Ergodic theorems for queuing systems with dependent inter-arrival times

Attila Lovas\textsuperscript{a} and Miklós Rásonyi\textsuperscript{b}

\textsuperscript{a}Corresponding author. Alfréd Rényi Institute of Mathematics, Reáltanoda utca 13-15, 1053 Budapest, Hungary; and Technical University, Budapest. E-mail: lovas@renyi.hu

\textsuperscript{b}Alfréd Rényi Institute of Mathematics, Reáltanoda utca 13-15, 1053 Budapest, Hungary. E-mail: rasonyi@renyi.hu

April 6, 2020

Abstract

We show that, under appropriate conditions, the waiting time in a queuing system converges to a stationary distribution as time tends to infinity, even in the case where inter-arrival times are dependent random variables. A convergence rate is given and a law of large numbers is established for functionals of the waiting time. These results provide tools for the statistical analysis of such systems, transcending the standard case where inter-arrival times are assumed independent.

Keywords: Queuing; Dependent random variables; Inter-arrival times; Ergodic theorem

1 Introduction

Let $\mathbb{R}_+ := \{x \in \mathbb{R} : x \geq 0\}$. We consider a strongly stationary sequence $(S_n, Z_n)_{n \in \mathbb{N}} \in \mathbb{R}_+^2$. A single-server queuing model will be considered where customers are numbered by $n \in \mathbb{N}$. The time between the arrival of customers $n+1$
and \( n \) is described by the random variable \( Z_{n+1} \), for each \( n \in \mathbb{N} \). The service time for customer \( n \) is given by the random variable \( S_n \), for \( n \in \mathbb{N} \).

The waiting time \( W_n \) of customer \( n \) satisfies the Lindley recursion

\[
W_{n+1} = (W_n + S_n - Z_{n+1})_+, \quad n \in \mathbb{N},
\]

where, for simplicity, we assume \( W_0 := 0 \) (we start with an empty queue).

The ergodic theory of general state space Markov chains (see e.g. [8]) allows to treat the case where \( (S_n, Z_n)_{n \in \mathbb{N}} \) are i.i.d. The independence assumption, however, seems too strong for applications and we wish to provide theoretical foundations for the statistical analysis of queuing systems also in cases where one of the two sequences is merely stationary.

The mathematics for such a setting is an order of magnitude more difficult as \( (W_n)_{n \in \mathbb{N}} \) fails to be a Markovian process. We will rely on the recent advances made by [7] in the theory of Markov chains in random environments. The case where both sequences are only stationary cannot be treated within this framework and requires further study. More complex (e.g. multiserver) queuing systems could be analysed along similar lines but we do not pursue such ramifications here.

As far as we know, ergodicity results in the general, stationary setting can only be found in [2] (in Russian), see also Example 14.1 on page 189 of [1]. Actually, Law\( (W_n) \) is known to converge to a limiting distribution under rather mild conditions. However, no convergence rate or law of large numbers is provided. Also, the approach of [2] works only for unbounded \( Z_0 \) while we are able to treat the bounded case as well.

Throughout this paper we will be working on a probability space \((\Omega, \mathcal{F}, P)\). For a Polish space \( \mathcal{X} \), we denote by \( B(\mathcal{X}) \) its Borel sigma-algebra. We denote by \( \mathbb{E}[Z] \) the expectation of a real-valued random variable \( Z \). For a \( \mathcal{X} \)-valued random variable \( X \) we will denote by \( \mathcal{L}(X) \) its law on on \( B(\mathcal{X}) \). The set of probability measures on \( B(\mathcal{X}) \) is denoted by \( \mathcal{M}_1(\mathcal{X}) \). The total variation metric on \( \mathcal{M}_1(\mathcal{X}) \) is defined by

\[
d_{TV}(\mu_1, \mu_2) = |\mu_1 - \mu_2|(\mathcal{X}), \quad \mu_1, \mu_2 \in \mathcal{M}_1(\mathcal{X}),
\]

where \( |\mu_1 - \mu_2| \) denotes the total variation of the signed measure \( \mu_1 - \mu_2 \). We do not indicate the dependence of the metric \( d_{TV} \) on \( \mathcal{X} \) since the latter will always be clear from the context.

We now present our standing assumptions. In a stable system service times should be shorter on average than inter-arrival times. In our approach we also
need that the service time sequence is independent of inter-arrival times. So we formulate the following hypothesis.

**Assumption 1.1.** We stipulate that $\mathbb{E}[S_0] < \mathbb{E}[Z_0]$ (where the latter may be infinity). The sequences $(S_n)_{n \in \mathbb{N}}$ and $(Z_n)_{n \in \mathbb{N}}$ are independent.

**Definition 1.2.** We say that a sequence $(Y_n)_{n \in \mathbb{N}}$ of real-valued random variables satisfies a Gärtner-Ellis type condition if there is $\eta > 0$ such that the limit

$$
\Gamma(\alpha) := \lim_{n \to \infty} \frac{1}{n} \ln \mathbb{E} \left[ e^{\alpha(Y_1 + \ldots + Y_n)} \right] 
$$

exists for all $\alpha \in (-\eta, \eta)$ and $\Gamma$ is differentiable on $(-\eta, \eta)$.

**Remark 1.3.** The notion above is inspired by the Gärtner-Ellis theorem, see [4], and it holds in a large class of models, well beyond the i.i.d. case. For instance, let $Y_n = \phi(H_n)$ for a measurable $\phi : \mathbb{R}^m \to \mathbb{R}$ satisfying a suitable growth condition and let $(H_n)_{n \in \mathbb{N}}$ be an $\mathbb{R}^m$-valued sufficiently regular Markov chain started from its invariant distribution. Then (2) holds true for all $\eta > 0$, see Theorem 3.1 of [6] for a precise formulation.

We mention another example: let $Y_n = \sum_{i=\infty}^{\infty} a_i \xi_i$, where $\xi_i$, $i \in \mathbb{Z}$ are independent and identically distributed $\mathbb{R}$-valued random variables with finite exponential moments of all orders and $\sum_{|i|=\infty}^{\infty} |a_i| < \infty$. Then (2) is satisfied for this process by Theorem 2.1 of [3].

We now present a result on the ergodic behaviour of queuing systems with dependent service time which was obtained in Section 3 of [7].

**Theorem 1.4.** Let $(Z_n)_{n \in \mathbb{N}}$ be an i.i.d. sequence and let $(S_n)_{n \in \mathbb{N}}$ be uniformly bounded and ergodic, satisfying a Gärtner-Ellis type condition. Let us assume that $P(Z_0 > z) > 0$ for all $z > 0$. Then there exists a probability $\mathbb{P}$ on $\mathcal{B}(\mathbb{R}_+)$ such that

$$
d_{TV}(\mathcal{L}(W_0), \mu) \leq c_1 \exp \left( -c_2 n^{1/3} \right),
$$

for some $c_1, c_2 > 0$. Furthermore, for an arbitrary measurable and bounded $\Phi : \mathbb{R}_+ \to \mathbb{R}$,

$$
\frac{\Phi(W_1) + \ldots + \Phi(W_n)}{n} \to \int_{\mathbb{R}_+} \Phi(z) \mu(\mathrm{d}z),
$$

in probability.

In the present article we concentrate on the (arguably) more interesting case where service times are independent but inter-arrival times may well be dependent.
**Theorem 1.5.** Let \((S_n)_{n\in\mathbb{N}}\) be an i.i.d. sequence and let \((Z_n)_{n\in\mathbb{N}}\) be bounded and ergodic, satisfying a Gärtner-Ellis type condition. Let us assume that \(\mathbb{E}[e^{\beta_0 S_0}] < \infty\) for some \(\beta_0 > 0\) and \(\mathcal{L}(S_0)\) has a density \(s \rightarrow f(s)\) (w.r.t. the Lebesgue measure) which is bounded away from 0 on compact subsets of \(\mathbb{R}_+\). Then the conclusions of Theorem 1.4 hold.

**Remark 1.6.** The mathematical setting of Theorem 1.5 is significantly more involved than that of Theorem 1.4. In Theorem 1.4, one may profit from the fact that, freezing the values of the process \((S_n)_{n\in\mathbb{N}}\), the waiting time becomes an inhomogeneous Markov chain with a particular state (the point 0) which is a reachable atom. In the proof of Theorem 1.5 one needs to guarantee ergodicity using a deeper coupling construction.

We can somewhat relax the boundedness condition on \(Z_0\) in Theorem 1.5 above at the price of more stringent assumptions on \(S_0\). Namely, we assume an exponential-like tail for \(S_0\) and for \(Z_0\) a very light tail, like that of the Gumbel distribution at \(-\infty\). It will become clear from the proof that requiring a thinner tail for \(S_0\) (e.g. Gaussian) would necessitate even more stringent tail assumptions for \(Z_0\), hence we do not strive for further generality here.

**Theorem 1.7.** Let \((S_n)_{n\in\mathbb{N}}\) be an i.i.d. sequence and let \((Z_n)_{n\in\mathbb{N}}\) be ergodic, satisfying a Gärtner-Ellis type condition and \(P(Z_0 \geq z) \leq C_1 \exp(-C_2 e^{C_3 z})\) with some \(C_1, C_2, C_3 > 0\). Let us assume that \(\mathbb{E}[e^{\beta_0 S_0}] < \infty\) for some \(\beta_0 > 0\) and the law of \(S_0\) has a nonincreasing density \(s \rightarrow f(s)\) such that \(f(s) \geq C_4 e^{-C_5 s}, s \geq 0\) for some \(C_4, C_5 > 0\). Then there exists a probability \(\mu\), on \(B(\mathbb{R}_+)\) such that

\[
d_{TV}(\mathcal{L}(W_n), \mu) \leq c_1 \exp(-c_2 n^{c_3}),
\]

holds for some \(c_1, c_2, c_3 > 0\). Furthermore, for an arbitrary measurable and bounded \(\Phi : \mathbb{R}_+ \rightarrow \mathbb{R}\), \((7)\) holds in probability.

Theorems 1.5 and 1.7 open the door for the statistical analysis of such queuing systems. For example, choosing \(\Phi(\omega) := 1_{\{\omega \geq \omega_0\}}\), \(\omega \in \mathbb{R}_+\) for some \(\omega_0 > 0\), the above results guarantee that we can consistently estimate the probability of the waiting time exceeding \(\omega_0\) in the stationary state by \((\Phi(W_1) + \ldots + \Phi(W_n))/n\) for large \(n\).

In Section 2 we recall the notion of Markov chains in random environments and certain results of [7]. Sections 3 and 4 contain the proofs of our two new theorems.
Remark 1.8. Our results are also closely related to those of [5]. If $S_n, n \in \mathbb{N}$ is a bounded sequence of i.i.d. random variables and $Z_n, n \in \mathbb{N}$ is bounded with a certain strong mixing property (see (10) and (11) in [5]) then Theorem 2 of [5] shows that $W_n, n \in \mathbb{N}$ satisfies a mixing property implying, in particular, a law of large numbers for averages of $\Phi(W_n)$, for Lipschitz functionals $\Phi$. Our results thus complement those of [5]: we establish the existence of a limiting law and even allow certain cases where the $Z_n$ are unbounded.

2 Markov chains in random environments

Let $\mathcal{Y}, \mathcal{X}$ be two Polish spaces and $Y : \mathbb{Z} \times \Omega \to \mathcal{Y}$ a strongly stationary $\mathcal{Y}$-valued stochastic process. Let $Q : \mathcal{Y} \times \mathcal{X} \times B(\mathcal{X}) \to [0, 1]$ be a mapping such that for all $B \in B(\mathcal{X})$ the function $(y, x) \mapsto Q(y, x, B)$ is $B(\mathcal{Y}) \otimes B(\mathcal{X})$-measurable and for all $(y, x) \in \mathcal{Y} \times \mathcal{X}, B \mapsto Q(y, x, B)$ is a probability measure on $B(\mathcal{X})$.

We will consider $\mathcal{X}$-valued process $X_t, t \in \mathbb{N}$ such that $X_0 = x_0 \in \mathcal{X}$ is fixed and

$$P(X_{t+1} \in B \mid \sigma(X_s, 0 \leq s \leq t; Y_s, s \in \mathbb{Z})) = Q(Y_t, X_t, B) \quad P - \text{a.s., } t \in \mathbb{N}. \quad (4)$$

We interpret the process $X$ as a Markov chain in a random environment described by the process $Y$.

Definition 2.1. Let $R : \mathcal{X} \times B(\mathcal{X}) \to [0, 1]$ be a probabilistic kernel. For a measurable function $\phi : \mathcal{X} \to \mathbb{R}_+$, we define

$$[R\phi](x) = \int_{\mathcal{X}} \phi(z)R(x, dz), x \in \mathcal{X}.$$

Consistently with Definition 2.1, for $y \in \mathcal{Y}, Q(y)\phi$ will refer to the action of the kernel $Q(y, \cdot)$ on $\phi$. First, a Foster-Lyapunov type drift condition is formulated.

Assumption 2.2. Let $V : \mathcal{X} \to \mathbb{R}_+$ be a measurable function. We consider measurable functions $K, \gamma : \mathcal{Y} \to \mathbb{R}_+$ with $K(\cdot) \geq 1$. We assume that, for all $x \in \mathcal{X}$ and $y \in \mathcal{Y},$

$$[Q(y)V](x) \leq \gamma(y)V(x) + K(y).$$

Here $\gamma(y) \geq 1$ may well occur, but in the next assumption we require that the system dynamics, on long-time average, is contracting.

Assumption 2.3. We assume that

$$\bar{\gamma} := \limsup_{n \to \infty} \mathbb{E}^{1/n} \left( K(Y_0) \prod_{k=1}^n \gamma(Y_k) \right) < 1.$$
We stipulate the existence of suitable “small sets”, which are familiar notions in Markov chain theory, see [8].

**Assumption 2.4.** Let \( \gamma(\cdot), K(\cdot) \) be as in Assumption 2.2. We assume that for some \( 0 < \varepsilon < 1/5^{1/2} - 1 \), there is a measurable function \( \alpha : \mathcal{Y} \to (0, 1) \) and a probability kernel \( \kappa : \mathcal{Y} \times \mathcal{B} \to [0, 1] \) such that, for all \( y \in \mathcal{Y} \) and \( A \in \mathcal{B} \),

\[
\inf_{x \in V^{-1}([0, R(y)])} Q(y, x, A) \geq (1 - \alpha(y))\kappa(y, A), \quad \text{where } R(y) = \frac{2K(y)}{\gamma(y)}
\]

and \( V^{-1}([0, R(y)]) \neq \emptyset \).

We also need to control the probability of \( \alpha(Y_0) \) approaching 1. For the purposes of the present paper it is enough to require a simplification of Assumption 2.2 of [7].

**Assumption 2.5.**

\[
\lim_{n \to \infty} E_{1/n^{\theta}} [\alpha(Y_0)^n] = 0
\]

holds for some \( 0 < \theta < 1 \).

We now recall results of [7]: with the above presented assumptions, the law of \( X_n \) converges to a limiting law as \( n \to \infty \), moreover, bounded functionals of the process \( X \) show ergodic behavior provided that \( Y \) is ergodic.

**Theorem 2.6.** Under Assumptions 2.2, 2.3, 2.4 and 2.5, there exists a probability law \( \mu \), such that

\[
d_{TV}(\mathcal{L}(X_n), \mu) \leq C_0 \left( e^{-\gamma_1 n^{1/3}} + \sum_{k=n}^{\infty} E_{1/n^{\theta}} [\alpha^{k-1}(Y_0)] \right)
\]

holds for all \( n \in \mathbb{N} \), with some \( C_0, \gamma_1, \gamma_2 > 0 \), where \( \gamma := \frac{1}{3(1-\theta)} \). If \( Y \) is ergodic, then for any bounded and measurable \( \Phi : \mathcal{X} \to \mathbb{R} \)

\[
\frac{\Phi(X_1) + \ldots + \Phi(X_n)}{n} \to \int_{\mathcal{X}} \Phi(z) \mu(dz), \quad n \to \infty
\]

holds in probability.

**Proof.** This follows from Theorems 2.8, 2.10 and Lemma 5.4 of [7].
3 Proof in the unbounded case

Throughout this section the assumptions of Theorem 1.7 are in force. We will use results of the previous section in the setting \( \mathcal{X} = \mathcal{Y} = \mathbb{R}_+; Y_n := Z_n, n \in \mathbb{N} \). We can easily extend the process \( Y \) on the negative time axis in such a way that \( Y_n, n \in \mathbb{Z} \) is stationary and ergodic. Define the parametrized kernel \( Q \) as follows:

\[
Q(z, w, A) := \mathbb{P} \left[ \left( w + S_0 - z \right)_+, z \in \mathcal{Y}, w \in \mathcal{X}, A \in B(\mathcal{X}) \right].
\]

We now turn to the verification of Assumptions 2.2 and 2.3.

Lemma 3.1. For some \( \bar{\eta} > 0 \) define

\[
V(w) := e^{\bar{\beta}w} - 1, \quad w \geq 0, \quad K(z) := e^{-\beta z} \mathbb{E} \left[ e^{\beta S_0} \right], \quad z \geq 0.
\]

We have

\[
[Q(z)V](w) \leq y(z)V(w) + K(z)
\]

for all \( z \in \mathcal{Y}, w \in \mathcal{X} \). Furthermore, choosing \( \bar{\beta} \) small enough,

\[
\bar{y} := \limsup_{n \to \infty} \mathbb{E}^{1/n} \left( K(Z_0) \prod_{k=1}^{n} y(Z_k) \right) < 1.
\]

Proof. Let us estimate

\[
[Q(z)V](w) = \mathbb{E}[e^{\bar{\beta}(w+S_0-z)}] - 1 \leq \mathbb{E}[e^{\bar{\beta}(w+S_0-z)}] + 1 - 1
\]

\[
= y(z)e^{\bar{\beta}w} = y(z)V(w) + y(z),
\]

so (8) holds. Define \( \lambda(\beta) := \Gamma(-\beta) + \ln(\mathbb{E}[e^{\beta S_0}]), \beta \in (-\bar{\eta}, \bar{\eta}) \) where \( \bar{\eta} := \min \{ \eta, \beta_0/2 \} \). Note that \( Z_0 \) has finite exponential moments of all orders. Hence the functions

\[
\lambda_n(\beta) := \frac{1}{n} \ln \mathbb{E} \left[ e^{\beta \sum_{i=1}^{n}(S_i-Z_i)} \right], \beta \in (-\bar{\eta}, \bar{\eta}), \quad n \geq 1
\]

are finite. They are also clearly convex. Define

\[
\psi_n(\beta) := \mathbb{E} \left[ \frac{e^{\beta \sum_{i=1}^{n}(S_i-Z_i)} - 1}{\beta} \right], \beta \in (0, \bar{\eta}), \quad n \geq 1.
\]

By the Lagrange mean value theorem and measurable selection, there exists a random variable \( \xi_n(\beta) \in [0, \beta] \) such that

\[
\psi_n(\beta) = \mathbb{E} \left[ \left( \sum_{j=1}^{n} (S_j - Z_j) \right) e^{\xi_n(\beta) \sum_{i=1}^{n}(S_i-Z_i)} \right].
\]
Here
\[
\left( \sum_{j=1}^{n} (S_j - Z_j) \right) e^{\lambda_0} \sum_{j=1}^{n} (S_j - Z_j) \leq \left( \sum_{j=1}^{n} S_j \right) e^{\lambda_0} \sum_{j=1}^{n} S_j,
\]
which is integrable. Hence reverse Fatou’s lemma shows that
\[
\limsup_{\beta \to 0^+} \psi_n(\beta) \leq \mathbb{E} \left[ \sum_{j=1}^{n} (S_j - Z_j) \right] = n \mathbb{E} [S_0 - Z_0].
\]
This implies that, for all \( n \geq 1 \), \( \lambda_n(0) = \frac{1}{n} \lim_{\beta \to 0^+} \psi_n(\beta) = \mathbb{E} [S_0 - Z_0] < 0 \).

Since \( \lambda_n(\beta) \to \lambda(\beta) \) for \( \beta \in (\tilde{\eta}, \eta) \) by the Gärtner-Ellis-type property of \((Z_n)_{n \in \mathbb{N}}\). It follows from Theorem 25.7 of [9] that also \( \lambda'_n(0) \to \lambda'(0) \) hence \( \lambda'(0) < 0 \). By Corollary 25.5.1 of [9], differentiability of \( \lambda \) implies its continuous differentiability, too. Hence from \( \lambda(0) = 0 \) and \( \lambda'(0) < 0 \) we obtain that there exists \( \tilde{\beta} > 0 \) satisfying
\[
\lim_{n \to \infty} \frac{1}{n} \ln \mathbb{E} e^{\tilde{\beta} (S_1 + \ldots + S_n) - \tilde{\beta} (Z_1 + \ldots + Z_n)} < 0. \tag{10}
\]
By (10), the long-time contractivity condition holds:
\[
\limsup_{n \to \infty} \mathbb{E}^{1/n} [K(Z_0)Y(Z_1) \ldots Y(Z_n)] < 1 \tag{11}
\]
since \( K(z) \leq \mathbb{E} [e^{\tilde{\beta} S_0}] \) for all \( z \in \mathcal{Y} \).

Choose \( \varepsilon := (1/\gamma^{1/2} - 1)/2 \). Notice that \( R := R(z) := 2K(z)(\varepsilon Y(z)) = 2/\varepsilon \) does not depend on \( z \). Now let us turn to the verification of the minorization condition. Let
\[
h := \ln \left( \frac{2}{\varepsilon} + 1 \right) / \tilde{\beta}.
\]

**Lemma 3.2.** For \( z \in \mathcal{Y}, A \in B(\mathcal{X}) \) define \( \kappa(A) := \kappa(z, A) := \text{Leb}(A \cap [h, h + 1]) \), where \( \text{Leb} \) denotes the Lebesgue measure. Then
\[
\inf_{w \in V^{-1}([0, R])} Q(z, w, A) \geq (1 - \alpha(z)) \kappa(A),
\]
holds for
\[
\alpha(z) := \frac{\sqrt{\varepsilon}}{2} e^{-\frac{\varepsilon (\varepsilon h + 1)}{2}}.
\]

**Proof.** Notice that \( V^{-1}([0, R]) = [0, h] \). For each \( z \in \mathcal{Y}, A \in B(\mathcal{X}) \) and \( w \in [0, h] \),
\[
Q(z, w, A) = P ([w + S_0 - z], \in A) \geq P (w + S_0 - z \in A \cap [h, h + 1])
\geq P (S_0 \in (A + z - w) \cap [h + z - w, h + 1 + z - w])
\geq f(z + h + 1) \kappa(A + z - w) = f(z + h + 1) \kappa(A)
\geq \frac{f(z + h + 1)}{2} \kappa(A) \geq \sqrt{\varepsilon} e^{-\frac{\varepsilon (\varepsilon h + 1)}{2}} \kappa(A),
\]

\[ 8 \]
by translation invariance of the Lebesgue measure and by the elementary inequality
\[ 1 - x = \sqrt{\frac{e}{2}} e^{-x}, \ x \in [0, 1/2], \]
noting also that \( f(z + h + 1) \leq 1 \). We may conclude.

Proof of Theorem 1.7. Let \( \mu_0 \) denote the law of \( Z_0 \). Let \( H > 0 \) be a constant to be chosen later. From Lemma 3.2

\[
E[\alpha^n(Z_0)] \\
\lesssim \int_0^\infty e^{-\frac{z}{C_2} - C_2 z} \mu_0(dz) \\
\lesssim \int_0^{H \ln(n)} e^{-\frac{z}{C_2} - C_2 z} \mu_0(dz) + P(Z_0 \geq H \ln(n)) \\
\leq H \ln(n) e^{-\frac{C_4}{C_2} - C_2 H \ln(n)} + C_1 \exp\left(-C_2 e^{C_2 H \ln(n)}\right) \\
\leq H \ln(n) \exp \left\{ -\frac{C_4}{2} n^{1-C_2 H}\right\} + C_1 \exp\{-C_2 n^{C_2 H}\}. \tag{12}
\]
Choosing \( H \) so small that \( HC_2 < 1 \) we get that Assumption 2.5 holds for \( \theta \) small enough. The claimed convergence rate also follows from (12) and (6).

\[ Q(z, w, A) \geq (1 - \alpha) \kappa(A), \]
holds for a constant \( \alpha > 0 \).

Proof. Let \( M > 0 \) be such that \( |Z_0| \leq M \) almost surely. As in Lemma 3.1

\[
Q(z, w, A) \geq P (S_0 \in (A + z - w) \cap [h + z - w, h + 1 + z - w]) \\
\geq f(z + h + 1) \kappa(A + z - w) \geq \inf_{z \in [0, M, h+1]} f(v) \kappa(A),
\]
which proves the statement since \( f \) is bounded away from 0 on compacts.

Proof of Theorem 1.5 Assumption 2.4 holds by Lemma 4.1 and Assumption 2.5 holds trivially. Lemma 3.1 implies Assumptions 2.2 and 2.3. We can conclude from Theorem 2.6 with the choice \( \theta = 1/2 \), noting that, in the present case,

\[
E^{\alpha^{\theta}}[\alpha^{2\alpha - 1}(Y_0)] \leq \frac{1}{\alpha} \alpha^{\theta} k^{1/3} \text{ and } \sum_{k=0}^{\infty} \alpha^{\gamma k} k^{1/3} \leq C_\gamma \alpha^{\gamma k^{1/3}} \text{ for some } C_\gamma, \gamma > 0.
\]
References

[1] A. A. Borovkov. *Egodicity and stability of stochastic processes*. Wiley & Sons, New York, 1998.

[2] A. A. Borovkov. *Asymptotic Methods in Queuing Theory*. Nauka, Moscow, 1980. In Russian.

[3] R. Burton and H. Dehling. Large deviations for some weakly dependent random processes. *Statistics Probab. Letters*, 9:397–401, 1990.

[4] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, Springer, 1998.

[5] L. Gerencsér and C. Prosdocimi. Input-output properties of the Hinkley detector. *Systems and Control Letters*, 60:486–491, 2011.

[6] I. Kontoyiannis and S. Meyn. Large deviations asymptotics and the spectral theory of multiplicative regular processes. *Electronic J. Probab.*, 10:61–123, 2005.

[7] A. Lovas and M. Rásonyi. Markov chains in random environment with applications in queuing theory and machine learning. *Preprint*, 2019. [arXiv:1911.04377](arXiv:1911.04377)

[8] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Springer-Verlag, 1993.

[9] R. T. Rockafellar. *Convex analysis*. Princeton University Press, Princeton, N. J., 1972.