Теорія Ймовір. та Матем. Статист.
Вип. 91, 2014

Theor. Probability and Math. Statist.
No. 91, 2015, Pages 17–27
http://dx.doi.org/10.1090/tpms/963
Article electronically published on February 3, 2016

# MAXIMAL COUPLING AND STABILITY OF DISCRETE NON-HOMOGENEOUS MARKOV CHAINS

UDC 519.21

V. V. GOLOMOZYĬ AND M. V. KARTASHOV

Abstract. We consider two time non-homogeneous discrete Markov chains whose one-step transition probabilities are close in the uniform total variation norm. The problem of stability of the transition probabilities for an arbitrary number of steps is investigated. The main assumption is the uniform mixing. We prove that the uniform difference between the distributions of the chains after an arbitrary number of steps does not exceed $\varepsilon/(1-\rho)$, where $\varepsilon$ is the uniform distance between transition matrices and $\rho$ is the uniform mixing coefficient. The proofs are based on the maximal coupling procedure that maximize the one-step coupling probabilities.

## 1. Introduction

The results of [28, 29] are generalized and estimates for the stability of transient probabilities over $n$ steps for two different non-homogeneous Markov chains are obtained in the current paper. We prove analogs of Theorem 1 and Corollary 1 of the paper [28] for the non-homogeneous case.

It is worth mentioning that the conditions and estimates below are analogous to those in [28] and the proofs follow a similar idea but take into account some specific properties of the non-homogeneous case. As in [28, 29], the proof uses the maximal coupling method. A brief account of the coupling procedure is given in the current paper and this is sufficient for the proof of results below. An interested reader may consult [28] for an extended description of the maximal coupling procedure.

The first mention of the coupling method can be found in a 1938 paper of Doeblin [1]. The coupling procedure as a whole is described in [8]. The first monograph devoted to the coupling method and its applications is [12]. Another known monograph of this kind is [17]. A number of papers where the coupling method (or similar analogs) is used for the analysis of the stability appeared since then; see, in particular, [11, 10, 14, 15, 18]. All these papers deal with homogeneous Markov chains and study their stability with respect to the initial measure.

The authors of the current paper obtained several results concerning the stability of two different Markov chains whose transient probabilities are close in a certain sense. The results of [24, 25] use methods of [22, 21, 19, 20, 23] and generalize the results therein to the case of different Markov chains; in particular, to non-homogeneous Markov chains.

The papers [24, 25] use the so-called $C$-coupling, which means that the chains coincide with a certain probability in the case where both chains hit a certain set $C$. In contrast, the current paper uses another coupling model called the maximal coupling.

2010 *Mathematics Subject Classification.* Primary 60J45; Secondary 60A05, 60K05.

*Key words and phrases.* Coupling theory, coupling method, maximal coupling, discrete Markov chains, stability of distributions.

General results of the theory of coupling such as the existence of coupling and bounds for the coupling moments for both homogeneous and non-homogeneous chains can be found in [27, 30].

## 2. Basic conditions and main results

Consider a discrete space $E = \{i, j, k, \dots\}$ equipped with the $\sigma$-algebra of all its subsets $\mathcal{E} = 2^E$. Consider two non-homogeneous Markov chains defined by a pair of families of stochastic matrices $P^{(t)} = \big(P^{(t)}_{ij},\ i, j \in E\big)$ and $P'^{(t)} = \big(P'^{(t)}_{ij}, i, j \in E\big)$. Denote by $\mathsf{P}_i$ and $\mathsf{E}_i$ the conditional probabilities and expectations in the probability space where a Markov chain $X = (X_n, n \geq 0)$ leaves; this chain has the transient matrix $P^{(t)}$ for the $t^{\text{th}}$ step and its initial value is $X_0 = i$. The notation $\mathsf{P}'_i$ and $\mathsf{E}'_i$ for a chain $X' = (X'_n, n \geq 0)$ is similar to that if the transient matrices are $P'^{(t)}$.

In what follows, if we do not specify the set of indices when evaluating a sum or limit superior, then these operations are extended to the whole space $E$. For a matrix $Q = (Q_{ij})$, the symbol $Q^{(n)}$ denotes its $n^{\text{th}}$ power and $Q_{i\bullet} = (Q_{ij}, j \in E)$ means its $i^{\text{th}}$ row. We denote by $x^{\pm}$ the positive or negative part of a number $x$. Throughout the paper, $\delta_{ij} = 1_{i=j}$ are Kronecker symbols and $\wedge$ means the minimum.

**1. Uniform stability.** Denote by $\|\mu\| = \sum_j |\mu_j|$ the total variation norm in the space of summable sequences $l_1(E) = \{\mu = (\mu_j, j \in E)\colon \|\mu\| < \infty\}$ and by $\mu Q_j = \sum_i \mu_i Q_{ij}$ the product of the measure $\mu$ by the matrix $Q$.

*Condition of the uniform one-step stability* means that the transient matrices $P^{(t)}$ and $P'^{(t)}$ become closer in the uniform metrics. More precisely,

$$\exists\, \varepsilon \in (0,1),\ \forall t \geq 0\colon \quad r\left(P^{(t)}, P'^{(t)}\right) \equiv \sup\nolimits_i \left\|P^{(t)}_{i\bullet} - P'^{(t)}_{i\bullet}\right\|/2 \leq \varepsilon. \tag{1}$$

Condition (1) is equivalent to the inequality

$$\left\|\mu P^{(t)} - \mu P'^{(t)}\right\| \leq 2\varepsilon \|\mu\|$$

for all measures $\mu$.

In the scheme of series, one can assume that the difference between $P'^{(t)}$ and $P^{(t)}$ approaches zero. Note however that the results below are stated in the form of inequalities and thus can be used for a fixed $\varepsilon > 0$ as well.

*Condition of the uniform mixing* means that the mutual mixing coefficient is separated from unity. More precisely,

$$\exists\, \rho \in (0,1), \forall t \geq 0\colon \quad m\left(P^{(t)}, P'^{(t)}\right) \equiv \sup_{i \neq k} \left\|P^{(t)}_{i\bullet} - P'^{(t)}_{k\bullet}\right\|/2 \leq \rho. \tag{2}$$

*Remark* 1. Condition (2) for small $\varepsilon$ holds if and only if the ordinary uniform mixing condition holds for $P^{(t)}$, namely if

$$m\big(P^{(t)}, P^{(t)}\big) \leq \rho_0 < 1,$$

since $\big|m(P'^{(t)}, P^{(t)}) - m(P^{(t)}, P^{(t)})\big| \leq r\big(P^{(t)}, P'^{(t)}\big)$ and $m\big(P^{(t)}, P'^{(t)}\big) \leq \rho_0 + \varepsilon < 1$ for sufficiently small $\varepsilon > 0$. In turn, the above condition imposed on $P^{(t)}$ is equivalent to the following operator contraction inequality:

$$\|\mu P\| \leq \rho_0 \|\mu\|, \qquad \forall \mu \in l_1^0(E) \equiv l_1(E) \cap \{\mu\colon \mu(E) = 0\}. \tag{3}$$

**Theorem 1.** *Let conditions of the one-step stability* (1) *and of the uniform mixing* (2) *hold. Then*

$$\sup_{B \subset E} \big|\mathsf{P}_i(X_n \in B) - \mathsf{P}'_i(X'_n \in B)\big| \leq \varepsilon(1 - \rho^n)/(1-\rho) < \varepsilon/(1-\rho) \tag{4}$$

*for any $n \geq 1$ and uniformly in $i \in E$.*

**Theorem 2.** *If assumptions of Theorem 1 hold, then*

$$\sup_{i,k}\sup_{B\subset E}\left|\mathsf{P}_i(X_n\in B)-\mathsf{P}'_k(X'_n\in B)\right|\le\rho^n+\varepsilon(1-\rho^n)/(1-\rho) \tag{5}$$

*for every* $n\ge 1$.

## 3. Example

Consider the following two non-homogeneous Markov chains defined by their transient probabilities in a $t^{\text{th}}$ step:

$$P^{(t)}=\begin{pmatrix} 0 & \alpha_1^{(t)} & \alpha_2^{(t)} & \alpha_3^{(t)} & \dots \\ \beta_1^{(t)} & 1-\beta_1^{(t)} & 0 & 0 & \dots \\ \beta_2^{(t)} & 0 & 1-\beta_2^{(t)} & 0 & \dots \\ \beta_3^{(t)} & 0 & 0 & 1-\beta_3^{(t)} & \dots \\ \dots \end{pmatrix},$$

$$P'^{(t)}=\begin{pmatrix} 0 & \alpha_1'^{(t)} & \alpha_2'^{(t)} & \alpha_3'^{(t)} & \dots \\ \beta_1'^{(t)} & 1-\beta_1'^{(t)} & 0 & 0 & \dots \\ \beta_2'^{(t)} & 0 & 1-\beta_2'^{(t)} & 0 & \dots \\ \beta_3'^{(t)} & 0 & 0 & 1-\beta_3'^{(t)} & \dots \\ \dots \end{pmatrix}.$$

Below we specify the conditions under which Theorems 1 and 2 are valid for these chains.

The condition for the uniform one-step stability is given by

$$\sum_{i\ge 1}\left|\alpha_i^{(t)}-\alpha_i'^{(t)}\right|\le 2\varepsilon, \tag{6}$$

$$\left|\beta_i^{(t)}-\beta_i'^{(t)}\right|\le\varepsilon. \tag{7}$$

The uniform mixing condition is given by

$$\left|\beta_i^{(t)}-\beta_k'^{(t)}\right|\le\rho, \tag{8}$$

$$\beta_i^{(t)}+1+\left|\alpha_i'^{(t)}+1-\beta_i^{(t)}\right|-\alpha_i'^{(t)}\le 2\rho, \tag{9}$$

$$\beta_i'^{(t)}+1+\left|\alpha_i^{(t)}+1-\beta_i'^{(t)}\right|-\alpha_i^{(t)}\le 2\rho. \tag{10}$$

Note that condition (8) follows from the classical condition $\inf\{\beta_i^{(t)},\beta_i'^{(t)}\}>0$.

Therefore conditions (6), (7), (8), (9), and (10) imply that the transient probabilities over $n$ steps for two such chains are closed irrespective of their initial values.

## 4. Maximal coupling of chains

Let

$$D=\{0,1\}.$$

Consider the Markov chain $Z_n=\left(Z_n^{(1)},Z_n^{(2)},d_n\right)$ defined as follows in the space $E\times E\times D$ equipped with the $\sigma$-algebra of all its subsets:

$$d_0\in\{0,1\},\qquad Z_0^{(1)}=X_0,\qquad Z_0^{(2)}=X'_0.$$

By $\mathbb{P}$ we denote the probability generated by this chain. We also use the abbreviation $(E,E,D)=E\times E\times D$.

Then we define $Z_{n+1}$ recursively with the help of $Z_n$.

In the above notation, $d_n$ characterizes whether or not the chains $n$ are coupled at a given moment, namely if $d_n=1$; then the chains are coupled, otherwise they are decoupled. Our current goal is to determine the transient probabilities in such a way that the marginal probabilities coincide with the initial probabilities.

First we determine the transient probabilities for *coupled chains*:

$$\mathbb{P}(Z_{n+1} = (j,j,1) \mid Z_n = (i,i,1)) = P_{ij}^{(n)} \wedge P_{ij}^{\prime(n)}, \qquad i,j \in E, \tag{11}$$

$$\mathbb{P}(Z_{n+1} = (j,k,1) \mid Z_n = (i,i,1)) = 0, \qquad k \neq j, \tag{12}$$

$$\begin{aligned} &\mathbb{P}(Z_{n+1} = (j,k,0) \mid Z_n = (i,i,1)) \\ &\quad = \frac{\left(P_{ij}^{(n)} - P_{ij}^{(n)} \wedge P_{ij}^{\prime(n)}\right)\left(P_{ik}^{\prime(n)} - P_{ik}^{(n)} \wedge P_{ik}^{\prime(n)}\right)}{1 - \sum_{l\in E} P_{il}^{(n)} \wedge P_{il}^{\prime(n)}}. \end{aligned} \tag{13}$$

Now we consider the marginal distributions:

$$\begin{aligned} &\mathbb{P}(Z_{n+1} \in (j,E,D) \mid Z_n = (i,i,1)) \\ &\quad = P_{ij}^{(n)} \wedge P_{ij}^{\prime(n)} + \sum_{k\in E} \left( \frac{\left(P_{ij}^{(n)} - P_{ij}^{(n)} \wedge P_{ij}^{\prime(n)}\right)\left(P_{ik}^{\prime(n)} - P_{ik}^{(n)} \wedge P_{ik}^{\prime(n)}\right)}{1 - \sum_{l\in E} P_{il}^{(n)} \wedge P_{il}^{\prime(n)}} \right) \\ &\quad = P_{ij}^{(n)} \wedge P_{ij}^{\prime(n)} + \left(P_{ij}^{(n)} - P_{ij}^{(n)} \wedge P_{ij}^{\prime(n)}\right) \frac{\sum_{k\in E}\left(P_{ik}^{\prime(n)} - P_{ik}^{(n)} \wedge P_{ik}^{\prime(n)}\right)}{1 - \sum_{l\in E} P_{il}^{(n)} \wedge P_{il}^{\prime(n)}} = P_{ij}^{(n)}. \end{aligned}$$

Therefore

$$\mathbb{P}(Z_{n+1} \in (j,E,D) \mid Z_n = (i,i,1)) = P_{ij}^{(n)}, \qquad \forall i,j \in E. \tag{14}$$

Similarly,

$$\mathbb{P}(Z_{n+1} \in (E,j,D) \mid Z_n = (i,i,1)) = P_{ij}^{\prime(n)}, \qquad \forall i,j \in E. \tag{15}$$

Then we turn to the transient probabilities for *decoupled chains*:

$$\mathbb{P}(Z_{n+1} = (j,j,1) \mid Z_n = (i,k,0)) = P_{ij}^{(n)} \wedge P_{kj}^{\prime(n)}, \qquad i,k,j \in E,$$

$$\mathbb{P}(Z_{n+1} = (j,l,1) \mid Z_n = (i,k,0)) = 0, \qquad k \neq l,$$

$$\begin{aligned} &\mathbb{P}(Z_{n+1} = (j,l,0) \mid Z_n = (i,k,0)) \\ &\quad = \frac{\left(P_{ij}^{(n)} - P_{ij}^{(n)} \wedge P_{kj}^{\prime(n)}\right)\left(P_{kl}^{\prime(n)} - P_{il}^{(n)} \wedge P_{kl}^{\prime(n)}\right)}{1 - \sum_{s\in E} P_{is}^{(n)} \wedge P_{ks}^{\prime(n)}}, \qquad i,k,j,l \in E. \end{aligned} \tag{16}$$

Considering the marginal distributions in the case of $d_n = 0$, we see that

$$\mathbb{P}\big(Z_{n+1} \in (j,E,D) \mid Z_n = (i,k,0)\big) = P_{ij}^{(n)}, \tag{17}$$

$$\mathbb{P}\big(Z_{n+1} \in (E,j,D) \mid Z_n = (i,k,0)\big) = P_{ij}^{\prime(n)} \tag{18}$$

which can be proved similarly to the case of coupled chains.

*Remark* 2. It follows from the definition of the process $Z_n$ that $d_n = 1$ if and only if $Z_n^{(1)} = Z_n^{(2)}$, that is, $d_n$ is a functional depending on both $Z_n^{(1)}$ and $Z_n^{(2)}$. There are two reasons to consider such an index in the current paper. The first reason is that we want to highlight that, in contrast to the classical coupling where two copies of the same chain couple and after this moment never decouple, the chains in the case under consideration may couple and decouple several times.

The second reason is that the property of $d_n$ indicated above is valid only for the maximal coupling of chains. In the papers of other authors (see, for example, [24, 25]) where the $C$-coupling is studied, one may face the situation where components of the chain coincide but the coupling does not appear. Hence it is reasonable to keep the general notation for the case under consideration to exhibit an analogy with other kinds of coupling.

## 5. Auxiliary results

**Lemma 5.1.** *For all $i, j \in E$,*

$$P_i(X_n = j) = \mathbb{P}_{ii1}\left(Z_n^{(1)} = j\right),$$
$$P_i'(X_n' = j) = \mathbb{P}_{ii1}\left(Z_n^{(2)} = j\right).$$

*Proof.* We will prove Lemma 5.1 by induction. The base of induction follows from the marginal and initial probabilities, namely from (14) and (15).

For the inductive step, assume that the statement holds for $n$, that is,

$$P_i(X_n = j) = \mathbb{P}_{ii1}\left(Z_n^{(1)} = j\right),$$

and prove it for $n+1$:

$$\begin{aligned} P_i(X_{n+1} = j) &= \sum_k P(X_{n+1} = j \mid X_n = k) P_i(X_n = k) = \sum_k P_{kj}^{(n)} \mathbb{P}_{ii1}\left(Z_n^{(1)} = k\right) \\ &= \sum_k \mathbb{P}_{ii1}(Z_n \in (k, E, 0)) P_{kj}^{(n)} + \sum_k \mathbb{P}_{ii1}(Z_n \in (k, E, 1)) P_{kj}^{(n)}. \end{aligned}$$

Using (14) and (17) we get

$$\begin{aligned} &\sum_k \mathbb{P}_{ii1}(Z_n \in (k, E, 0)) P_{kj}^{(n)} + \sum_k \mathbb{P}_{ii1}(Z_n \in (k, E, 1)) P_{kj}^{(n)} \\ &\quad = \sum_k \sum_l \mathbb{P}_{ii1}(Z_n = (k, l, 0)) \mathbb{P}\big(Z_{n+1} \in (j, E, D) \mid Z_n = (k, l, 0)\big) \\ &\quad\quad + \sum_k \mathbb{P}_{ii1}(Z_n = (k, k, 1)) \mathbb{P}\big(Z_n \in (j, E, D) \mid Z_n = (k, k, 1)\big) \\ &\quad = \int_{z \in E \times E \times D} \mathbb{P}_{ii1}(Z_n = dz) \mathbb{P}(Z_{n+1} \in (j, E, D) \mid Z_n = z) \\ &\quad = \mathbb{P}_{ii1}(Z_{n+1} \in (j, E, D) = \mathbb{P}_{ii1}\left(Z_{n+1}^{(1)} = j\right). \end{aligned}$$

The equality

$$P_i'(X_n' = j) = \mathbb{P}_{ii1}\left(Z_n^{(2)} = j\right)$$

is proved analogously. □

The following result establishes similar formulas for the initial decoupled state.

**Lemma 5.2.** *For all $i, j, s \in E$,*

$$P_i(X_n = j) = \mathbb{P}_{is0}\left(Z_n^{(1)} = j\right),$$
$$P_s'(X_n' = j) = \mathbb{P}_{is0}\left(Z_n^{(2)} = j\right).$$

*Proof.* The proof of this result is similar to that of Lemma 5.1, so we provide it without details.

As in Lemma 5.1 we use the induction. The result for the base of induction follows from (17) and (18).

To prove the result for the inductive step, assume it holds for $n$, namely let

$$P_i(X_n = j) = \mathbb{P}_{is0}\left(Z_n^{(1)} = j\right),$$

and prove it for $n+1$:

$$\begin{aligned}
P_i(X_{n+1}=j) &= \sum_k P(X_{n+1}=j \mid X_n=k)P_i(X_n=k) = \sum_k P_{kj}^{(n)}\mathbb{P}_{is0}\left(Z_n^{(1)}=k\right)\\
&= \sum_k \mathbb{P}_{is0}(Z_n\in(k,E,0))P_{kj}^{(n)} + \sum_k \mathbb{P}_{is0}(Z_n\in(k,E,1))P_{kj}^{(n)}\\
&= \sum_k\sum_l \mathbb{P}_{is0}(Z_n=(k,l,0))\mathbb{P}\big(Z_{n+1}\in(j,E,D)\mid Z_n=(k,l,0)\big)\\
&\quad + \sum_k \mathbb{P}_{is0}(Z_n=(k,k,1))\mathbb{P}\big(Z_n\in(j,E,D)\mid Z_n=(k,k,1)\big)\\
&= \mathbb{P}_{is0}(Z_{n+1}\in(j,E,D)) = \mathbb{P}_{is0}\left(Z_{n+1}^{(1)}=j\right).
\end{aligned}$$

The second equality

$$P_i'(X_n'=j) = \mathbb{P}_{is0}\left(Z_n^{(2)}=j\right)$$

is proved analogously. □

**Lemma 5.3.** *For an arbitrary set $B\subset E$ and all $i\in E$,*

$$|P_i(X_n\in B) - P_i'(X_n'\in B)| \le \mathbb{P}_{ii1}(d_n=0).$$

*Proof.* Using Lemma 5.1 we establish

$$|P_i(X_n\in B) - P_i'(X_n'\in B)| = \big|\mathbb{P}_{ii1}(Z_n\in(B,E,D)) - \mathbb{P}_{ii1}(Z_n\in(E,B,D))\big|,$$

whence we get for $d_n=0$ and $d_n=1$ that

$$\begin{aligned}
&|\mathbb{P}_{ii1}(Z_n\in(B,E,D)) - \mathbb{P}_{ii1}(Z_n\in(E,B,D))|\\
&\quad = \big|\mathbb{P}_{ii1}(Z_n\in(B,E,1)) + \mathbb{P}_{ii1}(Z_n\in(B,E,0))\\
&\qquad\qquad - \mathbb{P}_{ii1}(Z_n\in(E,B,1)) - \mathbb{P}_{ii1}(Z_n\in(E,B,0))\big|.
\end{aligned}$$

Note that

$$\begin{aligned}
&\mathbb{P}_{ii1}(Z_n\in(B,E,1)) - \mathbb{P}_{ii1}(Z_n\in(E,B,1))\\
&\quad = \sum_{k\in B}\big(\mathbb{P}_{ii1}(Z_n\in(k,E,1)) - \mathbb{P}_{ii1}(Z_n\in(E,k,1))\big).
\end{aligned}$$

By (12) the latter expression is equal to

$$\sum_{k\in B}\big(\mathbb{P}_{ii1}(Z_n=(k,k,1)) - \mathbb{P}_{ii1}(Z_n=(k,k,1))\big) = 0,$$

whence

$$\begin{aligned}
&\big|\mathbb{P}_{ii1}(Z_n\in(B,E,D)) - \mathbb{P}_{ii1}(Z_n\in(E,B,D))\big|\\
&\quad = \big|\mathbb{P}_{ii1}(Z_n\in(B,E,0)) - \mathbb{P}_{ii1}(Z_n\in(E,B,0))\big|\\
&\quad \le \max\big\{\mathbb{P}_{ii1}(Z_n\in(B,E,0),\mathbb{P}_{ii1}(Z_n\in(E,B,0))\big\}\\
&\quad \le \mathbb{P}_{ii1}(Z_n\in(E,E,0)) = \mathbb{P}_{ii1}(d_n=0).
\end{aligned}$$
□

**Lemma 5.4.** *For an arbitrary set $B\subset E$ and all $i,k\in E$,*

$$|P_i(X_n\in B) - P_k'(X_n'\in B)| \le P_{ik0}(d_n=0).$$

*Proof.* The proof follows the lines of that of the preceding lemma and thus we provide it without details. We have

$$\begin{aligned}|P_i(X_n \in B) - P'_k(X'_n \in B)| &= |\mathbb{P}_{ik0}(Z_n \in (B,E,D)) - \mathbb{P}_{ik0}(Z_n \in (E,B,D))| \\ &= \big|\mathbb{P}_{ik0}(Z_n \in (B,E,1)) + \mathbb{P}_{ik0}(Z_n \in (B,E,0)) - \mathbb{P}_{ik0}(Z_n \in (E,B,1)) \\ &\qquad - \mathbb{P}_{ik0}(Z_n \in (E,B,0))\big| \\ &\le \max\big\{\mathbb{P}_{ik0}(Z_n \in (B,E,0)), \mathbb{P}_{ik0}(Z_n \in (E,B,0))\big\} \\ &\le \mathbb{P}_{ik0}(Z_n \in (E,E,0)) = \mathbb{P}_{ik0}(d_n = 0).\end{aligned}$$

□

**Lemma 5.5.** *Assume that condition* (1) *of the uniform one-step stability holds for the chains $X_n$ and $X'_n$. Then*

$$\mathbb{P}(d_{n+1} = 0 \mid Z_n = (i,i,1)) \le \varepsilon$$

*for all $i \in E$ and $n \ge 0$.*

*Proof.* We use (13):

$$\begin{aligned}\mathbb{P}(d_{n+1} = 0 \mid Z_n = (i,i,1)) &= \sum_{j,k} \frac{\left(P_{ij}^{(n)} - P_{ij}^{(n)} \wedge P_{ij}'^{(n)}\right)\left(P_{ik}'^{(n)} - P_{ik}^{(n)} \wedge P_{ik}'^{(n)}\right)}{1 - \sum_{l\in E} P_{il}^{(n)} \wedge P_{il}'^{(n)}} \\ &= 1 - \sum_{l\in E} P_{il}^{(n)} \wedge P_{il}'^{(n)} \le \varepsilon.\end{aligned}$$

The latter inequality follows from

$$\begin{aligned}\varepsilon &\ge \sup_i \sum_j \left|P_{ij}^{(n)} - P_{ij}'^{(n)}\right| \Big/ 2 = \sup_i \sum_j \left(\left(P_{ij}^{(n)} - P_{ij}'^{(n)}\right)^+ + \left(P_{ij}'^{(n)} - P_{ij}^{(n)}\right)^+\right) \Big/ 2 \\ &= \sup_i \sum_j \left(P_{ij}^{(n)} - P_{ij}^{(n)} \wedge P_{ij}'^{(n)} + P_{ij}'^{(n)} - P_{ij}^{(n)} \wedge P_{ij}'^{(n)}\right) \Big/ 2 \\ &= \sup_i \left(2 - 2\sum_j P_{ij}^{(n)} \wedge P_{ij}'^{(n)}\right) \Big/ 2 = \sup_i \left(1 - \sum_j P_{ij}^{(n)} \wedge P_{ij}'^{(n)}\right).\end{aligned}$$

Above we used the equality $x = x \wedge y + (x-y)^+$. □

**Lemma 5.6.** *Assume that condition* (2) *of the uniform mixing holds for the chains $X_n$ and $X'_n$. Then*

$$\mathbb{P}\big(d_{n+1} = d_{n+2} = \dots d_{n+m} = 0 \mid Z_n = (i,j,0)\big) \le \rho^m$$

*for all $i, j \in E$ and $n \ge 0$.*

*Proof.* It follows from (16) that

$$\begin{aligned}\mathbb{P}\big(d_{n+1} = 0 \mid Z_n = (i,k,0)\big) &= \sum_{j,l} \frac{\left(P_{ij}^{(n)} - P_{ij}^{(n)} \wedge P_{kj}'^{(n)}\right)\left(P_{kl}'^{(n)} - P_{il}^{(n)} \wedge P_{kl}'^{(n)}\right)}{1 - \sum_{s\in E} P_{is}^{(n)} \wedge P_{ks}'^{(n)}} \\ &= 1 - \sum_{j\in E} P_{ij}^{(n)} \wedge P_{kj}'^{(n)}.\end{aligned}$$

At the same time

$$\begin{aligned}\rho &= \sup_{i\neq k}\sum_j \left|P_{ij}^{(n)} - P_{kj}^{\prime(n)}\right| \Big/ 2 = \sup_{i\neq k}\sum_j \left(\left(P_{ij}^{(n)} - P_{kj}^{\prime(n)}\right)^+ + \left(P_{kj}^{\prime(n)} - P_{ij}^{(n)}\right)^+\right)\Big/2 \\ &= \sup_{i\neq k}\sum_j \left(P_{ij}^{(n)} - P_{ij}^{(n)}\wedge P_{kj}^{\prime(n)} + P_{kj}^{\prime(n)} - P_{ij}^{(n)}\wedge P_{kj}^{\prime(n)}\right)\Big/2 \\ &= \sup_{i\neq k}\left(2 - 2\sum_j P_{ij}^{(n)}\wedge P_{kj}^{\prime(n)}\right)\Big/2 = \sup_{i\neq k}\left(1 - \sum_j P_{ij}^{(n)}\wedge P_{kj}^{\prime(n)}\right).\end{aligned}$$

Thus

$$\mathbb{P}\big(d_{n+1} = 0 \mid Z_n = (i,k,0)\big) \le \rho$$

for all $n \ge 0$ and $i, k \in E$. Now the statement of the lemma follows from

$$\begin{aligned}&\mathbb{P}\big(d_{n+1} = d_{n+2} = \dots d_{n+m} = 0 \mid Z_n = (i,j,0)\big) \\ &\quad = \sum_{k,l\in E}\mathbb{P}\big(d_{n+1} = d_{n+2} = \dots d_{n+m-1} = 0, Z_{n+m-1} = (k,l,0) \mid Z_n = (i,j,0)\big) \\ &\qquad\qquad \times \mathbb{P}\big(d_{n+m} = 0 \mid Z_{n+m-1} = (k,l,0)\big) \\ &\quad \le \rho\mathbb{P}\big(d_{n+1} = d_{n+2} = \dots d_{n+m-1} = 0 \mid Z_n = (i,j,0)\big) \le \dots \le \rho^m.\end{aligned}$$

□

The following is the key result for the proof of Theorems 1 and 2. Its proof is relied upon the so-called method of the last decoupling decomposition.

**Lemma 5.7.** *Assume that conditions* (1) *of the uniform one-step stability and* (2) *of the uniform mixing hold for the chains* $X_n$ *and* $X'_n$*. Then*

$$\mathbb{P}_{ii1}(d_n = 0) \le \varepsilon\frac{1-\rho^n}{1-\rho}$$

*for all* $i \in E$ *and* $n \ge 0$.

*Proof.* First we use the last decoupling decomposition. Since the chain is coupled at the zero moment and is decoupled at the moment $n$, the moment of the last decoupling occurs in the interval between 1 and $n$:

$$\begin{aligned}\mathbb{P}_{ii1}(d_n = 0) &= \sum_{t=1}^n \mathbb{P}_{ii1}(d_n = d_{n-1} = \dots d_t = 0, d_{t-1} = 1) \\ &= \sum_{t=1}^n\sum_j \mathbb{P}_{ii1}(d_n = d_{n-1} = \dots d_t = 0, Z_{t-1} = (j,j,1)) \\ &= \sum_{t=1}^n\sum_{j,k,l} \mathbb{P}_{ii1}(d_n = d_{n-1} = \dots d_t = 0, Z_t = (k,l,0), Z_{t-1} = (j,j,1)) \\ &= \sum_{t=1}^n\sum_{j,k,l} \mathbb{P}\big(d_n = d_{n-1} = \dots d_t = 0 \mid Z_t = (k,l,0)\big) \\ &\qquad\qquad \times \mathbb{P}\big(Z_t = (k,l,0) \mid Z_{t-1} = (j,j,1)\big)\mathbb{P}_{ii1}(Z_{t-1} = (j,j,1)).\end{aligned}$$

Then we use Lemma 5.6 and obtain

$$\mathbb{P}\big(d_n = d_{n-1} = \dots d_t = 0 \mid Z_t = (k,l,0)\big) \le \rho^{n-t}.$$

The latter equality together with the preceding reasoning implies that

$$\begin{aligned}\mathbb{P}_{ii1}(d_n = 0) &\le \sum_{t=1}^{n} \sum_{j,k,l} \rho^{n-t} \mathbb{P}\big(Z_t = (k,l,0) \mid Z_{t-1} = (j,j,1)\big) \mathbb{P}_{ii1}(Z_{t-1} = (j,j,1)) \\ &= \sum_{t=1}^{n} \sum_{j} \rho^{n-t} \mathbb{P}\big(d_t = 0 \mid Z_{t-1} = (j,j,1)\big) \mathbb{P}_{ii1}(Z_{t-1} = (j,j,1)) \\ &\le \varepsilon \sum_{t=1}^{n} \left( \rho^{n-t} \sum_{j} \mathbb{P}_{ii1}(Z_{t-1} = (j,j,1)) \right) = \varepsilon \sum_{t=1}^{n} \rho^{n-t} \mathbb{P}_{ii1}(d_{t-1} = 1) \\ &\le \varepsilon \sum_{t=1}^{n} \rho^{n-t} = \varepsilon \sum_{t=0}^{n-1} \rho^{t} = \varepsilon \frac{1-\rho^n}{1-\rho}.\end{aligned}$$

Here we used the equality

$$\mathbb{P}\big(d_t = 0 \mid Z_{t-1} = (j,j,1)\big) \le \varepsilon$$

being valid by Lemma 5.5. □

## 6. Proof of Theorem 1

The statement of Theorem 1 follows directly from Lemmas 5.1 and 5.7:

$$|P_i(X_n \in B) - P'_i(X'_n \in B)| \le \mathbb{P}_{ii1}(d_n = 0) \le \varepsilon \frac{1-\rho^n}{1-\rho} < \frac{\varepsilon}{1-\rho}.$$

## 7. Proof of Theorem 2

First we use Lemma 5.2:

$$|P_i(X_n \in B) - P'_k(X'_n \in B)| \le \mathbb{P}_{ik0}(d_n = 0).$$

In contrast to the proof of Theorem 1, now the chain is decoupled at both moments 0 and $n$. The idea of the proof is to split the random event $\{d_n = 0\}$ into two disjoint events

$$\{\forall t \le n, d_t = 0\} \quad \text{and} \quad \{\exists\, 0 < t < n, d_t = 1, d_s = 0, s < t\}.$$

Then we obtain

$$\begin{aligned}\mathbb{P}_{ik0}(d_n = 0) = \mathbb{P}_{ik0}(d_t = 0, t = 0, \dots, n) \\ + \sum_{t=1}^{n-1} \sum_{j \in E} \mathbb{P}_{ik0}(d_0 = \ldots = d_{t-1} = 0, Z_t = (j,j,1)) \\ \times \mathbb{P}\big(d_n = 0 \mid Z_t = (j,j,1)\big).\end{aligned}$$

We further use the inequality

$$\mathbb{P}\big(d_n = 0 \mid Z_t = (j,j,1)\big) \le \varepsilon \frac{1-\rho^{n-t}}{1-\rho} < \varepsilon(1-\rho)^{-1}$$

which follows from Lemma 5.7 and finally obtain

$$\begin{aligned}\mathbb{P}_{ik0}(d_n = 0) &\le \mathbb{P}_{ik0}(d_t = 0, t = 0, \dots, n) \\ &\quad + \varepsilon(1-\rho)^{-1} \sum_{t=1}^{n-1} \mathbb{P}_{ik0}(d_0 = \ldots = d_{t-1} = 0, d_t = 1) \\ &= \mathbb{P}_{ik0}(d_t = 0, t = 0, \dots, n) + \varepsilon(1-\rho)^{-1}(1 - \mathbb{P}_{ik0}(d_t = 0, t = 0, \dots, n)) \\ &= \varepsilon(1-\rho)^{-1} + \mathbb{P}_{ik0}(d_t = 0, t = 0, \dots, n)\left(1 - \varepsilon(1-\rho)^{-1}\right) \\ &\le \varepsilon(1-\rho)^{-1} + \rho^n \left(1 - \varepsilon(1-\rho)^{-1}\right) = \rho^n + \varepsilon \frac{1-\rho^n}{1-\rho},\end{aligned}$$

where Lemma 5.6 is used in the latter inequality.

## Acknowledgement

The authors are indebted to an anonymous referee for valuable suggestions and remarks.

## Bibliography

1. W. Doeblin, *Expose de la theorie des chaines simples constantes de Markov a un nomber fini d'estats*, Mathematique de l'Union Interbalkanique **2** (1938), 77–105.
2. W. Feller, *An Introduction to Probability Theory and its Applications*, vol. 1, John Wiley & Sons, New York, 1966. MR0210154 (35:1048)
3. N. V. Kartashov, *Strong Stable Markov Chains*, VSP/TViMS, Utrecht/Kiev, The Netherlands/Ukraine, 1996. MR1451375 (99e:60150)
4. N. V. Kartashov, *Exponential asymptotics of matrices of the Markov renewal*, Asymptotic Problems for Stochastic Processes, Preprint 77–24, Institute of Mathematics of Academy of Science of Ukraine, Kiev, 1977, pp. 2–43. (Russian)
5. E. Nummelin, *A splitting technique for Harris recurrent chains*, Z. Wahrscheinlichkeitstheorie verw. Gebiete **43** (1978), 309–318. MR0501353 (58:18732)
6. E. Nummelin and R. L. Tweedie, *Geometric ergodicity and R-positivity for general Markov chains*, Ann. Probab. **6** (1978), 404–420. MR0474504 (57:14143)
7. T. Lindvall, *On coupling of discrete renewal sequences*, Z. Wahrscheinlichkeitstheorie verw. Gebiete **48** (1979), 57–70. MR533006 (80g:60091)
8. I. N. Kovalenko and N. Ju. Kuznecov, *Construction of the embedded renewal process for essentially multidimensional processes in the queueing theory and its application to deriving of limit theorems*, Preprint 80, vol. 12, no. 80-12, Akad. Nauk Ukrain. SSR, Inst. Kibernet., Kiev, 1980 (Russian). MR612478 (82i:60142)
9. P. Ney, *A refinement of the coupling method in renewal theory*, Stoch. Process. Appl. **11** (1981), 11–26. MR608004 (82d:60169)
10. E. Numemelin and P. Tuominen, *Geometric ergodicity of Harris recurrent Markov chains with applications to renewal theory*, Stoch. Process. Appl. **12** (1982), 187–202. MR651903 (83f:60089)
11. E. Nummelin, *General Irreducible Markov Chains and Nonnegative Operators*, Cambridge University Press, Cambridge, 1984. MR776608 (87a:60074)
12. T. Lindvall, *Lectures on the Coupling Method*, John Wiley and Sons, 1991. MR1180522 (94c:60002)
13. S. P. Meyn and R. L. Tweedie, *Markov chains and Stochastic Stability*, Springer-Verlag, London, 1993. MR1287609 (95j:60103)
14. P. Tuominen and R. Tweedie, *Subgeometric rates of convergence of $f$-ergodic Markov chains*, Adv. Appl. Probab. **26** (1994), 775–798. MR1285459 (95m:60097)
15. P. Tuominen and R. L. Tweedie, *Subgeometric rates of convergence of f-ergodic Markov chains*, Adv. Appl. Probab. **26** (1994), 775–798. MR1285459 (95m:60097)
16. R. L. Tweedie and J. N. Corcoran, *Perfect sampling of ergodic Harris chains*, Ann. Appl. Probab. **11** (2001), no. 2, 438–451. MR1843053 (2002g:60111)
17. H. Thorisson, *Coupling, Stationarity, and Regeneration*, Springer, New York, 2000. MR1741181 (2001b:60003)

18. S. F. Jarner and G. O. Roberts, *Polynomial convergence rates of Markov chains*, Ann. Appl. Probab. **12** (2001), 224–247. MR1890063 (2003c:60117)
19. R. Douc, E. Moulines, and J. S. Rosenthal, *Quantitative bounds for geometric convergence rates of Markov chains*, Ann. Appl. Probab. **14** (2004), 1643–1664. MR2099647 (2005i:60146)
20. R. Douc, E. Moulines, and J. S. Rothenthal, *Quantitative bounds on convergence of time-inhomogeneous Markov chains*, Ann. Appl. Probab. **14** (2004), no. 4, 1643–1665. MR2099647 (2005i:60146)
21. R. Douc, E. Moulines, and P. Soulier, *Practical drift conditions for subgeometric rates of convergence*, Ann. Appl. Probab. **14** (2004), no. 4, 1353–1377. MR2071426 (2005e:60156)
22. R. Douc, E. Moulines, and P. Soulier, *Computable convergence rates for subgeometrically ergodic Markov chains*, Bernoulli **13** (2007), no. 3, 831–848. MR2348753 (2008j:60172)
23. R. Douc, G. Fort, and A. Guillin, *Subgeometric rates of convergence of f-ergodic strong Markov processes*, Stoch. Process. Appl. **119** (2009), no. 3, 897–923. MR2499863 (2010j:60184)
24. V. V. Golomozyĭ, *Stability of non-homogeneous Markov chains*, Visnyk Kyiv Univ., Ser. Fiz. Mat. Nauk **4** (2009), 10–15. (Ukrainian)
25. V. V. Golomozyĭ, *A subgeometric estimate of the stability for time-homogeneous Markov chains*, Teor. Imovir. Matem. Statist. **81** (2010), 31–46; English transl. in Theor. Probability and Math. Statist. **81** (2010), 35–50. MR2667308 (2011c:60232)
26. M. V. Kartashov, *Boundedness, limits, and stability of solutions of a perturbation of a nonhomogeneous renewal equation on a semiaxis*, Teor. Imovir. Matem. Statist. **81** (2009), 65–75; English transl. in Theor. Probability and Math. Statist. **81** (2010), 71–83. MR2667311 (2011f:60154)
27. M. V. Kartashov and V. V. Golomozyĭ, *The mean coupling time for independent discrete renewal processes*, Teor. Imovir. Matem. Statist. **84** (2011), 78–85; English transl. in Theor. Probability and Math. Statist. **84** (2012), 79–86. MR2857418 (2012f:60306)
28. M. V. Kartashov and V. V. Golomozyĭ, *Maximal coupling procedure and stability of discrete Markov chains. I*, Teor. Imovir. Matem. Statist. **86** (2012), 81–92; English transl. in Theor. Probability and Math. Statist. **86** (2013), 93–104. MR3241447
29. M. V. Kartashov and V. V. Golomozyĭ, *Maximal coupling procedure and stability of discrete Markov chains. II*, Teor. Imovir. Matem. Statist. **87** (2012), 58–70; English transl. in Theor. Probability and Math. Statist. **87** (2013), 65–78. MR3241447
30. M. V. Kartashov and V. V. Golomozyĭ, *On coupling moment integrability for time-inhomogeneous Markov chains*, Teor. Imovir. Matem. Statist. **89** (2014), 1–12; English transl. in Theor. Probability and Math. Statist. **89** (2014), 1–12. MR3235170
31. V. V. Golomozyĭ, *An inequality for the coupling moment in the case of two inhomogeneous Markov chains*, Teor. Imovir. Matem. Statist. **90** (2014), 39–51; English transl. in Theor. Probability and Math. Statist. **90** (2015), 43–56.

Department of Probability Theory, Statistics, and Actuarial Mathematics, Faculty for Mechanics and Mathematics, National Taras Shevchenko University, Volodymyrs'ka Street, 64, Kyiv 01601, Ukraine

*E-mail address*: mailtower@gmail.com

Department of Probability Theory, Statistics, and Actuarial Mathematics, Faculty for Mechanics and Mathematics, National Taras Shevchenko University, Volodymyrs'ka Street, 64, Kyiv 01601, Ukraine

*E-mail address*: mailtower@gmail.com

Received 11/OCT/2011
Translated by N. SEMENOV