Chapter 5
Computable Bounds of Exponential Moments of Simultaneous Hitting Time for Two Time-Inhomogeneous Atomic Markov Chains

Vitaliy Golomoziy

Abstract In this paper, we study the first simultaneous hitting of the atom by two discrete-time, inhomogeneous Markov chains with values in general phase space. We establish conditions for the existence and find computable bounds for the hitting time's exponential moment using a geometric drift condition adapted for time-inhomogeneous Markov chains.

Keywords Markov chain · Hitting time

MSC 2020 60J10

5.1 Introduction

Properties of hitting moments play an important role in the Markov chains theory, and various drift conditions are practical tools used in applications when dealing with such moments. The theory of hitting moments and convergence of homogeneous Markov chains is well developed. Many books are devoted to that topic, see, for example, [7, 23, 30]. The first work includes a good overview of the recent results, and we will refer to this book repeatedly.

Hitting moments play such an extraordinary role in the homogeneous Markov chains theory because of two important methods that are used in research nowadays: splitting and coupling. The splitting method was introduced in the seminal work of Nummelin [25] and was further developed by other authors. A famous book [23] presents the comprehensive theory of homogeneous Markov chains developed using the splitting technique.

The coupling method was first used by Doeblin [6] and became very popular afterward. The essence of the coupling method is covered in books [20, 31]. In the last years coupling method was extensively used (see [1, 8, 14–19, 27]).

V. Golomoziy (✉)
Taras Shevchenko National University of Kyiv, 64 Volodymyrska st., 01033 Kyiv, Ukraine
e-mail: vitaliy.golomoziy@knu.ua

A. Malyarenko et al. (eds.), *Stochastic Processes, Statistical Methods, and Engineering Mathematics*, Springer Proceedings in Mathematics & Statistics 408,
https://doi.org/10.1007/978-3-031-17820-7_5

At the same time majority of the papers were devoted to the homogeneous Markov chains, while the inhomogeneous theory is not so well developed. Inhomogeneous chains play an important role in applications, for example, in actuarial mathematics [16], risk theory [2], stochastic optimization [4, 22], etc. In addition, there is a great theoretical interest to this area of research. Foundational results in the theory of inhomogeneous Markov chains were established by Dobrushin [5], and following works [21, 26, 28]. Most of the aforementioned works are related to the ergodicity of time-inhomogeneous chains. However, the coupling technique can be applied to the study of the stability of inhomogeneous chains. By stability, we mean not only stability of the same chain with regard to different initial distributions, but proximity in some sense, of two different time-inhomogeneous chains. Stability results for inhomogeneous chains could be found in [14–19].

The essential tool in the application of the coupling method to stability research in the inhomogeneous case is the renewal theory. This theory is well-developed for the homogeneous renewal sequences, and such classical results as Blackwell, Kendall theorems, and Key Renewal Theorem play an important role. However, there are no such strong results for inhomogeneous renewal sequences (such sequences are generated by the inhomogeneous Markov chains). We can highlight the works of Chow and Robbins [3], and Smith [29] in this domain. Properties of renewal sequences (such as estimation of the expectation of the simultaneous hitting time) in the context of inhomogeneous Markov chains were studied in the papers [9–13]. The present paper can be considered as the contribution to the renewal theory of time-inhomogeneous Markov chains.

An important aspect of any research is the ability to verify the conditions in practical application. The standard tool that is used for that purpose in the Markov chains theory is drift conditions. In this paper, we develop a drift condition that is sufficient to ensure the existence of the exponential moment of the return time and evaluate its bounds. Such conditions extensively used in the theory of both homogeneous and inhomogeneous Markov chains. See [1, 8], as an example of drift condition being used for studying ergodicity of the inhomogeneous Markov chains.

The paper is organized as follows. In Sect. 5.2 we construct probability space and introduce notation used in the rest of the paper. Section 5.3 includes drift conditions adapted to time-inhomogeneous chains that ensure an exponential moment's existence. Section 5.4 presents theorems that guarantee the existence of an exponential moment of simultaneous return time of the couple of different inhomogeneous chains, computable bounds for such moment and application to ergodicity. Section 5.5 includes auxiliary lemmas used in the proof of the main results. In Appendix we provide the well-known Comparison Theorem for adapted processes with discrete time.

5.2 Notation

Let $(E, \mathcal{E})$ be a measurable space, $\mathbb{M}_1(\mathcal{E})$, $\mathbb{F}_+(\mathcal{E})$, $\mathbb{F}_b(\mathcal{E})$ be the spaces of all probability measures, positive and bounded measurable functions on $(E, \mathcal{E})$, respectively. Denote by $\mathbb{N}_0$ a set of all nonnegative integers, $\mathbb{N}_0 := \{0, 1, 2, \ldots\}$. In this paper, we study time-inhomogeneous Markov chains, taking values in the space $(E, \mathcal{E})$. We associate time-inhomogeneous Markov chain with a sequence of Markov kernels $P_t\colon E \times \mathcal{E} \to [0, 1]$, where $t \in \mathbb{N}_0$. Markov kernel $P_t(x, A)$ stands for the probability for the chain to be at time t at the state $x \in E$ and hit the set $A \in \mathcal{E}$ at time $t + 1$.

We introduce a special notation for a product of transition kernels:

$$P^{t,n}(x, A) = \left(\prod_{k=0}^{n-1} P_{t+k}\right)(x, A) = \int_E \cdots \int_E P_t(x, dx_1) \ldots P_{t+n-1}(x_{n-1}, A),\ n \geq 1,$$

$$P^{t,0}(x, A) = 1_A(x).$$

It is very well known that a sequence of Markov kernels (P_t) together with the starting measure $\lambda \in \mathbb{M}_1(\mathcal{E})$ defines a time-inhomogeneous Markov chain (see [24], Theorem 5.1). Our main goal in this section is to introduce a notation that is not overwhelmed with indexes and allows us to use intuition from the homogeneous Markov chains theory and emphasize the difference brought by time-inhomogeneity.

Let $\Omega = E^\infty$ be the set of all infinite sequences $\omega = (\omega_0, \omega_1, \ldots),\ \omega_j \in E$ and $\mathcal{F} = \mathcal{E}^\infty$ be the sigma-field generated by all cylinder sets. For each fixed $t \in \mathbb{N}_0$ and $\lambda \in \mathbb{M}_1(\mathcal{E})$ there exists unique probability measure P^t_λ (see [24], Chap. 5 for details of the construction), such that for any cylinder set $A_0 \times A_1 \ldots A_{t+n} \times E^\infty \in \mathcal{F}$:

$$\mathsf{P}^t_\lambda \{A_0 \times \cdots \times A_{t+n}\} = \int_{A_t} \int_{A_{t+1}} \cdots \int_{A_{t+n}} \lambda(dx_0) P_t(x_0, dx_1) \ldots P_{t+n-1}(x_{n-1}, dx_n).$$

Define a sequence of random variables $X_{t,n}(\omega) = \omega_{t+n},\ \ n \geq 0$, such that for all $A_0, \ldots, A_n \in \mathcal{E}$: $\mathsf{P}^t_\lambda \left\{X_{t,0} \in A_0, \ldots, X_{t,n} \in A_n\right\} = \mathsf{P}^t_\lambda\{E^t \times A_0 \times \ldots \times A_n\}$. Denote $\mathcal{F}_{t,n} = \sigma\left[X_{t,k},\ 0 \leq k \leq n\right]$, a natural filtration associated with random sequence $(X_{t,n},\ \ n \geq 0)$, and use a notation $\mathbb{E}^t\left[f(X_{t,n+m})|\mathcal{F}_{t,n}\right]$ for the conditional expectation associated with $X_{t,n}, n \geq 0$ (here $f \in \mathbb{F}_+(\mathcal{E})$, $n, m \geq 0$). Then the Markov property holds true: $\mathbb{E}^t\left[f(X_{t,n+m})|\mathcal{F}_{t,n}\right] = \mathbb{E}^t\left[f(X_{t,n+m})|X_{t,n}\right]$.

So far, we defined a space $(\Omega, \mathcal{F})$ and a sequence P^t_λ of probability measures on that space, as well as double-indexed sequence $X_{t,n}\colon \Omega \to E$, $t, n \geq 0$. Naturally, we associate $X_{t,n}$ with a probability space $(\Omega, \mathcal{F}, \mathsf{P}^t_\lambda)$. Now, we establish how P^t_λ and $X_{t,n}$ are connected for different t.

Let us introduce the shift operators $\theta_n : \Omega \to \Omega,\ \ n \geq 1$, where $\forall \omega = (\omega_0, \omega_1, \ldots) \in \Omega$: $\theta(\omega) = \theta_1(\omega) = (\omega_1, \omega_2, \ldots)$, $\theta_n(\omega) = (\theta_{n-1} \circ \theta)(\omega) = (\omega_n, \omega_{n+1}, \ldots)$, $n > 1$. It is clear that, for any $t, s, n \in \mathbb{N}_0$: $X_{t+s,n} = X_{t,n+s} = X_{t,n} \circ \theta_s$, but, for $B \in \mathcal{E}$: $\mathsf{P}^{t+s}_\lambda \left\{X_{t+s,n} \in B\right\} \neq \mathsf{P}^t_\lambda \left\{X_{t,n+s} \in B\right\}$. For a set $C \in \mathcal{E}$

define hitting and return times by $\tau_{t,C} = \inf\{n \geq 0 : X_{t+n} \in C\}$, $\sigma_{t,C} = \inf\{n \geq 1 : X_{t+n} \in C\}$. In order to simplify further transformation and use intuition originated by the time-homogeneous case, we make the following agreement: *We always use random element $X_{t,n}$ in context of probability P^t_λ and never in context of P^s_λ, $s \neq t$. So, we will omit lower index t for $X_{t,n}$, $\mathcal{F}_{t,n}$, $\tau_{t,C}$ and $\sigma_{t,C}$ in context of P^t_λ.*

For example, we can write

$$\mathsf{P}^t_\lambda\{\sigma_C > n\} = \mathsf{P}^t_\lambda\left\{\sigma_{t,C} > n\right\} = \mathsf{P}^t_\lambda\{X_1 \notin C, \dots X_n \notin C\}$$
$$= \mathsf{P}^t_\lambda\left\{X_{t,1} \notin C, \dots X_{t,n} \notin C\right\} = \mathsf{P}^t_\lambda\{\omega \in \Omega : \omega_{t+1} \notin C, \dots, \omega_{t+n} \notin C\}.$$

Similarly, for $f \in \mathbb{F}_+(\mathcal{E})$, using Markov property, we get

$$\mathsf{E}^t\left[f(X_{n+m})|\mathcal{F}_n\right] = \mathsf{E}^t\left[f(X_{t,n+m})|\mathcal{F}_{t,n}\right] = \mathsf{E}^t\left[f(X_{n+m})|X_n\right] = P^{t+n,m}f(X_{t,n})$$
$$= \int_E\int_E \dots \int_E P_{t+n}(\omega_{t+n}, dx_1)P_{t+n+1}(x_1, dx_2)\dots P_{t+n+m-1}(x_{m-1}, dx_m)f(x_m).$$

Note, that in the formula above the index t must be specified in $P^{t+n,m}f(X_{t,n})$. On the other hand, it is obvious that for $x \in E$,

$$\mathsf{E}^t\left[f(X_{n+m})|X_n = x\right] = \mathsf{E}^{t+n}_x\left[f(X_m)\right],$$

which is a typical expression in the theory of homogeneous Markov chains.

We conclude this section with the definition of an atom.

Definition 1 We say that a set $\alpha \in \mathcal{E}$ is an atom for the sequence of Markov kernels $(P_t,\ t \in \mathbb{N}_0)$, if there exists a sequence of probability measures $\mu_t \in \mathbb{M}_1(\mathcal{E})$ such that for any $t \in \mathbb{N}_0$, $A \in \mathcal{E}$ and $x \in \alpha$: $P_t(x, A) = \mu_t(A)$.

We say that atom α is *aperiodic* if there exists $m \geq 1$ such that

$$\inf_t\{P^{t,m}(\alpha,\alpha),\ P^{t,m+1}(\alpha,\alpha)\} > 0. \tag{5.1}$$

Remark 1 For a homogeneous Markov chain with kernel P aperiodic atom satisfies $P^n(\alpha,\alpha) > 0$, for all $n \geq m$, where m some positive integer. Note that condition (5.1) implies that there exists $m \geq 1$, such that $P^{t,m+n}(\alpha,\alpha) > 0$, for all $n \geq 0$ In contrast to the homogeneous case in the inhomogeneous case it is possible that $P^{t,m+n}(\alpha,\alpha) \to 0$, $t \to \infty$. That is why we require in (5.1) that $\inf_t\{P^{t,m}(\alpha,\alpha),\ P^{t,m+1}(\alpha,\alpha)\} > 0$. Since the greatest common divisor of m and $m+1$ equals to 1, the latter condition can be rewritten in the following form, which we use in the proof of the main result:

$$\exists m \geq 1,\ \forall n \geq 0, \exists \gamma_n > 0 : \inf_{t,0\leq k\leq n}\{P^{t,m+k}(\alpha,\alpha) \geq \gamma_n > 0\}. \tag{5.2}$$

5.3 Geometric Drift Condition

5.3.1 Drift Condition for Inhomogeneous Markov Chains

In this section, we construct a time-inhomogeneous analog of the very well known result for homogeneous Markov chains, regarding geometric drift conditions and existence of the exponential moments. It worth mentioning that the standard homogeneous drift condition in the form $PV(x) \leq \lambda V(x) + b1_C(x)$ is not useful in the inhomogeneous case. This is related to the fact that the inhomogeneous chain properties do not necessarily coincide with that of each particular P_t. In other words, the whole chain may have the finite exponential moment, while the homogeneous chains generated by most of P_t will not. In time-inhomogeneous case, constant λ and test function V have to be dependent on t. One other peculiarity of inhomogeneous chains is that it could be convenient to analyze the chain in terms of blocks $P^{t,k}$ rather than in terms of individual transition probabilities P_t. That is why we state the drift condition in such "blocks" form.
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 exist a sequence of positive integers $\{n_k, k \geq 1\}$, a sequence of measurable functions $V_k\colon E \to [1, \infty]$ and two sequences of positive constants $\{\lambda_k,\ k \geq 0\}$, and $\{b_k,\ k \geq 0\}$ such that for all $x \in E$

$$P^{N_k,n_{k+1}} V_{k+1}(x) \leq \lambda_{k+1} V_k(x) + b_k 1_C(x), \tag{5.3}$$

where $N_k = \sum_{j=1}^{k} n_j,\ k \geq 1$.
2. 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 $a^+ = \max\{a, 0\}$.

We find it convenient to use the following notation

$$\begin{aligned} &k(t) = \min\{k :\ N_k \geq t\}, \quad N(t) = N_{k(t)}, \\ &\tau = \inf\{j \geq 1 :\ X_{t,N_{k(t)+j}-t} \in C\},\ \text{where } t \in \mathbb{N}_0. \end{aligned} \tag{5.4}$$

Variable τ here depends on selection of t, which should be clear from the context.

Theorem 1 *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 $P^{t,N(t)-t} V_{N(t)}(x) < \infty$:

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

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

$$\mathsf{E}_x^t\left[\prod_{j=1}^{\tau}\lambda_{N(t)+j}^{-1}\right] \le P^{t,N(t)-t}V_{N(t)}(x) + \lambda_{N_{k(t)+1}}^{-1}b_{N(t)}P^{t,N(t)-t}(x, C),$$

where $k(t)$, $N(t)$ and τ are defined in (5.4).

Proof In the proof we use the notation from Sect. 5.2. The key tool of the proof is the Comparison Theorem (see Appendix). Assume that $t \in \mathbb{N}_0$ is fixed.

First, for readability purposes, we define an increasing sequence of positive numbers $\{m_k,\ k \ge 0\}$, inhomogeneous Markov chain Z_k and filtration $\mathcal{F}_n^*$ (all depending on t) in the following way:

$$m_j = N_{k(t)+j},\quad j \ge 0,\ Z_j = X_{t,m_j-t},\ j \ge 0,\quad \mathcal{F}_n^* = \mathcal{F}_{t,m_n}.$$

So, $m_0 \ge t$ is the first number N_k that is greater or equal than t, m_1 is the second such number and so on. It is also clear that τ is a stopping time for filtration $\mathcal{F}_n^*$ and it can be written as $\tau = \inf\{j \ge 1 :\ Z_j \in C\}$.

First we show that if $x \in E$ is such that $P^{t,N(t)-t}V_{N(t)}(x) < \infty$ then

$$\mathsf{P}_x^t\{\tau < \infty\} = \mathsf{P}_x^t\{\sigma_C < \infty\} = 1. \tag{5.5}$$

Define $A_n = \left(\prod_{j=0}^{n}\lambda_{m_j} \vee 1\right)^{-1}$, $\mathcal{V}_n = A_n V_{m_n}(Z_n)$,
$\mathcal{Z}_n = A_{n+1}(1-\lambda_{m_{n+1}})^+ V_{m_n}(Z_n)$ and $\mathcal{Y}_n = b_{N(t)}A_{n+1}1_C(Z_n)$. Then, we can write

$$\begin{aligned}
&\mathsf{E}^t\left[\mathcal{V}_{n+1}|\mathcal{F}_n^*\right] + \mathcal{Z}_n = A_{n+1}\left(\mathsf{E}^t\left[V_{m_{n+1}}(Z_{n+1})|\mathcal{F}_{m_n}\right] + (1-\lambda_{m_{n+1}})^+V_{m_n}(Z_n)\right)\\
&= A_{n+1}\left(P^{m_n,m_{n+1}-m_n}V_{m_{n+1}}(Z_n) + (1-\lambda_{m_{n+1}})^+V_{m_n}(Z_n)\right)\\
&\le A_{n+1}\left(\lambda_{m_{n+1}}V_{m_n}(Z_n) + (1-\lambda_{m_{n+1}})^+V_{m_n}(Z_n) + b_{N(t)}1_C(Z_n)\right).
\end{aligned}$$

Consider now two cases. If $\lambda_{m_{n+1}} \le 1$, then $A_{n+1} = A_n$ and

$$\lambda_{m_{n+1}}V_{m_n}(Z_n) + (1-\lambda_{m_{n+1}})^+V_{m_n}(Z_n) = V_{m_n}(Z_n),$$

which yields for $\lambda_{m_{n+1}} \le 1$:

$$\mathsf{E}^t\left[\mathcal{V}_{n+1}|\mathcal{F}_n^*\right] + \mathcal{Z}_n \le A_nV_{m_n}(Z_n) + A_{n+1}b_{N(t)}1_C(Z_n) = \mathcal{V}_n + \mathcal{Y}_n. \tag{5.6}$$

Now, let $\lambda_{m_{n+1}} > 1$. In this case $A_{n+1}\lambda_{m_{n+1}} = A_n$ and $(1-\lambda_{m_{n+1}})^+ = 0$. Then, for $\lambda_{m_{n+1}} > 1$,

$$\mathsf{E}^t\left[\mathcal{V}_{n+1}|\mathcal{F}_n^*\right]+\mathcal{Z}_n \le A_n V_{m_n}(Z_n)+A_{n+1}b_{N(t)}1_C(Z_n)=\mathcal{V}_n+\mathcal{Y}_n. \quad (5.7)$$

Combining (5.6) and (5.7), we conclude that these inequalities holds for all $\lambda_{m_{n+1}} > 0$.

Comparison Theorem then yields:

$$\mathsf{E}_x^t\left[\mathcal{V}_\tau 1_{\tau<\infty}\right]+\mathsf{E}_x^t\left[\sum_{k=0}^{\tau-1}\mathcal{Z}_k\right] \le \mathsf{E}_x^t\left[\mathcal{V}_0\right]+\mathsf{E}_x^t\left[\sum_{k=0}^{\tau-1}\mathcal{Y}_k\right].$$

Using the last inequality, and the fact that $V_k \ge 1$, for all $k \ge 0$, we can establish the finiteness of the series

$$\sum_{n\ge0}A_n(1-\lambda_{m_n})^+\mathsf{P}_x^t\{\tau>n\}=\mathsf{E}_x^t\left[\sum_{n=0}^{\tau-1}A_n(1-\lambda_{m_n})^+\right]$$

$$\le \mathsf{E}_x^t\left[\sum_{n=0}^{\tau-1}A_n(1-\lambda_{m_n})^+V_{m_n}(Z_n)\right]=\mathsf{E}_x^t\left[\sum_{n=0}^{\tau-1}\mathcal{Z}_n\right]\le \mathsf{E}_x^t\left[\mathcal{V}_0\right]+\mathsf{E}_x^t\left[\sum_{k=0}^{\tau-1}\mathcal{Y}_k\right]$$

$$=A_0P^{t,N(t)-t}V_{N(t)}(x)+\mathsf{E}_x^t\left[A_1b_{N(t)}1_C(X_{N(t)-t})\right]<\infty.$$

Therefore, we get the relation $\sum\limits_{n\ge0}A_n(1-\lambda_{m_n})^+\mathsf{P}_x^t\{\tau>n\}<\infty$. It follows from the **Condition (D)** that $\sum\limits_{n\ge0}A_n(1-\lambda_{m_n})^+=\infty$ which implies $\mathsf{P}_x^t\{\tau>n\}\to0$, which proves (5.5), since $\mathsf{P}_x^t\{\sigma_C<\infty\}\ge\mathsf{P}_x^t\{\tau<\infty\}=1$. The rest of the proof of the theorem follows the arguments from the [7], Proposition 4.3.3 (ii). We apply the Comparison Theorem once again. Put

$$\Lambda_0=1,\ \Lambda_n=\prod_{k=1}^n\lambda_{m_k}^{-1},\ n\ge1, \mathcal{V}_n=\Lambda_nV_{m_n}(Z_n),\ n\ge0,$$
$$\mathcal{Z}_n=0,\quad \mathcal{Y}_n=\Lambda_{n+1}b_{N(t)}1_C(Z_n),\ n\ge0.$$

Then, for all $n\ge0$,

$$\mathsf{E}^t\left[\mathcal{V}_{n+1}|\mathcal{F}_n^*\right]+Z_n=\Lambda_{n+1}P^{m_n,m_{n+1}-m_n}V_{m_{n+1}}(Z_n)$$
$$\le\Lambda_{n+1}\lambda_{m_{n+1}}V_{m_n}(Z_n)+\Lambda_{n+1}b_{N(t)}1_C(Z_n)=\Lambda_nV_{m_n}(Z_n)+\mathcal{Y}_n=\mathcal{V}_n+\mathcal{Y}_n.$$

Assume that $x\in E$ satisfies inequality $P^{t,N(t)-t}V_{N(t)}(x)<\infty$. Taking into account (5.5), the Comparison Theorem yields:

$$\mathsf{E}_x^t\left[\Lambda_\tau\right] \le \mathsf{E}_x^t\left[\mathcal{V}_\tau\right] \le \mathsf{E}_x^t\left[\mathcal{V}_0\right] + \mathsf{E}_x^t\left[\sum_{k=0}^{\tau-1}\mathcal{Y}_k\right]$$
$$= P^{t,N(t)-t}V_{N(t)}(x) + \lambda_{N_{k(t)+1}}^{-1} b_{N(t)} P^{t,N(t)-t}(x,C),$$

which completes the proof.

Corollary 1 *It follows from Theorem 1 that sufficient conditions for the existence of the moment of σ_C for given $t \in \mathbb{N}_0$ and $x \in E$ are the following:*

1. ***Condition (D)** holds true.*
2. *There exist $\beta > 1$ and $C_\beta > 0$ such that $\forall n, k \ge 0$: $\beta^k \le C_\beta \prod_{j=1}^{k} \lambda_{n+j}^{-1}$.*
3. *$x \in E$ is such that $P^{t,N(t)-t}V_{N(t)}(x) < \infty$.*
   Then, the following inequality is valid,

$$C_\beta^{-1}\mathsf{E}_x^t\left[\beta^{\sigma_C}\right] \le P^{t,N(t)-t}V_{N(t)}(x) + \lambda_{N_{k(t)+1}}^{-1} b_{N(t)} P^{t,N(t)-t}(x,C).$$

Proof Since for all $\omega \in \Omega$, return time $\sigma_C(\omega)$ satisfies the inequality $\sigma_C(\omega) \le \tau(\omega)$, we can conclude that $\mathsf{E}_x^t\left[\beta^{\sigma_C}\right] \le \mathsf{E}_x^t\left[\beta^{\tau}\right] \le C_\beta \mathsf{E}_x^t\left[\prod_{j=1}^{\tau}\lambda_{N(t)+j}^{-1}\right]$. Required statement then follows from Theorem 1.

Remark 2 In the case where $n_k = 1$, for every $k \ge 0$, the drift condition and exponential moment bound could be rewritten in the simpler form:

$$P_t V_{t+1}(x) \le \lambda_{t+1} V_t(x) + b_t 1_C(x), \quad C_\beta^{-1}\mathsf{E}_x^t\left[\beta^{\sigma_C}\right] \le V_t(x) + \lambda_{t+1}^{-1} b_t 1_C(x),$$

assuming that conditions of Corollary 1 are satisfied. In fact, **Condition (D)** implies one-step drift condition with the special functions V_t under some additional assumptions. The following proposition is a straightforward analog of a well-known homogeneous result (see [7], Proposition 4.3.3 (i)).

Proposition 1 *Let $\{\lambda_t,\ t \in \mathbb{N}_0\}$ be a set of positive constants such that $b := \sup_{t\in\mathbb{N}_0, x\in C} \mathsf{E}_x^t\left[\prod_{k=0}^{\sigma_C}\lambda_{t+k}^{-1}\right] < \infty$. Then drift condition $P_t V_{t+1}(x) \le \lambda_t V_t(x) + b 1_C(x)$ holds true for the function $V_t(x) = \mathsf{E}_x^t\left[\prod_{k=0}^{\tau_C}\lambda_{t+k}^{-1}\right]$.*

Proof Markov property yields:

$$
\begin{aligned}
P_t V_{t+1}(x) &= \mathsf{E}_x^t\left[V_{t+1}(X_1)\right] = \mathsf{E}_x^t\left[\mathsf{E}_{X_{t,1}}^{t+1}\left[\prod_{k=0}^{\tau_C}\lambda_{t+1+k}^{-1}\right]\right] \\
&= \mathsf{E}_x^t\left[\prod_{k=0}^{\tau_{t+1,C}}\lambda_{t+1+k}^{-1}\right] = \mathsf{E}_x^t\left[\prod_{k=0}^{\tau_C\circ\theta}\lambda_{t+1+k}^{-1}\right] = \sum_{j\geq 1}\mathsf{E}_x^t\left[\prod_{k=0}^{\tau_C\circ\theta}\lambda_{t+1+k}^{-1}1_{\sigma_C=j}\right] \\
&= \sum_{j\geq 1}\mathsf{E}_x^t\left[\prod_{k=0}^{j-1}\lambda_{t+1+k}^{-1}1_{\sigma_C=j}\right] = \sum_{j\geq 1}\mathsf{E}_x^t\left[\prod_{k=1}^{j}\lambda_{t+k}^{-1}1_{\sigma_C=j}\right] \\
&= \sum_{j\geq 1}\lambda_t\mathsf{E}_x^t\left[\prod_{k=0}^{j}\lambda_{t+k}^{-1}1_{\sigma_C=j}\right] = \lambda_t\mathsf{E}_x^t\left[\prod_{k=0}^{\sigma_C}\lambda_{t+k}^{-1}\right].
\end{aligned}
$$

For $x \notin C$ we have $\mathsf{P}_x^t\{\sigma_C = \tau_C\} = 1$, which means that $P_t V_{t+1}(x) = \lambda_t V_t(x)$. Additionally, for any $x \in C$:

$$
P_t V_{t+1}(x) = \lambda_t\mathsf{E}_x^t\left[\prod_{k=0}^{\sigma_C}\lambda_{t+k}^{-1}\right] \leq \lambda_t \sup_{x\in C}\mathsf{E}_x^t\left[\prod_{k=0}^{\sigma_C}\lambda_{t+k}^{-1}\right] \leq \lambda_t\lambda_t^{-1}b.
$$

Combining the inequalities for $x \in C$ and $x \notin C$ we get:

$$
P_t V_{t+1}(x) \leq \lambda_t V_t(x)1_{C^c}(x) + b1_C(x) \leq \lambda_t V_t(x) + b1_C(x).
$$

So, the statement Remark 2 is proved.

5.3.2 *Constructing a Sequence that Dominates Return Time*

Existence of a dominating sequence, i.e., such sequence of positive, real numbers $\{\hat{G}_n,\ n \geq 0\}$ that $\hat{G}_n(x) \geq \mathsf{P}_x^t\{\sigma_C > n\}$, plays an important role in a series of results for inhomogeneous Markov chains (see, [9–13]).

Practically, however, it is not always easy to find such a sequence. We will show that the drift condition could be used to address this problem.

Lemma 1 *Consider inhomogeneous Markov chain defined by a series of Markov kernels* $(P_t,\ t \in \mathbb{N}_0)$. *Assume that conditions of Corollary 1 are satisfied. Then*

$$
\mathsf{P}_x^t\{\sigma_C > n\} \leq C_\beta \frac{P^{t,N(t)-t}V_{N(t)}(x) + \lambda_{N_{k(t)+1}}^{-1} b_{N(t)} P^{t,N(t)-t}(x, C)}{e^{(n+1)\ln\beta}}. \tag{5.8}
$$

In particular, when conditions of Remark 2 are satisfied, (5.8) is equal to

$$
\mathsf{P}_x^t\{\sigma_\alpha > n\} \leq C_\beta \frac{V_t(x) + \frac{b_t}{\lambda_{t+1}}}{e^{(n+1)\ln\beta}}. \tag{5.9}
$$

Proof is a trivial application of the Chernoff inequality

$$\mathsf{P}_x^t\left\{\sigma_C > n\right\} = \mathsf{P}_x^t\left\{\sigma_C \geq n+1\right\} \leq \frac{\mathsf{E}_x^t\left[e^{\sigma_C \ln \beta}\right]}{e^{(n+1)\ln \beta}}.$$

Formulas (5.8) and (5.9) follows from Corollary 1 and Remark 2.

Using Lemma 1, we may construct dominating sequences in the assumption that right-hand sides of (5.8) or (5.9) are bounded as functions of t. The nice property of such dominating sequences is that they admit finite exponential moments.

5.4 Main Result

In this section we consider a pair of sequences of Markov kernels $(P_{0,t},\ t \in \mathbb{N}_0)$ and $(P_{1,t},\ t \in \mathbb{N}_0)$ defined on the $E \times \mathcal{E}$. Let $\mathcal{E} \otimes \mathcal{E}$ be a sigma-field generated by all products $A \times B,\ A, B \in \mathcal{E}$ and $\forall \lambda, \lambda' \in \mathbb{M}_1(\mathcal{E})$ denote as $\lambda \otimes \lambda'$ a product measure defined on the $\mathcal{E} \otimes \mathcal{E}$. We may construct the sequence of Markov kernels $\bar{P}_t \colon E^2 \times \mathcal{E} \otimes \mathcal{E} \to [0, 1]$, such that for all $t \in \mathbb{N}_0,\ x, y \in E,\ A \in \mathcal{E} \otimes \mathcal{E}$:

$$\bar{P}_t((x, y), A) = \int_{(z_0, z_1)\in A} P_{0,t}(x, dz_0) P_{1,t}(x, dz_1).$$

We can build the canonical space $(\bar{\Omega}, \bar{\mathcal{F}})$ and a series of probability measures $\overline{\mathsf{P}}^t_{\lambda_0 \otimes \lambda_1}$ $(\lambda_0, \lambda_1 \in \mathbb{M}_1(\mathcal{E}))$ using the same approach as in Sect. 5.2. It is clear that every $\bar{\omega} \in \bar{\Omega}$ can be written as $\bar{\omega} = (\bar{\omega}_0, \bar{\omega}_1, \ldots)$, where $\bar{\omega}_j = \left(\omega_j^{(0)}, \omega_j^{(1)}\right),\ \omega_j^{(i)} \in E, i \in \{0, 1\}$, $j \geq 0$. For each $t \in \mathbb{N}_0$ we have then a pair of time-inhomogeneous Markov chains $(X_{t,n}^{(0)}, X_{t,n}^{(1)},\ n \geq 0)$, such that $X_{t,n}^{(0)}(\bar{\omega}) = \omega_{t+n}^{(0)}$ and similarly $X_{t,n}^{(1)}(\bar{\omega}) = \omega_{t+n}^{(1)}$.

It follows from the construction, that $\forall A \in \mathcal{E}, i \in \{0, 1\}$:

$$\overline{\mathsf{P}}^t_{\lambda_0 \otimes \lambda_1}\left\{X_n^{(i)} \in A\right\} = \int_E \lambda_i(dx) P_i^{t,n}(x, A),$$

where, for $i \in \{0, 1\}$: $P_i^{t,n}(x, A) = \left(\prod_{k=0}^{n-1} P_{i,t+k}\right)(x, A)$. For a given set $C \in \mathcal{E}$ we define hitting and return times to $C \times C$:

$$\bar{\tau}_{t,C\times C} = \inf\{n \geq 0 : \left(X_{t,n}^{(0)}, X_{t,n}^{(1)}\right) \in C \times C\},$$

$$\bar{\sigma}_{t,C\times C} = \inf\{n \geq 1 : \left(X_{t,n}^{(0)}, X_{t,n}^{(1)}\right) \in C \times C\},$$

an the shift operator on $\bar{\Omega}$: $\bar{\theta}((\bar{\omega}_0, \bar{\omega}_1, \ldots)) = (\bar{\omega}_1, \bar{\omega}_2, \ldots)$.

We will also need indviduall probabilities and expectations $\mathsf{P}^t_{i,\lambda}$, $\mathsf{E}^t_{i,\lambda}$, $i \in \{0, 1\}$, $\lambda \in \mathbb{M}_1(\mathcal{E})$. They should be understood as canonical probabilities and expectations generated by the sequences $(P_{0,t},\ t \in \mathbb{N}_0)$, or $(P_{1,t},\ t \in \mathbb{N}_0)$ separately.

Going forward we will drop the bottom index t in the context of $\overline{\mathsf{P}}^t_{\lambda_0\otimes\lambda_1}$ as described in the Sect. 5.2.

We will need the following conditions:

Condition A: There exists set $\alpha \in \mathcal{E}$ which is an aperiodic atom for both $(P_{0,t})$, and $(P_{1,t})$.

Condition D1: Assume that Condition (D) holds true for each of the sequences $(P_{i,t},\ t \in \mathbb{N}_0, i \in \{0, 1\})$ with $V_t^{(i)}$, $\lambda_t^{(i)}$ and $\beta_t^{(i)}$. Assume also, there exists $\beta > 1$ and constants $C_\beta^{(i)} > 0$ such that for $i \in \{0, 1\}$, $t, n \geq 0$, and $\lambda_t^{(i)}$

$$\beta^n \leq C_\beta^{(i)} \left(\prod_{k=1}^{n} \lambda_{t+k}^{(i)}\right)^{-1}. \tag{5.10}$$

Now we introduce a notation specific for the proof of the main result.

Let **Condition (A)** hold true, so that α is an aperiodic atom for both chains. Then, we can assume, without loss of generality, that there are $m > 0$ and $\gamma_0 > 0$, such that

$$\gamma_0 = \inf_{t\in\mathbb{N}_0, i\in\{0,1\}} \{P_i^{t,m}(\alpha,\alpha),\ P_i^{t,m+1}(\alpha,\alpha), \ldots,\ P_i^{t,2m-1}(\alpha,\alpha)\} > 0. \tag{5.11}$$

Let us define a sequence of “coupling trials” $\nu_{t,k}$:

$$\begin{aligned}
&\nu_{t,-1} = \min\{\bar{\sigma}_{t,\alpha\times E}, \bar{\sigma}_{t,E\times\alpha}\}, \quad \nu_{t,0} = \max\{\bar{\sigma}_{t,\alpha\times E}, \bar{\sigma}_{t,E\times\alpha}\},\\
&\nu_{t,n+1} = \begin{cases} \infty, \text{ if } \nu_n = \infty, \\ \min\{k \geq \nu_{t,n} + m,\ X_{t,k}^{(1)} \in \alpha\}, \text{ if } X_{t,\nu_{t,n}}^{(0)} \in \alpha, \\ \min\{k \geq \nu_{t,n} + m,\ X_{t,k}^{(0)} \in \alpha\}, \text{ if } X_{t,\nu_{t,n}}^{(1)} \in \alpha, \end{cases}
\end{aligned} \tag{5.12}$$

where $n \geq 0$ and m is from (5.11). We also introduce the following notation

$$U_{t,n} = \nu_{t,n} - \nu_{t,n-1},\ n \geq 0, \quad \tau_t = \min\{k \geq 0 : \nu_{t,k-1} = \nu_{t,k}\}. \tag{5.13}$$

$U_{t,n}$ can be understood as a next after time m hit of α by $X^{(1-i)}$ if $X^{(i)}_{t,\nu_{t,n}} \in \alpha$, τ_t is a number of the first successfull coupling trial, and ν_{t,τ_t} is an index, such that $\left(X^{(0)}_{t,\nu_{t,\tau_t}}, X^{(1)}_{t,\nu_{t,\tau_t}}\right) \in \alpha\times\alpha$ for the first time. The main reason, why we added m steps of delay, is to ensure that renewal probabilities are separated out from 0, which is the critical element for the proof. Let us also define a family of sigma-fields:

$$\mathcal{B}_{t,n} = \sigma\left[\bar{\mathcal{F}}_{\nu_{t,n-1}}, U_{t,n}\right],\ n \geq 0. \tag{5.14}$$

Theorem 2 *Let $(P_{0,t},\ t \in \mathbb{N}_0)$ and $(P_{1,t},\ t \geq 1)$ be two sequences of Markov kernels. Assume that* ***Condition (A)*** *is satisfied and there exist constant $\beta > 1$ and sets*

$\tilde{E}_0, \tilde{E}_1 \in \mathcal{E}$, *such that* $\alpha \subset \tilde{E}_0 \cap \tilde{E}_1$ *and for all* $x, y \in \tilde{E}_0 \times \tilde{E}_1$:

$$\sup_t \left(\mathsf{E}^t_{0,x}\left[\beta^{\sigma_\alpha}\right] + \mathsf{E}^t_{1,y}\left[\beta^{\sigma_\alpha}\right]\right) < \infty.$$

Then there exists constant $M > 0$ *such that the following inequality holds true* :

$$\overline{\mathsf{E}}^t_{x,y}\left[\delta^{\bar{\sigma}_{\alpha\times\alpha}}\right] \le M\left(\mathsf{E}^t_{0,x}\left[\beta^{\sigma_\alpha}\right] + \mathsf{E}^t_{1,y}\left[\beta^{\sigma_\alpha}\right]\right). \tag{5.15}$$

Constant M *could be expressed as*

$$M = 1 + \frac{1}{1 - \sqrt{(1-\gamma)(1+\varepsilon)}}, \tag{5.16}$$

where $\gamma,\ \varepsilon > 0$ *some constants, such that* $(1-\gamma)(1+\varepsilon) < 1$.

Proof Since for every $\bar{\omega} \in \bar{\Omega}$ we have the inequality $\bar{\sigma}_{t,\alpha\times\alpha}(\bar{\omega}) \le \nu_{t,\tau_t}(\bar{\omega})$ then for all $x, y \in \tilde{E}_0 \times \tilde{E}_1$ we get:

$$\begin{aligned}
\overline{\mathsf{E}}^t_{x,y}\left[\beta^{\bar{\sigma}_{\alpha\times\alpha}}\right] &\le \overline{\mathsf{E}}^t_{x,y}\left[\beta^{\nu_\tau}\right] = \sum_{k=0}^{\infty} \overline{\mathsf{E}}^t_{x,y}\left[1_{\tau=k}\beta^{\nu_k}\right] \\
&\le \overline{\mathsf{E}}^t_{x,y}\left[\beta^{\nu_0}\right] + \sum_{k=1}^{\infty} \overline{\mathsf{E}}^t_{x,y}\left[1_{\tau>k-1}\beta^{\nu_k}\right] \\
&\le \overline{\mathsf{E}}^t_{x,y}\left[\beta^{\nu_0}\right] + \sum_{k=0}^{\infty}\left(\overline{\mathsf{P}}^t_{x,y}\left\{\tau > k\right\}\overline{\mathsf{E}}^t_{x,y}\left[\beta^{2\nu_{k+1}}\right]\right)^{\frac{1}{2}}.
\end{aligned} \tag{5.17}$$

Last inequality is due to the Cauchy-Schwarz inequality. By Lemma 6 with $r(k) = 1$ there exists $\gamma \in (0, 1)$ such that $\bar{\mathbb{P}}^t\left\{\tau > j | \bar{\mathcal{F}}_{\nu_{j-1}}\right\} \le (1-\gamma)1_{\tau_t > j-1}$, which entails:

$$\overline{\mathsf{P}}^t_{x,y}\left\{\tau > k\right\} < (1-\gamma)^k. \tag{5.18}$$

Note that $\nu_{t,k+1} = \nu_{t,k} + U_{t,k+1}$ and $\nu_{t,k}$ is $\mathcal{B}_{t,k}$-measurable. Let us select ε, such that $(1+\varepsilon)(1-\gamma) < 1$. Since $\sup_{t,i} \mathsf{E}^t_{i,\alpha}\left[\beta^{\sigma_\alpha}\right] < \infty$, we can apply Lemma 4 and find such $\delta \in (1, \beta)$ that

$$\overline{\mathsf{E}}^t_{x,y}\left[\delta^{2\nu_{k+1}}\right] = \overline{\mathsf{E}}^t_{x,y}\left[\delta^{2\nu_k}\overline{\mathsf{E}}^t\left[\delta^{2U_{k+1}}|\mathcal{B}_k\right]\right] \le (1+\varepsilon)\overline{\mathsf{E}}^t_{x,y}\left[\delta^{2\nu_k}\right]. \tag{5.19}$$

Applying (5.19) recursively we obtain the following estimate:

$$\overline{\mathsf{E}}^t_{x,y}\left[\delta^{2\nu_{k+1}}\right] \le (1+\varepsilon)^{k+1}\overline{\mathsf{E}}^t_{x,y}\left[\delta^{\nu_0}\right]. \tag{5.20}$$

Plugging (5.20) and (5.18) into (5.17), and taking into account that (5.17) remains true if we replace β with δ, we get:

$$\begin{aligned}\overline{\mathsf{E}}^t_{x,y}\left[\delta^{\bar{\sigma}_{\alpha\times\alpha}}\right] &\leq \overline{\mathsf{E}}^t_{x,y}\left[\delta^{\nu_0}\right]\left(1+\sum_{k=0}^{\infty}((1-\gamma)(1+\varepsilon))^{\frac{k}{2}}\right)\\ &\leq \overline{\mathsf{E}}^t_{x,y}\left[\delta^{\nu_0}\right]\left(1+\frac{1}{1-\sqrt{(1-\gamma)(1+\varepsilon)}}\right)\\ &\leq \left(\mathsf{E}^t_{0,x}\left[\delta^{\sigma_\alpha}\right]+\mathsf{E}^t_{1,y}\left[\delta^{\sigma_\alpha}\right]\right)\left(1+\frac{1}{1-\sqrt{(1-\gamma)(1+\varepsilon)}}\right).\end{aligned} \tag{5.21}$$

Since $\delta \leq \beta$, (5.21) renders:

$$\overline{\mathsf{E}}^t_{x,y}\left[\delta^{\bar{\sigma}_{\alpha\times\alpha}}\right] \leq \left(\mathsf{E}^t_{0,x}\left[\beta^{\sigma_\alpha}\right]+\mathsf{E}^t_{1,y}\left[\beta^{\sigma_\alpha}\right]\right)\left(1+\frac{1}{1-\sqrt{(1-\gamma)(1+\varepsilon)}}\right), \tag{5.22}$$

which proves the theorem with $M = 1+\frac{1}{1-\sqrt{(1-\gamma)(1+\varepsilon)}}$.

Theorem 2 establishes the existence of the exponential moment, however, it could be difficult to verify its conditions and find constants $\delta,\ \varepsilon, \gamma$, which are necessary to calculate M using (5.16). To address this problem, we state the next result.

Theorem 3 *Let* $(P_{i,t},\ \ i \in \{0,1\}, t \in \mathbb{N}_0)$ *be two sequences of Markov kernels. Assume that **Condition (A)** and **Condition (D1)** hold true. Assume additionally:*

1. *There exist constants* $\hat{C} > 0$, $\hat{\beta} > \beta$ *such that*

$$\mathsf{P}^t_{i,\alpha}\{\sigma_\alpha > n\} \leq \hat{C}\hat{\beta}^{-n},$$

so that $\hat{m} := \sum\limits_{n\geq 0}\hat{C}\hat{\beta}^{-n} = \frac{\hat{C}\hat{\beta}}{\hat{\beta}-1} < \infty.$

2. *There exist* $m > 0$ *and* $\gamma_0 > 0$ *such that for* $i \in \{0,1\}$,

$$\inf_{t\in\mathbb{N}_0}\{P_i^{t,m}(\alpha,\alpha),\ldots P_i^{t,2m-1}(\alpha,\alpha)\} \geq \gamma_0.$$

3. *There exist sets* $\mathcal{A}_i \in \mathcal{E}$, $\mathcal{A}_i \neq \emptyset$, $i \in \{0,1\}$ *such that for all* $x \in \mathcal{A}_i$,

$$\sup_t P_i^{t,N(t)-t}V^{(i)}_{N(t)}(x) < \infty.$$

Then the following inequality holds true for $x \in \mathcal{A}_0 \cup \alpha$, $y \in \mathcal{A}_1 \cup \alpha$,

$$\overline{\mathsf{E}}^t_{x,y}\left[\delta^{\bar{\sigma}_{\alpha\times\alpha}}\right] \leq M\left(C^{(0)}_\beta W^{(0)}_t(x)+C^{(1)}_\beta W^{(1)}_t(y)\right), \tag{5.23}$$

where

$$\begin{aligned} W_t^{(0)}(x) &= P_0^{t,N(t)-t} V_{N(t)}^{(0)}(x) + \frac{b_{N(t)}^{(0)}}{\lambda_{N(t)+1}^{(0)}} \mathsf{P}_{0,x}^t \left\{ X_{N(t)-t} \in \alpha \right\}, \\ W_t^{(1)}(y) &= P_1^{t,N(t)-t} V_{N(t)}^{(1)}(y) + \frac{b_{N(t)}^{(1)}}{\lambda_{N(t)+1}^{(1)}} \mathsf{P}_{1,x}^t \left\{ X_{N(t)-t} \in \alpha \right\}, \\ M = 1 + \frac{1}{1-\sqrt{(1-\gamma)(1+\varepsilon)}}, &\quad \gamma = \gamma_0 (1-\hat{G}_m)^{\frac{\hat{m}-\hat{G}_m}{\hat{G}_m}}, \quad \delta = (1+\varepsilon/2)^{\frac{1}{m+n_0}}, \\ n_0 &= \left\lfloor \ln\left(\frac{\varepsilon(\hat{\beta}-\beta)}{2\hat{C}\beta^{m+1}}\right) / \ln\left(\frac{\beta}{\hat{\beta}}\right) \right\rfloor + 3. \end{aligned} \tag{5.24}$$

Here ε is an arbitrary constant such that $\varepsilon < \frac{\gamma}{1-\gamma}$, and $\lfloor a \rfloor$ is an integer part of a real number a.

Proof Since **Condtion (D1)** is satisfied, we can apply Corollary 1 and get for every $x \in \mathcal{A}_i \setminus \alpha$, $i \in \{0,1\}$: $\sup_t \mathsf{E}_{i,x}^t [\beta^{\sigma_\alpha}] < \infty$. Condition 1 implies that $\sup_t \mathsf{E}_{i,\alpha}^t [\beta^{\sigma_\alpha}] < \infty$. So, conditions of Theorem 2 are satisfied with $\tilde{E}_i = \mathcal{A}_i \cup \alpha$.

Formulas for $W_t^{(0)}(x)$ and $W_t^{(1)}(y)$ follow from Theorem 1, the formula for the constant M is proven in Theorem 2, formulas for δ and n_0 are from Lemma 4. The formula for γ follows from Lemmas 2 and 3.

Remark 3 In the case when all n_k from **Condition (D)** are equal to 1, the formulas for $W_t^{(0)}(x)$ and $W_t^{(1)}(y)$ in (5.24) could be simplified to

$$W_t^{(0)}(x) = V_t^{(0)}(x) + \frac{b_t^{(0)}}{\lambda_{t+1}^{(0)}} 1_\alpha(x), \quad W_t^{(1)}(y) = V_t^{(1)}(y) + \frac{b_t^{(1)}}{\lambda_{t+1}^{(1)}} 1_\alpha(y). \tag{5.25}$$

Remark 4 Condition 1 in Theorem 3 seems more restrictive than **Condition (D1)**, but in fact, it can be derived from **Condition (D1)** as shown in Lemma 1. In this case we should find $\beta' \in (1, \beta)$ and set $\hat{\beta} = \beta$ and $\beta = \beta'$ which will satisfy conditions of Theorem 3. We stated condition 1 in Theorem 3 as a separate condition because $\hat{m}$ and $\hat{G}_m$ used in the obtained bounds. And of course, for some particular chains it is possible to find better $\hat{G}_n$ than provided by Lemma 1.

Next, we show how the bounds for an exponential moment could be applied to the ergodicity of inhomogeneous Markov chains. Conditions that guarantee strong and week ergodicity of inhomogeneous Markov chains are well known. Strong ergodicity was investigated in papers [5, 28] and criterion for weak ergodicity was established in [21, 26].

Condition (D) does not imply even weak ergodicity as defined in [21, 26] (unless $\sup_{x,t} V_t(x) < \infty$, which does not hold in practice), but this condition is sufficient for convergence in measure's norm for n-steps transition probabilities. Rates of such convergence have been studied in papers [1, 8]. In [1], geometric drift condition was used to establish convergence rates. However, **Condition (D)** in the present paper is less restrictive than one in [1] since we allow some λ_t to be greater than 1.

The main difference with the result in [1] is that we established bounds for geometric sums $\sum_{k=0}^{\infty} \delta^k ||P^{t,k}(x,\cdot) - P^{t,k}(y,\cdot)||$ while [1] is concerned with bounds for a single term $||P^{t,k}(x,\cdot) - P^{t,k}(y,\cdot)||$. In order to prove an estimate for the geometric sum defined above, in the next theorem, we apply the coupling method to two copies of the same inhomogeneous Markov chain started with different initial distributions. This allows us to show that the sum is bounded by the exponential moment of the simultaneous hitting time, and we can use Theorems 2 or 3 to obtain computable bounds in terms of exponential moments of each chain or test function from **Condition (D)**.

The next theorem is a well-known fact for homogeneous Markov chains, and the proof follows the same arguments as used for homogeneous chains (see. [7] Chaps. 8, 13). We state the theorem here to demonstrate one possible application of Theorems 2 and 3 and highlight the importance of the existence of exponential moment.

Theorem 4 *Let* $(P_t,\ t \in \mathbb{N}_0)$ *be a sequence of Markov kernels that admits an aperiodic atom* $\alpha \in \mathcal{E}$*, and* $\lambda, \lambda' \in \mathbb{M}_1(\mathcal{E})$ *two probability measures, such that for all* $t \in \mathbb{N}_0$, $\mathbb{P}^t_\lambda\{\sigma_\alpha < \infty\} = \mathbb{P}^t_{\lambda'}\{\sigma_\alpha < \infty\} = 1$. *Assume that there exists* $\beta > 1$ *such that* $\mathsf{E}^t_\alpha\left[\beta^{\sigma_\alpha}\right] < \infty$. *Then there exists* $\delta \in (1,\beta)$ *satisfying the following inequality*

$$\sum_{k \ge 0} \delta^k ||\lambda P^{t,k} - \lambda' P^{t,k}|| \le \frac{1}{\delta - 1}\left(\overline{\mathsf{E}}^t_{\lambda\otimes\lambda'}\left[\delta^{\bar{\sigma}_{\alpha\times\alpha}}\right] - 1\right).$$

Proof We conduct the proof using the standard coupling technique adapted for time-inhomogeneous chains.

Let $f : E \to \mathbb{R}$ be a bounded measurable function. Consider chains $X^{(0)}_{t,n}$ and $X^{(1)}_{t,n}$ as two copies of the same time-inhomogeneous chain with a sequence of Markov kernels $(P_t,\ t \in \mathbb{N}_0)$. Then

$$\begin{aligned}
\mathsf{E}^t_{0,\lambda}\left[f\left(X^{(0)}_n\right)\right] &= \overline{\mathsf{E}}^t_{\lambda\otimes\lambda'}\left[f\left(X^{(0)}_n\right)\right] \\
&= \sum_{k=0}^{n} \overline{\mathsf{E}}^t_{\lambda\otimes\lambda'}\left[f\left(X^{(0)}_n\right) 1_{\bar{\sigma}_{\alpha\times\alpha}=k}\right] + \overline{\mathsf{E}}^t_{\lambda\otimes\lambda'}\left[f\left(X^{(0)}_n\right) 1_{\bar{\sigma}_{\alpha\times\alpha}>n}\right] \\
&= \sum_{k=0}^{n} \overline{\mathsf{E}}^t_{\lambda\otimes\lambda'}\left[\overline{\mathsf{E}}^{t+k}_{\alpha\times\alpha}\left[f\left(X^{(0)}_{n-k}\right)\right] 1_{\bar{\sigma}_{\alpha\times\alpha}=k}\right] + \overline{\mathsf{E}}^t_{\lambda\otimes\lambda'}\left[f\left(X^{(0)}_n\right) 1_{\bar{\sigma}_{\alpha\times\alpha}>n}\right] \\
&= \sum_{k=0}^{n} \overline{\mathsf{P}}^t_{\lambda\otimes\lambda'}\left\{\bar{\sigma}_{\alpha\times\alpha} = k\right\} P^{t+k,n-k} f(\alpha) + \overline{\mathsf{E}}^t_{\lambda\otimes\lambda'}\left[f\left(X^{(0)}_n\right) 1_{\bar{\sigma}_{\alpha\times\alpha}>n}\right].
\end{aligned}$$

Similar inequality holds true for $\mathsf{E}^t_{1,\lambda'}\left[f\left(X^{(1)}_n\right)\right]$. By the bounds obtained above,

$$\begin{aligned}
\left|\mathsf{E}^t_{0,\lambda}\left[f\left(X^{(0)}_n\right)\right] - \mathsf{E}^t_{1,\lambda'}\left[f\left(X^{(1)}_n\right)\right]\right| &\le \overline{\mathsf{E}}^t_{\lambda\otimes\lambda'}\left[\left|f\left(X^{(1)}_n\right) - f\left(X^{(2)}_n\right)\right| 1_{\bar{\sigma}_{\alpha\times\alpha}>n}\right] \\
&\le \sup_{x,y\in E} |f(x) - f(y)| \overline{\mathsf{P}}^t_{\lambda\otimes\lambda'}\left\{\bar{\sigma}_{\alpha\times\alpha} > n\right\}.
\end{aligned}$$

Then, $||\lambda P^{t,n} - \lambda' P^{t,n}|| = \sup_{A \in \mathcal{E}} \left|\mathsf{E}^t_{0,\lambda}\left[1_A\left(X_n^{(1)}\right)\right] - \mathsf{E}^t_{1,\lambda'}\left[1_A\left(X_n^{(2)}\right)\right]\right|$ $\leq \overline{\mathsf{P}}^t_{\lambda\otimes\lambda'}\left\{\bar{\sigma}_{\alpha\times\alpha} > n\right\}$. Finally, we arrive to:

$$\sum_{n\geq 0} \delta^n ||\lambda P^{t,n} - \lambda' P^{t,n}|| \leq \sum_{n\geq 0} \delta^n \overline{\mathsf{P}}^t_{\lambda\otimes\lambda'}\left\{\bar{\sigma}_{\alpha\times\alpha} > n\right\} = \frac{1}{\delta - 1}\left(\overline{\mathsf{E}}^t_{\lambda\otimes\lambda'}\left[\delta^{\bar{\sigma}_{\alpha\times\alpha}}\right] - 1\right).$$

The theorem is proved.

5.5 Auxiliary Lemmas

Lemma 2 *For a sequence of Markov kernels* $(P_t,\ t \in \mathbb{N}_0)$, *let* α *be an aperiodic atom, and* $\gamma_0 = \inf_t\{P^{t,m}(\alpha,\alpha), P^{t,m+1}(\alpha,\alpha), \ldots, P^{t,2m-1}(\alpha,\alpha)\} > 0$. *Then, for all* $t, n \geq 0$,

$$\begin{aligned} P^{t,2m+n}(\alpha,\alpha) &\geq \gamma_0 \prod_{k=0}^{n} \mathbb{P}^{t+k}_{\alpha}\{\sigma_\alpha \leq m + n - k\} \\ &= \gamma_0 \prod_{k=0}^{n} \mathsf{P}^{t+n-k}_{\alpha}\left\{\sigma_\alpha \leq m + k\right\} > 0. \end{aligned} \tag{5.26}$$

Proof We prove the lemma by induction. Let us start with $n = 0$ in (5.26),

$$P^{t,2m}(\alpha,\alpha) = \sum_{k=1}^{2m} \mathsf{P}^t_\alpha\left\{\sigma_\alpha = k\right\} P^{t+k,2m-k}(\alpha,\alpha)$$

$$\geq \sum_{k=1}^{m} \mathsf{P}^t_\alpha\left\{\sigma_\alpha = k\right\} P^{t+k,2m-k}(\alpha,\alpha) \geq \gamma_0 \sum_{k=1}^{m} \mathsf{P}^t_\alpha\left\{\sigma_\alpha = k\right\} = \gamma_0 \mathsf{P}^t_\alpha\left\{\sigma_\alpha \leq m\right\}.$$

Assume that inequality (5.26) is true for all $t \in \mathbb{N}_0$, $k \leq n$, let's check if for $n + 1$. Using the first entrance decomposition (see [23], Chap. 8, p. 174). we can write

$$\begin{aligned} P^{t,2m+n+1}(\alpha,\alpha) &= \sum_{k=1}^{2m+n+1} \mathbb{P}^t_\alpha\{\sigma_\alpha = k\} P^{t+k,2m+n+1-k}(\alpha,\alpha) \\ &\geq \sum_{k=1}^{m+n+1} \mathbb{P}^t_\alpha\{\sigma_\alpha = k\} P^{t+k,2m+n+1-k}(\alpha,\alpha) \end{aligned}$$

$$= \sum_{k=1}^{n+1} \mathbb{P}^t_\alpha\{\sigma_\alpha = k\} P^{t+k,2m+n+1-k}(\alpha,\alpha) + \sum_{k=n+2}^{m+n+1} \mathbb{P}^t_\alpha\{\sigma_\alpha = k\} P^{t+k,2m+n+1-k}(\alpha,\alpha)$$

$$\geq \gamma_0 \sum_{k=1}^{n+1} \mathbb{P}^t_\alpha\{\sigma_\alpha = k\} \prod_{j=0}^{n+1-k} \mathbb{P}^{t+k+n+1-k-j}_\alpha\{\sigma_\alpha \leq m+j\} + \gamma_0 \mathbb{P}^t_\alpha\{n+2 \leq \sigma_\alpha \leq m+n+1\}$$

$$= \gamma_0 \sum_{k=1}^{n+1} \mathbb{P}^t_\alpha\{\sigma_\alpha = k\} \prod_{j=0}^{n+1-k} \mathbb{P}^{t+n+1-j}_\alpha\{\sigma_\alpha \leq m+j\} + \gamma_0 \mathbb{P}^t_\alpha\{n+2 \leq \sigma_\alpha \leq m+n+1\}$$

$$\geq \gamma_0 \sum_{k=1}^{n+1} \mathbb{P}^t_\alpha\{\sigma_\alpha = k\} \prod_{j=0}^{n} \mathbb{P}^{t+n+1-j}_\alpha\{\sigma_\alpha \leq m+j\} + \gamma_0 \mathbb{P}^t_\alpha\{n+2 \leq \sigma_\alpha \leq m+n+1\}$$

$$\geq \left(\gamma_0 \prod_{j=0}^{n} \mathbb{P}^{t+n+1-j}_\alpha\{\sigma_\alpha \leq m+j\}\right) \mathbb{P}^t_\alpha\{\sigma_\alpha \leq m+n+1\}$$

$$= \gamma_0 \prod_{j=0}^{n+1} \mathbb{P}^{t+n-j}_\alpha\{\sigma_\alpha \leq m+j+1\}.$$

Since for every $t \in \mathbb{N}_0$, $\mathbb{P}^t_\alpha\{\sigma_\alpha \leq m\} \geq P^{t,m}(\alpha,\alpha) > 0$, and sequence $\mathbb{P}^t_\alpha\{\sigma_\alpha \leq n\}$ is increasing in n, it implies that each term in the product (5.26) is positive. Therefore $P^{t,2m+n}(\alpha,\alpha) > 0$, for all $n \geq 0$.

Lemma 3 *Let conditions of Lemma 2 hold true, and assume there exists a sequence of decreasing, non-negative numbers $\{\hat{G}_n,\ n \geq 0\}$, such that $\hat{G}_n \geq \sup_t \mathbb{P}^t_\alpha\{\sigma_\alpha > n\}$, and $\sum_{n\geq 0} \hat{G}_n = M < \infty$. Assume also that $\hat{G}_m < 1$. Then for $\gamma = \gamma_0(1-\hat{G}_m)^{\frac{M-\hat{G}_m}{\hat{G}_m}} > 0$ and for all $t, n \geq 0$ we have the lower bound*

$$P^{t,2m+n}(\alpha,\alpha) \geq \gamma > 0. \tag{5.27}$$

Proof Lemma 2 yields:

$$P^{t,2m+n} \geq \gamma_0 \prod_{k=0}^{n} (1-\hat{G}_{m+k}). \tag{5.28}$$

The fact that (5.28) entails (5.27) is proved in Theorem 4.1 from [10].

Note that condition $\hat{G}_m < 1$ is not restrictive. Since $\sum_{n=0}^{\infty} G_n < \infty$ and $\{G_n,\ n \geq 0\}$ is nonincreasing, it is necessary that there exists n_0 such that $\hat{G}_k < 1$ for all $k > n_0$. In case $m \leq n_0$, Lemma 2 shows it is always possible to choose another, bigger m at the cost of smaller γ_0.

The next three lemmas are adjusted versions of the known results from homogeneous theory (see [7], Chap. 13 for more details). The main difference is that we study here two different inhomogeneous chains rather than two copies of the same homogeneous chain. In the next lemma, for example, we have conditions and estimates that are different from the homogeneous analog.

Lemma 4 *The following statements hold.*

1. *Let $(P_t,\ t \in \mathbb{N}_0)$ be a sequence of Markov kernels that admits an aperiodic atom α and there exist constant $\beta > 1$ such that*

$$\sup_t \mathbb{E}_\alpha^t[\beta^{\sigma_\alpha}] < \infty, \tag{5.29}$$

Then for every $m \geq 0$ and every $\varepsilon > 0$ there exists $\delta = \delta(m, \varepsilon) \in (1, \beta)$ such that

$$\sup_{t,n} \mathbb{E}_\alpha^t[\delta^{m+\tau_\alpha \circ \theta_n}] \leq 1 + \varepsilon. \tag{5.30}$$

2. *If additionally, there exists a dominating sequence $\hat{G}_n$ and constants $\hat{C} > 0$, $\hat{\beta} > \beta$, such that for all $t, k \geq 0$*

$$\mathsf{P}_\alpha^t\{\sigma_\alpha > k\} \leq \hat{G}_k \leq \hat{C}\hat{\beta}^{-k}, \tag{5.31}$$

then

$$\delta = (1 + \varepsilon/2)^{\frac{1}{m+n_0}}, \quad n_0 = \left\lfloor \ln\left(\frac{\varepsilon(\hat{\beta} - \beta)}{2\hat{C}\beta^{m+1}}\right) / \ln\left(\frac{\beta}{\hat{\beta}}\right) \right\rfloor + 3, \tag{5.32}$$

where $\lfloor a \rfloor$ is an integer part of a real number a.

Proof First we wish to establish the following inequality

$$\mathbb{P}_\alpha^t\{\tau_\alpha \circ \theta_n = k\} \leq \sum_{j=1}^n \mathbb{P}_\alpha^{t+n-1}\{\sigma_\alpha = k + j\}. \tag{5.33}$$

In order to do this we provide the next transformations

$$\mathbb{P}_\alpha^t\{\tau_\alpha \circ \theta_n = k\} = \sum_{j=0}^\infty \mathbb{P}^t\{\sigma_\alpha^{(j)} < n \leq \sigma_\alpha^{(j+1)}, \tau_\alpha \circ \theta_n = k\}$$
$$= \sum_{j=0}^\infty \sum_{i=0}^{n-1} \mathbb{P}_\alpha^t\{\sigma_\alpha^{(j)} = i, \sigma_\alpha \circ \theta_{\sigma_\alpha^{(j)}} = k + n - i\}$$
$$= \sum_{i=0}^{n-1} \sum_{j=0}^\infty \mathbb{P}_\alpha^t\{\sigma_\alpha^{(j)} = i\}\mathbb{P}_\alpha^{t+i}\{\sigma_\alpha = k + n - i\} = \sum_{i=0}^{n-1} \mathbb{P}_\alpha^{t+i}\{\sigma_\alpha = k + n - i\}\mathbb{P}_\alpha^t\{X_i \in \alpha\}$$
$$\leq \sum_{i=0}^{n-1} \mathbb{P}_\alpha^{t+i}\{\sigma_\alpha = k + n - i\} = \sum_{j=1}^n \mathbb{P}_\alpha^{t+n-j}\{\sigma_\alpha = k + j\}.$$

Now we can derive that for all $l \geq 0$

$$\begin{aligned} \sum_{k=l}^{\infty} \beta^k \mathbb{P}_\alpha^t\{\tau_\alpha \circ \theta_n = k\} &\leq \sum_{k=l}^{\infty} \sum_{j=1}^{n} \beta^k \mathbb{P}_\alpha^{t+n-j}\{\sigma_\alpha = k + j\} \\ &= \sum_{j=1}^{n} \beta^{-j} \sum_{k \geq l+j} \beta^k \mathbb{P}_\alpha^{t+n-j}\{\sigma_\alpha = k\}. \end{aligned} \tag{5.34}$$

1. We assume that (5.29) holds true. Let us denote $\beta_1 = \sqrt{\beta} > 0$, and $\xi_t = \beta_1^{\sigma_{t,\alpha}}$, note that random variables ξ_t defined on different probability spaces, as described in Sect. 5.2. Condition (5.29) implies $\sup_t \mathsf{E}_\alpha^t \left[|\xi_t|^2\right] = \sup_t \mathsf{E}_\alpha^t \left[\beta^{\sigma_\alpha}\right] < \infty$, which means that family of distributions of ξ_t is uniformly integrable. We introduce a special notation for its tails $a_n^{(t)} = \sum\limits_{k \geq n} \beta_1^k \mathsf{P}_\alpha^t \{\sigma_\alpha = k\}$. So, we have $\sup_t a_n^{(t)} \to 0,\ n \to \infty$. Then, (5.34) yields $\sup_n \sum\limits_{k=l}^{\infty} \beta_1^k \mathsf{P}_\alpha^t \{\tau_\alpha \circ \theta_n = k\} \leq \sum\limits_{j=1}^{\infty} \beta^{-j} \sup_t a_{l+j}^{(t)} \to 0,\ l \to 0$. The latter expression implies that we can find a number $n_0 > 0$ such that

$$\sum_{k > n_0}^{\infty} \beta_1^k \mathbb{P}_\alpha^t\{\tau_\alpha \circ \theta_n = k\} \leq \frac{\varepsilon}{2\beta_1^m}. \tag{5.35}$$

We now choose $\delta \in (1, \beta_1)$ such that $\delta^{m+n_0} \leq 1 + \varepsilon/2$. Then we have:

$$\begin{aligned} &\mathbb{E}_\alpha^t\left[\delta^{m+\tau_\alpha \circ \theta_n}\right] = \mathbb{E}_\alpha^t\left[\delta^{m+\tau_\alpha \circ \theta_n} 1_{\tau_\alpha \circ \theta_n \leq n_0}\right] + \mathbb{E}_\alpha^t\left[\delta^{m+\tau_\alpha \circ \theta_n} 1_{\tau_\alpha \circ \theta_n > n_0}\right] \\ &\leq \delta^{m+n_0} + \beta_1^m \sum_{k > n_0} \mathbb{E}_\alpha^t\left[\beta_1^k 1_{\tau_\alpha \circ \theta_n = k}\right] \\ &= \delta^{m+n_0} + \beta_1^m \sum_{k > n_0} \beta_1^k \mathbb{P}_\alpha^t\{\tau_\alpha \circ \theta_n = k\} \leq 1 + \varepsilon/2 + \frac{\varepsilon \beta_1^m}{2\beta_1^m} = 1 + \varepsilon. \end{aligned} \tag{5.36}$$

2. Assume that condition (5.31) holds true. Using trivial $\beta^k = (\beta - 1) \sum\limits_{i=0}^{k-1} \beta^i + 1$, we get for all $l \geq 1$,

$$\sum_{k=j+l}^{\infty} \beta^k \mathsf{P}_{\alpha}^{t+n-j}\{\sigma_{\alpha}=k\} \le (\beta-1)\sum_{k=l+j}^{\infty}\sum_{i=0}^{k-1}\beta^i \mathsf{P}_{\alpha}^{t+n-j}\{\sigma_{\alpha}=k\} + \hat{G}_{l+j-1}$$
$$\le (\beta-1)\left[\left(\sum_{i=0}^{j+l-2}\beta^i\right)\hat{G}_{j+l-1} + \sum_{i>j+l}\beta^i \mathsf{P}_{\alpha}^{t+n-j}\{\sigma_{\alpha}>i\}\right] + \hat{G}_{l+j-1}$$
$$\le (\beta-1)\left[\frac{\beta^{l+j-1}-1}{\beta-1}\hat{G}_{j+l-1} + \sum_{i>j+l}\beta^i \hat{G}_i\right] + \hat{G}_{l+j-1}$$
$$= \beta^{j+l-1}\hat{G}_{j+l-1} + (\beta-1)\sum_{i>j+l}\beta^i\hat{G}_i$$
$$\le \hat{C}\left(\frac{\beta}{\hat{\beta}}\right)^{j+l-1} + \hat{C}(\beta-1)\sum_{i>j+l}\left(\frac{\beta}{\hat{\beta}}\right)^i = \left(\frac{\beta}{\hat{\beta}}\right)^{j+l-1}\frac{\beta(\hat{\beta}-1)}{\hat{\beta}-\beta}\hat{C}.$$

Plugging this inequality into (5.34) we get

$$\sup_n \sum_{k=l}^{\infty}\mathsf{P}_{\alpha}^t\{\tau_{\alpha}\circ\theta_n=k\} \le \hat{C}\frac{\beta(\hat{\beta}-1)}{\hat{\beta}-\beta}\sum_{j=1}^{\infty}\beta^{-j}\left(\frac{\beta}{\hat{\beta}}\right)^{j+l-1}$$
$$= \hat{C}\frac{\beta(\hat{\beta}-1)}{\hat{\beta}-\beta}\left(\frac{\beta}{hat\beta}\right)^{l-1}\sum_{j=1}^{\infty}\hat{\beta}^{-j} = \hat{C}\frac{\beta}{\hat{\beta}-\beta}\left(\frac{\beta}{\hat{\beta}}\right)^{l-1}. \tag{5.37}$$

We can now find number $n_0 \ge 2$ such that

$$\sum_{k>n_0}\mathsf{P}_{\alpha}^t\{\tau_{\alpha}\circ\theta_n=k\} \le \hat{C}\frac{\beta}{\hat{\beta}-\beta}\left(\frac{\beta}{\hat{\beta}}\right)^{n_0-2} \le \frac{\varepsilon}{2\beta^m}. \tag{5.38}$$

From (5.38) we can derive a direct expression for n_0,

$$n_0 = \left\lfloor \ln\left(\frac{\varepsilon(\hat{\beta}-\beta)}{2\hat{C}\beta^{m+1}}\right) / \ln\left(\frac{\beta}{\hat{\beta}}\right)\right\rfloor + 3,$$

which proves the formula for n_0 in (5.32). The proof is completed by setting $\delta = (1+\varepsilon/2)^{\frac{1}{m+n_0}}$ and applying transformations (5.36) with β instead of β_1.

In the next two lemmas, we will use notation from Sect. 5.4 and assume that conditions of Theorem 2 hold true.

Lemma 5 *Let $h\colon E \to \mathbb{R}_+$ be a measurable function. Then $\forall t \in \mathbb{N}_0$:*

$$1_{\alpha}\left(X_{\nu_{j-1}}^{(i)}\right)\mathsf{E}^t\left[h\left(X_{\nu_j}^{(i)}\right)|\mathcal{B}_j\right] = 1_{\alpha}\left(X_{\nu_{j-1}}^{(i)}\right)P_i^{t,U_j}h(\alpha). \tag{5.39}$$

Proof We prove formula (5.39) using the definition of conditional expectation. The random variable $P_i^{t,U_j} h(\alpha)$ is $\mathcal{B}_j$-measurable by construction of $\mathcal{B}_j$. It is enough to prove that for any set $A \in \mathcal{F}_{\nu_{j-1}}$:

$$\mathsf{E}^t_{x,y}\left[1_A 1_{U_j=k} 1_\alpha(X^{(i)}_{\nu_{j-1}}) h(X^{(i)}_{\nu_j})\right] = \mathsf{E}^t_{x,y}\left[1_A 1_{U_j=k} 1_\alpha(X^{(i)}_{\nu_{j-1}}) P_i^{t,U_j} h(\alpha)\right]. \quad (5.40)$$

Using the definition of ν_j we get

$$\begin{aligned}
&\mathsf{E}^t_{x,y}\left[1_A 1_{U_j=k} 1_\alpha(X^{(i)}_{\nu_{j-1}}) h(X^{(i)}_{\nu_j})\right] = \mathsf{E}^t_{x,y}\left[1_A 1_\alpha(X^{(i)}_{\nu_{j-1}}) \mathsf{E}^t\left[1_{U_j=k} h(X^{(i)}_{\nu_j}) | \mathcal{F}_{\nu_{j-1}}\right]\right] \\
&= \mathsf{E}^t_{x,y}\left[1_A 1_\alpha(X^{(i)}_{\nu_{j-1}}) \mathsf{E}^t\left[1_{U_j=k} h(X^{(i)}_{\nu_{j-1}+k}) | \mathcal{F}_{\nu_{j-1}}\right]\right] \\
&= \mathsf{E}^t_{x,y}\left[1_A 1_\alpha(X^{(i)}_{\nu_{j-1}}) \mathsf{E}^t\left[1_{\nu_{j-1}+q+\tau^{(1-i)}\circ\theta_{\nu_{j-1}+q}=k} h(X^{(i)}_{\nu_{j-1}+k}) | \mathcal{F}_{\nu_{j-1}}\right]\right] \\
&= \mathsf{E}^t_{x,y}\left[1_A 1_\alpha(X^{(i)}_{\nu_{j-1}}) \mathsf{E}^t_{X^{(i)}_{\nu_{j-1}}, X^{(1-i)}_{\nu_{j-1}}}\left[1_{\tau^{(1-i)}\circ\theta_q=k-q} h(X^{(i)}_k)\right]\right] \\
&= \mathsf{E}^t_{x,y}\left[1_A 1_\alpha(X^{(i)}_{\nu_{j-1}}) \mathsf{E}^t_{i,\alpha}\left[h(X^{(i)}_k)\right] \mathbb{P}_{X^{(1-i)}_{\nu_{j-1}}}\{\tau^{(1-i)} \circ \theta_q = k-q\}\right] \\
&= \mathsf{E}^t_{i,\alpha}\left[h(X^{(i)}_k)\right] \mathsf{E}^t_{x,y}\left[1_A 1_\alpha(X^{(i)}_{\nu_{j-1}}) \mathbb{P}^t\{U_j = k | \mathcal{F}_{\nu_{j-1}}\}\right] \\
&= P_i^{t,k} h(\alpha) \mathsf{E}^t_{x,y}\left[1_A 1_{U_j=k} 1_\alpha(X^{(i)}_{\nu_{j-1}})\right] = \mathsf{E}^t_{x,y}\left[1_A 1_{U_j=k} 1_\alpha(X^{(i)}_{\nu_{j-1}}) P_i^{t,U_j} h(\alpha)\right].
\end{aligned}$$

So, formula (5.40) and thus (5.39) is proved.

Lemma 6 *Let $r(n)$, $n \geq 0$ be a nonnegative sequence. Then there is $\gamma < 1$ such that*

$$\mathbb{E}^t\left[1_{\tau>j} r(\nu_j) | \mathcal{F}_{\nu_{j-1}}\right] \leq (1-\gamma) 1_{\tau>j-1} \mathbb{E}^t\left[r(\nu_j) | \mathcal{F}_{\nu_{j-1}}\right]. \quad (5.41)$$

Proof In this proof all random variables outside $\mathbb{E}^t$ should be understood as having lower index t, that is, $1_\alpha\left(X^{(i)}_{\nu_{j-1}}\right) = 1_\alpha\left(X^{(i)}_{t,\nu_{t,j-1}}\right)$. We have

$$\begin{aligned}
1_\alpha\left(X^{(i)}_{\nu_{j-1}}\right) \mathbb{E}^t\left[1_{\tau>j} r(\nu_j) | \mathcal{F}_{\nu_{j-1}}\right] &= 1_\alpha\left(X^{(i)}_{\nu_{j-1}}\right) \mathbb{E}^t\left[1_{\tau>j-1} 1_{\alpha^c}\left(X^{(i)}_{\nu_j}\right) r(\nu_j) | \mathcal{F}_{\nu_{j-1}}\right] \\
&= 1_{\tau>j-1} 1_\alpha\left(X^{(i)}_{\nu_{j-1}}\right) \mathbb{E}^t\left[\mathbb{E}^t\left[1_{\alpha^c}\left(X^{(i)}_{\nu_j}\right) | \mathcal{B}_j\right] r(\nu_j) | \mathcal{F}_{\nu_{j-1}}\right].
\end{aligned}$$

Using Lemma 5 we get

$$\begin{aligned}
1_{\tau>j-1} 1_\alpha&\left(X^{(i)}_{\nu_{j-1}}\right) \mathbb{E}^t\left[\mathbb{E}^t\left[1_{\alpha^c}\left(X^{(i)}_{\nu_j}\right) | \mathcal{B}_j\right] r(\nu_j) | \mathcal{F}_{\nu_{j-1}}\right] \\
&= 1_{\tau>j-1} 1_\alpha\left(X^{(i)}_{\nu_{j-1}}\right) \mathbb{E}^t\left[P_i^{t,U_j}(\alpha, \alpha^c) r(\nu_j) | \mathcal{F}_{\nu_{j-1}}\right] \\
&= 1_{\tau>j-1} 1_\alpha\left(X^{(i)}_{\nu_{j-1}}\right) \mathbb{E}^t\left[(1 - P_i^{t,U_j}(\alpha, \alpha)) r(\nu_j) | \mathcal{F}_{\nu_{j-1}}\right].
\end{aligned}$$

By Lemma 3, $P_i^{t,2m+n}(\alpha,\alpha) \geq \gamma$, $\forall n \geq 0$, and since $U_j \geq 2m$:

$$1_{\tau>j-1} 1_\alpha\left(X^{(i)}_{\nu_{j-1}}\right) \mathbb{E}^t\left[(1-P_i^{t,U_j}(\alpha,\alpha)) r(\nu_j) | \mathcal{F}_{\nu_{j-1}}\right]$$

$$\leq (1-\gamma) 1_{\tau>j-1} 1_\alpha\left(X^{(i)}_{\nu_{j-1}}\right) \mathbb{E}^t\left[r(\nu_j)|\mathcal{F}_{\nu_{j-1}}\right].$$

It means, that we have established the following relation

$$1_\alpha\left(X^{(i)}_{\nu_{j-1}}\right) \mathbb{E}^t\left[1_{\tau>j} r(\nu_j)|\mathcal{F}_{\nu_{j-1}}\right] \leq (1-\gamma) 1_{\tau>j-1} 1_\alpha\left(X^{(i)}_{\nu_{j-1}}\right) \mathbb{E}^t\left[r(\nu_j)|\mathcal{F}_{\nu_{j-1}}\right]. \tag{5.42}$$

We may note that by the definition of ν_{j-1}:

$$\left[1_\alpha\left(X^{(0)}_{\nu_{j-1}}\right) + 1_\alpha\left(X^{(1)}_{\nu_{j-1}}\right)\right] 1_{\tau>j-1} = 1. \tag{5.43}$$

Now we sum inequalities (5.42) for $i \in \{0, 1\}$ and using (5.43) derive (5.41).

Appendix

We state here the Comparison Theorem, it is proved in [7], Theorem 4.3.1.

Theorem 5 *Let $\{\mathcal{V}_n,\ n \geq 0\}$, $\{\mathcal{Y}_n,\ n \geq 0\}$, and $\{Z_n,\ n \geq 0\}$ be three $\{\mathcal{F}_n,\ n \geq 0\}$-adapted nonnegative processes such that for all $n \geq 0$,*

$$\mathbb{E}\left[\mathcal{V}_{n+1}|\mathcal{F}_n\right] + \mathcal{Z}_n \leq \mathcal{V}_n + \mathcal{Y}_n,\ \mathbb{P}\text{– a.s.}.$$

Then for every $\{\mathcal{F}_n,\ n \geq 0\}$-stopping time τ,

$$\mathbb{E}\left[\mathcal{V}_\tau 1_{\tau<\infty}\right] + \mathbb{E}\left[\sum_{k=0}^{\tau-1} \mathcal{Z}_k\right] \leq \mathbb{E}[\mathcal{V}_0] + \mathbb{E}\left[\sum_{k=0}^{\tau-1} \mathcal{Y}_k\right].$$

References

1. Andrieu, C. Fort., G., Vihola, M.: Quantitive convergence rates for subgeometric Markov chains. J. Appl. Prob. **52**, 391–404 (2015)
2. Andrulyte, I.M., Bernackaite, E., Kievinaite, D., Siaulys, J.: A Lundberg-type inequality for an inhomogeneous renewal risk model. Modern Stoch. Theory Appl. **2**(2), 173–184 (2015)
3. Chow, Y.S., Robbins, H.: A renewal theorem for random variables which are dependent or non-identically distributed. Ann. Math. Statist. **34**(2), 390–395 (1963)

4. Connors, D.P., Kumar, P.R.: Simulated annealing and balance of recurrent order in time-inhomogeneous Markov chains. In: Proceedings of the 26th Conference on Decision and Control, pp. 2261–2263 (1987)
5. Dobrushin, R.: Central limit theorems for non-stationary Markov chains I, II. Theory Probab. Its Appl. **1**(1) 65-80, **1**(4) 329-383 (1956)
6. Doeblin, W.: Expose de la theorie des chaines simples constantes de Markov a un nomber fini d'estats. Mathematique de l'Union Interbalkanique **2**, 77–105 (1938)
7. Douc, R., Moulines, E., Priouret, P., Soulier, P.: Markov Chains. Springer, Cham (2018)
8. Douc, R., Moulines, E., Rosenthal, J.S.: Quantitive bounds on convergence of time-inhomogeneous Markov chains. Ann. Appl. Probab. **14**(4), 1643–1665 (2004)
9. Golomoziy, V., Kartashov, N.: On the integrability of the coupling moment for time-inhomogeneous Markov chains. Theory Probab. Math. Stat. **89**, 1–12 (2015)
10. Golomoziy, V.: An inequality for the coupling moment in the case of two inhomogeneous Markov chains. Theory of Probability and Mathematical Statistics **90**, 43–56 (2015)
11. Golomoziy, V.: An estimate for an expectation of the simultaneous renewal for time-inhomogeneous Markov chains. Modern Stoch. Theory Appl. **3**(4), 315–323 (2016)
12. Golomoziy, V.: An estimate for an expectation of the excess of the renewal sequence generated by the non-homogeneous Markov chain under a condition of existence square-integrable stochastic dominant. Theory Probab. Math. Stat. **94**, 50–59 (2016) (in Ukrainian)
13. Golomoziy, V.: On estimation of expectation of simultaneous renewal time of time-inhomogeneous Markov chains using dominating sequence. Modern Stoch. Theory Appl. **6**(3), 333–343 (2019)
14. Golomoziy, V., Mishura, Y.: Stability estimates for finite-dimensional distributions of time-inhomogeneous Markov chains. Mathematics **8**(2), 174 (2020)
15. Golomoziy, V.: Stability estimates for transition probabilities of time-inhomogeneous Markov chains under the condition of the minorization on the whole space. Theory Probab. Math. Stat. **2**(101), 78–92 (2019) (in Ukrainian)
16. Kartashov, Y., Golomoziy, V., Kartashov, N.: The impact of stress factor on the price of widow's pension, In: Silverstrov, D., Martin-Lof, A. (eds.) Modern Problems in Insurance Mathematics, E. A. A. Series, pp. 223–237. Springer, Berlin, Germany (2014)
17. Kartashov, N., Golomoziy, V.: Maximal coupling procedure and stability of discrete Markov chains. I. Theory Probab. Math. Stat. **86**, 81–92 (2012)
18. Kartashov, N., Golomoziy, V.: Maximal coupling procedure and stability of discrete Markov chains. II. Theory Probab. Math. Stat. **87**, 58–70 (2012)
19. Kartashov, N., Golomoziy, V.: Maximal coupling and stability of discrete non-homogeneous Markov chains. Theory Probab. Math. Stat **91**, 17–27 (2015)
20. Lindvall, T.: Lectures on the Coupling Method, John Wiley and Sons, 1992
21. Madsen, R.W.: A note on some ergodic theorems of A. Paz. Ann. Math. Stat. **42**(1), 405–408 (1971)
22. Maire, F., Douc, R., Olsson, J.: Comparison of asymptotic variances of inhomogeneous Markov chains with application to Markov chain Monte Carlo methods. Ann. Stat. **42**(4), 1483–1510 (2014)
23. Meyn, S., Tweedie, R.L.: Markov Chains and Stochastic Stability, 2nd edn., Cambridge University Press (2009)
24. Neveu, J.: Probability Theory, 2nd edn., pp. 45–56. Springer, New York (1996)
25. Nummelin, E.: A splitting technique for Harris recurrent Markov chains. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **43**, 309–318 (1978)
26. Paz, A.: Ergodic theorems for infinite probabilistic tables. Ann. Math. Statist. **41**, 539–550
27. Silvestrov, D.: Coupling for Markov renewal processes and the rate of convergence in ergodic theorems for processes with semi-Markov switchings. Acta Appl. Math. **34**, 109–124 (1994)
28. Sethuraman, S., Varadhan, S.R.S.: A martingale proof of Dobrushin's theorem for non-homogeneous Markov chains. Electron. J. Probab. **10**, 1221–1235 (2005)
29. Smith, W.: On the elementary renewal theorem for non-identically distributed variables. Pacif. J. Math. **14**(2), 673–699 (1964)
30. Revuz, D.: Markov Chains, North Holland (1984)
31. Thorisson, H.: Coupling, Stationarity, and Regeneration. Springer, New York (2000)