, under the assumptions of Theorem 10, for all λ with $\Re(\lambda) \leq \lambda_a + \varepsilon$, K_λ is a compact operator on the Banach space $X = C([0,r])$ equipped with the supremum norm.*

Furthermore, if $\lambda \in [0, \lambda_a + \varepsilon)$, then for every $X \ni f > 0$ (i.e. everywhere non-negative and somewhere strictly positive) and all $x \in [0,r]$, $K_\lambda f(x) > 0$.

Proof We have to show that, for fixed λ, $\{K_\lambda f : f \in C([0,r]), \|f\| \leq 1\}$ is equicontinuous. This will imply that K_λ maps to $C([0,r])$, in particular $\|K_\lambda 1\| < \infty$, and the proof can then be concluded by noting that $\sup\{\|K_\lambda f\| : f \in C([0,r]), \|f\| \leq 1\} = \|K_\lambda 1\| < \infty$, which, together with equicontinuity, yields compactness.

For equicontinuity, one has to bring the smoothing effect of the density into play. We write

$$\Phi_{x,k,l}(x_1,\ldots,x_{k+l}) := \varphi(x_1 - ax)\cdots\varphi(x_{k+l} - ax_{k+l-1})$$

for the common density of $X_1,\ldots,X_{k+l}$ and write

$$K_\lambda f(x) = \sum_{k,l=1}^{N} I_{f,k,l}(x) + R_{f,N,1}(x) + R_{f,N,2}(x)$$

and

$$\begin{aligned}&|K_\lambda f(y) - K_\lambda f(x)|\\ &\quad\leq \sum_{k,l=1}^{N}|I_{f,k,l}(y) - I_{f,k,l}(x)| + 2\sup_{x\in[0,r]}\big(|R_{1,N,1}(x)| + |R_{1,N,2}(x)|\big)\end{aligned} \tag{39}$$

with

$$\begin{aligned}I_{f,k,l}(x) &= \mathrm{e}^{\lambda(k+l)}\mathbb{E}_x\big[\mathbf{1}_{X_1,\ldots,X_{k-1}\in]0,r[,\ X_k\geq r,\ X_{k+1},\ldots,X_{k+l-1}>r,\ X_{k+l}\in]0,r]}f(X_{k+l})\big]\\ &= \mathrm{e}^{\lambda(k+l)}\int_0^r\cdots\int_0^r\int_r^\infty\cdots\int_r^\infty\int_0^r \Phi_{x,k,l}(x_1,\ldots,x_{k+l})f(x_{k+l})\mathrm{d}x_{k+l}\cdots\mathrm{d}x_1\,,\\ R_{f,N,1}(x) &= \mathbb{E}_x\big[\mathrm{e}^{\lambda\sigma_r}\mathbf{1}_{N<\sigma_r<T_0}\mathbb{E}_{X_{\sigma_r}}\big[\mathrm{e}^{\lambda T_r}\mathbf{1}_{T_r<T_0}f(X_{T_r})\big]\big]\end{aligned}$$

and

$$R_{f,N,2}(x) = \mathbb{E}_x\Big[\mathrm{e}^{\lambda\sigma_r}\mathbf{1}_{\{\sigma_r<T_0,\sigma_r\le N\}}\mathbb{E}_{X_{\sigma_r}}\Big[\mathrm{e}^{\lambda T_r}\mathbf{1}_{\{N<T_r<T_0\}}f(X_{T_r})\Big]\Big].$$

Since

$$\Phi_{y,k,l}(x_1,\ldots,x_{k+l}) = \Phi_{x,k,l}(x_1-a(y-x),\ldots,x_{k+l}-a^{k+l}(y-x)),$$

we can conclude

$$|I_{f,k,l}(y)-I_{f,k,l}(x)| \le \int_0^\infty\ldots\int_0^\infty |\Phi_{y,k,l}-\Phi_{x,k,l}|\mathrm{d}(x_1,\ldots,x_{k+l}) \xrightarrow{y\to x} 0,$$

independently of f. (This is clear for continuous Φ with compact support and follows easily for general Φ if one approximates them in L^1 by such ones.)

It follows from (37) and (38) that

$$\lim_{N\to\infty}\sup_{x\in[0,r]}|R_{1,N,1}(x)| = 0$$

and

$$\lim_{N\to\infty}\sup_{x\in[0,r]}|R_{1,N,2}(x)| = 0.$$

So, if we go back to Eq. (39), choose first N large and then y sufficiently close to x, equicontinuity follows.

We now prove the positivity of K_λ for real values of λ. If $f(z)\ge\varepsilon>0$ for $z\in I:=[z_0-\delta,z_0+\delta]$, then

$$K_\lambda f(x) \ge \varepsilon\mathbb{E}_x\left[\mathbf{1}_{\{T_0>\sigma_r\}}\mathbb{P}_{X_{\sigma_r}}\left(T_r<T_0,\,X_{T_r}\in I\right)\right].$$

The expression is bounded from below by $\varepsilon\mathbb{P}_x(X_1\in(r,r+\delta],X_2\in I)$, which is positive if the density φ is positive on all of $\mathbb{R}$. □

5.3 Fredholm Alternative and the Proof of Theorem 10

We will now make use of the so-called analytic Fredholm alternative. For the convenience of the reader, we formulate a suitable version of this result.

Theorem B (Theorem 1 in [22]). *Let $(X,\|\cdot\|_X)$ be a complex Banach space and let $\mathcal{B}(X)$ denote the space of bounded linear endomorphisms. Let D be an open connected subset of $\mathbb{C}$. Let $A: D\mapsto\mathcal{B}(X)$ be an operator-valued analytic function such that $A(z)$ is a compact operator for every $z\in D$. Then either:*

(i) *$(I-A(z))^{-1}$ does not exist for any $z\in D$, or*
(ii) *$(I-A(z))^{-1}$ exists for all $z\in D\backslash S$, where $S\subset D$ is discrete.*

We are now going to apply this result in order to deduce the subsequent proposition:

Proposition 16 *For $z \in \mathbb{C}$ with $\Re z < \lambda_a + \varepsilon$, the operator-valued function*

$$\{z \in \mathbb{C} \mid \Re z < \lambda_a + \varepsilon\} \ni z \mapsto R_z := (I - K_z)^{-1}$$

is meromorphic.

Proof It suffices to show that for some $z \in \mathbb{C}$ with $\Re z < \lambda_a + \varepsilon$ the inverse of $I - K_z$ exists and defines a bounded operator. This follows from the fact that for λ sufficiently small the operator norm of K_λ can be easily seen to be strictly smaller than one and the inverse thus can be shown to exist by the Neumann series. An application of Theorem B therefore shows that $(I - K_z)^{-1}$ is meromorphic on the required domain. □

Now, we are able to draw a conclusion analogously to Corollary 1 of [23]:

Corollary 17 *If for $z \in \mathbb{C}$ with $\Re z < \lambda_a + \varepsilon$ the function $u = u_z$ is a solution of*

$$u = F_z + K_z u,$$

then the X-valued function u is meromorphic in z. Therefore, we conclude that the function

$$z \mapsto L(x, z) := \mathbb{E}_x\left[\mathrm{e}^{zT_0}\right]$$

has a meromorphic continuation to $\{z \in \mathbb{C} \mid \Re z < \lambda_a + \varepsilon\}$.

We now aim to study the dominant poles.

Proposition 18 *The function $L(z)$ which meromorphically extends the Laplace transform of T_0 has a simple pole at $z = \lambda_a$ and no other poles on the interval $\{z = \lambda_a + i\psi,\ \psi \in (-\pi, \pi)\}$.*

Proof We first observe that

$$\mathbb{E}_x\left[z^{T_0}\right] = \sum_{n=0}^{\infty} \mathbb{P}_x(T = n) z^n$$

is a power series with non-negative coefficients and therefore by Pringsheim' s theorem (see, e.g. Theorem 4.1.2 in [16]) the radius of convergence $r_a = \mathrm{e}^{\lambda_a}$ is a singularity. Therefore, λ_a is a singularity of L and as L is meromorphic, we conclude that λ_a is a pole. Observe that this implies that the operator-valued function $(I - K_z)^{-1}$ has a pole at $z = \lambda_a$. Therefore, we conclude that K_{λ_a} has the eigenvalue 1.

Observe next that for $\lambda < \lambda_a + \varepsilon$ the operator K_λ is positive and compact on the Banach lattice X. Using the second part of Lemma 15, we conclude by the classical Krein–Rutman theorem as given in Theorem 4 of [20] that the spectral radius $r(K_\lambda) > 0$ is in fact an eigenvalue with algebraic multiplicity one and the associated eigenfunction can be chosen to be non-negative. Furthermore, by Theorem 2.1 in [10] the spectral radius is continuous in λ.

We now claim that 1 coincides with the spectral radius of K_{λ_a}. Assume contrary that the spectral radius $r(K_{\lambda_a})$ of K_{λ_a} is strictly bigger than 1. Under this assumption and the continuity of the spectral radius, we conclude that for $\lambda < \lambda_a$ we still have $\rho(K_\lambda) > 1$. We now claim that this contradicts that the equation

$$L_\lambda = F_\lambda + K_\lambda L_\lambda \tag{40}$$

holds. Observe that by Lemma 14 the function F_λ is lower-bounded by 1 and iterating equation (40) we conclude that for $n \geq 1$

$$L_\lambda \geq K_\lambda^n \mathbf{1}.$$

This contradicts the fact $\|K_\lambda \mathbf{1}\|_\infty \to \infty$ if $n \to \infty$ as a consequence of the spectral radius formula $\rho(K_\lambda) = \lim_{n\to\infty} \|K_\lambda^n\|^{1/n}$.

We now aim to show that the function L has a simple pole at λ_a. Here, we can observe that the derivative $\frac{\mathrm{d}}{\mathrm{d}\lambda} K_\lambda$ for real $\lambda < \lambda_a + \varepsilon$ defines a positive operator on the Banach space X. Therefore, we have shown that the properties $P1$, $P2$ and $P3$ in [23] are satisfied and, as shown on page 233 of [23], we can then conclude via Corollary 1 in [23] that the pole at λ_a is in fact simple.

Now assume that $L(z)$ has another pole at $z = \lambda_a + i\psi$, i.e. K_z has an eigenvalue 1 with some eigenfunction g there. Applying the triangle inequality to the definition of $K_z g$, one obtains $|g(x)| = |K_z g(x)| \leq K_{\lambda_a}|g(x)|$ for all $x \in [0, r]$. The positivity of the density φ implies that, for any $m, n \in \mathbb{N}$, it happens with positive probability that $\sigma_r = m$, $T_r = n$ and the whole expression inside the expectation value defining K_z is nonzero. Consequently, its phase is not a.s. constant unless ψ is a multiple of 2π, the triangle inequality is even strict and, by compactness of $[0, r]$,

$$\inf_{x\in[0,r]} \frac{K_{\lambda_a}|g(x)|}{|g(x)|} > 1\,.$$

Now, we look at K_{λ_a} as an operator on the real-valued continuous functions $C([0, r])$. Since $r(K_{\lambda_a}) = 1$, what we just found would mean that the min-max principle in the form

$$r(K_{\lambda_a}) = \max_{h\in\mathcal{K}} \inf_{\substack{h'\in\mathcal{H}' \\ \langle h,h'\rangle>0}} \frac{\langle K_{\lambda_a} h, h'\rangle}{\langle h, h'\rangle}$$

with $\mathcal{K}$ denoting the non-negative elements of $C([0, r])$ and $\mathcal{H}' := \{h \mapsto h(x) \mid x \in [0, r]\}$ is violated. However, all conditions from [15] ensuring its validity are satisfied:

- $C([0, r]) = \mathcal{K} - \mathcal{K}$ and the cone $\mathcal{K}$ is closed and normal ($\|g + h\| \geq \|g\|$ for all normalized $g, h \in \mathcal{K}$).
- $\mathcal{H}'$ is total, i.e. if $\langle h, h'\rangle \geq 0$ for all $h' \in \mathcal{H}'$, then $h \in \mathcal{K}$.
- K_{λ_a} is semi non-supporting: For all nonzero $h \in \mathcal{K}$, $K_{\lambda_a} h(x) > 0$ for all x as in Lemma 15 and therefore one can conclude for all nonzero $h' \in \mathcal{K}'$ (the dual cone, i.e. positive measures) $\langle K_{\lambda_a} h, h'\rangle > 0$.

- The resolvent $\lambda \mapsto (\lambda - K_{\lambda_a})^{-1}$ has only finitely many singularities with $|\lambda| = r(K_{\lambda_a})$, all of them poles, because K_{λ_a} is compact.

$\square$

Completion of the proof of Theorem 10 It follows from the Proposition 18 that the generating function $\mathbb{E}_x[z^{T_0}]$ has the unique simple pole at e^{λ_a} and that there are no further poles on the disc or radius $e^{\lambda_a+\gamma}$ with some $\gamma > 0$. Then, we have the following representation

$$\mathbb{E}_x[z^{T_0}] = \frac{V(x)}{e^{\lambda_a} - z} + g_x(z),$$

where g_x is analytical on the disc with radius $e^{\lambda_a+\gamma}$. Therefore,

$$\mathbb{P}_x(T_0 = n) = V(x)e^{-\lambda_a(n+1)} + O(e^{-(\lambda_a+\gamma)n}). \tag{41}$$

Since the left-hand side is positive, we infer that the function V is positive as well. Therefore, it remains to show that v is e^{λ_a}-harmonic for P_+. Let r be so large that $\mathbb{E}_y[e^{\lambda_a T_r}]$ is finite. For every $x \geq r$, one has the inequality

$$\mathbb{P}_x(T_0 \geq n) \leq \sum_{m=1}^{n} \mathbb{P}_x(T_r = m)\mathbb{P}_r(T_0 \geq n - m).$$

It follows from (41) that $\mathbb{P}_r(T_0 \geq k) \leq C(r)e^{-\lambda_a(k+1)}$, $k \geq 0$. Thus,

$$\mathbb{P}_x(T_0 \geq n) \leq C(r)e^{-\lambda_a n} \sum_{m=1}^{n} \mathbb{P}_x(T_r = m) \leq C(r)\mathbb{E}_x[e^{\lambda_a T_r}]e^{-\lambda_a(n+1)}. \tag{42}$$

It is immediate from (41) that

$$\mathbb{P}_x(T_0 \geq n) = \frac{V(x)}{1 - e^{-\lambda_a}} e^{-\lambda_a(n+1)} + O\left(e^{-(\lambda_a+\gamma)n}\right). \tag{43}$$

From this equality and from (42), we infer that

$$V(x) \leq \frac{C(r)}{1 - e^{-\lambda_a}} \mathbb{E}_x[e^{\lambda_a T_r}], \quad x \geq r. \tag{44}$$

Since

$$\mathbb{E}_x[e^{\lambda_a T_r}] = \int_0^r P(x, dy)e^{\lambda_a} + \int_r^\infty P(x, dy)e^{\lambda_a}\mathbb{E}_y[e^{\lambda_a T_r}],$$

we infer that

$$\int_r^\infty P(x, dy)\mathbb{E}_y[e^{\lambda_a T_r}] < \infty \tag{45}$$

and, in view of (44),

$$\int_0^\infty P(x,dy)V(y) < \infty. \tag{46}$$

Fix some $A > r$ and consider the equality

$$\begin{aligned}\mathbb{P}_x(T_0 \geq n+1) &= \int_0^\infty P(x,dy)\mathbb{P}_y(T_0 \geq n) \\ &= \int_0^A P(x,dy)\mathbb{P}_y(T_0 \geq n) + \int_A^\infty P(x,dy)\mathbb{P}_y(T_0 \geq n). \end{aligned} \tag{47}$$

Combining (42) and (45), we obtain

$$\lim_{A\to\infty} \limsup_{n\to\infty} \mathrm{e}^{\lambda_a(n+1)} \int_A^\infty P(x,dy)\mathbb{P}_y(T_0 \geq n) = 0. \tag{48}$$

Furthermore, by (43) and (46),

$$\lim_{A\to\infty} \lim_{n\to\infty} \mathrm{e}^{\lambda_a(n+1)} \int_0^A P(x,dy)\mathbb{P}_y(T_0 \geq n) = \int_0^\infty P(x,dy)V(y). \tag{49}$$

Plugging (48) and (49) into (47), we obtain

$$\lim_{n\to\infty} \mathrm{e}^{\lambda_a(n+1)}\mathbb{P}_x(T_0 \geq n+1) = \int_0^\infty P(x,dy)V(y).$$

According to (43), the limit on the left-hand side is equal to $\mathrm{e}^{-\lambda_a}V(x)$. As a result, we have the equality

$$\mathrm{e}^{-\lambda_a}V(x) = \int_0^\infty P(x,dy)V(y).$$

Therefore, the proof is complete. □

6 Innovations with Regularly Varying Tails

The main purpose of this section is to show that the finiteness of all moments of ξ_1^+ is necessary for getting purely exponential decay for the tail of T_0. More precisely, we are going to show that if the right tail is regularly varying, then, independent of the index of regular variation, the asymptotic behaviour of $\mathbb{P}_x(T_0 > n)$ depends on the slowly varying component of $\mathbb{P}(\xi_1 > x)$. In particular, it may happen that $\mathrm{e}^{\lambda_a n}\mathbb{P}_x(T_0 > n) \to 0$ as $n \to \infty$.

Proposition 19 *Assume that*

$$\mathbb{P}(\xi_1 > x) = x^{-r}L(x) \tag{50}$$

for some $r > 0$ and some slowly varying function L. Then, for all $M \geq 0$,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(T_M > n) \geq r \log a. \tag{51}$$

If, in addition,

$$L(x) = O(\log^{-2r-2} x), \tag{52}$$

then, for all M sufficiently large,

$$\lim_{n\to\infty} a^{-rn}\mathbb{P}(T_M > n) = 0. \tag{53}$$

In fact, one even has

$$\sum_{n=0}^{\infty} a^{-rn}\mathbb{P}(T_M > n) < \infty.$$

Remark 20 It is immediate from (51) and (53) that

$$\lambda_a = -r \log a$$

for innovations satisfying (50) and (52). We conjecture that the same holds under the assumption (50) and that the asymptotic behaviour of the slowly varying function L affects lower-order corrections only. □

Proof of Proposition 19 Fix some $A > a^{-1}$ and define $u_k = MA^k$, $k \geq 0$. Then,

$$\mathbb{P}_{u_k}(X_1 \geq u_{k-1}) = \mathbb{P}(au_k + \xi_1 > u_{k-1}) = \mathbb{P}(\xi_1 > -(aA-1)u_{k-1}).$$

Therefore, using the Markov property and the stochastic monotonicity of X_n, we get

$$\begin{aligned}\mathbb{P}_{u_k}(T_M > k-1) &\geq \mathbb{P}_{u_k}(X_1 \geq u_{k-1}, X_2 > u_{k-2}, \ldots, X_{k-1} \geq u_1)\\ &= \prod_{j=2}^{k} \mathbb{P}_{u_j}(X_1 \geq u_{j-1}) = \prod_{j=1}^{k-1} \mathbb{P}(\xi_1 > -(aA-1)u_j).\end{aligned}$$

It follows now from the assumption $\mathbb{E}\log(1 + |\xi_1|) < \infty$ that

$$\inf_{k\geq 2} \prod_{j=1}^{k-1} \mathbb{P}(\xi_1 > -(aA-1)u_j) =: p(M, A) > 0.$$

This implies that

$$\mathbb{P}_M(T_M > k) \geq p(M, A)\mathbb{P}(\xi_1 > u_k).$$

Since the tail of ξ_1 is regularly varying of index $-r$, we obtain

$$\liminf_{k\to\infty} \frac{1}{k} \log \mathbb{P}_M(T_M > k) \geq -r \lim_{k\to\infty} \log u_k = -r \log A.$$

Letting now $A \downarrow a^{-1}$, we arrive at (51).

In order to prove the second statement, we consider events

$$B_n := \left\{ \xi_k \leq h \frac{a^{-(n-k+1)}}{(n-k+1)^2} \text{ for all } k < T_M \wedge n \right\}.$$

On the event $\{T_M > n\} \cap B_n$, one has

$$X_n = a^n M + \sum_{k=1}^{n} a^{n-k} \xi_k \leq a^n M + \frac{h}{a} \sum_{k=1}^{n} \frac{1}{(n-k+1)^2} \leq aM + \frac{2h}{a}.$$

Thus, for every $M \geq \frac{2}{a(1-a)} h$ and all $n \geq 1$ one has $X_n \leq M$ and, consequently, $\{T_M > n\} \cap B_n = \emptyset$. This implies that

$$\begin{aligned}\mathbb{P}_M(T_M > n) &= \mathbb{P}_M(T_M > n;\, B_n^c)\\ &\leq \sum_{k=1}^{n} \mathbb{P}_M(T_M > k-1) \mathbb{P}\left(\xi_k > h \frac{a^{-(n-k+1)}}{(n-k+1)^2} \right).\end{aligned}$$

From the assumption 52 on $L(x)$, we get

$$\mathbb{P}\left(\xi_k > h \frac{a^{-(n-k+1)}}{(n-k+1)^2} \right) \leq \frac{c(a)}{h^r} \frac{a^{r(n-k+1)}}{(n-k+1)^2}.$$

Therefore,

$$\mathbb{P}_M(T_M > n) \leq \frac{c(a)}{h^r} \sum_{j=0}^{n-1} \mathbb{P}_M(T_M > j) \frac{a^{r(n-j)}}{(n-j)^2}, \quad n \geq 1.$$

Multiplying both sides with s^n and summing over all n, we obtain

$$\begin{aligned}\sum_{n=0}^{\infty} s^n \mathbb{P}_M(T_M > n) &\leq 1 + \sum_{n=1}^{\infty} s^n \frac{c(a)}{h^r} \sum_{j=0}^{n-1} \mathbb{P}_M(T_M > j) \frac{a^{r(n-j)}}{(n-j)^2}\\ &= 1 + \frac{c(a)}{h^r} \sum_{j=0}^{\infty} s^j \mathbb{P}_M(T_M > j) \sum_{n=j+1}^{\infty} \frac{(sa^r)^{n-j}}{(n-j)^2}\\ &= 1 + \frac{c(a)}{h^r} \sum_{j=0}^{\infty} s^j \mathbb{P}_M(T_M > j) \sum_{n=1}^{\infty} \frac{(sa^r)^n}{n^2}.\end{aligned}$$

In other words,

$$\sum_{n=0}^{\infty} s^n \mathbb{P}_M(T_M > n) \leq \left(1 - \frac{c(a)}{h^r} \sum_{n=1}^{\infty} \frac{(sa^r)^n}{n^2}\right)^{-1}.$$

If h is so large that $h^r > 4c(a)$, then we have

$$\sum_{n=0}^{\infty} a^{-rn} \mathbb{P}_M(T_M > n) \leq 2.$$

This yields (53). □

In order to see that spectral properties remain relevant in the situation at hand, we observe the following assertion formulated in terms of Tweedie's **R**-theory (see [24] and [25]).

Corollary 21 *Assume that the conditions* (50) *and* (52) *are satisfied. Then, for r large enough the operator K_{λ_a} is well defined and the spectral radius $r(K_{\lambda_a})$ belongs to* $[0, 1]$. *Under the assumption $r(K_{\lambda_a}) < 1$, the Laplace transform of T_0 remains bounded up to the critical line. In particular, the submarkovian transition operator P is R-transient in the sense of Tweedie.*

Until now, a complete analysis of persistence probabilities including the effects of polynomial decay factors as well as thorough investigation of the quasistationary behaviour of AR(1) sequences with heavy tailed innovations does not seem to exist and constitutes an interesting open problem.

7 Discussion

In this section, we summarize the general ideas of the two main analytic approaches used in this work and comment on their possible applicability to other models. As a matter of fact, both approaches use different tools but also share some structural similarities. For a Markov process $(X_n)_{n\in\mathbb{N}_0}$, we want to find the precise tail behaviour of the first hitting time T_B of a measurable subset B of the state space.

- The first approach is in essence spectral theoretic. We first analyse the decay rate

$$\theta_B = -\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}_x\big(T_B > n\big)$$

  and make sure that θ_B does not depend on the starting point. The second step consists in showing that for a strictly larger measurable set $B' \supset B$ one has

$$\theta_B < \theta_{B'}\,. \tag{54}$$

  This allows to introduce a suitable weighted Banach space and under appropriate conditions on the distribution of the innovations to prove the quasicompactness of the killed transition kernel. Application of a suitable result of Perron–Frobenius type allows to establish precise exponential decay of the tails of T_B.

- The second approach consists in analysing the Laplace transform

$$\{z \in \mathbb{C} \mid \Re z < \theta_B\} \ni \lambda \mapsto F_\lambda(x) := \mathbb{E}_x\left[e^{\lambda T_B}\right]$$

near the abscissa of convergence θ_B. We prove that $F_\lambda(x)$ has a meromorphic extension to $\{z \in \mathbb{C} \mid \Re z < \theta_B + \varepsilon\}$ for some $\varepsilon > 0$ with λ_B being a pole. This allows to deduce the precise exponential decay using a suitable Tauberian theorem. In order to prove the existence of a meromorphic extension, we derive a renewal-type equation in terms of the transition kernel K_λ. The existence of a meromorphic extension is shown using the analytic Fredholm alternative and suitable properties of K_λ. In particular, we need to show that the operators K_λ are compact for all λ with $\Re\lambda < \lambda_a + \varepsilon$ and satisfy the conditions of a suitable Perron–Frobenius theorem. Here, results of the type given in (54) are again essential as well as absolute continuity and strict positivity of the transition kernel.

We believe that these methods can be applied without big changes to some other Markov chain models. For example, to max-autoregressive processes or to random exchange processes, which are quite closely related to AR(1)-sequences, see [26].

In both approaches, we have assumed that the underlying distributions are absolutely continuous with almost everywhere positive density and have sufficiently light tails. The positivity on the whole axis has been used only to give very simple proofs of the positivity of compact operators. We believe that this positivity can be shown also in the case when the density is zero on significant parts of the axis. Probably, one will have the positivity of an appropriate power of the operator, but this does not restrict the applicability of Perron–Frobenius type results. The absolute continuity assumption seems to be really crucial for both approaches, and it is not clear whether one can replace it by the existence of an absolute continuous component. A very challenging problem is to study the case of discrete innovations.

The second approach might also be applicable to chains with a certain periodicity, presumably resulting in poles of higher order. A different behaviour of the pole(s) also can be expected in the presence of distributions with regularly varying tails. Getting completely rid of absolute continuity seems so be much more tricky. Moreover, the stochastic monotonicity of AR(1)-sequences was extensively used, and a lack of this property is expected to result in largely technical complications.

The authors of [6] are able to prove similar results using a different approach. Their conditions are of the following types (see page 8 in [6]):

- Local minorization of Doeblin type
- Two global Lyapunov criteria
- Local Harnack inequality
- Aperiodicity

The approach used by Champagnat and Villemonais is more related to coupling ideas of Doeblin type. The global Lypunov criteria are related to our condition (54), whereas the local minorization, the local Harnack inequality and the aperiodicity are closer to the condition used in our approaches to prove quasicompactness with a leading eigenvalue of multiplicity one in approach one and compactness of the renewal operator with a leading eigenvalue of multiplicity one in approach two. All three approaches have their merits. As far as possible extensions to AR(1)-processes with fat tailed innovations are concerned, the approach in Sect. 5 seems to be most promising.

Acknowledgements The authors would like to thank a referee for her or his useful report and for suggestion to add the final section. This project was initiated during the sabbatical stay of VW at the Technion, Israel. He thanks the Humboldt foundation and the Technion for their financial support. MK would like to thank his student P. Trykacz for critical reading of the final draft.

References

1. Aliprantis, C.D., Burkinshaw, O.: Positive Operators. Springer, Berlin (2006)
2. Aurzada, F., Mukherjee, S., Zeitouni, O.: Persistence exponents in Markov chains. ArXiv preprint: arXiv:1703.06447
3. Aurzada, F., Baumgarten, C.: Survival probabilities of weighted random walks. ALEA Lat. Am. J. Probab. Math. Stat. **8**, 235–258 (2011)
4. Baumgarten, C.: Persistence of sums of independent random variables, iterated processes and fractional Brownian motion. Doctoral Thesis, Technical University Berlin (2013)
5. Champagnat, N., Villemonais, D.: Exponential convergence to quasi-stationary distributions and Q-process. Probab. Theory Relat. Fields **164**, 243–283 (2016)
6. Champagnat, N., Villemonais, D.: General criteria for the study of quasi-stationarity. ArXiv preprint: arXiv:1712.08092
7. Christensen, S.: Phase-type distributions and optimal stopping for autoregressive processes. J. Appl. Prob. **49**, 22–39 (2012)
8. Collet, P., Martinez, S., San Martin, J.: Quasi-Stationary Distributions: Markov Chains, Diffusions and Dynamical Systems. Springer, Berlin (2013)
9. Davies, E.B.: Linear Operators and Their Spectra. Cambridge University Press, Cambridge (2007)
10. Degla, G.: An overview of semi-continuity results on the spectral radius and positivity. J. Math. Anal. Appl. **338**, 101–110 (2008)
11. Gosselin, F.: Aysmptotic behavior of absorbing markov chains conditional on nonabsorption for applications in conservation biology. Ann. Appl. Probab. **11**, 261–284 (2001)
12. Hennion, H., Hervé, L.: Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-Compactness. Lecture Notes in Mathematics. Springer, Berlin (2000)
13. Lalley, S.P.: Return probabilities for random walk on a half-line. J. Theor. Probab. **8**, 571–599 (1995)
14. Larralde, H.: A first passage time distribution for a discrete version of the Ornstein–Uhlenbeck process. J. Phys. A **37**, 3759–3767 (2004)
15. Marek, I.: Spektrale Eigenschaften der $\mathcal{K}$-positiven Operatoren und Einschließungssätze für den Spektralradius. Czechoslovak Math. J. **16**, 493–517 (1966)
16. Meyer-Nieberg, P.: Banach Lattices. Springer, Universitätstext, Berlin (1991)
17. Novikov, A.: Some remarks on the distribution of the first passage times and the optimal stopping of AR(1)-sequences. Theory Probab. Appl. **53**, 419–429 (2009)
18. Novikov, A., Kordzakhia, N.: Martingales and first passage times of AR(1) sequences. Stochastics **80**, 197–210 (2008)
19. Revuz, D.: Markov Chains. North Holland Mathematical Library, Amsterdam (1984)
20. Sasser, D.W.: Quasi-positive operators. Pac. J. Math. **14**, 1029–1037 (1964)
21. Schechter, M.: Principles of Functional Analysis. Academic Press, New York (1971)
22. Steinberg, S.: Meromorphic families of compact operators. Arch. Rational Mech. Anal. **31**, 372–380 (1968)
23. Shumitzky, A., Wenska, T.: An operator residue theorem with applications to branching processes and renewal type integral equations. SIAM J. Math. Anal. **6**, 229–235 (1975)
24. Tweedie, R.L.: R-theory for Markov chains on a general spate space I, solidítary properties and R-recurrent chains. Ann. Probab. **2**, 840–864 (1974)

25. Tweedie, R.L.: R-theory for Markov chains on a general spate space II, r-subinvariant measures and r-transient chains. Ann. Probab. **2**, 865–878 (1974)
26. Zerner, M.P.W.: Recurrence and transience of contractive autoregressive processes and related Markov chains. Electron. J. Probab. **23**, 1–24 (2018)