

On geometric recurrence for time-inhomogeneous autoregression

Vitaliy Golomoziy

Taras Shevchenko National University of Kyiv, Faculty of Mechanics and Mathematics, Department of Probability Theory, Statistics and Actuarial Mathematics, 60 Volodymyrska Street, City of Kyiv, Ukraine, 01033

vitaliy.golomoziy@knu.ua (V. Golomoziy)

Received: 8 October 2022, Revised: 18 March 2023, Accepted: 18 March 2023,
Published online: 3 May 2023

Abstract The time-inhomogeneous autoregressive model AR(1) is studied, which is the process of the form $X_{n+1} = \alpha_n X_n + \varepsilon_n$, where α_n are constants, and ε_n are independent random variables. Conditions on α_n and distributions of ε_n are established that guarantee the geometric recurrence of the process. This result is applied to estimate the stability of n -steps transition probabilities for two autoregressive processes $X^{(1)}$ and $X^{(2)}$ assuming that both $\alpha_n^{(i)}$, $i \in \{1, 2\}$, and distributions of $\varepsilon_n^{(i)}$, $i \in \{1, 2\}$, are close enough.

Keywords Coupling, renewal theory, inhomogeneous Markov chain, autoregressive model
2010 MSC 60J10, 60K05

1 Introduction

The classical autoregressive model $X_{n+1} = \alpha X_n + \varepsilon_n$, where α is a constant and ε_n are i.i.d. standard normal random variables, is well studied, and it is known, in particular, that for $\alpha \in (0, 1)$ the corresponding Markov chain is geometrically recurrent, positive and ergodic.

However, in real applications, we cannot always guarantee that all ε_n are standard normal i.i.d. random variables. In addition, parameter α may not always be the same for all X_n . From the theoretical standpoint, it is interesting to study the behavior of such a model when α is “oscillating” around unity and ε_n are independent but not necessarily normal and identically distributed.

© 2023 The Author(s). Published by VTeX. Open access article under the [CC BY](https://creativecommons.org/licenses/by/4.0/) license.

So, we come to the model $X_{n+1} = \alpha_n X_n + \varepsilon_n$, where α_n are some numbers “around” unity and ε_n are independent random variables. Our goal is to find the conditions which guarantee the geometric recurrence of such a chain and study its stability in terms of the stability of n -steps transition probability functions.

Usually, when studying the recurrence of Markov chains, we use techniques developed in the classical theory based on a drift condition (see, for example, [5, 24]). However, the process of interest is time-inhomogeneous, so the aforementioned techniques are not applicable. That’s why we use a particular time-inhomogeneous drift condition developed in the paper [13]. In order to establish geometric recurrence, we use results from [12].

The general theory for inhomogeneous Markov chains is much more involved than its homogeneous counterpart. One of the most popular instruments in research is the coupling method (see classical books by T. Lindvall [22] and H. Thorisson [25]). An interested reader may find an example of applying the coupling method to the stability of Markov chains on a countable phase space in both homogeneous and inhomogeneous cases in the papers [15, 16, 18, 20]. The papers [1, 3, 7] are devoted to the calculation of convergence rates for subgeometrically ergodic, homogeneous Markov chains, and [4] addresses the inhomogeneous case. All these works use the coupling method. We heavily rely on the coupling method in the present paper. We present a modified coupling construction that enables us to couple two different inhomogeneous chains. Such modified coupling generates an inhomogeneous renewal process which we use in our analysis. We can refer to the papers [14, 17] for more details about inhomogeneous renewal processes. Generally speaking, renewal processes have been used for a long time to study Markov chains. See, for example, [2]. The papers [6, 8, 9, 19, 21] are also related to the inhomogeneous case. Another interesting example of using the renewal theory for the analysis of Markov chains with applications in statistics is [23]. Since general inhomogeneous renewal theory is not well-developed, we use some techniques that are specially adapted to the problem under consideration. One of the key techniques that helps to handle inhomogeneity is stochastic domination. Building a dominating sequence is a key aspect of our development. See more details about this subject in [11] and [10].

Summarizing the methods described above, we can outline the plan for this paper. First, we use the modified drift condition from [13] to establish geometric recurrence of the autoregressive model. Then, we obtain an estimate for $\mathbb{E}_x[\beta^{\sigma_C}]$ for some $\beta > 1$, where σ_C is the first return time to some “recurrence” set C . Second, we use the result from [12] to establish a similar estimate for a pair of autoregressive processes, i.e. $\mathbb{E}_{x,y}[\beta^{\sigma_{C \times C}}]$. Third, we introduce the coupling schema for two inhomogeneous processes and employ the renewal theory for the stability analysis. Finally, we construct a domination sequence (see Lemmas 1–4) and use it to obtain the stability estimate. We will show that n -steps transition probabilities for inhomogeneous processes will be close, assuming that one-step transition probabilities are also close.

This paper is organized as follows. Section 2 contains main definitions and notations. In Section 3, we present the geometric recurrence result. Section 4 contains the main result, the stability estimate. Section 5 includes technical auxiliary lemmas. Finally, the appendix contains the example of calculating all the necessary constants in a practical application.

2 Definitions and notation

Throughout this paper, we assume that all random variables are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. In Section 3 we deal with an inhomogeneous Markov chain $(X_n)_{n \geq 0}$ with values in a general phase space (E, \mathcal{E}) . We denote the chain's transition probability by

$$P_n(x, A) = \mathbb{P}\{X_{n+1} \in A | X_n = x\}, \quad x \in E, A \in \mathcal{E}.$$

When it is convenient, we will understand P_n as an operator acting on an appropriate functional space as

$$P_n f(x) = \int_{\mathbb{R}} P_n(x, dy) f(y).$$

By $\mathbb{P}_x\{X \in \cdot\}$ we denote the probability measure generated by the process $(X_n)_{n \geq 0}$ which starts at $x \in E$. Thus, for example

$$\mathbb{P}_x\{X_n \in A, X_m \in B\} = \mathbb{P}\{X_n \in A, X_m \in B | X_0 = x\}.$$

We denote by \mathbb{E}_x the corresponding expectation. Strictly speaking, \mathbb{P}_x is defined on a measurable space $(E^\infty, \mathcal{E}^\infty)$, and for every $w = (w_0, w_1, \dots) \in E^\infty$ we have a canonical variable $X_n(w) = w_n$. Thus, rigorously speaking X_n , in the context of \mathbb{P} and \mathbb{P}_x , represents two different random variables defined on different probability spaces. However, such notation is very convenient and helps to develop the proper intuition.

Next, we need a notation for a process conditioned on $X_n = x$, i.e.

$$\mathbb{P}\{(X_{n+m})_{m \geq 0} \in \cdot | X_n = x\}.$$

Note that in the homogeneous case, this would be the same \mathbb{P}_x introduced above. But in our inhomogeneous case, we must maintain the index n in the notation since, generally speaking, probabilities

$$\mathbb{P}\{(X_{n+m})_{m \geq 0} \in \cdot | X_n = x\}, \text{ and } \mathbb{P}\{(X_{k+m})_{m \geq 0} \in \cdot | X_k = x\}$$

are not equal if $k \neq n$. At the same time, we would like to stick to the homogeneous notation as closely as possible. Thus, with some abuse of notation, we introduce probabilities $\mathbb{P}_x^n\{(X_k)_{k \geq 0} \in \cdot\}$, which mean

$$\mathbb{P}_x^n\{(X_k)_{k \geq 0} \in \cdot\} = \mathbb{P}\{(X_{n+k})_{k \geq 0} \in \cdot | X_n = x\}.$$

We will denote the corresponding expectation by \mathbb{E}_x^n . It is important to note that, unlike the homogeneous case, \mathbb{P}_x^n and \mathbb{P}_x^m are probabilities defined on a different probability spaces. Thus, we can have a ‘‘homogeneous-style’’ derivation like

$$\mathbb{E}_x[f(X_{n+k})] = \mathbb{E}_x[\mathbb{E}_{X_n}^n[f(X_k)]] .$$

In Section 4 and later we deal with a pair of inhomogeneous Markov chains $(X_n^{(1)}, X_n^{(2)})_{n \geq 0}$. In this case, we will add a subscript i to probabilities and expectations related to the chain $X^{(i)}$, $i \in \{1, 2\}$, thus having $P_{i,n}$, the transition probability,

$\mathbb{P}_{i,x}^n$ and $\mathbb{P}_{i,x}^m$, the canonical probabilities generated by the chain $X_n^{(i)}$. In case we need to operate with both chains simultaneously, we will use the notation

$$\begin{aligned} & \mathbb{P}_{x,y}^n \left\{ \left(X_m^{(1)}, X_m^{(2)} \right)_{m \geq 0} \in \cdot \right\} \\ &= \mathbb{P}_{x,y} \left\{ \left(X_{n+m}^{(1)}, X_{n+m}^{(2)} \right)_{m \geq 0} \in \cdot \mid \left(X_n^{(1)}, X_n^{(2)} \right) = (x, y) \right\}. \end{aligned}$$

As usual, we denote by $\mathbb{E}_{x,y}^n$ the corresponding expectation.

3 Geometric recurrence for inhomogeneous autoregression chains

In this section we deal with a single time-inhomogeneous Markov chain $(X_n)_{n \geq 1}$ with values in $(\mathbb{R}, \mathcal{B})$ where \mathcal{B} is a Borel σ -field.

The main result of this section is presented in the following theorem.

Theorem 1. *Consider a time-inhomogeneous Markov chain $(X_n)_{n \geq 0}$ with values in $(\mathbb{R}, \mathcal{B})$ starting from a fixed $x_0 \in \mathbb{R}$ and having the form*

$$X_{n+1} = \alpha_{n+1} X_n + W_{n+1}, \quad n \geq 0,$$

where $\alpha_n \geq 0$, W_n , $n \geq 1$, are independent and centered random variables defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Denote their distribution functions and tails by $\Gamma_n(x)$ and $\bar{\Gamma}_n(x) = 1 - \Gamma_n(x) = \int_x^\infty \Gamma_n(dy)$, respectively. Assume the following conditions hold.

1. There exists $\delta > 0$ such that

$$\sum_{k=1}^\infty \left(\prod_{j=1}^k \max\{\alpha_j + \delta, 1\} \right)^{-1} (1 - \alpha_k - \delta)^+ = \infty.$$

2. $G := \sup_{n \geq 1} \mathbb{E} W_n^+ < \infty$.

Then for all $x \in \mathbb{R}$ and for all $n \in \mathbb{N}$:

$$\mathbb{E}_x^n \left[\prod_{j=1}^{\sigma_C} \frac{1}{\alpha_j + \delta} \right] \leq 1 + |x| + (2G + 1 - \delta) \mathbb{1}_{[-c, c]}(x). \tag{1}$$

Here

$$c := \max \left\{ \frac{2G + 1}{\delta} - 1, 1 \right\}, \tag{2}$$

$\sigma_C = \inf \{n \geq 1 \mid X_n \in C\}$ (stipulating that $\inf \emptyset = \infty$) is the first return time to the set $C = [-c, c]$.

Proof. The main tool in the proof is the drift condition from the paper [13]. As a test function we will consider

$$V(x) = 1 + |x|.$$

We will use this function for all $n \geq 1$. We have then

$$\begin{aligned} P_n V(x) &= \int_{\mathbb{R}} (1 + |\alpha_n x + y|) \Gamma_n(dy) \\ &= 1 - \int_{(-\infty, -\alpha_n x]} (\alpha_n x + y) \Gamma_n(dy) + \int_{(-\alpha_n x, \infty)} (\alpha_n x + y) \Gamma_n(dy). \end{aligned}$$

For $x \geq 0$ we can derive

$$\begin{aligned} P_n V(x) &= 1 - 2 \int_{(-\infty, -\alpha_n x]} (\alpha_n x + y) \Gamma_n(dy) + \int_{\mathbb{R}} (\alpha_n x + y) \Gamma_n(dy) \\ &= 1 + \alpha_n x - 2\alpha_n x \Gamma_n(-\alpha_n x) + 2 \int_{(-\infty, -\alpha_n x]} (-y) \Gamma_n(dy) \\ &\leq \alpha_n (1 + |x|) + (1 - \alpha_n) + 2\mathbb{E}W_n^- = \alpha_n V(x) + (1 - \alpha_n) + 2G, \end{aligned}$$

where we used the fact $\mathbb{E}W_n^+ - \mathbb{E}W_n^- = \mathbb{E}W_n = 0$ and Condition 2.

Put

$$\lambda_n = \alpha_n + \delta,$$

and write

$$\begin{aligned} P_n V(x) &\leq \lambda_n V(x) + (1 - \alpha_n) + 2G - (\lambda_n - \alpha_n)V(x) \\ &= \lambda_n V(x) + (1 - \alpha_n) + 2G - \delta(1 + |x|). \end{aligned} \tag{3}$$

Now we select constant c such that

$$(1 - \alpha_n) + 2G - \delta(1 + c) \leq 0,$$

for example

$$c = \max \left\{ \frac{2G + 1}{\delta} - 1, 1 \right\}.$$

Thus, we arrived to the following inequality for all $x \geq 0$:

$$P_n V(x) \leq \lambda_n V(x) + (2G + 1 - \delta) \mathbb{1}_{[-c, c]}(x), \tag{4}$$

By exactly the same reasoning applied to $x \leq 0$ we see that inequality (4) is valid for all $x \in \mathbb{R}$.

Condition 1, along with inequality (4), ensures the conditions of Theorem 4 with the set $C = [-c, c]$. The desired result now follows from Theorem 4. \square

Remark 1. Condition 1 in Theorem 1 is a relaxed form of a separation from unity condition:

$$\sup_n \alpha_n < 1.$$

Clearly, in this case, we can select $\delta = (1 - \sup_n \alpha_n)/2$, so that $\alpha_n + \delta < 1$ for all n , and Condition 1 is equivalent to

$$\sum_{n \geq 1} (1 - \alpha_n - \delta) = \infty$$

which is obviously true since $1 - \alpha_n - \delta \geq (1 - \sup_n \alpha_n)/2 > 0$. The novelty of this result is that we allow α_n to be greater than 1 from time to time.

Remark 2. Since random variables W_n are centered, Condition 2 is equivalent to

$$\sup_n \mathbb{E}|W_n| < \infty.$$

The following immediate corollary could be useful in practical applications.

Corollary 1. *Assume conditions of Theorem 1 hold true and there exist two constants $\psi > 1$ and $C_0 > 0$ such that*

$$\prod_{k=1}^n (\alpha_k + \delta) \leq C_0 \psi^{-n}.$$

Then, for all $x \in \mathbb{R}$,

$$\mathbb{E}_x^n [\psi^{\sigma c}] \leq C_0 (1 + |x| + (2G + 1 - \delta) \mathbb{1}_{[-c,c]}(x)).$$

Remark 3. In the homogeneous theory, geometric recurrence implies a corresponding chain’s positivity and geometric ergodicity. However, in the time-inhomogeneous case, such a conclusion is wrong in general since essentially inhomogeneous chains (which are not asymptotically homogeneous) usually do not have a stationary distribution.

4 Stability of general autoregressive models

In this section we consider a set of independent random variables $W_n^{(i)} : \Omega \rightarrow \mathbb{R}$ with distributions $\Gamma_n^{(i)}(A) = \mathbb{P}\{W_n^{(i)} \in A\}$ and a pair of Markov chains

$$X_n^{(i)} = \alpha_n^{(i)} X_{n-1}^{(i)} + W_n^{(i)}, \tag{5}$$

$n \geq 1, i \in \{1, 2\}, x_0^{(i)} \in \mathbb{R}$.

Our goal is to demonstrate that

$$\sup_n \left\| \mathbb{P} \left\{ X_n^{(1)} \in \cdot \right\} - \mathbb{P} \left\{ X_n^{(2)} \in \cdot \right\} \right\|_{TV} \rightarrow 0, \quad \varepsilon \rightarrow 0,$$

where ε defines how “close” are families $\{\Gamma_n^{(1)}, n \geq 0\}$ and $\{\Gamma_n^{(2)}, n \geq 1\}$.

We define ε by

$$\varepsilon := \frac{1}{2} \sup_{x \in \mathbb{R}, n \geq 1, A \in \mathcal{B}} \left| \Gamma_n^{(1)}(A - \alpha_n^{(1)} x) - \Gamma_n^{(2)}(A - \alpha_n^{(2)} x) \right|. \tag{6}$$

It is important to emphasize that one-step transition probabilities for two chains do not coincide when $n \rightarrow \infty$ and remain different. Strictly speaking, ε is an internal characteristic of a pair of processes and “changing” ε means shifting to another pair of processes. This can be illustrated by the example that we will use across this paper which emphasizes the dependence of ε and assumes the following representation:

$$\begin{aligned} \Gamma_n^{(1)} &= (1 - \tilde{\varepsilon})\Gamma_n + \tilde{\varepsilon}R_n^{(1)}, \\ \Gamma_n^{(2)} &= (1 - \tilde{\varepsilon})\Gamma_n + \tilde{\varepsilon}R_n^{(2)}, \end{aligned} \tag{7}$$

where $\tilde{\varepsilon}$ is a small constant. Here we can think of Γ_n as a “common part” of $\Gamma_n^{(1)}$ and $\Gamma_n^{(2)}$, and $R_n^{(i)}$ states for a residual part. This particular example can also be considered as a two-steps model, in which we flip a coin with a probability of “success” $1 - \tilde{\varepsilon}$, and if the flip was “successful” we render the next $W_n^{(i)}$ from a common distribution Γ_n , otherwise from a residual distribution $R_n^{(i)}$. The connection between $\tilde{\varepsilon}$ and ε in (6) is immediate:

$$\varepsilon = \frac{1}{2} \sup_{x \in R, n \geq 1, A \in \mathcal{B}} \left| \tilde{\varepsilon}R_n^{(1)}(A - \alpha_n^{(1)}x) - \tilde{\varepsilon}R_n^{(2)}(A - \alpha_n^{(2)}x) \right| \leq \tilde{\varepsilon}.$$

Our goal is to find conditions which ensure proximity of n -steps transition probabilities for all n , given only one-step proximity. Additionally, we want to demonstrate that such proximity is approximately of order ε , and all involved constants can be calculated in practical applications.

For this purpose we construct a coupling for chains $X^{(1)}$ and $X^{(2)}$. First, we assume that conditions of Theorem 3 hold, let $C = [-c, c]$ to be a corresponding set and introduce the following condition.

Condition M (Minorization condition). Assume that there exists a sequence of real numbers $\{a_n, n \geq 1\}$, $a_n \in (0, 1)$, and a sequence of probability measures ν_n on $(\mathbb{R}, \mathcal{B})$ such that:

$$\begin{aligned} \inf_{x \in C} \Gamma_n(A - \alpha_n^{(i)}x) &\geq a_n \nu_n(A), \quad i \in \{1, 2\}, \\ \inf_n \nu_n(C) &> 0, \\ 0 < a_* &:= \inf_n a_n \leq a_n \leq a^* := \sup_n a_n < 1, \end{aligned}$$

for all $A \in \mathcal{B}$ and $n \geq 1$.

Remark 4. Going forward we will require Condition M for both $\{\Gamma_n^{(1)}, n \geq 1\}$ and $\{\Gamma_n^{(2)}, n \geq 1\}$ with the same $\{a_n, n \geq 1\}$ and ν_n . At the same time, in the scheme (7) it is sufficient to require Condition M only for the “common part” Γ_n . This illustrates that parameters $\{a_n, n \geq 1\}$ and measures ν_n do not depend on ε .

Second, we introduce substochastic kernels

$$Q_n(x, \cdot) = \min_i \left\{ \Gamma_n^{(i)}(\cdot - \alpha_n^{(i)}x) \right\}, \tag{8}$$

where min should be understood as a minimum of two measures (see [5], Proposition D.2.8 as an example of a formal definition).

Note that by definition of ε and elementary equality $a \wedge b = \frac{a+b-|a-b|}{2}$

$$Q_n(x, \mathbb{R}) \geq 1 - \varepsilon.$$

We denote the residual substochastic kernels by

$$R_n^{(i)}(x, A) = \Gamma_n^{(i)}(A - \alpha_n^{(i)}x) - Q_n(x, A),$$

so that

$$R_n^{(i)}(x, \mathbb{R}) \leq \varepsilon.$$

Here, Q_n is a general analogue of a “common part” in (7).

To prove the main result, we will need some regularity conditions on Q_n .

Condition T (Tails condition). Denote $A_m = \{y \in \mathbb{R} : |y| \in [m, m + 1)\}$. Assume that there exist sequences $\{\hat{S}_n, n \geq 1\}$, $\{\hat{r}_n, n \geq 1\}$, such that

$$\begin{aligned} Q^{t,k}(x, A_m) &\leq Q^{t,k}(x, \mathbb{R})\hat{S}_m, \quad x \in C, \\ \nu_t Q^{t,k}(A_m) &\leq \nu_t Q^{t,k}(x, \mathbb{R})\hat{S}_m, \\ \hat{m} &= \sum_{m \geq 1} \hat{S}_m < \infty, \\ \sup_{i,k,x \in A_m} \int_{\mathbb{R}} R_k^{(i)}(x, dy)|y| &\leq \hat{r}_m, \\ \Delta &= \sum_{m \geq 1} \hat{r}_m \hat{S}_m < \infty, \end{aligned}$$

The motivation for this condition is as follows. We may think of $\tilde{Q}^{t,k}(x, \cdot) = \frac{Q^{t,k}(x, \cdot)}{Q^{t,k}(x, \mathbb{R})}$ as a proper k -steps transition probability, which is true in the case of schema (7). Generally speaking, this is not true, and we need additional regularity conditions to ensure that function $x \mapsto Q_t(x, \mathbb{R})$ does not vary too much (in schema (7) $Q_t(x, \mathbb{R}) = 1 - \tilde{\varepsilon}$ and does not depend on x at all). We do not need $\tilde{Q}_t = \tilde{Q}^{t,1}$ to be a proper Markov kernel in our derivations, so we do not impose any additional conditions. But we can think of \tilde{Q}_t as a Markov kernel and $\tilde{Q}^{t,k}$ as k -steps transition probability function to develop the intuition behind Condition T. First, imagine $\tilde{Q}_t(x, \cdot)$ is a homogeneous Markov kernel that is driving “coupling chain”, so that $\tilde{Q}_t(x, \cdot) = \tilde{Q}(x, \cdot)$. Assume the kernel is irreducible and geometrically recurrent (and thus positive), so there exists an invariant probability measure, say π . Denote $\bar{\pi}_m = \pi(A_m)$. We can write

$$\begin{aligned} \sum_{m \geq 0} \tilde{Q}^k(x, A_m) &= \sum_{m \geq 0} \tilde{Q}^k(x, A_m) - \bar{\pi}_m + \bar{\pi}_m \leq \sum_{m \geq 0} |\tilde{Q}^k(x, \cdot) - \pi|(A_m) + \bar{\pi}_m \\ &\leq 1 + |\tilde{Q}^k(x, \cdot) - \pi|(\mathbb{R}) = 1 + \|\tilde{Q}^k(x, \cdot) - \pi\|_{TV} \\ &\leq 1 + \sum_{k \geq 1} \|\tilde{Q}^k(x, \cdot) - \pi\|_{TV} < \infty. \end{aligned}$$

Assuming that C is a geometrically recurrent set, we may conclude

$$\sup_{x \in C, k \geq 1} \sum_{m \geq 0} \tilde{Q}^k(x, A_m) \leq 1 + \sup_{x \in C} \sum_{k \geq 1} \|\tilde{Q}^k(x, \cdot) - \pi\|_{TV} < \infty.$$

This demonstrates that condition $\frac{Q^{t,k}(x, A_m)}{Q^{t,k}(x, \mathbb{R})} \leq \hat{S}_m, x \in C$, is reasonable.

Let us now discuss the condition on $R_k^{(i)}(x, dy)$. This condition requires that the “shifted” residual process does not go far away from its starting position. For example, if $R_k^{(i)}$ are also centred and have the “autoregressive” form, $R_k^{(i)}(x, dy) = F_k^{(i)}(dy - x)$, the derivation of Section 3 demonstrates that it is reasonable to expect $\hat{r}_n = O(n)$. In this case, condition $\Delta < \infty$ is reduced to

$$\sup_{x \in C} \sum_{n \geq 1} n \|Q^n(x, \cdot) - \pi(\cdot)\|_{TV} < \infty,$$

and

$$\int_{\mathbb{R}} |y| \pi(dy) < \infty,$$

which are also reasonable conditions that are satisfied in the case of the geometrically ergodic autoregressive model. Moreover, the homogeneous theory tells us that existing of a small set C (in the sense of Condition M) with a finite second moment of the return time is sufficient for the above conditions to hold.

It is easy to show that a similar motivation remains valid for inhomogeneous autoregressive models. Assume $\tilde{Q}_n(x, \cdot) = \tilde{\Gamma}_n(\cdot - \alpha_n x)$, so that we can construct the chain $(Y_n)_{n \geq 0}$ of the form $Y_{n+1} = \alpha_{n+1} Y_n + U_{n+1}, n \geq 0$, where $U_n \sim \tilde{\Gamma}_n$ are independent and centred random variables. The transition kernels of the chain (Y_n) are \tilde{Q}_n . Then the Chebyshev inequality turns into

$$\tilde{Q}^{t,k}(x, A_m) \leq \mathbb{P}_x^t \{|Y_k| \geq m\} \leq \frac{\mathbb{E}_x^t[|Y_k|^2]}{m^2}.$$

Next, we write

$$Y_n = \sum_{k=0}^n \left(\prod_{j=k+1}^n \alpha_j \right) U_k,$$

here $U_0 \in [-c, c]$ is a fixed (nonrandom) initial state and $\prod_{n+1}^n = 1$. Assume that $U_n, n \geq 1$, have the uniformly bounded second moment. Denote its bound by μ_2 . We can then set

$$\hat{S}_m = \frac{\max\{\mu_2, c^2\}}{m^2} \sup_n \sum_{k=0}^n \prod_{j=k+1}^n \alpha_j^2.$$

So, we can conclude that the existence of the uniformly bounded second moment (or any higher moments) for distributions $\tilde{\Gamma}_n$ along with the condition

$$\sup_n \sum_{k=0}^n \prod_{j=k+1}^n \alpha_j^2 < \infty \tag{9}$$

are the sufficient conditions for building $\{\hat{S}_n, n \geq 1\}$. Note that Condition 1 of Theorem 1 requires a set $\{n \geq 1 : \alpha_n > 1 - \delta\}$ (for some $\delta < 1$) to be very sparse, so that it is reasonable to expect

$$\sup_n \sum_{k=0}^n \prod_{j=k+1}^n \alpha_j^2 \leq \frac{M}{1 - \delta^\gamma},$$

for some constants M and $\gamma > 0$. Of course, in order to ensure $\sum_m \hat{r}_m \hat{S}_m < \infty$ we may need to impose a stronger condition on U_n , for example, to have a uniformly bounded fourth, or even exponential, moment.

Generally speaking, Condition T requires that the “residual mean” grows slower than the \tilde{Q}_n -chain “advances in space”, in other words, the probability of hitting distant regions of the phase space is small enough to compensate for the “residual mean’s” growth as the starting point x is moving away from the origin. It should be clear now that the most typical autoregression, the Gaussian one, satisfies this condition as long as all variances are uniformly bounded.

Remark 5. In the discussion above, we have shown that we may expect $\hat{r}_n = O(n)$, which in turn means that $\hat{m} < A\Delta$, where A is some constant. However, we have never formally proved that $\hat{r}_n \geq 1$, so we keep both conditions $\hat{m} < \infty$ and $\Delta < \infty$ even if, in usual circumstances, the latter implies the former.

Finally, we would like to mention that Δ is in fact $\Delta(\varepsilon)$ (we have shown that $R_k^{(i)}(x, \mathbb{R}) \leq \varepsilon$), which means it should be small as $\varepsilon \rightarrow 0$. In schema (7) we clearly have $\Delta = O(\varepsilon)$.

Assuming Condition M holds true, we define “noncoupling” operators

$$T_t^{(i)}(x, A) = \frac{\Gamma_t^{(i)}(A - a_t x) - a_t v_t(A)}{1 - a_t},$$

$$T_{xy}^{(i)}(A, B) = T_t^{(1)}(x, A)T_t^{(2)}(y, B).$$

We define the Markov chain $\bar{Z}_n = (Z_n^{(1)}, Z_n^{(2)}, d_n)$ with values in $(\mathbb{R}, \mathbb{R}, \{0, 1, 2\})$ by setting its transition probabilities

$$\bar{P}_n(x, y, 1; A \times B \times \{2\}) = \mathbb{1}_{x=y} Q_n(x, A \cap B),$$

$$\bar{P}_n(x, y, 1; A \times B \times \{0\}) = \mathbb{1}_{x=y} \frac{R_n^{(1)}(x, A)R_n^{(2)}(y, B)}{1 - Q_n(x, \mathbb{R})},$$

we assume the latter probability is equal to zero if $Q_n(x, \mathbb{R}) = 1$,

$$\bar{P}_n(x, y, 0; A \times B \times \{0\}) = (1 - a_n)\mathbb{1}_{C \times C}(x, y)T_n^{(1)}(x, A)T_n^{(2)}(y, B) + (1 - \mathbb{1}_{C \times C}(x, y))\Gamma_n^{(1)}(A - \alpha_n^{(1)}x)\Gamma_n^{(2)}(A - \alpha_n^{(2)}y),$$

$$\bar{P}_n(x, y, 0; A \times B \times \{1\}) = \mathbb{1}_{C \times C}(x, y)a_n v_n(A \cap B).$$

$$\bar{P}_n(x, y, 2; \cdot) = \bar{P}_n(x, y, 1; \cdot).$$

All other probabilities are equal to zero.

It is straightforward that marginal distributions of the process \bar{Z}_n equal to those of $X_n^{(i)}$.

We will use the canonical probability $\bar{\mathbb{P}}_{x,y,d}^t$ and the expectation $\bar{\mathbb{E}}_{x,y,d}^t$, $x, y \in \mathbb{R}$, $d \in \{0, 1, 2\}$ in the same sense as before.

Let us denote by

$$\begin{aligned} \bar{\sigma}_{C \times C} &= \bar{\sigma}_{C \times C}(1) = \inf \left\{ n \geq 1 : \left(Z_n^{(1)}, Z_n^{(2)} \right) \in C \times C \right\}, \\ \bar{\sigma}_{C \times C}(m) &= \inf \left\{ n \geq \bar{\sigma}_{C \times C}(m-1) : \left(Z_n^{(1)}, Z_n^{(2)} \right) \in C \times C \right\}, \quad m \geq 2, \end{aligned}$$

the first and m -th return times to $C \times C$ by the pair $\left(Z_n^{(1)}, Z_n^{(2)} \right)$.

We will also need a special notation for the sets

$$\begin{aligned} D_n &= \{d_1 = d_2 = \dots = d_n = 0\}, \\ D_{nk} &= \{d_1 = d_2 = \dots = d_n = 0, \bar{\sigma}_{C \times C}(k) = n\}, \\ B_{nk} &= \{d_k \in \{1, 2\}, d_{k+1} = 0, \dots, d_n = 0\}, \end{aligned} \tag{10}$$

and for the values

$$\rho_{nk} = \sup_{x,y \in C,t} \bar{\mathbb{P}}_{x,y,0}^t(D_{nk}). \tag{11}$$

Theorem 2. *Let $X^{(i)}$ be two Markov chains defined in (5) that simultaneously satisfy Condition M, Condition T and conditions of Corollary 1 with $\psi > 1$, and $C = [-c, c]$ be a corresponding set.*

Then there exist constants $M_1, M_2 \in \mathbb{R}$, such that for every $x \in C$

$$\left| \left| \mathbb{P}_x^t \left\{ X_n^{(1)} \in \cdot \right\} - \mathbb{P}_x^t \left\{ X_n^{(2)} \in \cdot \right\} \right| \right| \leq \varepsilon \hat{m} M_1 + \Delta M_2, \tag{12}$$

where \hat{m} and Δ are defined in Condition T.

For every $x \notin C$ the following inequality holds true:

$$\left| \left| \mathbb{P}_x^t \left\{ X_n^{(1)} \in \cdot \right\} - \mathbb{P}_x^t \left\{ X_n^{(2)} \in \cdot \right\} \right| \right| \leq \varepsilon(2\hat{m} M_1 + \hat{\mu}(x)) + 2\Delta M_2, \tag{13}$$

where

$$\hat{\mu}(x) = \sup_t \sum_{k \geq 1} \left(\prod_{j=0}^{k-1} Q_{t+j} \mathbb{1}_{\mathbb{R} \setminus C} \right) (x, \mathbb{R} \setminus C).$$

Remark 6. By $Q_t \mathbb{1}_A$ we understand a kernel that is equal to $Q_t(x, \cdot)$ if $x \in A$ and zero otherwise. So we have

$$\begin{aligned} & \left(\prod_{j=0}^{n-1} Q_{t+j} \mathbb{1}_A \right) (x, B) \\ &= \mathbb{1}_A(x) \int_A \dots \int_A Q_t(x, dy_1) Q_{t+1}(y_1, dy_2) \dots Q_{t+n-1}(y_{n-1}, B). \end{aligned}$$

Remark 7. Recall the process $\bar{Z} = \left(Z_n^{(1)}, Z_n^{(2)}, d_n \right)_{n \geq 0}$ that is defined above. We can interpret $\hat{\mu}(x)$ as

$$\hat{\mu}(x) = \sup_t \sum_{n \geq 1} \bar{\mathbb{P}}_{x,x,1}^t \{d_1 = \dots = d_n = 2, Z_j^{(1)} = Z_j^{(2)} \notin C, j \leq n\}.$$

We can think of $\hat{\mu}(x)$ as an “expectation” for the first time the “coupled chain” visits C . The difficulty of calculating $\hat{\mu}$ in a practical application depends on the exact model. However, it is typical that the “common part process” enjoys the same properties as each of the original ones. In the context of this paper, this means that the “common part process” will be geometrically recurrent.

We can demonstrate how to calculate $\hat{\mu}(x)$ in schema (7). In this case, we can factor out decoupling trials and the process’s advancement in space. Let $(Y_n)_{n \geq 0}$ be a canonical autoregressive process generated by the family $\{\Gamma_n, n \geq 1\}$ and $\tilde{\mathbb{P}}_x$ be the corresponding probability. Without loss of generality we assume $\tilde{\varepsilon} = \varepsilon$. Let

$$\tilde{\sigma} = \inf\{t \geq 1 : Y_t \in C\}$$

and assume there exists $\psi > 1$ such that

$$g(x) = \sup_t \tilde{\mathbb{E}}_x^t \left[\psi^{\tilde{\sigma}} \right] < \infty.$$

Then, we can write

$$\begin{aligned} \hat{\mu}(x) &= \sup_t \sum_{n \geq 1} \bar{\mathbb{P}}_{x,x,1}^t \{d_1 = \dots = d_n = 2, Z_j^{(1)} = Z_j^{(2)} \notin C, j \leq n\} \\ &= \sup_t \sum_{n \geq 1} (1 - \varepsilon)^n \tilde{\mathbb{P}}_x^t (\tilde{\sigma} \geq n) \leq \sup_t \sum_{n \geq 1} (1 - \varepsilon)^n \psi^{-n} g(x) \\ &= g(x) \frac{(1 - \varepsilon)\psi^{-1}}{1 - (1 - \varepsilon)\psi^{-1}} \rightarrow \frac{g(x)}{\psi - 1}, \quad \varepsilon \rightarrow 0. \end{aligned}$$

Finally, $g(x)$ can usually be estimated by constructing a Foster–Lyapunov function satisfying a drift condition for the process $(Y_n)_{n \geq 0}$.

Remark 8. It is possible to calculate all constants involved in (12) in terms of $G, \delta, a_*, a^*, \inf_n v_n(C), \psi$ that are defined in Theorem 1, Corollary 1 and Condition M. See Appendix A for an example of such calculation.

Proof. Using the standard coupling approach we first obtain

$$\begin{aligned} & \left| \mathbb{P}_x^t \left\{ X_n^{(1)} \in A \right\} - \mathbb{P}_x^t \left\{ X_n^{(2)} \in A \right\} \right| \\ &= \left| \bar{\mathbb{P}}_{x,x,1}^t \left\{ \bar{Z}_n \in (A, \mathbb{R}, \{0, 1, 2\}) \right\} - \bar{\mathbb{P}}_{x,x,1}^t \left\{ \bar{Z}_n \in (\mathbb{R}, A, \{0, 1, 2\}) \right\} \right| \\ &= \left| \bar{\mathbb{P}}_{x,x,1}^t \left\{ \bar{Z}_n \in (A, \mathbb{R}, \{0\}) \right\} - \bar{\mathbb{P}}_{x,x,1}^t \left\{ \bar{Z}_n \in (\mathbb{R}, A, \{0\}) \right\} \right| \\ &\leq \max \left\{ \bar{\mathbb{P}}_{x,x,1}^t \left\{ \bar{Z}_n \in (A, \mathbb{R}, \{0\}) \right\}, \bar{\mathbb{P}}_{x,x,1}^t \left\{ \bar{Z}_n \in (\mathbb{R}, A, \{0\}) \right\} \right\} \\ &\leq \bar{\mathbb{P}}_{x,x,1}^t \{d_n = 0\}. \end{aligned}$$

Let us denote a kernel

$$\Lambda_k^t(x, A) = \bar{\mathbb{P}}_{x,x,1}^t \left\{ d_k \in \{1, 2\}, Z_k^{(1)} = Z_k^{(2)} \in A \right\},$$

and a set

$$A_m = \{y \in \mathbb{R} : |y| \in [m, m + 1)\}.$$

We note that since $\bar{P}_n(x, y, 1; \cdot) = \bar{P}_n(x, y, 2; \cdot)$ we have

$$\begin{aligned} \bar{\mathbb{P}}_{x,x,1}^t(B_{nk}) &= \int_{\mathbb{R}} \bar{\mathbb{P}}_{x,x,1}^t \left\{ d_k = 1, Z_k^{(1)} = Z_k^{(2)} = dy \right\} \bar{\mathbb{P}}_{y,y,1}^{t+k}(D_{n-k}) \\ &+ \int_{\mathbb{R}} \bar{\mathbb{P}}_{x,x,1}^t \left\{ d_k = 2, Z_k^{(1)} = Z_k^{(2)} = dy \right\} \bar{\mathbb{P}}_{y,y,2}^{t+k}(D_{n-k}) \\ &= \Lambda_k^t(x, dy) \bar{\mathbb{P}}_{y,y,1}^{t+k}(D_{n-k}). \end{aligned}$$

We apply this equality to the last decoupling time and derive:

$$\begin{aligned} \bar{\mathbb{P}}_{x,x,1}^t\{d_n = 0\} &= \sum_{k=1}^{n-1} \bar{\mathbb{P}}_{x,x,1}^t(B_{nk}) = \sum_{k=1}^{n-1} \int_{\mathbb{R}} \Lambda_k^t(x, dy) \bar{\mathbb{P}}_{y,y,1}^{t+k}(D_{n-k}) \\ &= \sum_{k=1}^{n-1} \sum_{m \geq 0} \int_{A_m} \Lambda_k^t(x, dy) \bar{\mathbb{P}}_{y,y,1}^{t+k}(D_{n-k}) \\ &\leq \sum_{k=1}^{n-1} \sum_{m \geq 0} \Lambda_k^t(x, A_m) \sup_{y \in A_m} \bar{\mathbb{P}}_{y,y,1}^{t+k}(D_{n-k}). \end{aligned} \tag{14}$$

Next, we want to replace $\Lambda_k^t(x, A_m)$ with \hat{S}_m and apply Lemma 8. However, this can only be done for $x \in C$. Indeed:

$$\begin{aligned} \Lambda_k^t(x, A_m) &= \sum_{j=0}^{k-1} \bar{\mathbb{P}}_{x,x,1}^t \{d_j = 1\} \bar{\mathbb{P}}_{v_j,1}^{t+j} \{d_l = 2, l \in \overline{1, k-j}, Z_{k-j}^{(1)} = Z_{k-j}^{(2)} \in A_m\} \\ &= \sum_{j=0}^{k-1} \bar{\mathbb{P}}_{x,x,1}^t \{d_j = 1\} v_j Q^{t+j, k-j}(A_m) \leq \sum_{j=0}^{k-1} \bar{\mathbb{P}}_{x,x,1}^t \{d_j = 1\} v_j Q^{t+j, k-j}(\mathbb{R}) \hat{S}_m \\ &= \hat{S}_m \sum_{j=0}^{k-1} \bar{\mathbb{P}}_{x,x,1}^t \{d_j = 1, d_{j+1} = \dots = d_k = 2\} = \hat{S}_m \mathbb{P}_{x,x,1}^t \{d_k = 2\} \leq \hat{S}_m. \end{aligned}$$

Here $v_j(du, dv) = \mathbb{1}_{du=dv}(\mathbb{1}_{\{j=0\}}\delta_x + \mathbb{1}_{\{j>0\}}\nu(du))$ helps us to include the case of $j = 0$ (i.e. no decouplings from the beginning) into the previous derivations. Applying this bound and Lemma 8 to (14) we obtain for $x \in C$:

$$\bar{\mathbb{P}}_{x,x,1}^t\{d_n = 0\} \leq \sum_{m \geq 0} \hat{S}_m (\varepsilon M_1 + \hat{r}_m M_2) = \varepsilon \hat{m} M_1 + \Delta M_2. \tag{15}$$

Now we consider the case $x \notin C$. Let us denote

$$\bar{\sigma} = \inf\{t > 0 : Z_t^{(1)} = Z_t^{(2)} \in C, d_1 = \dots = d_t = 2\} \leq \infty.$$

For $x \notin C$ we have

$$\begin{aligned} \Lambda_k^t(x, A_m) &= \sum_{j=1}^{k-1} \bar{\mathbb{P}}_{x,x,1}^t \{d_j = 1\} \bar{\mathbb{P}}_{v,1}^t \{d_l = 2, l \in \overline{1, k-j}, Z_{k-j}^{(1)} = Z_{k-j}^{(2)} \in A_m\} \\ &\quad + \bar{\mathbb{P}}_{x,x,1}^t \{d_2 = \dots = d_{k-1} = 2, Z_{k-1}^{(1)} = Z_{k-1}^{(2)} \in A_m\} \leq \hat{S}_m \\ &\quad + \sum_{j=1}^{k-1} \bar{\mathbb{P}}_{x,x,1}^t \{d_2 = \dots = d_{k-1} = 2, \bar{\sigma} = j, Z_{k-1}^{(1)} = Z_{k-1}^{(2)} \in A_m\} \\ &\quad + \bar{\mathbb{P}}_{x,x,1}^t \{d_2 = \dots = d_{k-1} = 2, \bar{\sigma} \geq k, Z_{k-1}^{(1)} = Z_{k-1}^{(2)} \in A_m\} \\ &\leq \hat{S}_m + \hat{S}_m \bar{\mathbb{P}}_{x,x,1}^t (d_1 = \dots = d_{k-1} = 2, \bar{\sigma} < k) + q_{k,m}(x) \leq 2\hat{S}_m + q_{k,m}(x), \end{aligned}$$

where

$$q_{k,m}(x) = \bar{\mathbb{P}}_{x,x,1}^t \{d_2 = \dots = d_{k-1} = 2, \bar{\sigma} \geq k, Z_{k-1}^{(1)} = Z_{k-1}^{(2)} \in A_m\}.$$

Denote

$$q_k(x) = \sum_{m \geq 0} q_{k,m}(x).$$

Applying this result and the obvious inequality $\bar{\mathbb{P}}_{y,y,1}^t(D_{n-k}) \leq \varepsilon$ to (14), we get for $x \notin C$

$$\bar{\mathbb{P}}_{x,x,1}^t \{d_n = 0\} \leq 2\varepsilon \hat{m} M_1 + 2\Delta M_2 + \varepsilon \sum_{k \geq 0} q_k(x) \leq \varepsilon(2\hat{m} M_1 + \hat{\mu}(x)) + 2\Delta M_2.$$

Here we used the fact

$$\sum_{k \geq 0} q_k(x) = \hat{\mu}(x). \quad \square$$

5 Auxiliary results

Everywhere in this section we assume that conditions of Theorem 1 are valid as well as Condition C and will use the corresponding notation.

We start by introducing an important result that allows us to connect the expectation of the exponential moment of $\sigma_{C \times C}$ with individual exponential moments of $\sigma_C^{(i)}$.

Theorem 3. *Let*

$$X_{n+1}^{(1)} = \alpha_n^{(1)} X_n^{(1)} + W_n^{(1)},$$

and

$$X_{n+1}^{(2)} = \alpha_n^{(2)} X_n^{(2)} + W_n^{(2)},$$

be two Markov chains. Let $\Gamma_n^{(i)}$ be a distribution of $W_n^{(i)}$ and all $W_n^{(i)}$ be independent. Assume that both chains satisfy conditions of Corollary 1 and Condition M for the set $C = [-c, c]$ where c is defined in (2).

Then there exist $\psi_1 > 1$ and $\Xi_0, \Xi_1 \in \mathbb{R}$, such that for every $x, y \in \mathbb{R}$ and $n \geq 0$

$$\mathbb{E}_{x,y}^n [\psi_1^{\sigma_{C \times C}}] \leq \Xi_0(|x| + |y|) + \Xi_1.$$

Remark 9. Constants Ξ_0, Ξ_1 can be calculated using results from [12]. We will show how this can be done in Appendix A.

Proof. From [12], Theorem 4.2, we get

$$\mathbb{E}_{x,y}^n [\psi_1^{\sigma_C \times C}] \leq M \left(\mathbb{E}_x^n [\psi_0^{\sigma_C^{(1)}}] S_1(\psi_0) + \mathbb{E}_y^n [\psi_0^{\sigma_C^{(2)}}] S_2(\psi_0) \right),$$

where, $\psi_0, \psi_1 \in (1, \psi)$ and $M \in \mathbb{R}$ are some constants and

$$S_i(u) = \sup_{n,x \in C} \left\{ \frac{1}{1 - a_n} \left(\mathbb{E}_x^n [u^{\sigma_C^{(i)}}] - a_n \mathbb{E}_{v_n}^n [u^{\sigma_C^{(i)}}] \right) \right\}.$$

Since $\psi_0 < \psi$, Corollary 1 turns into

$$S_i(\psi_0) \leq \frac{1}{1 - a^*} \sup_{n,x \in C} \mathbb{E}_x^n [\psi^{\sigma_C^{(i)}}] \leq \frac{C_0(2 + 2G + c - \delta)}{1 - a^*}.$$

Next we apply Corollary 1 to $\mathbb{E}_x^n [\psi_0^{\sigma_C^{(i)}}]$ and get

$$\mathbb{E}_{x,y}^n [\psi_1^{\sigma_C \times C}] \leq \Xi_0(|x| + |y|) + \Xi_1, \tag{16}$$

$$\Xi_0 = \frac{MC_0^2(2 + 2G + c - \delta)}{1 - a^*}, \tag{17}$$

$$\Xi_1 = \Xi_0(2 + (1 + 2G - \delta)(\mathbb{1}_C(x) + \mathbb{1}_C(y))). \tag{18}$$

□

Next we introduce

$$h(x, y) = \sup_n \mathbb{E}_{x,y}^n [\psi_1^{\sigma_C \times C}] \tag{19}$$

and

$$H_t(x) = \int_{\mathbb{R}^2 \setminus C \times C} \frac{R_t^{(1)}(x, dy) R_t^{(2)}(x, dz)}{1 - Q_t(x, \mathbb{R})} h(y, z). \tag{20}$$

Going forward we will use the notation from Theorem 1, Corollary 1, Theorem 3 and Condition M. We now present a series of lemmas that are necessary to prove the main result.

Lemma 1. For all $(x, y) \notin C \times C$

$$\bar{\mathbb{P}}_{x,y,0}^t \{ \sigma_{C \times C} \geq n \} \leq \psi_1^{-n} h(x, y).$$

Proof. For all $(x, y) \notin C \times C$

$$\begin{aligned} \bar{\mathbb{P}}_{x,y,0}^t \{ \sigma_{C \times C} \geq n \} &= \mathbb{P} \left\{ \sigma_{C \times C} \geq n \mid X_t^{(1)} = x, X_t^{(2)} = y \right\} \leq \delta_1^{-n} \mathbb{E}_{x,y}^t \left[\delta_1^{-\sigma_{C \times C}} \right] \\ &= \psi_1^{-n} h(x, y). \end{aligned} \tag{21}$$

The first equality above is due to the definition $\bar{\mathbb{P}}_{x,y,0}^t$, then we used the Chernoff inequality and definition (19). □

Remark. Note, that it is important in the preceding derivations that $(x, y) \notin C \times C$. In this case the distribution of the pair $(Z^{(1)}, Z^{(2)})$ coincides with that of $(X^{(1)}, X^{(2)})$ on the time intervall $[0, \bar{\sigma}_{C \times C} - 1]$ which makes the first equality valid and allows us to move from the probability measure $\bar{\mathbb{P}}$ to \mathbb{P} , and use geometric recurrence result for a paired chain.

Lemma 2.

$$\bar{\mathbb{P}}_{x,x,1}^t \{d_1 = 0, \bar{\sigma}_{C \times C} \geq n\} \leq \psi_1^{n-1} H_t(x),$$

where $H_t(x)$ is defined in (20) and $x \in \mathbb{R}$.

Proof. We use Lemma 1 and definition of probability $\bar{\mathbb{P}}$ to derive

$$\begin{aligned} & \bar{\mathbb{P}}_{x,x,1}^t \{d_1 = 0, \bar{\sigma}_{C \times C} \geq n\} \\ &= \int_{\mathbb{R}^2 \setminus C \times C} \frac{R_t^{(1)}(x, dy) R_t^{(2)}(x, dz)}{1 - Q_t(x, \mathbb{R})} \bar{\mathbb{P}}_{y,z,0}^t \{\bar{\sigma}_{C \times C} \geq n - 1\} \\ &\leq \psi_1^{n-1} \int_{\mathbb{R}^2 \setminus C \times C} \frac{R_t^{(1)}(x, dy) R_t^{(2)}(x, dz)}{1 - Q_t(x, \mathbb{R})} h(y, z) \\ &= \psi_1^{n-1} H_t(x). \quad \square \end{aligned}$$

Now, we consider a situation when the pair $(Z_k^{(1)}, Z_k^{(2)})$ hits $C \times C$ exactly once.

Lemma 3.

$$\bar{\mathbb{P}}_{x,x,1}^t \{d_j = 0, j = \overline{1, n}, \bar{\sigma}_{C \times C} \leq n < \bar{\sigma}_{C \times C}(2)\} \leq \frac{S((n - 1)H_t(x) + \varepsilon)}{\psi_1^{n-1}}, \quad (22)$$

where

$$S = (1 - a_*) \sup_{x,y \in C, t \geq 0} \int_{\mathbb{R}^2 \setminus C \times C} T_{x,y}^{(t)}(du, dv) h(u, v).$$

Proof. For any $x \in \mathbb{R}$

$$\begin{aligned} & \bar{\mathbb{P}}_{x,x,1}^t \{d_j = 0, j = \overline{1, n}, \bar{\sigma}_{C \times C} \leq n < \bar{\sigma}_{C \times C}(2)\} \\ &= \sum_{k=1}^n \bar{\mathbb{P}}_{x,x,1}^t \{d_j = 0, j = \overline{1, n}, \bar{\sigma}_{C \times C} = k, \bar{\sigma}_{C \times C}(2) > n\} \\ &= \sum_{k=1}^n \int_{C \times C} \bar{\mathbb{P}}_{x,x,1}^t \{d_1 = 0, \bar{\sigma}_{C \times C} = k, (Z_k^{(1)}, Z_k^{(2)}, d_k) = (du, dv, 0)\} \times \\ &\times \bar{\mathbb{P}}_{u,v,0}^{t+k} \{\bar{\sigma}_{C \times C} > n - k\} \leq \sum_{k=1}^n \bar{\mathbb{P}}_{x,x,1}^t \{d_1 = 0, \bar{\sigma}_{C \times C} = k, d_k = 0\} \times \\ &\times \sup_{x,y \in C} \bar{\mathbb{P}}_{x,y,0}^{t+k} \{d_1 = 0, \bar{\sigma}_{C \times C} > n - k\}. \end{aligned}$$

Note that we cannot apply Lemma 1 to the last term, because initial values $(x, y) \in C \times C$, and so the distribution of the first step for a \bar{Z} -chain under the measure $\bar{\mathbb{P}}$ is different from that of $(X^{(1)}, X^{(2)})$ -chain under \mathbb{P} . So we have for all $(x_0, y_0) \in C \times C$

$$\bar{\mathbb{P}}_{x_0,y_0,0}^{t+k} \{d_1 = 0, \bar{\sigma}_{C \times C} > n - k\}$$

$$= (1 - a_{t+k}) \int_{\mathbb{R}^2 \setminus C \times C} T_{t+k}^{(1)}(x_0, dx) T_{t+k}^{(2)}(y_0, dy) \bar{\mathbb{P}}_{x,y,0}^{t+k+1} \{ \bar{\sigma}_{C \times C} \geq n - k \}.$$

Now we can apply Lemma 1 to $\bar{\mathbb{P}}_{x,y,0}^{t+k+1} \{ \bar{\sigma}_{C \times C} \geq n - k \}$ and obtain the inequality

$$\begin{aligned} & \sup_{x_0, y_0 \in C} \bar{\mathbb{P}}_{x_0, y_0, 0}^{t+k} \{ d_1 = 0, \bar{\sigma}_{C \times C} \geq n - k \} \\ & \leq \frac{1 - a_*}{\psi_1^{(n-k)}} \sup_{x_0, y_0 \in C} \int_{\mathbb{R}^2 \setminus C \times C} T_{x_0, y_0}^{(t+k)}(dx, dy) h(x, y) \\ & = S \psi_1^{-(n-k)}. \end{aligned} \tag{23}$$

Using Lemma 2 for $k \geq 2$ we get

$$\begin{aligned} \bar{\mathbb{P}}_{x,x,1}^t \{ d_1 = 0, d_k = 0, \bar{\sigma}_{C \times C} = k \} & \leq \bar{\mathbb{P}}_{x,x,1}^t \{ d_1 = 0, \bar{\sigma}_{C \times C} \geq k \} \\ & \leq \psi_1^{-(k-1)} H_t(x). \end{aligned} \tag{24}$$

It is clear that for $k = 1$

$$\bar{\mathbb{P}}_{x,x,1}^t \{ d_1 = 0, \bar{\sigma}_{C \times C} = 1 \} = \frac{R_t^{(1)}(x, C) R_t^{(2)}(x, C)}{1 - Q_t(x, \mathbb{R})} \leq \varepsilon. \tag{25}$$

Now combining (23)–(25) we get that

$$\begin{aligned} & \bar{\mathbb{P}}_{x,x,1}^t \{ d_j = 0, j = \overline{1, n}, \bar{\sigma}_{C \times C} \leq n < \bar{\sigma}_{C \times C}(2) \} \\ & \leq S \sum_{k=1}^n \bar{\mathbb{P}}_{x,x,1}^t \{ d_1 = 0, \bar{\sigma}_{C \times C} = k, d_k = 0 \} \psi_1^{-(n-k)} \\ & \leq S \left(\sum_{k=2}^n H_t(x) \psi_1^{-(k-1)} \psi_1^{-(n-k)} + \varepsilon \psi_1^{-(n-1)} \right) \\ & = S \left(\psi_1^{-(n-1)} (n-1) H_t(x) + \varepsilon \psi_1^{-(n-1)} \right) \\ & = S \psi_1^{-(n-1)} ((n-1) H_t(x) + \varepsilon). \end{aligned} \quad \square$$

Finally, we look at the situation when the pair $(Z_k^{(1)}, Z_k^{(2)})$ hits $C \times C$ exactly $k \geq 2$ times.

Lemma 4. For all $x \in \mathbb{R}$

$$\begin{aligned} & \bar{\mathbb{P}}_{x,x,1}^t \{ D_n, \bar{\sigma}_{C \times C}(k) \leq n < \bar{\sigma}_{C \times C}(k+1) \} \\ & \leq S H_t(x) \sum_{i=2}^{n-k+1} \delta_1^{-(i-1)} \sum_{j=k-1}^{n-i} \rho_{j,k-1} \delta_1^{-(n-i-j)} \\ & \quad + \varepsilon S \sum_{j=k-1}^{n-1} \rho_{j,k-1} \delta_1^{-(n-j-1)}, \end{aligned}$$

where ρ_{nk} is defined in (11), and S is from Lemma 3.

Proof. Using the first entrance – last exit decomposition and the Markov property, we get the following representation of the probability of interest

$$\begin{aligned} & \bar{\mathbb{P}}_{x,x,1}^t \{D_n, \bar{\sigma}_{C \times C}(k) \leq n < \bar{\sigma}_{C \times C}(k + 1)\} \\ &= \sum_{i=1}^{n-k+1} \sum_{j=i+k-1}^n \int_{C \times C} \bar{\mathbb{P}}_{x,x,1}^t \left\{ D_{i1}, \left(Z_i^{(1)}, Z_i^{(2)} \right) = (dy_1, dy_2) \right\} \times \\ & \times \int_{C \times C} \bar{\mathbb{P}}_{x,y,0}^{t+i} \left\{ D_{j-i,k-1}, \left(Z_{j-i}^{(1)}, Z_{j-i}^{(2)} \right) = (dz_1, dz_2) \right\} \\ & \times \bar{\mathbb{P}}_{z_1,z_2,0}^{t+j} \{d_1 = 0, \bar{\sigma}_{C \times C} > n - j\}. \end{aligned}$$

Here index i stands for the first hitting. Since there are k total hittings during the n timesteps, the first hitting may not occur later than $n - k$, otherwise there will be no “space” for the rest $k - 1$ hittings. Similar arguments applied to j , which is the time of the last hitting. Since there are exactly $k - 1$ hittings on the time interval $(i, j]$, j may not be closer than $k - 1$ to i , otherwise $k - 1$ hittings will not fit into the interval. Now we take supremum over $x, y \in C$ and $t \geq 0$ in the second and third terms and arrive to the inequality

$$\begin{aligned} & \bar{\mathbb{P}}_{x,x,1}^t \{D_n, \bar{\sigma}_{C \times C}(k) \leq n < \bar{\sigma}_{C \times C}(k + 1)\} \leq \sum_{i=1}^{n-k+1} \sum_{j=i+k-1}^n \bar{\mathbb{P}}_{x,x,1}^t \{D_{i1}\} \times \\ & \times \sup_{x,y \in C, t} \bar{\mathbb{P}}_{x,y,0}^t \{D_{j-i,k-1}\} \sup_{x,y \in C, t} \bar{\mathbb{P}}_{x,y,0}^t \{d_1 = 0, \bar{\sigma}_{C \times C} > n - j\} \end{aligned} \tag{26}$$

We immediately recognize the second multiplier as $\rho_{j-i,k-1}$.

Using the same arguments as in Lemma 3 (see inequality (23)), we get an estimate for the third multiplier

$$\sup_{x,y \in C, t} \bar{\mathbb{P}}_{x,y,0}^t \{d_1 = 0, \bar{\sigma}_{C \times C} > n - j\} \leq S\psi_1^{-(n-j)}.$$

Using Lemma 2 we obtain a bound for the first multiplier for $i \geq 2$

$$\begin{aligned} \mathbb{P}_{x,x,1}^t \{d_1 = d_i = 0, \bar{\sigma}_{C \times C} = i\} & \leq \sup_{x \in \mathbb{R}} \mathbb{P}_{x,x,1}^t \{d_1 = 0, \bar{\sigma}_{C \times C} \geq i\} \\ & \leq \psi_1^{-(i-1)} H_t(x). \end{aligned}$$

For $i = 1$ we have

$$\mathbb{P}_{x,x,1}^t \{d_1 = 0, \bar{\sigma}_{C \times C} = 1\} = \frac{R_t^{(1)}(x, C)R_t^{(2)}(x, C)}{1 - Q_t(x, \mathbb{R})} \leq \varepsilon. \tag{27}$$

Substituting the bounds for the first and third multipliers into (26), we get

$$\begin{aligned} & \bar{\mathbb{P}}_{x,x,1}^t \{D_n, \bar{\sigma}_{C \times C}(k) \leq n < \bar{\sigma}_{C \times C}(k + 1)\} \\ & \leq \sum_{i=2}^{n-k+1} \sum_{j=i+k-1}^n H_t(x)\psi_1^{-(i-1)} \rho_{j-i,k-1} S\psi_1^{-(n-j)} \end{aligned}$$

$$\begin{aligned}
 & + \sum_{j=k}^n \varepsilon \rho_{j-1,k-1} S \psi_1^{-(n-j)} \\
 & = SH_t(x) \sum_{i=2}^{n-k+1} \psi_1^{-(i-1)} \sum_{j=i+k-1}^n \rho_{j-i,k-1} \psi_1^{-(n-j)} \\
 & + \varepsilon S \sum_{j=k}^n \rho_{j-1,k-1} \psi_1^{-(n-j)} \\
 & = SH_t(x) \sum_{i=2}^{n-k+1} \delta_1^{-(i-1)} \sum_{j=k-1}^{n-i} \rho_{j,k-1} \delta_1^{-(n-i-j)} \\
 & + \varepsilon S \sum_{j=k-1}^{n-1} \rho_{j,k-1} \delta_1^{-(n-j-1)}.
 \end{aligned} \tag{28}$$

□

It is critical that the probability from Lemma 4 is double summable over n, k . One may notice that since $\rho_{jk} = 0$ if $j < k$ (because it is impossible that $\bar{\sigma}_{C \times C}(k) = j < k$), the double sum in the right hand side of the inequality from Lemma 4 is a convolution of two geometric sequences (ψ_1^{-n}) with ρ_{jk} (as function of j with fixed k) evaluated at $n - k + 1$. Thus in order to establish finiteness of the desired sum it is required to demonstrate that the double sequence $\{\rho_{nk}, n \geq 1, k = \overline{1, n}\}$ is summable. For doing so we will need the next three lemmas.

Lemma 5.

$$\rho_{nk} = \sup_{x,y \in C,t} \bar{\mathbb{P}}_{x,y,0}^t \{D_{nk}\} \leq (1 - a_*)^k. \tag{29}$$

Proof. We proceed by induction. Let $n = 1$ and $k = 1$. Then

$$\begin{aligned}
 \mathbb{P}_{x,y,0}^t \{D_{11}\} & = \mathbb{P}_{x,y,0}^t \{\bar{Z}_1 \in (C \times C \times 0)\} \\
 & = (1 - a_t) T_t^{(1)}(x, C) T_t^{(2)}(y, C) \leq (1 - a_*),
 \end{aligned}$$

since $T_t^{(i)}(x, C) \leq T_t^{(i)}(x, \mathbb{R}) = 1$.

Consider now $n \geq 2, k = 1$:

$$\begin{aligned}
 \bar{\mathbb{P}}_{x_0,y_0,0}^t \{D_{n1}\} & = (1 - a_t) \int_{\mathbb{R}^2 \setminus C \times C} T_t^{(1)}(x_0, dx) T_t^{(2)}(y_0, dy) \bar{\mathbb{P}}_{x,y,0}^{t+1} \{\bar{\sigma}_1 = n\} \\
 & \leq (1 - a_t) \int_{\mathbb{R}^2 \setminus C \times C} T_t^{(1)}(x_0, dx) T_t^{(2)}(y_0, dy) \leq (1 - a_*).
 \end{aligned}$$

So, we have shown that for all positive integers n and $k = 1$ inequality (29) is valid.

In what follows we will use a simplified notation

$$\bar{\sigma}_i = \bar{\sigma}_{C \times C}(i).$$

Assume it is valid for all positive n and k , let us check it for $k + 1$. Clearly, it is nothing to prove if $n < k + 1$ since the underlying probability is zero. Consider $n \geq k + 1$ and write

$$\bar{\mathbb{P}}^t_{x_0, y_0, 0} \{D_{n, k+1}\} = \sum_{s=1}^{n-k} \bar{\mathbb{P}}^t_{x_0, y_0, 0} \{\bar{\sigma}_1 = s, D_{n, k+1}\}.$$

Now we use the conditioning by \mathcal{F}_s , the Markov property and the fact that

$$\begin{aligned} \{\bar{\sigma}_1 = s, D_{n, k+1}\} &= \{D_{s1}, d_{s+1} = \dots = d_n = 0, \bar{\sigma}_{k+1} = n\} \\ &= D_{s1} \cap \theta_s \{d_1 = \dots = d_{n-s} = 0, \bar{\sigma}_k = n - s\} = D_{s1} \cap \theta_s D_{n-s, k}, \end{aligned}$$

where θ_s is a shift operator. The obvious inclusion $D_s \in \mathcal{F}_s$ and the induction assumption allow us to obtain

$$\begin{aligned} \sum_{s=1}^{n-k} \bar{\mathbb{P}}^t_{x_0, y_0, 0} \{\bar{\sigma}_1 = s, D_{n, k+1}\} &= \sum_{s=1}^{n-k} \bar{\mathbb{E}}^t_{x_0, y_0, 0} \left\{ D_{s1}, \bar{\mathbb{P}}^t_{\bar{Z}_s} \{D_{n-s, k}\} \right\} \\ &\leq \sum_{s=1}^{n-k} \sup_{x, y \in C, t} \bar{\mathbb{P}}^t_{x, y} \{D_{n-s, k}\} \bar{\mathbb{P}}^t_{x_0, y_0, 0}(D_{s1}) \\ &\leq (1 - a_*)^k \sum_{s=1}^{n-k} \bar{\mathbb{P}}^t_{x_0, y_0, 0}(D_{s1}) = (1 - a_*)^k \sum_{s=1}^{n-k} \bar{\mathbb{P}}^t_{x_0, y_0, 0}(d_1 = 0, \bar{\sigma}_1 = s) \\ &= (1 - a_*)^k \bar{\mathbb{P}}^t_{x_0, y_0, 0}(d_1 = 0, \bar{\sigma}_1 \leq n - k) \\ &= (1 - a_*)^k (1 - a_t) \int_{\mathbb{R}^2 \setminus C \times C} T_t^{(1)}(x_0, dx) T_t^{(2)}(y_0, dy) \times \\ &\quad \times \bar{\mathbb{P}}^t_{x, y, 0} \{\bar{\sigma}_1 \leq n - k - 1\} + (1 - a_*)^k (1 - a_t) T_t^{(1)}(x_0, C) T_t^{(2)}(y_0, C) \\ &\leq (1 - a_*)^{k+1} \int_{\mathbb{R}^2} T_t^{(1)}(x_0, dx) T_t^{(2)}(y_0, dy) = (1 - \alpha)^{k+1}. \end{aligned} \tag{30}$$

Thus the inequality (29) is proved for all n, k . □

Lemma 6. *Let $m \geq 2$ be an integer, k an arbitrary positive integer and $j \in \{0, \dots, k - 1\}$. Then*

$$\rho_{mk+j, k} \leq (1 - a_*)^{k-1} S \psi_1^{-(m-1)},$$

where S is defined in Lemma 3.

Proof. In this lemma we use the following notation which is aimed to simplify further formulas:

$$\begin{aligned} n &= mk + j, \\ \bar{\sigma}_0 &= 0, \\ \bar{\sigma}_i &= \bar{\sigma}_{C \times C}(i), \quad i \geq 0, \\ \bar{\Delta}_i &= \bar{\sigma}_i - \bar{\sigma}_{i-1}, \quad i \geq 1, \end{aligned}$$

$$\begin{aligned} \bar{\zeta}_k &= \max_{1 \leq i \leq k} \{\bar{\Delta}_i\}, \\ \bar{\tau}_n &= \min\{l \geq 0 : \bar{\Delta}_l = \bar{\zeta}_k\} \leq k, \\ D &= D_{mk+j,k}. \end{aligned}$$

It is easy to see that

$$D \subset \{\bar{\zeta}_k \geq m\}. \tag{31}$$

Indeed, assume $D \cap \{\bar{\zeta}_k < m\} \neq \emptyset$ and let $\omega \in D \cap \{\bar{\zeta}_k < m\}$. Then

$$\bar{\sigma}_k(\omega) = \sum_{j=1}^k \bar{\Delta}_j(\omega) \leq k\bar{\zeta}_k(\omega) < mk.$$

This contradicts to the definition of $D = D_{mk+j,k}$, so (31) is proved.

Assume first that $\bar{\tau}_n = 1$. This means that the period before the first hitting is the longest, or, in other words, all subsequent intervals between hitting $C \times C$ are smaller or equal to the first one. Note, that since $m \geq 2$ and $\bar{\zeta}_k(\omega) \geq m$, $\omega \in D$, we conclude that hitting will not occur at the first timestep. This observation allows us to write the following decomposition for $x_0, y_0 \in C$:

$$\begin{aligned} &\bar{\mathbb{P}}_{x_0, y_0, 0}^t \{D_{nk}, \bar{\tau}_n = 1\} \\ &= (1 - a_t) \int_{\mathbb{R}^2 \setminus C \times C} T_{x_0 y_0}^{(t)}(dx, dy) \bar{\mathbb{P}}_{x, y, 0}^{t+1} \{\bar{\sigma}_1 \geq m - 1, D_{n-1, k}\}, \end{aligned} \tag{32}$$

where we used the notation $T_{xy}^{(t)}(A, B) = T_t^{(1)}(x, A)T_t^{(2)}(y, B)$ introduced in the previous section.

Now, using the same arguments as in the proof of Lemma 5 we notice that for $s \geq m$

$$\begin{aligned} &\bar{\mathbb{P}}_{x, y, 0}^{t+1} \{\bar{\sigma}_1 = s, D_{n-1, k}\} = \bar{\mathbb{E}}_{x, y, 0}^{t+1} \left[\bar{\sigma}_1 = s, D_s, \bar{\mathbb{P}}_{\bar{Z}_s}^{t+s} \{D_{n-s, k-1}\} \right] \\ &\leq \rho_{n-s, k-1} \bar{\mathbb{P}}_{x, y, 0}^{t+1} \{\bar{\sigma}_1 = s, D_s\} \leq (1 - a_*)^{k-1} \bar{\mathbb{P}}_{x, y, 0}^{t+1} \{\bar{\sigma}_1 = s, D_s\}, \end{aligned}$$

where we used Lemma 5.

Substituting this upper bound into (32), we get

$$\begin{aligned} \bar{\mathbb{P}}_{x_0, y_0, 0}^t \{D_{nk}, \bar{\tau}_n = 1\} &\leq (1 - a_*)^k \sum_{s \geq m-1} \int_{\mathbb{R}^2 \setminus C \times C} T_{x_0 y_0}^{(t)}(dx, dy) \bar{\mathbb{P}}_{x, y, 0}^{t+1} [\bar{\sigma}_1 = s, D_s] \\ &= (1 - a_*)^k \int_{\mathbb{R}^2 \setminus C \times C} T_{x_0 y_0}(dx, dy) \bar{\mathbb{P}}_{x, y, 0}^{t+1} \{\bar{\sigma}_1 \geq m - 1\} \\ &\leq \frac{(1 - a_*)^k}{\psi_1^{m-1}} \int_{\mathbb{R}^2 \setminus C \times C} T_{x_0 y_0}(dx, dy) h(x, y) \\ &= \frac{(1 - a_*)^{k-1} S}{\psi_1^{m-1}}. \end{aligned}$$

So, finally we get the estimate

$$\bar{\mathbb{P}}_{x_0, y_0, 0}^t \{D_{nk}, \bar{\tau}_n = 1\} \leq \frac{S_1(1 - a_*)^k}{\psi_1^{m-1}}. \tag{33}$$

Here we introduced $S_1 = S/(1 - a_*)$ because we would like to maintain the term $(1 - a_*)^k$.

Now we use (31) and (33) to write a decomposition with respect to $\bar{\tau}_n$:

$$\begin{aligned} \bar{\mathbb{P}}_{x, y, 0}^t \{D_{nk}\} &= \bar{\mathbb{P}}_{x, y, 0}^t \{D_{nk}, \bar{\zeta}_k \geq m\} = \sum_{j=1}^k \bar{\mathbb{P}}_{x, y, 0}^t \{D_{nk}, \bar{\zeta}_k \geq m, \bar{\tau}_n = j\} \\ &= \sum_{j=1}^k \sum_s \bar{\mathbb{E}}_{x, y, 0}^t \left[\bar{\mathbb{P}}_{Z_j^{(1)}, Z_j^{(2)}, 0}^{t+j} \{\bar{\tau}_{n-s} = 1, D_{n-s, k-j+1}\}; D_{s, j-1}, \bar{\tau}_n = j, \bar{\sigma}_{j-1} = s \right] \\ &\leq \frac{S_1}{\psi_1^{m-1}} \sum_{j=1}^k \sum_s (1 - a_*)^{k-j+1} \bar{\mathbb{P}}_{x, y, 0}^t [D_{s, j-1}, \bar{\tau}_n = j, \bar{\sigma}_{j-1} = s] \\ &= \frac{S_1}{\psi_1^{m-1}} \sum_{j=1}^k \sum_s (1 - a_*)^{k-j+1} \bar{\mathbb{P}}_{x, y, 0}^t [D_{s, j-1}, \bar{\tau}_n = j, \bar{\sigma}_{j-1} = s] \\ &\leq \frac{S_1}{\psi_1^{m-1}} \sum_{j=1}^k \sum_s (1 - a_*)^{k-j+1} \bar{\mathbb{P}}_{x, y, 0}^t [\bar{\tau}_n = j, \bar{\sigma}_{j-1} = s] \sup_{u, v \in C, t} \bar{\mathbb{P}}_{u, v, 0}^t \{D_{s, j-1}\} \\ &= \frac{S_1}{\psi_1^{m-1}} \sum_{j=1}^k \sum_s (1 - a_*)^{k-j+1} \bar{\mathbb{P}}_{x, y, 0}^t [\bar{\tau}_n = j, \bar{\sigma}_{j-1} = s] \rho_{s, j-1} \\ &\leq \frac{S_1}{\psi_1^{m-1}} \sum_{j=1}^k (1 - a_*)^{k-j+1} (1 - a_*)^{j-1} \bar{\mathbb{P}}_{x, y, 0}^t [\bar{\tau}_n = j] = \frac{S_1(1 - a_*)^k}{\psi_1^{m-1}}. \end{aligned}$$

Note that we used Lemma 5 in the last inequality. □

Finally we can show that $\{\rho_{nk}\}$ is summable.

Lemma 7.

$$\rho := \sum_{n=1}^{\infty} \sum_{k=1}^n \rho_{nk} \leq \frac{1 - a_*}{(a_*)^2} \left(1 + \frac{S}{(1 - a_*)(\psi_1 - 1)} \right) < \infty.$$

Proof. As in the previous lemma, we denote by $S_1 = S/(1 - a_*)$, so that we can maintain the term $(1 - a_*)^k$. Then we obtain the following upper bound:

$$\begin{aligned} \sum_{n=1}^{\infty} \sum_{k=1}^n \rho_{nk} &= \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} \rho_{nk} = \sum_{k=1}^{\infty} \sum_{m=1}^{\infty} \left(\sum_{j=0}^{k-1} \rho_{mk+j, k} \right) \\ &= \sum_{k=1}^{\infty} \left(\sum_{j=0}^{k-1} \rho_{k+j, k} \right) + \sum_{k=1}^{\infty} \sum_{m=2}^{\infty} \left(\sum_{j=0}^{k-1} \rho_{mk+j, k} \right) \end{aligned}$$

$$\begin{aligned}
 &\leq \sum_{k=1}^{\infty} \left(\sum_{j=0}^{k-1} \rho_{k+j,k} \right) + \sum_{k=1}^{\infty} \sum_{m=2}^{\infty} k(1-a_*)^k \psi_1^{-(m-1)} S_1 \\
 &\leq \sum_{k=1}^{\infty} (1-a_*)^k k + \sum_{k=1}^{\infty} \sum_{m=1}^{\infty} k(1-a_*)^k \psi_1^{-m} S_1 \\
 &= \frac{1-a_*}{(a_*)^2} + S_1 \sum_{k=1}^{\infty} k(1-a_*)^k \sum_{m=1}^{\infty} \psi_1^{-m} \\
 &= \frac{1-a_*}{(a_*)^2} + S_1 \frac{1-a_*}{(\psi_1-1)(a_*)^2} = \frac{1-a_*}{(a_*)^2} \left(1 + S_1 \frac{1}{\psi_1-1} \right).
 \end{aligned}$$

Note that we used Lemma 6 in the first inequality. □

Lemma 8.

$$\sum_{n \geq 1} \sup_{x \in \{m, m+1\}, t} \mathbb{P}_{x,x,1}^t \{D_n\} \leq \varepsilon M_1 + \hat{r}_m M_2$$

where ρ is defined in Lemma 7 and

$$\begin{aligned}
 M_1 &= \mathbb{E}_0 \left(\frac{\psi_1 S(1 + \rho \psi_1)}{(\psi_1 - 1)^2} + \frac{\psi_1(1 + S)}{\psi_1 - 1} \right), \\
 M_2 &= 2 \mathbb{E}_0 \left(S \frac{\psi_1(1 + \rho + \Xi_1(1 + S))}{\psi_1 - 1} + \mathbb{E}_1 \frac{\psi_1 S(1 + \rho \psi_1)}{(\psi_1 - 1)^2} \right),
 \end{aligned}$$

$\mathbb{E}_0, \mathbb{E}_1$ are from Theorem 3.

Proof. Let us introduce

$$H_t^{(m)}(x) = \sup_{|x| \in \{m, m+1\}} H_t(x).$$

Using Lemmas 2, 3, 4 and 6, we get

$$\begin{aligned}
 \sum_{n \geq 1} \sup_{|x| \in \{m, m+1\}} \mathbb{P}_{x,x,1}^t \{D_n\} &\leq \sum_{n \geq 1} \left(\frac{H_t^{(m)}(x)}{\psi_1^{n-1}} + \frac{S((n-1)H_t^{(m)}(x) + \varepsilon)}{\psi_1^{n-1}} \right) \\
 &+ \sum_{n \geq 1} \sum_{k \geq 2} S H_t^{(m)}(x) \sum_{i=2}^{n-k+1} \psi_1^{-(i-1)} \sum_{j=k-1}^{n-i} \rho_{j,k-1} \psi_1^{-(n-i-j)} \\
 &+ \sum_{n \geq 1} \sum_{k \geq 2} \varepsilon S \sum_{j=k-1}^{n-1} \rho_{j,k-1} \psi_1^{-(n-j-1)} = A_1 + A_2 + \varepsilon S A_3.
 \end{aligned}$$

We start with A_1 .

$$A_1 = \sum_{n \geq 1} \left(\frac{H_t^{(m)}(x)}{\psi_1^{n-1}} + \frac{S((n-1)H_t^{(m)}(x) + \varepsilon)}{\psi_1^{n-1}} \right) = \frac{\psi_1(H_t^{(m)}(x) + \varepsilon S)}{\psi_1 - 1}$$

$$+ \frac{\psi_1 SH_t^{(m)}(x)}{(\psi_1 - 1)^2}.$$

Before moving to A_2 let us first look at the sum

$$\begin{aligned} & \sum_{n=1}^{\infty} \sum_{k \geq 2} \sum_{i=1}^{n-k+1} \psi_1^{-(i-1)} \sum_{j=k-1}^{n-i} \rho_{j,k-1} \psi_1^{-(n-i-j)} \\ &= \sum_{n=1}^{\infty} \sum_{k \geq 1} \sum_{i=1}^{n-k} \psi_1^{-(i-1)} \sum_{j=k}^{n-i} \rho_{jk} \psi_1^{-(n-i-j)}. \end{aligned}$$

It is clear that k may not exceed n , as well as $\rho_{jk} = 0$ if $j < k$. Put $p_n = \psi_1^{-(n-1)}$, $n \geq 1$, $p_0 = 0$ and $q_j(k) = \rho_{j+k,k}$ for a fixed k , $j \geq 0$. We can then swap n and k and rewrite the previous sums as

$$\begin{aligned} & \sum_{k=1}^{\infty} \sum_{n \geq k} \sum_{i=1}^{n-k} \psi_1^{-(i-1)} \sum_{j=k}^{n-i} \rho_{jk} \psi_1^{-(n-i-j)} \\ &= \sum_{k=1}^{\infty} \sum_{n \geq k} \sum_{i=1}^{n-k} \psi_1^{-(i-1)} \sum_{j=0}^{n-k-i} q_j(k) \psi_1^{-(n-k-i-j)} \\ &= \sum_{k=1}^{\infty} \sum_{n \geq k} (p \star q(k) \star p)_{n-k} = \sum_{k=1}^{\infty} \left(\sum_{n=1}^{\infty} p_n \right)^2 \left(\sum_{j \geq 0} q_j(k) \right) \\ &= \left(\sum_{n=1}^{\infty} \psi_1^{-(n-1)} \right)^2 \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \rho_{jk} = \left(\frac{\psi_1}{\psi_1 - 1} \right)^2 \rho. \end{aligned}$$

Now we can calculate A_2 .

$$\begin{aligned} A_2 &= \sum_{n \geq 1} \sum_{k \geq 2} SH_t^{(m)}(x) \sum_{i=2}^{n-k+1} \psi_1^{-(i-1)} \sum_{j=k-1}^{n-i} \rho_{j,k-1} \psi_1^{-(n-i-j)} \\ &= SH_t^{(m)}(x) \rho \left(\frac{\psi_1}{\psi_1 - 1} \right)^2 - SH_t^{(m)}(x) A_3. \end{aligned}$$

For A_3 we have

$$\begin{aligned} A_3 &= \sum_{n \geq 1} \sum_{k \geq 2} \sum_{j=k-1}^{n-1} \rho_{j,k-1} \psi_1^{-(n-j-1)} = \sum_{n \geq 1} \sum_{k \geq 1} \sum_{j=k}^{n-1} \rho_{j,k} \psi_1^{-(n-j-1)} \\ &= \left(\sum_{n \geq 0} \psi_1^{-n} \right) \left(\sum_{j,k} \rho_{j,k} \right) = \frac{\rho \psi_1}{\psi_1 - 1}. \end{aligned}$$

Now we have

$$\sum_{n \geq 1} \sup_{|x| \in [m, m+1)} \mathbb{P}'_{x,x,1} \{D_n\} \leq \frac{\psi_1 (H_t^{(m)}(x) + \varepsilon S)}{\psi_1 - 1}$$

$$\begin{aligned}
 & + \frac{\psi_1 S H_t^{(m)}(x)}{(\psi_1 - 1)^2} + S H_t^{(m)}(x) \rho \left(\frac{\psi_1}{\psi_1 - 1} \right)^2 \\
 & + (\varepsilon S - S H_t^{(m)}(x)) \frac{\rho \psi_1}{\psi_1 - 1} = \frac{\psi_1 S H_t^{(m)}(x) (1 + \rho \psi_1)}{(\psi_1 - 1)^2} \\
 & + \frac{\psi_1 (H_t^{(m)}(x) + \varepsilon S + S H_t^{(m)}(x) + \rho \varepsilon S - \rho S H_t^{(m)}(x))}{\psi_1 - 1} \\
 & \leq \varepsilon S \frac{\psi_1 (1 + \rho)}{\psi_1 - 1} + H_t^{(m)}(x) \left(\frac{\psi_1 S (1 + \rho \psi_1)}{(\psi_1 - 1)^2} + \frac{\psi_1 (1 + S)}{\psi_1 - 1} \right).
 \end{aligned}$$

We can use result of Theorem 3 and Condition T to obtain

$$\begin{aligned}
 H_t^{(m)}(x) & \leq \sup_{|x| \in [m, m+1]} \int_{\mathbb{R}^2} \frac{R_t^{(1)}(x, dy) R_t^{(2)}(x, dz)}{1 - Q_t(x, \mathbb{R})} (\Xi_0(|y| + |z|) + \Xi_1) \\
 & \leq \Xi_0 \left(\sup_{|x| \in [m, m+1]} \int_{\mathbb{R}} R_t^{(1)}(x, dy) |y| + \sup_{|x| \in [m, m+1]} \int_{\mathbb{R}} R_t^{(2)}(x, dy) |y| \right) + \varepsilon \Xi_1 \\
 & \leq 2 \Xi_0 \hat{r}_m + \varepsilon \Xi_1.
 \end{aligned}$$

Finally we can write

$$\begin{aligned}
 \sum_{n \geq 1} \sup_{|x| \in [m, m+1]} \mathbb{P}_{x, x, 1}^f \{D_n\} & \leq 2 \hat{r}_m \Xi_0 \left(\frac{\psi_1 S (1 + \rho \psi_1)}{(\psi_1 - 1)^2} + \frac{\psi_1 (1 + S)}{\psi_1 - 1} \right) \\
 & + \varepsilon \left(S \frac{\psi_1 (1 + \rho + \Xi_1 (1 + S))}{\psi_1 - 1} + \Xi_1 \frac{\psi_1 S (1 + \rho \psi_1)}{(\psi_1 - 1)^2} \right). \quad \square
 \end{aligned}$$

Appendix A

In this appendix we want to demonstrate how to calculate the constants involved in the bound in Theorem 2.

First, we note that the bound in Theorem 2 depends on constants M_1 and M_2 that are defined in Lemma 8. These constants in turn, depend on Ξ_0, Ξ_1 defined in Theorem 3 and S defined in Lemma 3. Constant Ξ_1 is known as far as we know Ξ_0 (see (18)). Thus, we have to calculate Ξ_0 and S .

We start with Ξ_0 . From (17)

$$\Xi_0 = \frac{M C_0^2 (2 + 2G + c - \delta)}{1 - a^*},$$

and we see that the only unknown constant is M . It is defined in [12], Theorem 4.2, as

$$M = \left(1 + \frac{1}{1 - \sqrt{(1 + \varepsilon)(1 - \gamma_1)}} \right) \left(\frac{1 + \gamma_0}{1 - \gamma_0} \right),$$

where ε is an arbitrary constant such that $(1 + \varepsilon)(1 - \gamma_1) < 1$. Thus, we only need to calculate γ_0 and γ_1 . This can be done using Theorem 4.3 from [12]. The essential

condition in Theorem 4.3 reduces to the exponentially small bound on the tails of $\sigma_C^{(i)}$:

$$\mathbb{P}_x^t \left\{ \sigma_C^{(i)} > n \right\} \leq (V(x)\mathbb{1}_{\mathbb{R} \setminus C}(x) + D\mathbb{1}_C(x))\psi^{-n},$$

for some $\psi > 0$ and $V(x)$. Under the conditions of Theorem 2 (and thus Corollary 1) this is true, with $V(x) = C_0(1 + |x|)$ and $D = C_0(c + 2 + 2G - \delta)$.

We start with calculating

$$C_0 \sup_{t,x \in C, i} \int_{\mathbb{R}} \frac{P_{t,i}(x, dy) - a_t v_t(dy)}{1 - a_t} (1 + |y|) \leq C_0 \left(1 + \frac{1 + 2G + c - \delta}{1 - a^*} \right).$$

We denote the last expression by \hat{Q} , so that

$$\hat{Q} = C_0 \left(1 + \frac{1 + 2G + c - \delta}{1 - a^*} \right).$$

Theorem 4.3 from [12] now gives us

$$\delta_0 = \sqrt{1 + \theta \frac{a_*(\psi - 1)}{(1 - a_*)\psi \hat{Q} + a_*}},$$

where θ is an arbitrary number in $(0, 1)$.

$$\begin{aligned} \gamma_0 &= \left(\frac{\theta a_*}{\psi + a_*(1 - \theta)(\psi \hat{Q}(1 - a_*))^{-1}} \right)^{\frac{1}{2}}, \\ \hat{D} &= D \frac{1 + \gamma_0}{1 - \gamma_0} \left(1 + \frac{\delta_0 - 1}{\psi - \delta_0} \psi \hat{Q} \right) \left(1 + \frac{\delta_0(\psi - 1)}{\psi - \delta_0} \right), \\ m &= \min\{n \geq 1 \mid \hat{D}\delta_0^{-1} < 1\}, \\ \gamma_1 &= (a_* \inf_t v_t(C))^m \exp \left(\ln \left(1 - \hat{D}\delta_0^{-m} \right) \left(\frac{\delta_0^{m+1}}{\delta_0 - 1} - 1 \right) \right). \end{aligned}$$

So, we calculated γ_0 and γ_1 which allows us to express M and thus Ξ_0 and Ξ_1 in terms of known values.

Finally, we should calculate S from Lemma 3.

$$\begin{aligned} S &= (1 - a_*) \sup_{x,y \in C, t} \int_{\mathbb{R}^2 \setminus C \times C} T_{xy}^{(t)}(du, dv)h(u, v) \\ &\leq (1 - a_*) \sup_t \int_{\mathbb{R}} \left(T_t^{(1)}(x, du)\Xi_0|u| + T_t^{(2)}(y, dv)\Xi_0|v| + \Xi_1 \right) \\ &\leq \frac{1 - a_*}{1 - a^*} (2\Xi_0(1 + 2G + c - \delta) + \Xi_1(1 - a^*)). \end{aligned}$$

Appendix B

Condition D. We say that a sequence of Markov kernels $(P_t, t \in \mathbb{N}_0)$ satisfies Condition D with the set $C \in \mathcal{E}$ if:

- (1) There exists a sequence of measurable functions $V_k: E \rightarrow [1, \infty]$ and two sequences of positive constants $\{\lambda_k, k \geq 0\}$, and $\{b_k, k \geq 0\}$ such that for all $x \in E$

$$P_k V_{k+1}(x) \leq \lambda_{k+1} V_k(x) + b_k \mathbb{1}_C(x); \tag{34}$$

- (2) The sequence $\{\lambda_k, k \geq 0\}$ defined in item (1), satisfies

$$\sum_{k=0}^{\infty} \left(\prod_{j=0}^k (\lambda_j \vee 1) \right)^{-1} (1 - \lambda_k)^+ = \infty.$$

Here $a \vee b = \max\{a, b\}$, and $a^+ = \max\{a, 0\}$.

Theorem 4. *Let (P_t) be a sequence of Markov transition kernels, $C \in \mathcal{E}$ be some set and Condition D hold true.*

Then the following two statements hold true.

1. *For any $t \in \mathbb{N}_0$ and $x \in E$ such that $P_t V_{t+1}(x) < \infty$:*

$$\mathbb{P}_x^t \{ \sigma_C < \infty \} = 1.$$

2. *For any $x \in E, t \in \mathbb{N}_0$:*

$$\mathbb{E}_x^t \left[\prod_{j=1}^{\sigma_C} \lambda_j^{-1} \right] \leq V_t(x) + \lambda_{t+1}^{-1} b_t \mathbb{1}_C(x).$$

The theorem is proven in [13], Theorem 1.

References

[1] Andrieu, C., Fort, G., Vihola, M.: Quantitative convergence rates for subgeometric Markov chains. *Ann. Appl. Probab.* **52**, 391–404 (2015) [MR3372082](#). <https://doi.org/10.1239/jap/1437658605>

[2] Baxendale, P.: Renewal theory and computable convergence rates for geometrically ergodic Markov chains. *Ann. Appl. Probab.* **15**, 700–738 (2005) [MR2114987](#). <https://doi.org/10.1214/105051604000000710>

[3] Douc, R., Moulines, E., Soulier, P.: Practical drift conditions for subgeometric rates of convergence. *Ann. Appl. Probab.* **14**, 1353–1377 (2004) [MR2071426](#). <https://doi.org/10.1214/105051604000000323>

[4] Douc, R., Moulines, E., Soulier, P.: Quantitative bounds on convergence of time-inhomogeneous Markov chains. *Ann. Appl. Probab.* **14**, 1643–1665 (2004) [MR2099647](#). <https://doi.org/10.1214/105051604000000620>

[5] Douc, R., Moulines, E., Priouret, P., P., S.: *Markov Chains*. Springer, Switzerland (2018). [MR3889011](#). <https://doi.org/10.1007/978-3-319-97704-1>

[6] D’Amico, G., Gismondi, F., Janssen, J., Manca, R., Petroni, F., Volpe di Prignano, E.: The study of basic risk processes by discrete-time non-homogeneous markov processes. *Theory Probab. Math. Stat.* **96**, 27–43 (2018). [MR3666870](#). <https://doi.org/10.1090/tpms/1032>

- [7] Fort, G., Roberts, G.O.: Subgeometric ergodicity of strong Markov processes. *Ann. Appl. Probab.* **15**, 1565–1589 (2005) [MR2134115](#). <https://doi.org/10.1214/105051605000000115>
- [8] Gismondi, F., Janssen, J., Manca, R.: Non-homogeneous time convolutions, renewal processes and age-dependent mean number of notorcar accidents. *Ann. Actuar. Sci.* **9**, 36–57 (2015)
- [9] Golomoziy, V.: An inequality for the coupling moment in the case of two inhomogeneous markov chains. *Theory Probab. Math. Stat.* **90**, 43–56 (2015). [MR3241859](#). <https://doi.org/10.1090/tpms/948>
- [10] Golomoziy, V.: An estimate of the expectation of the excess of a renewal sequence generated by a time-inhomogeneous markov chain if a square-integrable majorizing sequence exists. *Theory Probab. Math. Stat.* **94**, 53–62 (2017). [MR3553453](#). <https://doi.org/10.1090/tpms/1008>
- [11] Golomoziy, V.: On estimation of expectation of simultaneous renewal time of time-inhomogeneous markov chains using dominating sequence. *Mod. Stoch. Theory Appl.* **6(3)**, 333–343 (2019). [MR4028080](#). <https://doi.org/10.15559/19-vmsta138>
- [12] Golomoziy, V.: Exponential moments of simultaneous hitting time for non-atomic markov chains. *Glas. Mat.* **57(1)**, 129–147 (2022). [MR4450636](#). <https://doi.org/10.3336/gm.57.1.09>
- [13] Golomoziy, V.: Computable bounds of exponential moments of simultaneous hitting time for two time-inhomogeneous atomic markov chains. In: *Stochastic Processes, Statistical Methods, and Engineering Mathematics. SPAS 2019. Springer Proceedings in Mathematics and Statistics 408*, pp. 97–119. Springer, Cham (2023). [MR4450636](#). https://doi.org/10.1007/978-3-031-17820-7_5
- [14] Golomoziy, V., Kartashov, M.: The mean coupling time for independent discrete renewal processes. *Theory Probab. Math. Stat.* **84**, 79–86 (2012). [MR2857418](#). <https://doi.org/10.1090/S0094-9000-2012-00855-7>
- [15] Golomoziy, V., Kartashov, M.: Maximal coupling procedure and stability of discrete markov chains. i. *Theory Probab. Math. Stat.* **86**, 93–104 (2013). [MR3241447](#). <https://doi.org/10.1090/S0094-9000-2014-00905-9>
- [16] Golomoziy, V., Kartashov, M.: Maximal coupling procedure and stability of discrete markov chains. ii. *Theory Probab. Math. Stat.* **87**, 65–78 (2013). [MR3241447](#). <https://doi.org/10.1090/S0094-9000-2014-00905-9>
- [17] Golomoziy, V., Kartashov, M.: On the integrability of the coupling moment for time-inhomogeneous markov chains. *Theory Probab. Math. Stat.* **89**, 1–12 (2014). [MR3235170](#). <https://doi.org/10.1090/S0094-9000-2015-00930-3>
- [18] Golomoziy, V., Kartashov, M.: Maximal coupling and stability of discrete non-homogeneous markov chains. *Theory Probab. Math. Stat.* **91**, 17–27 (2015). [MR2986452](#). <https://doi.org/10.1090/S0094-9000-2013-00891-6>
- [19] Golomoziy, V., Kartashov, M.: Maximal coupling and v -stability of discrete nonhomogeneous markov chains. *Theory Probab. Math. Stat.* **93**, 19–31 (2016). [MR3553437](#). <https://doi.org/10.1090/tpms/992>
- [20] Golomoziy, V., Mishura, Y.: Stability estimates for finite-dimensional distributions of time-inhomogeneous markov chains. *Mathematics* **8(2)**, 1–13 (2020). <https://doi.org/10.3390/math8020174>
- [21] Golomoziy, V., Kartashov, M., Kartashov, Y.: Impact of the stress factor on the price of widow’s pensions. *proofs. Theory Probab. Math. Stat.* **92**, 17–22 (2016). [MR3330687](#). <https://doi.org/10.1090/tpms/979>

- [22] Lindvall, T.: Lectures on Coupling Method. John Wiley and Sons (1991) [MR1180522](#)
- [23] Melfi, V.: Nonlinear Markov renewal theory with statistical applications. *Ann. Probab.* **20**, 753–771 (1992) [MR1159572](#)
- [24] Meyn, S., Tweedie, R.: *Markov Chains and Stochastic Stability*. Springer, Switzerland (1993). [MR1287609](#). <https://doi.org/10.1007/978-1-4471-3267-7>
- [25] Thorisson, H.: *Coupling, Stationarity, and Regeneration*. Springer, New York (2000) [MR1741181](#). <https://doi.org/10.1007/978-1-4612-1236-2>