Sub-exponential mixing rate for a class of Markov chains*

S. A. KLOKOV† AND A. YU. VERETENNIKOV‡

Abstract. *We establish sub-exponential bounds for the β-mixing rate and for the rate of convergence to invariant measures for discrete time Markov processes under recurrence type conditions weaker than used for exponential inequalities and stronger than for polynomial ones.*

Key words: *Markov processes, recurrence, invariant measure, mixing coefficients, sub-exponential convergence*

AMS subject classifications: 60J10, 60F25

Received November 28, 2003 Accepted January 21, 2004

1. Introduction

The goal of this paper is to establish sub-exponential mixing and convergence rates to equilibrium for a class of Markov chains, that is, bounds of the type

$$\beta_{n,x} \le C(x)e^{-cn^{\delta}}, \qquad c, \delta, C(x) > 0,\ n \ge 0. \tag{1}$$

Here $\beta_{n,x}$ is a β-mixing coefficient, see definition (6) below; certain other mixing coefficients and distances between current and limiting distributions might stand on the left-hand side of (1) either.

Assumptions will include two standard groups of conditions: recurrence type conditions and local mixing. The former are formulated for model (2) (see below) in terms of "drift and diffusion", similar to continuous time setting, cf. [39, 43, 44] where drift and diffusion are the most natural objects for assumptions. Other possibilities provide assumptions in terms of Lyapunov functions, as in the majority of papers. However, they often leave an open question how to check them for a wide class of processes. A habit to use "true" Lyapunov functions clearly delayed studying of weak mixing bounds.

*Supported by EPSRC grant GR/R40746/01 (UK), grant NFGRF 2301863 (University of Kansas) for both authors, grant RFBR 03-01-00045 for the first author and grants INTAS 99-0559 and RFBR-02-01-00444 for the second author.

†Omsk Branch of Institute of Mathematics SB RAS, Omsk, Russia, e-mail: klokov@iitam.omsk.net.ru

‡School of Mathematics, University of Leeds, UK & Institute of Information Transmission Problems, Moscow, Russia, e-mail: veretenn@maths.leeds.ac.uk

Classical uniform exponential convergence to equilibrium for Markov processes was established by Doeblin, Kolmogorov, Doob in the middle of the 20th century under various conditions, one of which is now known as a "Doeblin type condition", the name and final formulation belongs to Doob [10]. In 70ies, an interest moved to non-uniform convergence, see [26, 34], et al., simply because a lot of useful processes do not satisfy any uniform convergence bounds. It became clear that this non-uniform convergence relates to "local" Doeblin type condition and to hitting time bounds for some "petite sets". The latter bounds were established in [17, 26, 34] et al.

At first, only exponential convergence rates were treated, using fairly simple "Lyapunov functions", see [17, 28, 39, 34], et al. Recently polynomial convergence bounds were studied [38, 43, 44, 36, 22, 11, 5], which considerably expanded the class of examples. The same idea of Lyapunov functions was used in most of these works, although not straightforward. In [38] a special sequence of auxiliary functions was constructed. In [41, 15] polynomial bounds were obtained under assumptions on hitting times.

In [43, 44], for the same aim under weaker assumptions, "quasi-Lyapunov" functions imitating their main properties up to a small controllable discrepancy were used. In [22, 11] a simplified idea from [38] was treated. In a broader sense all cited papers on non-uniform bounds explore extensions of the original Lyapunov idea [31]. When we speak of "quasi-Lyapunov" functions, we do not mean that exact Lyapunov functions do not exist: e.g., for diffusions, $\mathbb{E}_x\tau$ (τ being some hitting time) is an example of such a function. However, it is not an explicit formula, and "approximate" or "quasi-Lyapunov" functions may be preferable.

Other important results related to polynomial inequalities for hitting times can be found in [30, 25, 33, 37, 3]. The first two papers in this list were pioneering, and both were not appreciated at their time. The use of both papers could have accelerated a lot the investigation of polynomial non-uniform convergence and mixing bounds. Hitting time inequalities constitute the part of the technique used in [43, 44], as well as in this paper. It is impossible to provide a complete list of all related works here, they can be found in other references.

Essentially, the approach from [43, 44] was repeated in [22] and [12] in a more abstract form using an assumption $PV \le V - cV^\alpha + b\mathbf{1}_C$. Analogous methods based on [38] and the same assumptions on V were used in [11, 24] to get polynomial convergence rates for the Metropolis–Hastings algorithm and for another related Markov chain model.

There was a parallel process of exploration of the so-called mixing coefficients, a notion more general than convergence to equilibrium, very useful in limit theorems for "weakly dependent" random variables, cf. [20, 13], et al. Most of works employed various mixing coefficients assuming certain decreasing rate for them. Only a few papers were devoted to verifying different mixing conditions for stationary or Markov processes, see [18, 19, 6, 7, 1]. In [15, 39, 41, 43, 44] both mixing and convergence rate were studied using the coupling method, see [2, 35]; its source is attributed to Doeblin and Kolmogorov.

Our results include inequalities for β-mixing coefficient which was introduced in [46] under the name "complete regularity coefficient" for stationary processes, and now in a slightly extended form (for non-stationary processes) is known as Kolmogorov's or β-coefficient; later this coefficient was studied in [21] et al. This is a very useful intermediate coefficient between Rosenblatt's α-mixing and Ibragimov's φ-mixing ones. It works in situations where there is no φ-mixing (cf. [20, 8]), which is common.

Assumptions in [38] which provided polynomial convergence turned out to be non-optimal. In [43, 44] for certain classes of Markov processes, similar polynomial bounds for convergence in total variation and also for β-mixing were established under much less restrictive assumptions. On the other hand, under original conditions used in [38], better "sub-exponential" bounds like (1) were proved in [32]. For rather specific diffusions and under more restrictive assumptions sub-exponential convergence (as well as exponential and polynomial) was also established in [14].

In this paper we establish "real" sub-exponential bounds for a class of Markov chains under recurrence type conditions similar to those in [32], and improve them in the following sense: the power degree coefficient δ should approach 1 if recurrence type assumptions "approach" those implying the exponential convergence rate (cf. [39, 34]). This natural property fails for bounds in [32].

The sub-exponential bounds may be useful in moderate deviations for various processes, see [4, 9, 16, 27]. In more practical applications to the MCMC and insurance theory, there is also an interest in intermediate cases between exponential and polynomial ones, corresponding to "normal" and "catastrophic" according to the insurance theory terminology.

2. The problem setting

Consider a homogeneous Markov process $(X_n,\ n \ge 0)$ in $\mathbb{R}^d$ with the usual scalar product $\langle \cdot,\cdot \rangle$, the norm $|\cdot|$, and the family of Borel sets $\mathcal{B}(\mathbb{R}^d)$. We define $\mathcal{F}^X_I = \sigma\,(X_k,\, k \in I)$, and write $\mathcal{F}_n$ instead of $\mathcal{F}^X_{\le n}$. Denote by $\mathbb{E}_x$ expectation of the process (X_n) with the initial value $X_0 = x$, $\mathbb{P}_x(\cdot) = \mathbb{E}_x \mathbf{1}(\cdot)$, and $\mu_{n,x} = \mathcal{L}(X_n|X_0 = x)$. We write $\mathbb{E}_\mu$ and $\mathbb{P}_\mu$, if $\mathcal{L}(X_0) = \mu$. Symbols $\wedge$ and $\vee$ denote operations of taking minimum and maximum of real numbers, respectively.

We will use the representation of X_n in the form of non-linear autoregression,

$$X_{n+1} = g(X_n) + V_{n+1}, \tag{2}$$

where g is a Borel function satisfying $g(X_n) = \mathbb{E}\{X_{n+1}|\mathcal{F}_n\}$ (a.s.), and $V_{n+1} = X_{n+1} - \mathbb{E}\{X_{n+1}|\mathcal{F}_n\}$. The definition of V_{n+1} implies that

$$\mathbb{E}\{V_{n+1}|\mathcal{F}_n\} = 0. \tag{3}$$

Suppose that the following *recurrence type conditions* are true:

- there exist positive constants R_0, C_0, r, and $0 < p < 1$ such that

$$|g(x)| \le \begin{cases} C_0, & |x| \le R_0, \\ |x|(1 - r/|x|^{1+p}), & |x| > R_0, \end{cases} \tag{4}$$

- there are $K > 0$ and $0 < \alpha \le 1 - p$ such that

$$\sup_n \operatorname{ess\,sup} \mathbb{E}\{e^{k|V_{n+1}|^\alpha}|\mathcal{F}_n\} < \infty, \quad 0 \le k < K. \tag{5}$$

Condition (4) provides an "attraction" to the ball $\{x : |x| \le R\}$; the parameter p regulates the force of this attraction. The cases $p = 0$ and $p = 1$ were considered in [39] and [44], respectively; $p = 0$ leads to an exponential β-mixing, while $p = 1$ provides a polynomial one (depending also on r). Condition (5) implies the existence of certain sub-exponential moments of X_n. Additional assumptions on (X_n) will be formulated later.

The β-mixing coefficient is defined by the formula

$$\beta_{n,x} = \sup_{m\ge 0} \mathbb{E}_x \operatorname*{var}_{B\in\mathcal{F}^X_{\ge n+m}} (\mathbb{P}(B|\mathcal{F}_m) - \mathbb{P}(B)), \tag{6}$$

where $\operatorname{var}\nu$ is the total variation of ν. Also, the following "average" version of this coefficient is often useful,

$$\bar\beta_n = \int \beta_{n,x}\, \mu_\infty(dx). \tag{7}$$

Here μ_∞ stands for the invariant measure of our Markov chain. Assumptions of our theorems will guarantee that μ_∞ exists and is unique.

Fix $B \in \mathcal{B}(\mathbb{R}^d)$ and let $\tau_0 = 0$, $\tau_{n+1} = \inf\{t > \tau_n : X_t \in B\}$. Define "the process on B", $X^B_n = X_{\tau_n}$, and denote by $P^B(x, dy)$ its transition probability.

We say that the process (X_n) satisfies *the local Doeblin condition*, if for every R large enough and $B = \{x \in \mathbb{R}^d : |x| \le R\}$,

$$\inf_{x,x'\in B} \int \min\left\{\frac{P^B(x, dy)}{P^B(x', dy)}, 1\right\} P^B(x', dy) = \kappa(R) > 0; \tag{8}$$

it provides non-singularity of the measures within the ball B, and (together with (4)–(5)) implies irreducibility. It is also possible in terms of n-step transition probabilities.

There is another equivalent and popular description of non-singularity based on "petite sets". It could be used in this problem as an alternative. However, for uniformly ergodic Markov chains, this approach provides exponential convergence rates with worse constants under the exponent. Condition (8) is an extension of Dobrushin's ergodic condition; it gives better constants which are in some cases optimal. It is likely that in the sub-exponential case the local Doeblin condition serves better, too. Cf. [12] concerning the use of this type of condition. One can also notice that in positive operators approach to large deviations exactly this kind of nonsingularity condition is in use, too, cf. [42]. There are other advantages of this integral condition in comparision to the "petite sets" one, e.g., in situations when any non-singularity type condition must be *verified,* the integral one may turn to be much more simple; this work is now in progress for certain models related to discretisation.

3. Main result. Examples

Theorem 1. *Let (X_n) satisfy conditions (4)–(5), (8), and $0 < \delta < \alpha/(1+p)$. Then there exist c_0, c_1, $C > 0$, such that for any $x = X_0$ the following bounds hold true,*

$$\operatorname{var}(\mu_{n,x} - \mu_\infty) \le Ce^{c_1|x|^\alpha - c_0 n^\delta}, \tag{9}$$

$$\beta_{n,x} \le Ce^{c_1|x|^\alpha - c_0 n^\delta}, \tag{10}$$

$$\bar\beta_n \le Ce^{-c_0 n^\delta}. \tag{11}$$

Example 1. *This is a situation when mixing is not exponential, and the power under the sub-exponential cannot be improved. Consider the one-dimensional process (2) with*

$$g(x) = \begin{cases} 0, & x \le 1, \\ x - x^{-p}, & x > 1, \end{cases}$$

and the parameter p as in (4) and (5). Let (V_n) be a sequence of i.i.d. random variables having distribution with just one atom on the negative side of the real line, a sub-exponential tail on the positive side and zero mean.

It is proved in [29] that the conditions of Theorem 1 are satisfied, and for the stationary regime β-mixing holds with a truly sub-exponential rate, i.e. for any $\varepsilon > 0$,

$$c_1 \exp\left\{-c_2 n^{\frac{\alpha}{1+p}}\right\} \le \beta_n \le c_3 \exp\left\{-c_4 n^{\frac{\alpha}{1+p} - \varepsilon}\right\}, \qquad c_1, \ldots, c_4 > 0.$$

Example 2. *Consider the Metropolis algorithm on $\mathbb{R}$ (see, e.g. [11, 23]) with target and proposal densities*

$$\pi(x) \propto e^{-|x|^\delta}, \qquad q(x,y) = q(|x-y|) = \frac{1}{2\rho}\mathbf{1}(|x-y| \le \rho),$$

with $0 < \delta < 1$, $\rho > 0$, which is described as follows. Suppose the chain is in state x. A possible move from x to y is generated accordingly to the proposal density. The move is accepted with probability

$$\alpha(x,y) = \begin{cases} 1 \wedge \frac{\pi(y)}{\pi(x)}, & \pi(x)q(|x-y|) > 0, \\ 1, & \pi(x)q(|x-y|) = 0. \end{cases}$$

If the move is not accepted, the chain stays in x. Under very relaxed conditions on q and π the chain is time reversible and has $\pi(x)$ as its invariant density.

To implement the algorithm on a computer one may use the identity,

$$X_{n+1} = X_n + W_{n+1}\mathbf{1}(U_{n+1} \le \alpha(X_n, X_n + W_{n+1})),$$

where W_{n+1} has conditional density $q(X_n, X_n + w)$ given X_n and (U_n) is a sequence of independent uniformly distributed on $[0,1]$ random variables. Representing X_{n+1}

in the form of non-linear autoregression, we get (2) with

$$g(x) = x + \int w\alpha(x, x+w)\, q(|w|)dw,$$
$$V_{n+1} = W_{n+1}\mathbf{1}(U_{n+1} \le \alpha(X_n, X_n + W_{n+1})) - \mathbb{E}\{W_{n+1}\alpha(X_n, X_n + W_{n+1})|\mathcal{F}_n\}.$$

Then $\mathbb{E}\{e^{k|V_{n+1}|^\alpha}|\mathcal{F}_n\} \le \mathbb{E}\{e^{2k|W_{n+1}|^\alpha}|\mathcal{F}_n\} \le C < \infty$. *Therefore, (3) and (5) are satisfied. For* $x > 0$ *large enough one writes,*

$$\int w\alpha(x, x+w)q(|w|)dw = \int_0^\infty w\left(\frac{\pi(x+w)}{\pi(x)} - 1\right) q(|w|)dw.$$

Since $\pi(x+w)/\pi(x) - 1 = e^{-(x+w)^\delta + x^\delta} \sim -\delta w x^{-1+\delta}, \quad x \to \infty$, *we get*

$$\int w\alpha(x, x+w)q(|w|)dw \le -\frac{r}{x^{1-\delta}}, \quad r > 0.$$

It implies (4) with $p = 1 - \delta$. *The* n*-step version of (8) with* some n *is straightforward. By Theorem 1, sub-exponential* β*-mixing and convergence follow.*

4. Preliminaries

Denote $B_R = \{x \in \mathbb{R}^d : |x| \le R\}$, $r_2 = \sup_n \operatorname{ess\,sup} \mathbb{E}\{|V_{n+1}|^2|\mathcal{F}_n\}$, and $\tau = \inf\{n \ge 0 : |X_n| \le R\}$. We establish some inequalities on (X_n) (called Lyapunov's drift condition) and on τ.

Lemma 1. *Let conditions (4), (5) hold true (with any* $0 < k < K$*), and* $0 < \alpha < 1 - p$. *Then for every* $0 < r' < r$ *the value of* $R_1 \ge R_0$ *can be chosen in such a way that for any* $R \ge R_1$ *and* $|x| > R$,

$$\mathbb{E}_x\left(e^{k|X_{n+1}|^\alpha} - e^{k|X_n|^\alpha}\right)\mathbf{1}(n < \tau) \le -k\alpha r' \mathbb{E}_x \frac{e^{k|X_n|^\alpha}}{|X_n|^{1+p-\alpha}}\mathbf{1}(n < \tau), \qquad (12)$$

$$\mathbb{E}_x e^{k|X_n|^\alpha}\mathbf{1}(n < \tau) \le e^{k|x|^\alpha}, \qquad (13)$$

and there is $C > 0$ *such that*

$$\mathbb{E}_x \tau \le C e^{k|x|^\alpha}. \qquad (14)$$

If $\alpha = 1 - p$, $k < 2r/(\alpha r_2)$, *and* $0 < r' < r - k\alpha r_2/2$, *then (12) remains true.*

Inequality (14) will be significantly improved in *Lemma 2.* However, it suffices for the existence of a unique invariant measure (see Theorem 10.0.1 in [34]).

Lemma 2. *Let hypothesis of* Lemma 1 *hold true,* $0 < \delta < \alpha/(1+p)$, *and* R *is chosen in accordance with* Lemma 1. *Then there is* $C > 0$ *such that*

$$\mathbb{E}_x e^{\tau^\delta} \le C e^{k|x|^\alpha}.$$

Consider the direct product of two identical probability spaces with two independent copies of our Markov process (X_n) and (X'_n) with $X_0 = x$ and $X'_0 = x'$.

Let $\gamma = \inf\{n \ge 1 : |X_n| \vee |X'_n| \le \tilde R\}$. We study the behavior of the process $|X_n| \vee |X'_n|$, and establish some properties of γ similar to those of τ.

Lemma 3. *Let conditions (4), (5) hold true, $0 < \alpha < 1 - p$. Then for every $0 < r'' < r$ the value of $\tilde R_1 \ge R_0$ can be chosen in such a way that for any $\tilde R \ge \tilde R_1$ and any $|x| \vee |x'| > \tilde R$,*

$$\begin{aligned}
&\mathbb{E}_{x,x'} \left(e^{k|X_{n+1}|^\alpha} + e^{k|X'_{n+1}|^\alpha}\right) \mathbf{1}(n < \gamma) \\
&\quad \le \mathbb{E}_{x,x'} \left(e^{k|X_n|^\alpha} + e^{k|X'_n|^\alpha}\right) \mathbf{1}(n < \gamma) \\
&\qquad - k\alpha r'' \mathbb{E}_{x,x'} e^{k(|X_n|\vee|X'_n|)^\alpha} (|X_n| \vee |X'_n|)^{-1-p+\alpha} \mathbf{1}(n < \gamma),
\end{aligned} \tag{15}$$

and for every n

$$\mathbb{E}_{x,x'} \left(e^{k|X_n|^\alpha} + e^{k|X'_n|^\alpha}\right) \mathbf{1}(n < \gamma) \le e^{k|x|^\alpha} + e^{k|x'|^\alpha}. \tag{16}$$

If $\alpha = 1 - p$, $k < 2r/(\alpha r_2)$, and $0 < r'' < r - k\alpha r_2/2$, then (15) remains true.

Lemma 4. *Let hypothesis of* Lemma 3 *hold true, $0 < \delta < \alpha/(1+p)$, and $\tilde R$ is chosen in accordance with the same* Lemma 3. *Then there is $C > 0$ such that*

$$\mathbb{E}_{x,x'} e^{\gamma^\delta} \le C(e^{k|x|^\alpha} + e^{k|x'|^\alpha}).$$

Lemma 5. *Let hypothesis of* Lemma 1 *and condition (8) hold true. Then for any $0 \le k < K$,*

$$\mathbb{E}_{\mu_\infty} e^{k|X_n|^\alpha} = \int e^{k|x|^\alpha} \, \mu_\infty(dx) < \infty.$$

Lemma 6. *Let hypothesis of* Lemma 2 *and condition (8) hold true. Then for every $0 \le k < K$,*

$$\sup_n \mathbb{E}_x e^{k|X_n|^\alpha} < \infty.$$

5. Proofs

We use the scheme from [44] adapted for the sub-exponential case.

PROOF of *Lemma 1.* The main idea is to use Taylor's expansion of the second order and condition of "zero conditional mean value" (3). However, this is hard to implement directly because of $|x|$ under the sub-exponential, except for rather special cases like $d = 1$ and $g(x) = x - rx^{-p}$. We suggest inequalities in the spirit of Taylor's expansions instead.

0. To simplify formulas let $X = X_n$, $V = V_{n+1}$, $g = g(X_n)$, then $X_{n+1} = g + V$.

1. Due to condition (5), random variables from the sequence (V_n) have finite conditional expectations, for every $0 \le k < K$ and $m \ge 0$,

$$\sup_n \operatorname{ess\,sup} \mathbb{E}\{e^{k|V_{n+1}|^\alpha} |V_{n+1}|^m | \mathcal{F}_n\} < \infty. \tag{17}$$

This implies (with non-random $o_N(1)$) that

$$\sup_n \operatorname{ess\,sup} \mathbb{E}\{e^{k|V_{n+1}|^\alpha}|V_{n+1}|^m \mathbf{1}(|V_{n+1}|>N)|\mathcal{F}_n\} = o_N(1),\ N\to\infty. \tag{18}$$

2. We will estimate $\mathbb{E}_x\left(e^{k|g+V|^\alpha}-e^{k|X|^\alpha}\right)\mathbf{1}(n<\tau)$. For an arbitrary $\varepsilon\in(0,1]$ split the expectation in two terms:

$$\begin{aligned} I_1 &:= \mathbb{E}_x\left(e^{k|g+V|^\alpha}-e^{k|X|^\alpha}\right)\mathbf{1}(n<\tau)\mathbf{1}(|V|>\varepsilon|X|^{1-\alpha}),\\ I_2 &:= \mathbb{E}_x\left(e^{k|g+V|^\alpha}-e^{k|X|^\alpha}\right)\mathbf{1}(n<\tau)\mathbf{1}(|V|\le\varepsilon|X|^{1-\alpha})). \end{aligned}$$

Each term will be estimated separately.

3. By (4) we have $|g|\le|X|$, if $n<\tau$. Using inequalities $(x+y)^\alpha\le x^\alpha+y^\alpha$ for $x,y\ge 0$, $0<\alpha\le 1$, and

$$\mathbf{1}(|V|>\varepsilon|X|^{1-\alpha})\le\frac{|V|^q}{\varepsilon^q|X|^{q(1-\alpha)}}\mathbf{1}(|V|>\varepsilon|X|^{1-\alpha}),\quad q>0, \tag{19}$$

one writes,

$$\begin{aligned} I_1 &\le \mathbb{E}_x e^{k|X|^\alpha}e^{k|V|^\alpha}\mathbf{1}(n<\tau)\mathbf{1}(|V|>\varepsilon|X|^{1-\alpha})\\ &\le \varepsilon^{-q}\mathbb{E}_x\frac{e^{k|X|^\alpha}}{|X|^{q(1-\alpha)}}\mathbf{1}(n<\tau)\mathbb{E}\{e^{k|V|^\alpha}|V|^q\mathbf{1}(|V|>\varepsilon R^{1-\alpha})|\mathcal{F}_n\}. \end{aligned}$$

If we take $q=(1+p-\alpha)/(1-\alpha)$ and use (18), then

$$I_1\le o_R(1)\cdot\varepsilon^{-q}\mathbb{E}_x\frac{e^{k|X|^\alpha}}{|X|^{1+p-\alpha}}\mathbf{1}(n<\tau),\quad R\to\infty. \tag{20}$$

4. We start estimating I_2. Assuming $|X|>R$ and $|V|\le\varepsilon|X|^{1-\alpha}$, one writes,

$$\begin{aligned} |g+V|^\alpha-|X|^\alpha &= |X|^\alpha\left[\left\langle\frac{g+V}{|X|},\frac{g+V}{|X|}\right\rangle^{\alpha/2}-1\right]\\ &= |X|^\alpha\left[\left(1+2\left\langle\frac{g}{|X|},\frac{V}{|X|}\right\rangle+\frac{|g|^2+|V|^2-|X|^2}{|X|^2}\right)^{\alpha/2}-1\right]\\ &\le \alpha\left[\left\langle\frac{g}{|X|},\frac{V}{|X|^{1-\alpha}}\right\rangle+\frac{|g|^2-|X|^2}{2|X|^{2-\alpha}}+\frac{|V|^2}{2|X|^{2-\alpha}}\right]\\ &=: \alpha[J_1+J_2+J_3], \end{aligned}$$

due to the inequality $(1+z)^{\alpha/2}\le 1+\alpha z/2$, $z\ge -1$. Notice that

$$|J_1|\le\varepsilon,\qquad 0\le J_3\le\frac{\varepsilon^2}{2|X|^\alpha},$$

$$J_2\le\frac{1}{2|X|^{-\alpha}}\left[\left(1-\frac{r}{|X|^{1+p}}\right)^2-1\right]\le-\frac{\tilde r}{|X|^{1+p-\alpha}}=:J_4,$$

where $\tilde r$ can be taken arbitrarily close to r, if R is increased. Here, the key negative term is J_4, and J_1 will be estimated using (3). We proceed,

$$e^{k|g+V|^\alpha} - e^{k|X|^\alpha} = e^{k|X|^\alpha}\left[e^{k|g+V|^\alpha - k|X|^\alpha} - 1\right] \le e^{k|X|^\alpha}\left[e^{k\alpha(J_1+J_4+J_3)} - 1\right].$$

For any $\delta > 0$ the following inequality is true, $e^x - 1 \le x + \frac{1+\delta}{2}x^2$ for x small enough. Applying it with $x = k\alpha(J_1 + J_4 + J_3)$ together with $(J_1 + J_4 + J_3)^2 \le (1+\delta)J_1^2 + (2+\frac{2}{\delta})(J_4^2 + J_3^2)$, we get the upper bound for $e^{k|g+V|^\alpha} - e^{k|X|^\alpha}$,

$$k\alpha e^{k|X|^\alpha}\left[J_1 + J_4 + J_3 + \tfrac{1+\delta}{2}k\alpha\left((1+\delta)J_1^2 + \left(2+\tfrac{2}{\delta}\right)(J_3^2 + J_4^2)\right)\right]. \qquad (21)$$

5. Using the identity $\mathbf{1}(|V| \le \varepsilon|X|^{1-\alpha}) = 1 - \mathbf{1}(|V| > \varepsilon|X|^{1-\alpha})$, relation (19) with any $q > 0$ and (18), we obtain,

$$\begin{aligned}
&\mathbb{E}_x e^{k|X|^\alpha} J_4 \mathbf{1}(n<\tau)\mathbf{1}(|V| \le \varepsilon|X|^{1-\alpha})\\
&\qquad = -\tilde r\mathbb{E}_x \frac{e^{k|X|^\alpha}}{|X|^{1+p-\alpha}}\mathbf{1}(n<\tau)\\
&\qquad\quad + \tilde r\mathbb{E}_x \frac{e^{k|X|^\alpha}}{|X|^{1+p-\alpha}}\mathbf{1}(n<\tau)\mathbb{E}\{\mathbf{1}(|V| > \varepsilon|X|^{1-\alpha})|\mathcal{F}_n\}\\
&\qquad = (-\tilde r + o_R(1))\mathbb{E}_x \frac{e^{k|X|^\alpha}}{|X|^{1+p-\alpha}}\mathbf{1}(n<\tau).
\end{aligned} \qquad (22)$$

6. Similarly, using condition (3), (19) with $q = p/(1-\alpha)$ and (18), we have,

$$\begin{aligned}
&\left|\mathbb{E}_x e^{k|X|^\alpha} J_1 \mathbf{1}(n<\tau)\mathbf{1}(|V| \le \varepsilon|X|^{1-\alpha})\right|\\
&\qquad = \left| - \mathbb{E}_x e^{k|X|^\alpha} J_1 \mathbf{1}(n<\tau)\mathbf{1}(|V| > \varepsilon|X|^{1-\alpha})\right|\\
&\qquad \le \mathbb{E}_x e^{k|X|^\alpha}\frac{|V|}{|X|^{1-\alpha}}\cdot\frac{|V|^q}{\varepsilon^q|X|^{q(1-\alpha)}}\mathbf{1}(n<\tau)\mathbf{1}(|V| > \varepsilon|X|^{1-\alpha})\\
&\qquad = o_R(1)\cdot\varepsilon^{-q}\mathbb{E}_x\frac{e^{k|X|^\alpha}}{|X|^{1+p-\alpha}}\mathbf{1}(n<\tau), \quad R\to\infty.
\end{aligned} \qquad (23)$$

7. By definition of r_2, one writes,

$$\begin{aligned}
\mathbb{E}_x e^{k|X|^\alpha} J_1^2 \mathbf{1}(n<\tau) &\le \mathbb{E}_x \frac{e^{k|X|^\alpha}}{|X|^{2-2\alpha}}\mathbf{1}(n<\tau)\mathbb{E}\{|V|^2|\mathcal{F}_n\}\\
&\le \frac{r_2}{R^{1-p-\alpha}}\mathbb{E}_x\frac{e^{k|X|^\alpha}}{|X|^{1+p-\alpha}}\mathbf{1}(n<\tau).
\end{aligned} \qquad (24)$$

Similarly,

$$\mathbb{E}_x e^{k|X|^\alpha} J_3 \mathbf{1}(n<\tau) \le O(R^{-1+p})\mathbb{E}_x\frac{e^{k|X|^\alpha}}{|X|^{1+p-\alpha}}\mathbf{1}(n<\tau), \qquad (25)$$

$$\mathbb{E}_x e^{k|X|^\alpha} J_3^2 \mathbf{1}(n<\tau) \le O(R^{-3+p+\alpha})\mathbb{E}_x\frac{e^{k|X|^\alpha}}{|X|^{1+p-\alpha}}\mathbf{1}(n<\tau), \qquad (26)$$

$$\mathbb{E}_x e^{k|X|^\alpha} J_4^2 \mathbf{1}(n<\tau) \le O(R^{-1-p+\alpha})\mathbb{E}_x\frac{e^{k|X|^\alpha}}{|X|^{1+p-\alpha}}\mathbf{1}(n<\tau), R\to\infty. \qquad (27)$$

8. From upper bound (21) and estimates (22)–(27) we derive that

$$I_2 \le -k\alpha\left(\tilde r - \frac{k\alpha r_2(1+\delta)^2}{2R^{1-p-\alpha}} + o_{\varepsilon,R}(1)\right)\mathbb{E}_x \frac{e^{k|X|^\alpha}}{|X|^{1+p-\alpha}}\mathbf{1}(n<\tau). \tag{28}$$

Estimates (20) and (28) give (12) by taking ε, δ small and R large enough.

Since $\{n<\tau\} \subset \{n-1<\tau\}$ we obtain (13) from (12) by iterations,

$$\mathbb{E}_x e^{k|X_n|^\alpha}\mathbf{1}(n<\tau) \le \mathbb{E}_x e^{k|X_{n-1}|^\alpha}\mathbf{1}(n-1<\tau) \le \cdots \le e^{k|x|^\alpha}.$$

Letting $c = \inf_{x>R} e^{k|x|^\alpha}|x|^{-1-p+\alpha} > 0$, we deduce from (12),

$$r' c\mathbb{E}_x\mathbf{1}(n<\tau) \le \mathbb{E}_x e^{k|X_n|^\alpha}\mathbf{1}(n<\tau) - \mathbb{E}_x e^{k|X_{n+1}|^\alpha}\mathbf{1}(n+1<\tau).$$

Hence, inequality (14) holds with $C = 1/(cr')$, summing over n. □

Remark 1. *Estimate (24) shows that in the case $\alpha > 1-p$ our method does not give bounds better than those for $\alpha = 1-p$. It is likely that for p fixed, the "attraction force" to the origin does not suffice to provide recurrence in terms of sub-exponential moments of order α greater than $1-p$.*

PROOF of *Lemma 2.* Defining $c := k\alpha r' \inf_{x\ge R} e^{kx^\alpha}/x^{1+p-\alpha} > 0$, one writes the following corollary of inequality (12),

$$\frac{c}{2}\mathbb{E}_x\mathbf{1}(n<\tau) \le \mathbb{E}_x(e^{k|X_n|^\alpha} - e^{k|X_{n+1}|^\alpha})\mathbf{1}(n<\tau) - \frac{k\alpha r'}{2}\mathbb{E}_x\frac{e^{k|X_n|^\alpha}\mathbf{1}(n<\tau)}{|X_n|^{1+p-\alpha}}.$$

Multiplying it by e^{n^δ} implies,

$$\begin{aligned}\frac{c}{2}e^{n^\delta}\mathbb{E}_x\mathbf{1}(n<\tau) \le\ & \mathbb{E}_x e^{(n-1)^\delta + k|X_n|^\alpha}\mathbf{1}(n<\tau) - \mathbb{E}_x e^{n^\delta + k|X_{n+1}|^\alpha}\mathbf{1}(n+1<\tau)\\ &+ \mathbb{E}_x(e^{n^\delta} - e^{(n-1)^\delta})e^{k|X_n|^\alpha}\mathbf{1}(n<\tau)\\ &- \frac{k\alpha r'}{2}\mathbb{E}_x\frac{e^{n^\delta + k|X_n|^\alpha}}{|X_n|^{1+p-\alpha}}\mathbf{1}(n<\tau).\end{aligned} \tag{29}$$

Since $e^{n^\delta} - e^{(n-1)^\delta} \sim \delta n^{-1+\delta}e^{n^\delta}$ as $n\to\infty$, we have $e^{n^\delta} - e^{(n-1)^\delta} \le An^{-1+\delta}e^{n^\delta}$, $A>0$, for any $n\ge 1$, hence,

$$\mathbb{E}_x(e^{n^\delta} - e^{(n-1)^\delta})e^{k|X_n|^\alpha}\mathbf{1}(n<\tau) \le A\mathbb{E}_x\frac{e^{n^\delta + k|X_n|^\alpha}}{n^{1-\delta}}\mathbf{1}(n<\tau). \tag{30}$$

Let $0<\varepsilon\le k\alpha r'/(2A)$, $D = \{|X_n|^{1+p-\alpha} < \varepsilon n^{1-\delta}\}$. Using (30), compare the two last terms in (29) on D:

$$\begin{aligned}&A\mathbb{E}_x\frac{e^{n^\delta + k|X_n|^\alpha}}{n^{1-\delta}}\mathbf{1}(n<\tau)\mathbf{1}(D) - \frac{k\alpha r'}{2}\mathbb{E}_x\frac{e^{n^\delta + k|X_n|^\alpha}}{|X_n|^{1+p-\alpha}}\mathbf{1}(n<\tau)\mathbf{1}(D)\\ &\qquad\le \left(-\frac{k\alpha r'}{2} + \varepsilon A\right)\mathbb{E}_x\frac{e^{n^\delta + k|X_n|^\alpha}}{|X_n|^{1+p-\alpha}}\mathbf{1}(n<\tau)\mathbf{1}(D) \le 0.\end{aligned} \tag{31}$$

Now compare the same terms on D^c. Since $\mathbf{1}(a \ge b) \le e^{a^q - b^q}$, $q > 0$, one obtains,

$$\mathbb{E}_x \frac{e^{n^\delta + k|X_n|^\alpha}}{n^{1-\delta}} \mathbf{1}(n<\tau)\mathbf{1}(D^c) \le e^{n^\delta - \varepsilon^q n^{(1-\delta)q}} \mathbb{E}_x e^{k|X_n|^\alpha + |X_n|^{(1+p-\alpha)q}} \mathbf{1}(n<\tau).$$

Choose the value of q such that $(1-\delta)q > \delta$, and $(1+p-\alpha)q < \alpha$, which is always possible when $\delta < \alpha/(1+p)$, then

$$e^{k|X_n|^\alpha + |X_n|^{(1+p-\alpha)q}} \le e^{k'|X_n|^\alpha}, \text{ and } e^{n^\delta - \varepsilon^q n^{(1-\delta)q}} \le B e^{-n^{\delta'}}, \; B > 0,$$

where $k' > k$ but can be taken arbitrarily close to k, if we increase R (recall that $|X_n| > R$ when $n < \tau$), and $\delta' > 0$. Therefore, by *Lemma 1*,

$$\begin{aligned} A\mathbb{E}_x \frac{e^{n^\delta + k|X_n|^\alpha}}{n^{1-\delta}} \mathbf{1}(n < \tau)\mathbf{1}(D^c) - \frac{k\alpha r'}{2} \mathbb{E}_x \frac{e^{n^\delta + k|X_n|^\alpha}}{|X_n|^{1+p-\alpha}} \mathbf{1}(n < \tau)\mathbf{1}(D^c) \\ \le B e^{-n^{\delta'}} \mathbb{E}_x e^{k'|X_n|^\alpha} \mathbf{1}(n < \tau) \le B e^{-n^{\delta'}} e^{k'|x|^\alpha}. \end{aligned} \tag{32}$$

Using inequalities (31) and (32) in (29), summing over all $n \ge 1$ and applying *Lemma 1*, we get with $C > 0$,

$$\frac{c}{2} \sum_{n=1}^{\infty} e^{n^\delta} \mathbb{E}_x \mathbf{1}(n<\tau) \le \mathbb{E}_x e^{k|X_1|^\alpha} \mathbf{1}(n<\tau) + B e^{k'|x|^\alpha} \sum_{n=1}^{\infty} e^{-n^{\delta'}} \le C e^{k'|x|^\alpha},$$

which implies the desired inequality with a new constant $C > 0$, and k' instead of k, because of the bound, $\mathbb{E}_x e^{\tau^\delta} \le 1 + e + e \sum_{n=1}^{\infty} e^{n^\delta} \mathbb{E}_x \mathbf{1}(n < \tau)$. $\square$

Remark 2. *The case $\delta = \alpha/(1-p)$ can be included, too, with the help of more accurate estimations. We do not consider it here.*

PROOF of *Lemma 3* is based on the lines of *Lemma 1*. Let R be so large that *Lemma 1* is applicable with such r' that $r'' < r' < r$. The value of $\tilde{R} > R$ will be chosen later. One writes,

$$\begin{aligned} \mathbf{1}(n < \gamma) = \mathbf{1}(\{n < \gamma\} \cap \{|X_n| > R, |X'_n| > R\}) \\ + \mathbf{1}(\{n < \gamma\} \cap \{|X_n| \le R, |X'_n| > R\}) \\ + \mathbf{1}(\{n < \gamma\} \cap \{|X_n| > R, |X'_n| \le R\}), \end{aligned}$$

and one represents $\mathbb{E}_{x,x'} \left(e^{k|X_{n+1}|^\alpha} + e^{k|X'_{n+1}|^\alpha} \right) \mathbf{1}(n < \gamma)$ as a sum of three terms, I_1, I_2 and I_3, respectively.

Since (X_n) and (X'_n) are independent and inclusions

$$\begin{aligned} \{n < \gamma\} \cap \{|X_n| > R, |X'_n| > R\} \subset \{|X_n| > R\}, \\ \{n < \gamma\} \cap \{|X_n| > R, |X'_n| > R\} \subset \{|X'_n| > R\} \end{aligned}$$

hold true, then the calculus in the proof of *Lemma 1* can be used which provides (12).

One estimates,

$$\begin{aligned}
I_1 &\le \mathbb{E}_{x,x'}\left(e^{k|X_n|^\alpha} + e^{k|X_n'|^\alpha}\right)\mathbf{1}(\{n<\gamma\}\cap\{|X_n|>R, |X_n'|>R|\}) \\
&\quad - r'\mathbb{E}_{x,x'}\left(e^{k|X_n|^\alpha}|X_n|^{-1-p+\alpha} + e^{k|X_n'|^\alpha}|X_n'|^{-1-p+\alpha}\right) \\
&\qquad \times\mathbf{1}(\{n<\gamma\}\cap\{|X_n|>R, |X_n'|>R|\}) \\
&\le \mathbb{E}_{x,x'}\left(e^{k|X_n|^\alpha} + e^{k|X_n'|^\alpha}\right)\mathbf{1}(\{n<\gamma\}\cap\{|X_n|>R, |X_n'|>R|\}) \\
&\quad - r'\mathbb{E}_{x,x'}e^{k(|X_n|\vee|X_n'|)^\alpha}(|X_n|\vee|X_n'|)^{-1-p+\alpha} \\
&\qquad \times\mathbf{1}(\{n<\gamma\}\cap\{|X_n|>R, |X_n'|>R|\})
\end{aligned}$$

with the same r' as in *Lemma 1.*

In the case of I_2 we may write,

$$\begin{aligned}
&\mathbb{E}_{x,x'}e^{k|X_{n+1}|^\alpha}\mathbf{1}(\{n<\gamma\}\cap\{|X_n|\le R, |X_n'|>R\}) \\
&\qquad \le \mathbb{E}_{x,x'}e^{k|X_n|^\alpha+k|V_{n+1}|^\alpha}\mathbf{1}(\{n<\gamma\}\cap\{|X_n|\le R, |X_n'|>R\}) \\
&\qquad \le e^{kR^\alpha}\mathbb{E}_{x,x'}\mathbf{1}(\{n<\gamma\}\cap\{|X_n|\le R, |X_n'|>R\})\mathbb{E}\{e^{k|V_{n+1}|^\alpha}|\mathcal{F}_n\} \\
&\qquad \le C_R\mathbb{E}_{x,x'}\mathbf{1}(\{n<\gamma\}\cap\{|X_n|\le R, |X_n'|>R\}),
\end{aligned}$$

due to (17). Since $|X_n'|>\tilde R$, if $n<\gamma$, let us take $\tilde R$ so large that

$$(r'-r'')e^{k|X_n'|^\alpha}|X_n'|^{-1-p+\alpha} > C_R.$$

Then

$$\begin{aligned}
I_2 &\le \mathbb{E}_{x,x'}e^{k|X_n'|^\alpha}\mathbf{1}(\{n<\gamma\}\cap\{|X_n|\le R, |X_n'|>R\}) \\
&\quad - r''\mathbb{E}_{x,x'}e^{k|X_n'|^\alpha}|X_n'|^{-1-p+\alpha}\mathbf{1}(\{n<\gamma\}\cap\{|X_n|\le R, |X_n'|>R\}) \\
&\le \mathbb{E}_{x,x'}\left(e^{k|X_n|^\alpha} + e^{k|X_n'|^\alpha}\right)\mathbf{1}(\{n<\gamma\}\cap\{|X_n|\le R, |X_n'|>R\}) \\
&\quad - r''\mathbb{E}_{x,x'}e^{k(|X_n|\vee|X_n'|)^\alpha}(|X_n|\vee|X_n'|)^{-1-p+\alpha} \\
&\qquad \times\mathbf{1}(\{n<\gamma\}\cap\{|X_n|\le R, |X_n'|>R\}).
\end{aligned}$$

The calculus for I_3 is similar. Summing up the estimates for I_1, I_2 and I_3, one gets (15). This relation is quite similar to (12), so that the last step, obtaining (16), is the same as the corresponding one in *Lemma 1.* □

PROOF of *Lemma 4* follows the lines of the proof of *Lemma 2* with some modifications. Multiplying (15) by e^{n^δ}, we proceed as in *Lemma 2* and instead of (29) obtain the following inequality,

$$\begin{aligned}
\frac{\tilde c}{2}e^{n^\delta}\mathbb{E}_{x,x'}\mathbf{1}(n<\gamma) &\le \mathbb{E}_{x,x'}e^{(n-1)^\delta}(e^{k|X_n|^\alpha}+e^{k|X_n'|^\alpha})\mathbf{1}(n<\gamma) \\
&\quad - \mathbb{E}_{x,x'}e^{n^\delta}(e^{k|X_{n+1}|^\alpha}+e^{k|X_{n+1}'|^\alpha})\mathbf{1}(n<\gamma) \\
&\quad + \mathbb{E}_{x,x'}(e^{n^\delta}-e^{(n-1)^\delta})(e^{k|X_n|^\alpha}+e^{k|X_n'|^\alpha})\mathbf{1}(n<\gamma) \\
&\quad - \frac{k\alpha r''}{2}\mathbb{E}_{x,x'}\frac{e^{k(|X_n|\vee|X_n'|)^\alpha}}{(|X_n|\vee|X_n'|)^{1+p-\alpha}}\mathbf{1}(n<\gamma),
\end{aligned}$$

where $\tilde c := k\alpha r'' \inf_{x\ge \tilde R} e^{kx^\alpha}/x^{1+p-\alpha} > 0$. Further, instead of (30) we get,

$$\mathbb{E}_{x,x'}(e^{n^\delta} - e^{(n-1)^\delta})(e^{k|X_n|^\alpha} + e^{k|X'_n|^\alpha})\mathbf{1}(n<\gamma) \\ \le 2A\mathbb{E}_{x,x'}\frac{e^{n^\delta + k(|X_n|\vee|X'_n|)^\alpha}}{n^{1-\delta}}\mathbf{1}(n<\gamma).$$

Let $0<\varepsilon\le k\alpha r''/(4A)$ and $\tilde D = \{(|X_n|\vee|X'_n|)^{1+p-\alpha} < \varepsilon n^{1-\delta}\}$, and proceed as in *Lemma 2* with obvious changes like writing $|X_n|\vee|X'_n|$ instead of $|X_n|$ and applying *Lemma 3* instead of *Lemma 1*, to obtain,

$$\frac{\tilde c}{2}\sum_{n=1}^\infty e^{n^\delta}\,\mathbb{E}_{x,x'}\mathbf{1}(n<\gamma) \le \tilde C(e^{k'|x|^\alpha} + e^{k'|x'|^\alpha}),\ \tilde C>0,$$

which implies the estimate required in the Lemma. □

Remark 3. *Analyzing the proofs of* Lemmas 1–4*, we see that all the lines can be rewritten in a slightly more general case, which indeed will be useful in the sequel. E.g., let τ_0 and τ be two stopping-times, $0\le\tau_0<\tau$ and $\{\tau_0\le n<\tau\}\subset\{|X_n|>R\}$ for every n. Then the conclusion of* Lemma 1 *would be*

$$\mathbb{E}_x e^{k|X_n|^\alpha}\mathbf{1}(\tau_0\le n<\tau) \le \mathbb{E}_x e^{k|X_{\tau_0}|^\alpha}\mathbf{1}(\tau_0\le n).$$

Moreover, if $\mathcal{L}(X_0)=\mu$, then one may write a similar inequality with $\mathbb{E}_\mu$ instead of $\mathbb{E}_x$, having integrated the last inequality with respect to the measure μ. Integration is possible once μ has finite sub-exponential moments of the same order as in (5).

PROOF of *Lemma 5*. Consider the process on $B=\{x\in\mathbb{R}^d : |x|\le R\}$. Because of condition (8) this process has an invariant measure μ_∞^B. Let $\tau=\inf\{n\ge 1 : X_n\in B\}$. Due to the Harris representation for the invariant measure μ_∞ via μ_∞^B, for any non-negative function h,

$$\int_{\mathbb{R}^d} h(x)\,\mu_\infty(dx) = \frac{1}{c(B)}\int_B \mu_\infty^B(dx)\mathbb{E}_x\sum_{n=0}^{\tau-1}h(X_n)$$

with $c(B)=\displaystyle\int_B \mu_\infty^B(dx)\mathbb{E}_x\tau$ (equivalent to Proposition 10.4.8 and Theorem 10.4.9 in [34]). In our case $h(x)=e^{k|x|^\alpha}$. Since $e^{k|x|^\alpha}\le e^{(k+\varepsilon)|x|^\alpha}|x|^{-1-p+\alpha}$ for $|x|>R$ and R large, we get by *Lemma 1* with $k+\varepsilon$ instead of k, where $\varepsilon<(K-k)/2$,

$$\mathbb{E}_x\sum_{n=0}^{\tau-1}h(X_n)\le e^{kR^\alpha}+\sum_{n=1}^\infty\mathbb{E}_x\frac{e^{(k+\varepsilon)|X_n|^\alpha}}{|X_n|^{1+p-\alpha}}\mathbf{1}(n<\tau)\le e^{kR^\alpha}+\frac{e^{(k+\varepsilon)|x|^\alpha}}{k\alpha r'}. \qquad \square$$

Remark 4. *In the following proofs we will refer to* the coupling method. *See [35, 43, 44, 45] for details.*

Consider the pair of independent copies of the Markov process (X_n, X'_n) on an appropriate probability space, with initial data $\mathcal{L}(X_0)=\mu$ and $\mathcal{L}(X'_0)=\mu'$. Let m be an arbitrary non-negative integer, and moments $\gamma_1<\gamma_2<\dots$ defined by the formulas

$$\gamma_1 = \inf\{n\ge m : |X_n|\vee|X'_n|\le\tilde R\},$$
$$\gamma_{j+1} = \inf\{n>\gamma_j : |X_n|\vee|X'_n|\le\tilde R\},$$

where $\tilde R$ is chosen in Lemma 3.

If the process (X_n, X'_n) is positive recurrent with respect to the set $\{(x,y) : |x| \vee |y| \le \tilde R\}$, then (see e.g. [43, 45]) one can extend the probability space (without changing notations for probability and expectation) and define a random variable $L_m \ge m$ on this extension, such that, uniformly with respect to μ, μ',

$$\mathbb{E}_{\mu,\mu'} \operatorname*{var}_{B\in\mathcal{F}^X_{\ge n+m}} (\mathbb{P}_\mu(B|\mathcal{F}_m) - \mathbb{P}_{\mu'}(B)) \le \mathbb{P}_{\mu,\mu'}(L_m > n+m), \tag{33}$$

and with the same $\kappa = \kappa(\tilde R)$ as in (8),

$$\sup_m \mathbb{P}_{\mu,\mu'}(L_m > \gamma_j) \le (1-\kappa)^{j-1}, \qquad j \ge 1. \tag{34}$$

PROOF of the inequality (9) in *Theorem 1* is similar to the proof of Theorem 1 in [44]. We will use the coupling method with μ concentrated in $x \in \mathbb{R}^d$, $\mu' = \mu_\infty$, and $m = 0$, accordingly to *Remark 4*, to establish (9).

Due to *Lemmas 4* and *5*,

$$\mathbb{E}_{\mu,\mu'} e^{\gamma_1^\delta} \le C(x) + C(\mu_\infty) < \infty, \qquad C(x) = Ce^{k|x|^\alpha}, \quad k < K.$$

Further,

$$\begin{aligned}\mathbb{E}_{\mu,\mu'} e^{(\gamma_{j+1}-\gamma_j)^\delta} &\le e\mathbb{E}_{\mu,\mu'}\mathbf{1}(\gamma_{j+1} = \gamma_j + 1) \\ &\quad + e\mathbb{E}_{\mu,\mu'} e^{(\gamma_{j+1}-(\gamma_j+1))^\delta}\mathbf{1}(\gamma_{j+1} > \gamma_j + 1).\end{aligned}$$

Since $|X_{\gamma_j}| \vee |X'_{\gamma_j}| \le \tilde R$, and due to (4)–(5), and *Lemma 4*, one can choose $M, C > 0$ such that uniformly with respect to μ, μ'

$$\mathbb{E}_{\mu,\mu'}\left(e^{k|X_{\gamma_j+1}|^\alpha} + e^{k|X'_{\gamma_j+1}|^\alpha}\right) \le M, \quad \text{and} \quad \mathbb{E}_{\mu,\mu'} e^{(\gamma_{j+1}-\gamma_j)^\delta} \le C.$$

By induction and strong Markov property,

$$\mathbb{E}_{\mu,\mu'} e^{\gamma_{j+1}^\delta} \le \mathbb{E}_{\mu,\mu'} e^{(\gamma_{j+1}-\gamma_j)^\delta} e^{(\gamma_j-\gamma_{j-1})^\delta} \dots e^{\gamma_1^\delta} \le (C(x) + C(\mu_\infty))C^j,$$

so due to Bienaimé-Chebyshev's inequality

$$\mathbb{P}_{\mu,\mu'}(\gamma_{j+1} > n) \le e^{-n^\delta}\mathbb{E}_{\mu,\mu'} e^{\gamma_{j+1}^\delta} \le (C(x) + C(\mu_\infty))C^j e^{-n^\delta}. \tag{35}$$

Hölder's inequality with $1/a + 1/b = 1$, $a > 1$, $b > 1$, (35) and (34) imply

$$\begin{aligned}\operatorname{var}(\mu_{n,x} - \mu_\infty) &\le \mathbb{P}_{x,\mu_\infty}(L_0 > n) \\ &= \sum_{j=1}^\infty \mathbb{E}_{x,\mu_\infty}\mathbf{1}(L_0 > n)\mathbf{1}(\gamma_j \le n < \gamma_{j+1}) \\ &\le \sum_{j=1}^\infty \mathbb{P}^{1/a}_{x,\mu_\infty}(L_0 > \gamma_j)\,\mathbb{P}^{1/b}_{x,\mu_\infty}(\gamma_{j+1} > n) \\ &\le C(x,\mu_\infty)e^{-n^\delta/b}\sum_{j=1}^\infty (1-\kappa)^{j/a}C^{j/b},\end{aligned}$$

with $C(x,\mu_\infty) \le Ce^{k|x|^\alpha}$. Fix large b, such that $(1-\kappa)^{1/a}C^{1/b} < 1$. Then

$$\mathrm{var}(\mu_{n,x} - \mu_\infty) \le C(\kappa, b)e^{k|x|^\alpha}\, e^{-n^\delta/b}, \quad k < K.$$

□

PROOF of *Lemma 6.* We apply the coupling method similarly to the first part of the proof of *Theorem 1.* One writes,

$$\mathbb{E}_x e^{k|X_n|^\alpha} \le \mathbb{E}_{x,\mu_\infty} \left| e^{k|X_n|^\alpha} - e^{k|X'_n|^\alpha} \right| + \mathbb{E}_{\mu_\infty} e^{k|X'_n|^\alpha}.$$

The first term on the right-hand side which is obviously equal to

$$\mathbb{E}_{x,\mu_\infty} \left| e^{k|X_n|^\alpha} - e^{k|X'_n|^\alpha} \right| \mathbf{1}(n < L_0) \le \mathbb{E}_{x,\mu_\infty} e^{k|X_n|^\alpha} \mathbf{1}(n < L_0) + \mathbb{E}_{\mu_\infty} e^{k|X'_n|^\alpha}.$$

Since $\mathcal{L}(X'_n) = \mu_\infty$, then $\mathbb{E}_{\mu_\infty} e^{k|X'_n|^\alpha} < \infty$ due to *Lemma 5.*

Take $a > 1$ to provide $ka < K$. By virtue of *Lemma 3*,

$$\mathbb{E}_{x,\mu_\infty} e^{ka|X_n|^\alpha} \mathbf{1}(\gamma_i \le n < \gamma_{i+1}) \le M = M(x,\mu_\infty),$$

and by Hölder's inequality with $1/a + 1/b = 1$, $a > 1$, $b > 1$,

$$\begin{aligned}
\mathbb{E}_{x,\mu_\infty} e^{k|X_n|^\alpha} \mathbf{1}(n < L_0) &\le \sum_{i=1}^{\infty} \mathbb{E}_{x,\mu_\infty} e^{k|X_n|^\alpha} \mathbf{1}(\gamma_i \le n < \gamma_{i+1}) \mathbf{1}(\gamma_i < L_0) \\
&\le \sum_{i=1}^{\infty} \mathbb{E}^{1/a}_{x,\mu_\infty} e^{ka|X_n|^\alpha} \mathbf{1}(\gamma_i \le n < \gamma_{i+1}) \mathbb{P}^{1/b}_{x,\mu_\infty}(L_0 > \gamma_i) \\
&\le \sum_{i=1}^{\infty} M^{1/a} (1-\kappa)^{(i-1)/b} < \infty.
\end{aligned}$$

□

PROOF of inequality (10) in *Theorem 1.* To prove the rate of β-mixing, we repeat the first part of the proof with an arbitrary $m \ge 0$. The only new feature is the estimate for $\sup_m \mathbb{E}_{\mu,\mu'} e^{(\gamma_1 - m)^\delta}$ (recall that γ_1 depends on m). *Lemma 6* along with *Lemma 4* show that this value is finite. Due to (33),

$$\beta_{n,x} \le \sup_m \mathbb{P}_{x,x}(L_m > n + m).$$

The latter probability is estimated using (34), similarly to the proof of the bound (9). Thus, (10) holds true. Finally, (11) follows from (10) after integration, due to *Lemma 5.* □

Remark 5. *Detailed inspection of the proofs in this article shows us that the results of [44] are valid under weaker restrictions on the sequence of noise. Namely, instead of i.i.d. random variables one can take a sequence satisfying (4) with finite moments of the order required in lemmas and theorems in [44]. Notice that in [44, 43, 39] process (X_n) was constructed directly by defining the function $g(x)$ and by introducing the i.i.d. sequence (V_n).*

References

[1] P. ANGO NZÉ, *Critères d'ergodicité de quelques modèles à représentation markovienne,* (French) *[Criteria of ergodicity for some models with a Markovian representation]* C. R. Acad. Sci. Paris Sér. I Math. **315**(1992), 1301–1304.

[2] S. ASMUSSEN, *Applied probability and queues,* John Wiley & Sons, Ltd., Chichester, 1987.

[3] S. ASPANDIIAROV, R. IASNOGORODSKI, M. MENSHIKOV, *Passage-time moments for nonnegative stochastic processes and application to reflected random walks in a quadrant,* Ann. Probab. **24**(1996), 932–960.

[4] P. BERTAIL, A. GAMST, D. N. POLITIS, *Moderate deviations in subsampling distribution estimation,* Proc. Amer. Math. Soc. **129**(2001), 551–557.

[5] A. A. BOROVKOV, *Ergodicity and stability of stochastic processes,* John Wiley & Sons, Ltd., Chichester, 1998.

[6] R. C.BRADLEY, *Basic properties of strong mixing conditions*, Prog. Probab. Stat. **11**(1986), 165-192; also published in: Proceedings of the Conference *Dependence in probability and statistics* (E.Eberlein, M.S.Taqqu , Eds.) 1985), Birkhäuser Boston, Boston, MA, 1986, 165–192.

[7] R. C. BRADLEY, *Counterexamples to the central limit theorem under strong mixing conditions. II,* Colloq. Math. Soc. János Bolyai **57**(1990), 59-67.

[8] JU. A. DAVYDOV, *Mixing conditions for Markov chains* (in Russian), Teor. Verojatnost. i Primenen. **18**(1973), 321–338.

[9] H. DJELLOUT, A. GUILLIN, *Moderate deviations for Markov chains with atom,* Stochastic Process. Appl. **95**(2001), 203–217.

[10] J. L. DOOB, *Stochastic processes,* John Wiley & Sons, Inc., New York; Chapman & Hall, Limited, London, 1953.

[11] G. FORT, E. MOULINES, *V-subgeometric ergodicity for a Hastings-Metropolis algorithm,* Statist. Probab. Lett. **49**(2000), 401–410.

[12] G. FORT, E. MOULINES, *Computable bounds for polynomial ergodicity,* Stoch. Proc. Appl., to be published

[13] M. I. FREIDLIN, A. D. WENTZELL, *Random perturbations of dynamical systems*, Springer-Verlag, New York, 1998.

[14] H. GANIDIS, B. ROYNETTE, F. SIMONOT, *Convergence rate of some semigroups to their invariant probability,* Stoch. Proc. Appl. **79**(1999), 243–263.

[15] O. V. GULINSKY, A. YU. VERETENNIKOV, *Large deviations for discrete-time processes with averaging,* VSP, Utrecht, 1993.

[16] A. Guillin, *Moderate deviations of inhomogeneous functionals of Markov processes and application to averaging,* Stoch. Proc. Appl. **92**(2001), 287–313.

[17] R. Z. Has'minskii, *Stochastic stability of differential equations,* Sijthoff & Noordhoff, Alphen aan den Rijn–Germantown, Md., 1980.

[18] I. A. Ibragimov, *On stationary Gaussian processes with a strong mixing property* (in Russian), Dokl. Akad. Nauk SSSR **147**(1962), 1282–1284.

[19] I. A. Ibragimov, *Complete regularity of multidimensional stationary processes with discrete time* (in Russian), Sov. Math. Dokl. **6**(1965), 783–785; English translation: Dokl. Akad. Nauk SSSR **162**(1965), 983-985.

[20] I. A. Ibragimov, Ju. V. Linnik, *Nezavisimye i statsionarno svyazannye velichiny* (in Russian), Izdat. "Nauka", Moscow, 1965. English translation: *Independent and stationary sequences of random variables*, Wolters-Noordhoff Publishing, Groningen, 1971.

[21] I. A. Ibragimov, Y. A. Rozanov, Y *Gaussian random processes,* Springer-Verlag, New York-Berlin, 1978.

[22] S. Jarner, G. O. Roberts, *Polynomial convergence rates of Markov chains,* Ann. Appl. Probab. **12**(2002), 224–247.

[23] S. Jarner, G. O. Roberts, *Convergence of heavy tailed MCMC algorithms,* submitted.

[24] S. Jarner, R. L. Tweedie, *Convergence rate and moments of Markov chain associated with the mean of Dirichlet processes,* Stoch. Proc. Appl. **101**(2002), 257–271.

[25] V. V. Kalashnikov, *The property of γ-reflexivity for Markovian sequences*, Sov. Math. Dokl. **14**(1973), 1869-1873; English translation: Dokl. Akad. Nauk SSSR **213**(1973), 1243–1246.

[26] V. V. Kalashnikov, *Qualitative analysis of the behavior of complex systems by the method of test functions* (in Russian), Moskva: Nauka, 1978.

[27] W. C. M. Kallenberg, *On moderate deviation theory in estimation,* Ann. Statist. **11**(1983), 498–504.

[28] N. V. Kartashov, *Strong stable Markov chains,* VSP, Utrecht; TBiMC Scientific Publishers, Kiev, 1996.

[29] S. A. Klokov, *On lower bounds for mixing rates for a class of Markov processes*, submitted.

[30] J. Lamperti, *Criteria for stochastic processes. II. Passage-time moments,* J. Math. Anal. Appl. **7**(1963), 127–145.

[31] A. M. Lyapunov, *The general problem of the stability of motion*, (Translated by A. T. Fuller from Edouard Davaux's French translation (1907) of the 1892 Russian original.) Internat. J. Control **55**(1992), 521–790.

[32] M. N. MALYSHKIN, *Subexponential estimates of the convergence rate to the invariant measure for stochastic differential equations* (in Russian), Teor. Veroyatnost. i Primenen. **45**(2000), 489–504.

[33] M. MENSHIKOV, R. J. WILLIAMS, *Passage-time moments for continuous non-negative stochastic processes and applications,* Adv. in Appl. Probab. **28**(1996), 747–762.

[34] S. P. MEYN, R. L. TWEEDIE, *Markov Chains and Stochastic Stability*, Springer-Verlag, Berlin, 1993.

[35] E. NUMMELIN, *General irreducible Markov chains and non-negative operators,* Cambridge University Press, Cambridge, 1984.

[36] R. B. SCHINAZI, *Classical and spatial stochastic processes,* Birkäuser Boston, Inc., Boston, MA, 1999.

[37] D. S. SILVESTROV, *Recurrence relations for generalized hitting times for semi-Markov processes*, Ann. Appl. Probab. **6**(1996), 617–649.

[38] P. TUOMINEN, R. TWEEDIE, *Subgeometric rates of convergence of f-ergodic Markov chains*, Adv. in Appl. Probab. **26**(1994), 775–798.

[39] A. YU. VERETENNIKOV, *Estimates of the mixing rate for stochastic equations* (in Russian), Teor. Veroyatnost. i Primenen. **32**(1987), 299–308.

[40] A. YU. VERETENNIKOV, *The mixing rate and the averaging principle for hypoelliptic stochastic differential equations* (in Russian), Izv. Akad. Nauk SSSR Ser. Mat. **52**(1988), 899–908; English translation: Math. USSR-Izv. **33**(1988), 221–231.

[41] A. YU. VERETENNIKOV, *Estimates for the mixing rate for Markov processes* (in Russian), Litovsk. Mat. Sb. **31**(1991), 40–49; English translation: Lithuanian Math. J. **31**(1991), 27–34.

[42] A. YU. VERETENNIKOV, *On large deviations for additive functionals of Markov processes. I*, Theory Probab. Appl. **38**(1993), 706-719; translation from Teor. Veroyatn. Primen. **38**(1993), 758-774.

[43] A. YU. VERETENNIKOV, *On polynomial mixing bounds for stochastic differential equations*, Stoch. Proc. Appl. **70**(1997), 115–127.

[44] A. YU. VERETENNIKOV, *On polynomial mixing and the rate of convergence for stochastic differential and difference equations* (in Russian), Teor. Veroyatnost. i Primenen. **44**(1999), 312–327; English translation: Theory Probab. Appl. **44**(2000), 361–374.

[45] A. YU. VERETENNIKOV, *Coupling method for Markov chains under integral Doeblin type condition*, Theory of Stochastic Processes, **8**(**24**)(2002), 384–391.

[46] V. A. VOLKONSKII, YU. A. ROZANOV, *Some limit theorems for random functions I* (in Russian), Teor. Veroyatnost. i Primenen **4**(1959), 186–207; English translation: Theor. Probability Appl. **4**(1959), 178–197.