# ERGODICITY COEFFICIENTS OF SEVERAL MATRICES: NEW RESULTS AND APPLICATIONS

UDREA PĂUN

We give new results about and applications of ergodicity coefficients of several matrices. The results refer to the improvement of some results from [7], [8], and [9] (Section 1). The applications refer to the approximate computation for products of stochastic matrices and probability distributions of finite Markov chains (Section 2) and the new proofs of some known results (Section 3; some results are from homogeneous finite Markov chain theory related to convergence and speed of convergence and one is related to backward products).

*AMS 2010 Subject Classification:* 60J10, 15A15, 15A60, 60E99.

*Key words:* finite Markov chain, $\Delta$-ergodic theory, ergodicity coefficient of several matrices, approximate computation, ergodic approximation, product of stochastic matrices, probability distribution, convergence, speed of convergence, backward product.

## 1. ERGODICITY COEFFICIENTS OF SEVERAL MATRICES

In this section (see also [7] and [9]) we consider some ergodicity coefficients of one or two matrices. Then we improve some results from [7], [8], and [9] related to these ergodicity coefficients. Moreover, we give some new results.

Set

$$\mathrm{Par}(E) = \{\Delta \mid \Delta \text{ is a partition of } E\},$$

where $E$ is a nonempty set. We shall agree that the partitions do not contain the empty set.

Set

$$\begin{aligned}
&R_{m,n} = \{T \mid T \text{ is a real } m \times n \text{ matrix}\},\\
&N_{m,n} = \{T \mid T \text{ is a nonnegative } m \times n \text{ matrix}\},\\
&S_{m,n} = \{T \mid T \text{ is a stochastic } m \times n \text{ matrix}\},\\
&R_m = R_{m,m},\ N_m = N_{m,m}, \text{ and } S_m = S_{m,m}.
\end{aligned}$$

REV. ROUMAINE MATH. PURES APPL., **55** (2010), *1*, 53–77

Let $P = (P_{ij}) \in R_{m,n}$. Let $\emptyset \neq U \subseteq \{1, 2, \ldots, m\}$ and $\emptyset \neq V \subseteq \{1, 2, \ldots, n\}$. Define

$$P_U = (P_{ij})_{i\in U,\, j\in\{1,2,\ldots,n\}}\,, \quad P^V = (P_{ij})_{i\in\{1,2,\ldots,m\},\, j\in V}\,,$$

and

$$P_U^V = (P_{ij})_{i\in U,\, j\in V}\,.$$

Let $P \in R_{m,n}$. Below we give a list of coefficients associated with $P$ (i.e., they are coefficients of one matrix) which are called ergodicity coefficients if $P$ is a stochastic matrix.

$$\alpha(P) = \min_{1\leq i,j\leq m} \sum_{k=1}^{n} \min\left(P_{ik}, P_{jk}\right)$$

(if $P \in S_{m,n}$, then $\alpha(P)$ is called *Dobrushin's ergodicity coefficient of* $P$ (see, e.g., [4, p. 56])),

$$\overline{\alpha}(P) = \frac{1}{2} \max_{1\leq i,j\leq m} \sum_{k=1}^{n} |P_{ik} - P_{jk}|$$

and, for $\Delta \in \mathrm{Par}\left(\{1, 2, \ldots, m\}\right)$,

$$\gamma_\Delta(P) = \min_{\substack{K\in\Delta\\ i,j\in K}} \sum_{k=1}^{n} \min\left(P_{ik}, P_{jk}\right)$$

and

$$\overline{\gamma}_\Delta(P) = \frac{1}{2} \max_{\substack{K\in\Delta\\ i,j\in K}} \sum_{k=1}^{n} |P_{ik} - P_{jk}|$$

(see [7] for $\gamma_\Delta$ and $\overline{\gamma}_\Delta$; obviously, we have

$$\gamma_\Delta(P) = \min_{K\in\Delta} \alpha\left(P_K\right) \quad \text{and} \quad \overline{\gamma}_\Delta(P) = \max_{K\in\Delta} \overline{\alpha}\left(P_K\right)$$

and, if $\Delta = (\{1, 2, \ldots, m\})$, then $\gamma_\Delta = \alpha$ and $\overline{\gamma}_\Delta = \overline{\alpha}$).

We can consider (following [7]) a coefficient which generalizes $\gamma_\Delta$ (also $\alpha$) and other which generalizes $\overline{\gamma}_\Delta$ (also $\overline{\alpha}$). For this, let $P \in R_{m,n}$, $\emptyset \neq X \subseteq \{1, 2, \ldots, m\}^2$, and the coefficients

$$a_X(P) = \min_{(i,j)\in X} \sum_{k=1}^{n} \min\left(P_{ik}, P_{jk}\right)$$

and

$$\overline{a}_X(P) = \frac{1}{2} \max_{(i,j)\in X} \sum_{k=1}^{n} |P_{ik} - P_{jk}|\,.$$

*Remark* 1.1. If $P \in R_{m,n}$ then $\overline{a}_X\left(P\right) = 0$ (equivalently, $a_X(P) = 1$ if $P \in S_{m,n}$) if and only if $P_{\{i,j\}}$ is a stable matrix (i.e., a matrix with identical rows), $\forall(i,j) \in X$. In particular, $\overline{\alpha}(P)=0$ (equivalently, $\alpha(P)=1$ if $P \in S_{m,n}$) if and only if $P$ is a stable matrix.

*Remark* 1.2. If $P \in R_{m,n}$ and $\emptyset \neq K \subseteq \{1,2,\ldots,m\}$, then $a_{K^2}(P) = \alpha(P_K)$ and $\overline{a}_{K^2}(P) = \overline{\alpha}(P_K)$. In particular, for $K=\{1,2,\ldots,m\}$, we have $a_{K^2}(P)=\alpha(P)$ and $\overline{a}_{K^2}(P)=\overline{\alpha}(P)$.

Let $P \in R_{m,n}$ and $\emptyset \neq X \subseteq \{1,2,\ldots,m\}^2$. Define

$S_X(P) = \{k \mid k \in \{1,2,\ldots,n\} \text{ and } \forall(i,j) \in X \text{ we have } P_{ik} = P_{jk}\}.$

We remark that $k \in S_X(P)$ implies that $|P_{ik}-P_{jk}|=0$, $\forall(i,j) \in X$. Therefore,

$$\overline{a}_X(P) = \frac{1}{2}\max_{(i,j)\in X}\sum_{k\in\mathcal{C}S_X(P)}|P_{ik}-P_{jk}|$$

if $\mathcal{C}S_X(P) \neq \emptyset$ ($\mathcal{C}S_X(P)$ is the complement of $S_X(P)$; if $\mathcal{C}S_X(P)=\emptyset$ (equivalently, $S_X(P)=\{1,2,\ldots,n\}$), then $\overline{a}_X(P)=0$). It is possible that we only know a set $Y$, $\emptyset \neq Y \subseteq S_X(P)$. In this case we have

$$\overline{a}_X(P) = \frac{1}{2}\max_{(i,j)\in X}\sum_{k\in V}|P_{ik}-P_{jk}|$$

if $V \neq \emptyset$, where $V := \mathcal{C}Y$ (this implies that $\mathcal{C}S_X(P) \subseteq V \subseteq \{1,2,\ldots,n\}$).

Theorem 1.3. *Let $P \in S_{m,n}$, $Q \in S_{n,p}$, and $\emptyset \neq X \subseteq \{1,2,\ldots,m\}^2$. Then*
(i) $\overline{a}_X(P) = 1 - a_X(P)$;
(ii) $\overline{a}_X(P) = \max_{(i,j)\in X}\max_{I\in\mathcal{P}(\{1,2,\ldots,n\})}\sum_{k\in I}(P_{ik}-P_{jk})$;
(iii) $\overline{a}_X(P) = \overline{a}_X(P^V)$, $\forall V, V \neq \emptyset$ *and* $\mathcal{C}S_X(P) \subseteq V \subseteq \{1,2,\ldots,n\}$;
(iv) $\overline{a}_X(PQ) = \overline{a}_X(P^V Q_V)$, $\forall V, V \neq \emptyset$ *and* $\mathcal{C}S_X(P) \subseteq V \subseteq \{1,2,\ldots,n\}$.

*Proof.* (i) and (ii) See [7, Proposition 1.9].
(iii) Let $V \neq \emptyset$ and $\mathcal{C}S_X(P) \subseteq V \subseteq \{1,2,\ldots,n\}$. We have

$$\overline{a}_X(P) = \frac{1}{2}\max_{(i,j)\in X}\sum_{k\in\mathcal{C}S_X(P)}|P_{ik}-P_{jk}| = \overline{a}_X\left(P^{\mathcal{C}S_X(P)}\right) = \overline{a}_X\left(P^V\right).$$

(iv) Let $V \neq \emptyset$ and $\mathcal{C}S_X(P) \subseteq V \subseteq \{1,2,\ldots,n\}$. By (ii) we have

$$\overline{a}_X(PQ) = \max_{(i,j)\in X}\max_{I\in\mathcal{P}(\{1,2,\ldots,p\})}\sum_{k\in I}\sum_{l=1}^{n}(P_{il}-P_{jl})Q_{lk} =$$

$$= \max_{(i,j)\in X}\max_{I\in\mathcal{P}(\{1,2,\ldots,p\})}\sum_{k\in I}\sum_{l\in\mathcal{C}S_X(P)}(P_{il}-P_{jl})Q_{lk} =$$

$$= \max_{(i,j)\in X}\max_{I\in\mathcal{P}(\{1,2,\ldots,p\})}\sum_{k\in I}\sum_{l\in V}(P_{il}-P_{jl})Q_{lk} = \overline{a}_X\left(P^V Q_V\right). \quad \square$$

Theorem 1.4. *Let* $P \in S_{m,n}, Q \in S_{n,p}$, *and* $\emptyset \neq X \subseteq \{1, 2, \ldots, m\}^2$. *Then*

$$\overline{a}_X (PQ) \leq \overline{a}_X (P) \overline{\alpha} (Q).$$

*Proof.* See [7, Theorem 1.10]. $\square$

Theorem 1.4 can be generalized as follows.

Theorem 1.5. *Let* $P \in S_{m,n}$, $Q \in S_{n,p}$, *and* $\emptyset \neq X \subseteq \{1, 2, \ldots, m\}^2$. *Then* $\overline{a}_X (PQ) \leq \overline{a}_X \left(P^V\right) \overline{\alpha} (Q_V)$, $\forall V, V \neq \emptyset$ *and* $\mathcal{C}S_X(P) \subseteq V \subseteq \{1, 2, \ldots, n\}$.

*Proof.* Let $V \neq \emptyset$ and $\mathcal{C}S_X (P) \subseteq V \subseteq \{1, 2, \ldots, n\}$. Let $U = \mathrm{pr}_1 X \cup \mathrm{pr}_2 X$, where

$$\mathrm{pr}_1 X := \{j \mid j \in \{1, 2, \ldots, m\} \text{ and } \exists k \in \{1, 2, \ldots, m\} \text{ such that } (j, k) \in X\}$$

and

$$\mathrm{pr}_2 X := \{j \mid j \in \{1, 2, \ldots, m\} \text{ and } \exists k \in \{1, 2, \ldots, m\} \text{ such that } (k, j) \in X\}.$$

Obviously, $\exists c \geq 0, \exists R \in S_{|U|,|V|}$ such that $P_U^V = cR$ (i.e., $P_U^V$ is a generalized stochastic matrix (cf. Definition 1.17)). Then

$$\overline{a}_X (PQ) =$$

(by Theorem 1.3(iv))

$$= \overline{a}_X \left(P^V Q_V\right) =$$

(because $\emptyset \neq X \subseteq U^2$ and the labels of rows and columns of matrices are kept when we use the operators $(\cdot)_U$, $(\cdot)^V$, and $(\cdot)_U^V$ defined at the beginning of this section)

$$= \overline{a}_X \left(P_U^V Q_V\right) = \overline{a}_X (cRQ_V) = c\overline{a}_X (RQ_V) \leq$$

(by Theorem 1.4)

$$\leq c\overline{a}_X(R)\overline{\alpha} (Q_V) = \overline{a}_X (cR) \overline{\alpha} (Q_V) =$$

$$= \overline{a}_X \left(P_U^V\right) \overline{\alpha} (Q_V) = \overline{a}_X \left(P^V\right) \overline{\alpha} (Q_V). \quad \square$$

*Remark* 1.6. By Remark 1.2, the inequality from Theorem 1.5 can be written as

$$\overline{a}_X (PQ) \leq \overline{a}_X \left(P^V\right) \overline{a}_{V^2} (Q), \ \forall V, V \neq \emptyset \text{ and } \mathcal{C}S_X(P) \subseteq V \subseteq \{1, 2, \ldots, n\}.$$

Theorem 1.7. *Let* $P \in S_{m,n}, Q \in S_{n,p}$, *and* $\emptyset \neq X \subseteq \{1, 2, \ldots, m\}^2$. *If* $a_X(P) > 0$, *then* $a_X (PQ) > 0$ (*equivalently, if* $\overline{a}_X(P) < 1$, *then* $\overline{a}_X (PQ) < 1$).

*Proof.* By Theorem 1.3(i), since $a_X(P) > 0$, we have $\overline{a}_X(P) < 1$. Further, by Theorem 1.4 we have

$$\overline{a}_X (PQ) \leq \overline{a}_X (P) \overline{\alpha} (Q) \leq \overline{a}_X(P) < 1.$$

Now, using the fact that $\overline{a}_X(PQ) < 1$ and Theorem 1.3(i), we obtain $a_X(PQ) > 0$. $\square$

Let $P, Q \in R_{m,n}$. Below we give a list of coefficients (see [9]) associated with $P$ and $Q$ (i.e., they are coefficients of two matrices) which we call ergodicity coefficients if $P, Q \in S_{m,n}$.

$$\overline{N}_\infty(P,Q) = \frac{1}{2}\,|||P-Q|||_\infty\,,$$
$$N_\infty(P,Q) = 1 - \overline{N}_\infty(P,Q) \quad \text{if } P,Q \in S_{m,n}$$

(*the ergodicity coefficients induced by the matrix norm* $|||\cdot|||_\infty$ (if $T \in R_{m,n}$, then $|||T|||_\infty := \max\limits_{1\le i\le m} \sum\limits_{j=1}^{n} |T_{ij}|$)),

$$\zeta(P,Q) = \min_{1\le i,j\le m} \sum_{k=1}^{n} \min(P_{ik}, Q_{jk})\,,$$
$$\overline{\zeta}(P,Q) = \frac{1}{2} \max_{1\le i,j\le m} \sum_{k=1}^{n} |P_{ik} - Q_{jk}|$$

($X$ should be equal to $\{1,2,\ldots,m\}^2$ in [9]) and, for $\Delta \in \mathrm{Par}(\{1,2,\ldots,m\})$,

$$\theta_\Delta(P,Q) = \min_{\substack{K\in\Delta\\ i,j\in K}} \sum_{k=1}^{n} \min(P_{ik}, Q_{jk})$$

and

$$\overline{\theta}_\Delta(P,Q) = \frac{1}{2} \max_{\substack{K\in\Delta\\ i,j\in K}} \sum_{k=1}^{n} |P_{ik} - Q_{jk}|$$

($X$ should be equal to $\{(i,j) \mid (i,j) \in \{1,2,\ldots,m\}^2$ and $\exists K \in \Delta$ such that $i,j \in K\}$ in [9]; obviously (see [9]), we have

$$\theta_\Delta(P,Q) = \min_{K\in\Delta} \zeta(P_K, Q_K)$$

and

$$\overline{\theta}_\Delta(P,Q) = \max_{K\in\Delta} \overline{\zeta}(P_K, Q_K)\,).$$

*Remark* 1.8. Obviously, we have ($P \in R_{m,n}$)

$$\zeta(P,P) = \alpha(P), \quad \overline{\zeta}(P,P) = \overline{\alpha}(P),$$
$$\theta_\Delta(P,P) = \gamma_\Delta(P), \quad \overline{\theta}_\Delta(P,P) = \overline{\gamma}_\Delta(P),$$
$$\theta_{(\{1,2,\ldots,m\})} = \zeta \quad \text{and} \quad \overline{\theta}_{(\{1,2,\ldots,m\})} = \overline{\zeta}.$$

Also, as in the case of $\gamma_\Delta$ and $\overline{\gamma}_\Delta$, we can consider a coefficient which generalizes $\theta_\Delta$ (also $\zeta$ (see Remark 1.8)) and another one which generalizes $\overline{\theta}_\Delta$ (also $\overline{\zeta}$). For this, let $P, Q \in R_{m,n}$ and $\emptyset \neq X \subseteq \{1, 2, \ldots, m\}^2$. Define

$$b_X(P, Q) = \min_{(i,j)\in X} \sum_{k=1}^{n} \min(P_{ik}, Q_{jk})$$

and

$$\overline{b}_X(P, Q) = \frac{1}{2} \max_{(i,j)\in X} \sum_{k=1}^{n} |P_{ik} - Q_{jk}|.$$

Note that $b_X(P, P) = a_X(P)$ and $\overline{b}_X(P, P) = \overline{a}_X(P)$, $\forall P \in R_{m,n}$, $\forall X$, $\emptyset \neq X \subseteq \{1, 2, \ldots, m\}^2$. Also, note that if $P, Q \in R_{m,n}$, then $\overline{b}_X(P, Q) = 0$ (equivalently, $b_X(P, Q) = 1$ if $P, Q \in S_{m,n}$) if and only if $P_{\{i\}} = Q_{\{j\}}$, $\forall (i, j) \in X$. In particular, $\overline{\zeta}(P, Q) = 0$ (equivalently, $\zeta(P, Q) = 1$ if $P, Q \in S_{m,n}$) if and only if there exists a stable matrix $\Pi$ such that $P = Q = \Pi$.

Let $\sigma$ and $\tau$ be two probability distributions on $\{1, 2, \ldots, m\}$. The *total variation distance between* $\sigma$ *and* $\tau$, denoted $\|\sigma - \tau\|$, is defined as

$$\|\sigma - \tau\| = \frac{1}{2} \sum_{i=1}^{m} |\sigma_i - \tau_i|$$

(see, e.g., [1, pp. 109−110]). (Therefore, $\|\sigma - \tau\| = \frac{1}{2} \|\sigma - \tau\|_1$.)

This notion suggests the next definition.

*Definition* 1.9. Let $P, Q \in S_{m,n}$ and $\emptyset \neq X \subseteq \{1, 2, \ldots, m\}^2$. We say that $\overline{b}_X(P, Q)$ is the *total* $X$*-variation distance between* $P$ *and* $Q$. *In particular, for* $X = \{1, 2, \ldots, m\}^2$, $\overline{b}_X(P, Q) = \overline{\zeta}(P, Q)$ *and we say that* $\overline{\zeta}(P, Q)$ *is the total variation distance between* $P$ *and* $Q$ *for short.*

Let $P, Q \in R_{m,n}$ and $\emptyset \neq X \subseteq \{1, 2, \ldots, m\}^2$. Define

$$\Sigma_X(P, Q) = \{k \mid k \in \{1, 2, \ldots, n\} \text{ and } \forall (i, j) \in X \text{ we have } P_{ik} = Q_{jk}\}.$$

We have, obviously, $\Sigma_X(P, P) = S_X(P)$ and

$$\overline{b}_X(P, Q) = \frac{1}{2} \max_{(i,j)\in X} \sum_{k \in \mathcal{C}\Sigma_X(P,Q)} |P_{ik} - Q_{jk}|$$

if $\mathcal{C}\Sigma_X(P, Q) \neq \emptyset$ (if $\mathcal{C}\Sigma_X(P, Q) = \emptyset$ (equivalently, $\Sigma_X(P, Q) = \{1, 2, \ldots, n\}$), then $\overline{b}_X(P, Q) = 0$). It is possible that we only know a set $Z$, $\emptyset \neq Z \subseteq \Sigma_X(P, Q)$. In this case we have

$$\overline{b}_X(P, Q) = \frac{1}{2} \max_{(i,j)\in X} \sum_{k \in W} |P_{ik} - Q_{jk}|$$

if $W \neq \emptyset$, where $W := \mathcal{C}Z$ (this implies that $\mathcal{C}\Sigma_X(P,Q) \subseteq W \subseteq \{1,2,\ldots,n\}$).

THEOREM 1.10. *Let* $P,Q \in S_{m,n},\ R \in S_{n,p},$ *and* $\emptyset \neq X \subseteq \{1,2,\ldots,m\}^2$. *Then*

(i) $\overline{b}_X(P,Q) = 1 - b_X(P,Q)$;

(ii) $\overline{b}_X(P,Q) = \max\limits_{(i,j)\in X} \max\limits_{I\in\mathcal{P}(\{1,2,\ldots,n\})} \sum\limits_{k\in I} (P_{ik} - Q_{jk})$;

(iii) $\overline{b}_X(P,Q) = \overline{b}_X\left(P^W, Q^W\right)$, $\forall W, W \neq \emptyset$ and $\mathcal{C}\Sigma_X(P,Q) \subseteq W \subseteq \{1,2,\ldots,n\}$;

(iv) $\overline{b}_X\left(PR,QR\right) = \overline{b}_X\left(P^W R_W, Q^W R_W\right)$, $\forall W, W \neq \emptyset$ and $\mathcal{C}\Sigma_X(P,Q) \subseteq W \subseteq \{1,2,\ldots,n\}$.

*Proof.* (i) and (ii) See [9, Proposition 1.4].

(iii) Let $W \neq \emptyset$ and $\mathcal{C}\Sigma_X(P,Q) \subseteq W \subseteq \{1,2,\ldots,n\}$. We have

$$\overline{b}_X(P,Q) = \frac{1}{2}\max_{(i,j)\in X} \sum_{k\in\mathcal{C}\Sigma_X(P,Q)} |P_{ik} - Q_{jk}| = \overline{b}_X\left(P^W, Q^W\right).$$

(iv) Let $W \neq \emptyset$ and $\mathcal{C}\Sigma_X(P,Q) \subseteq W \subseteq \{1,2,\ldots,n\}$. By (ii) we have

$$\overline{b}_X\left(PR,QR\right) = \max_{(i,j)\in X} \max_{I\in\mathcal{P}(\{1,2,\ldots,p\})} \sum_{k\in I}\sum_{l=1}^{n} \left(P_{il} - Q_{jl}\right) R_{lk} =$$

$$= \max_{(i,j)\in X} \max_{I\in\mathcal{P}(\{1,2,\ldots,p\})} \sum_{k\in I}\sum_{l\in\mathcal{C}\Sigma_X(P,Q)} \left(P_{il} - Q_{jl}\right) R_{lk} =$$

$$= \max_{(i,j)\in X} \max_{I\in\mathcal{P}(\{1,2,\ldots,p\})} \sum_{k\in I}\sum_{l\in W} \left(P_{il} - Q_{jl}\right) R_{lk} = \overline{b}_X\left(P^W R_W, Q^W R_W\right). \quad \square$$

THEOREM 1.11 ([9]). *Let* $P,Q \in S_{m,n}$, $R \in S_{n,p}$, *and* $\emptyset \neq X \subseteq \{1,2,\ldots,m\}^2$. *Then*

$$\overline{b}_X\left(PR,QR\right) \leq \overline{b}_X\left(P,Q\right)\overline{\alpha}(R).$$

*Proof.* See [9]. $\square$

Theorem 1.11 can be improved as follows.

THEOREM 1.12. *Let* $P,Q \in S_{m,n}$, $R \in S_{n,p}$, *and* $\emptyset \neq X \subseteq \{1,2,\ldots,m\}^2$. *Then*

$$\overline{b}_X\left(PR,QR\right) \leq \overline{b}_X\left(P^W, Q^W\right)\overline{\alpha}\left(R_W\right),$$

$$\forall W,\ W \neq \emptyset \text{ and } \mathcal{C}\Sigma_X\left(P,Q\right) \subseteq W \subseteq \{1,2,\ldots,n\}.$$

*Proof.* Let $W \neq \emptyset$ and $\mathcal{C}\Sigma_X(P,Q) \subseteq W \subseteq \{1,2,\ldots,n\}$. As in the proof of Theorem 1.5, setting $U = \mathrm{pr}_1 X \cup \mathrm{pr}_2 X$, we have $P_U^W = cE$ and $Q_U^W = cF$, where $c \geq 0$ and $P_U^W, Q_U^W \in S_{|U|,|W|}$. By Theorems 1.10(iv) and 1.11 we have

$$\overline{b}_X(PR, QR) = \overline{b}_X\left(P^W R_W, Q^W R_W\right) = \overline{b}_X\left(P_U^W R_W, Q_U^W R_W\right) =$$
$$= \overline{b}_X(cER_W, cFR_W) = c\overline{b}_X(ER_W, FR_W) \leq c\overline{b}_X(E,F)\,\overline{\alpha}(R_W) =$$
$$= \overline{b}_X(cE, cF)\,\overline{\alpha}(R_W) = \overline{b}_X\left(P_U^W, Q_U^W\right)\overline{\alpha}(R_W) = \overline{b}_X\left(P^W, Q^W\right)\overline{\alpha}(R_W). \quad \square$$

In this paper, a vector $x \in \mathbf{R}^n$ is a row vector and $x'$ denotes its transpose. Set $e = e(n) = (1, 1, \ldots, 1) \in \mathbf{R}^n$.

THEOREM 1.13. (i) *If* $\xi \in \mathbf{R}^n$ *such that* $\xi e' = 0$ *and* $R \in S_{n,p}$, *then*

$$\|\xi R\|_1 \leq \|\xi\|_1\,\overline{\alpha}(R)$$

(*an inequality of Dobrushin* (*see, e.g.,* [4, p. 59], *or* [5, p. 147], *or* [9])).

(ii) (*a generalization of* (i)) *If* $T \in R_{m,n}$ *such that* $Te' = 0$ *and* $R \in S_{n,p}$, *then*

$$|||TR|||_\infty \leq |||T|||_\infty\,\overline{\alpha}(R)$$

(*see, e.g.,* [5, p. 147] *or* [9]).

(iii) *If* $P, Q \in S_{m,n}$ *and* $R \in S_{n,p}$, *then*

$$\overline{N}_\infty(PR, QR) \leq \overline{N}_\infty(P,Q)\overline{\alpha}(R)$$

(*see* [9]; *therefore* $\overline{N}_\infty$, *too, enjoys a property as in Theorem* 1.11).

*Proof.* See [9] or, for (i), see, e.g., [4, p. 59], or [5, p. 147], for (ii), see, e.g., [5, p. 147], and for (iii), use (ii) taking $T = P - Q$. $\square$

Let $T \in R_{m,n}$. Define

$$Z(T) = \left\{ j \mid j \in \{1, 2, \ldots, n\} \text{ and } T^{\{j\}} = 0 \right\}$$

(the *set of zero columns of* $T$).

Theorem 1.13 can be improved as follows.

THEOREM 1.14. (i) *If* $\xi \in \mathbf{R}^n$ *such that* $\xi e' = 0$ *and* $R \in S_{n,p}$, *then*

$$\|\xi R\|_1 \leq \left\|\xi^W\right\|_1\overline{\alpha}(R_W),\ \forall W,\ W \neq \emptyset \text{ and } \mathcal{C}Z(\xi) \subseteq W \subseteq \{1, 2, \ldots, n\}.$$

(ii) (*a generalization of* (i)) *If* $T \in R_{m,n}$ *such that* $Te' = 0$ *and* $R \in S_{n,p}$, *then*

$$|||TR|||_\infty \leq \left|\left\|T^W\right\|\right|_\infty\overline{\alpha}(R_W),\ \forall W,\ W \neq \emptyset \text{ and } \mathcal{C}Z(T) \subseteq W \subseteq \{1, 2, \ldots, n\}.$$

(iii) *If* $P, Q \in S_{m,n}$ (*more generally,* $P, Q \in R_{m,n}$ *with* $(P - Q)e' = 0$) *and* $R \in S_{n,p}$, *then*

$$\overline{N}_\infty(PR, QR) \leq \overline{N}_\infty\left(P^W, Q^W\right)\overline{\alpha}(R_W),$$
$$\forall W,\ W \neq \emptyset \text{ and } \mathcal{C}Z(P - Q) \subseteq W \subseteq \{1, 2, \ldots, n\}.$$

*Proof.* This is left to the reader. $\square$

*Remark* 1.15. By Theorem 1.12, taking $X = \{1, 2, \ldots, m\}^2$, and Theorem 1.14(iii) we obtain two analogous inequality, namely,

$$\overline{\zeta}\,(PR, QR) \leq \overline{\zeta}\left(P^W, Q^W\right)\overline{\alpha}\,(R_W)$$

and

$$\overline{N}_\infty\,(PR, QR) \leq \overline{N}_\infty\left(P^W, Q^W\right)\overline{\alpha}\,(R_W)\,,$$
$$\forall W, W \neq \emptyset \text{ and } \mathcal{C}\Sigma_X\,(P, Q) \subseteq W \subseteq \{1, 2, \ldots, n\}$$

(here $\Sigma_X(P, Q) \subseteq Z\,(P - Q)$). Obviously, the first inequality contains more information than the second one (cf. the definitions of $\overline{\zeta}$ and $\overline{N}_\infty$). Moreover,

$$\overline{N}_\infty \leq \overline{\zeta}$$

(see [9]). If $P$ or $Q$ is a stable matrix, then

$$\overline{N}_\infty(P, Q) = \overline{\zeta}\,(P, Q)\,.$$

This means that in cases such as this it does not matter if we use $\overline{N}_\infty$ or $\overline{\zeta}$. The converse is not true. Indeed, if

$$P = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

then $\overline{N}_\infty(P, Q) = \overline{\zeta}(P, Q) = 1$, but neither $P$ nor $Q$ is a stable matrix.

In the next example we compare Theorems 1.13(i) and 1.14(i).

*Example* 1.16. Let

$$P = \begin{pmatrix} 0 & 0 & \frac{2}{4} & \frac{2}{4} \\ 0 & 0 & \frac{1}{4} & \frac{3}{4} \\ \frac{1}{4} & \frac{3}{4} & 0 & 0 \\ \frac{3}{4} & \frac{1}{4} & 0 & 0 \end{pmatrix}$$

and let $\xi_1$ and $\xi_2$ be two probability distributions on $\{1, 2, 3, 4\}$ with $\operatorname{supp}\xi_1$, $\operatorname{supp}\xi_2 \subseteq \{1, 2\}$. (If $\pi$ is a probability distribution on $\{1, 2, \ldots, r\}$, then

$$\operatorname{supp}\pi := \{i \mid i \in \{1, 2, \ldots, r\} \text{ and } \pi_i > 0\}\,.)$$

By Theorem 1.13(i) we have

$$\|\xi_1 P - \xi_2 P\|_1 \leq \|\xi_1 - \xi_2\|_1\,\overline{\alpha}(P) = \|\xi_1 - \xi_2\|_1$$

while by Theorem 1.14(i) we have

$$\|\xi_1 P - \xi_2 P\|_1 \leq \|\xi_1 - \xi_2\|_1\,\overline{\alpha}\left(P_{\{1,2\}}\right) = \frac{1}{4}\,\|\xi_1 - \xi_2\|_1\,.$$

Therefore, in the last case we have an upper bound four times smaller than in the first case.

*Definition* 1.17. Let $P \in N_{m,n}$. We say that $P$ is a *generalized stochastic matrix* if $\exists a \geq 0$, $\exists Q \in S_{m,n}$ such that $P = aQ$.

Let $\Delta_1 \in \mathrm{Par}(\{1,2,\ldots,m\})$ and $\Delta_2 \in \mathrm{Par}(\{1,2,\ldots,n\})$. Define

$$G_{\Delta_1,\Delta_2} = \{P \mid P \in S_{m,n} \text{ and } \forall K \in \Delta_1,\ \forall L \in \Delta_2,\ P_K^L \text{ is a generalized stochastic matrix}\}.$$

In particular, if $m = n$ and $\Delta_1 = \Delta_2 := \Delta$, we set $G_\Delta = G_{\Delta,\Delta}$ (as in [9] (see also [8])).

In the next theorem we give a result similar to Theorem 1.11, but using $\overline{\gamma}_\Delta$ in place of $\overline{\alpha}$. It improves Theorem 1.9 in [9] (Note that Theorem 1.9 in [9] is too restrictive and contains a mistake, namely, '$i \leq j$', that should be replaced by '$(i,j) \in \{1,2,\ldots,m\}^2$'.)

THEOREM 1.18. *Let $P, Q \in S_{m,n}$, $R \in S_{n,p}$, and $\emptyset \neq X \subseteq \{1,2,\ldots,m\}^2$. Let $\Delta_2 \in \mathrm{Par}(\{1,2,\ldots,n\})$. If we have $\sum\limits_{k\in L} P_{ik} = \sum\limits_{k\in L} Q_{jk}$, $\forall (i,j) \in X$, $\forall L \in \Delta_2$, then*

$$\overline{b}_X(PR, QR) \leq \overline{b}_X(P,Q)\,\overline{\gamma}_{\Delta_2}(R).$$

*In particular, for $X = \{(i,j) \mid (i,j) \in \{1,2,\ldots,m\}^2$ and $\exists K \in \Delta_1$ such that $i,j \in K\}$ (in this case, $P, Q \in G_{\Delta_1,\Delta_2}$), where $\Delta_1 \in \mathrm{Par}(\{1,2,\ldots,m\})$, we have*

$$\overline{\theta}_{\Delta_1}(PR, QR) \leq \overline{\theta}_{\Delta_1}(P,Q)\overline{\gamma}_{\Delta_2}(R),$$

*which for $P = Q := C$ and $R := D$ yields*

$$\overline{\gamma}_{\Delta_1}(CD) \leq \overline{\gamma}_{\Delta_1}(C)\,\overline{\gamma}_{\Delta_2}(D).$$

*Proof.* See the proof of Theorem 1.9 in [9]. □

Let $\Delta \in \mathrm{Par}(E)$ and $\emptyset \neq D \subseteq E$, where $E$ is a nonempty set. Define

$$\Delta \cap D = \{K \cap D \mid K \in \Delta\}.$$

Obviously, $\Delta \cap D \in \mathrm{Par}(D)$ and this is the *partition induced on $D$ by $\Delta$*.

Let $P, Q \in R_{m,n}$, $\emptyset \neq X \subseteq \{1,2,\ldots,m\}^2$, and $\Delta \in \mathrm{Par}(\{1,2,\ldots,m\})$. If $X = \{1,2,\ldots,m\}^2$, then

$$\overline{b}_X(P,Q) = \overline{\zeta}(P,Q) \quad \text{and} \quad \Sigma_X(P,Q) \subseteq Z(P-Q)$$

(moreover, $P^{Z(P-Q)} = Q^{Z(P-Q)}$). If $X = \{(i,j) \mid (i,j) \in \{1,2,\ldots,m\}^2$ and $\exists K \in \Delta$ such that $i,j \in K\}$, then

$$\overline{b}_X(P,Q) = \overline{\theta}_\Delta(P,Q) \text{ and } \Sigma_X(P,Q) = \bigcap_{K\in\Delta} \Sigma_{K^2}(P_K,Q_K) \subseteq \bigcap_{K\in\Delta} Z((P-Q)_K)$$

(moreover, $P_K^{Z((P-Q)_K)} = Q_K^{Z((P-Q)_K)}$, $\forall K \in \Delta$).

Theorem 1.18 can be improved as follows.

THEOREM 1.19. *Let* $P, Q \in S_{m,n}$, $R \in S_{n,p}$, *and* $\emptyset \neq X \subseteq \{1,2,\ldots,m\}^2$. *Let* $\Delta_2 \in \operatorname{Par}(\{1,2,\ldots,n\})$. *If we have* $\sum\limits_{k\in L} P_{ik} = \sum\limits_{k\in L} Q_{jk}$, $\forall (i,j) \in X$, $\forall L \in \Delta_2$, *then*

$$\bar{b}_X(PR, QR) \leq \bar{b}_X\left(P^W, Q^W\right)\overline{\gamma}_{\Delta_2\cap W}(R_W),$$
$$\forall W,\ W \neq \emptyset \text{ and } \mathcal{C}\Sigma_X(P,Q) \subseteq W \subseteq \{1,2,\ldots,n\}.$$

*In particular, for* $X = \{(i,j) \mid (i,j) \in \{1,2,\ldots,m\}^2$ *and* $\exists K \in \Delta_1$ *such that* $i,j \in K\}$ *(in this case,* $P, Q \in G_{\Delta_1,\Delta_2}$*), where* $\Delta_1 \in \operatorname{Par}(\{1,2,\ldots,m\})$, *we have*

$$\overline{\theta}_{\Delta_1}(PR, QR) \leq \overline{\theta}_{\Delta_1}\left(P^W, Q^W\right)\overline{\gamma}_{\Delta_2\cap W}(R_W),$$
$$\forall W,\ W \neq \emptyset \text{ and } \mathcal{C}\Sigma_X(P,Q) \subseteq W \subseteq \{1,2,\ldots,n\},$$

*which for* $P = Q := C$ *and* $R := D$ *yields*

$$\overline{\gamma}_{\Delta_1}(CD) \leq \overline{\gamma}_{\Delta_1}\left(C^W\right)\overline{\gamma}_{\Delta_2\cap W}(D_W),$$
$$\forall W,\ W \neq \emptyset \text{ and } \mathcal{C}S_X(C) \subseteq W \subseteq \{1,2,\ldots,n\}.$$

*Proof.* Let $W \neq \emptyset$ and $\mathcal{C}\Sigma_X(P,Q) \subseteq W \subseteq \{1,2,\ldots,n\}$. By Theorems 1.10(iv) and 1.18 we have

$$\bar{b}_X(PR, QR) = \bar{b}_X\left(P^W R_W, Q^W R_W\right) \leq$$

(as in the proof of Theorem 1.12)

$$\leq \bar{b}_X\left(P^W, Q^W\right)\overline{\gamma}_{\Delta_2\cap W}(R_W).$$

If $X = \{(i,j) \mid (i,j) \in \{1,2,\ldots,m\}^2$ and $\exists K \in \Delta_1$ such that $i,j \in K\}$ (in this case, $P^W, Q^W \in G_{\Delta_1,\Delta_2\cap W}$), then $\bar{b}_X = \overline{\theta}_{\Delta_1}$. □

Theorem 1.18 or its generalization, Theorem 1.19, can be applied, e.g., to the case $P, Q \in S_{\Delta_1,\Delta_2,f}$, where $\Delta_1 = (K_1, K_2, \ldots, K_u) \in \operatorname{Par}(\{1,2,\ldots,m\})$, $\Delta_2 = (L_1, L_2, \ldots, L_v) \in \operatorname{Par}(\{1,2,\ldots,n\})$, $f : \{1,2,\ldots,u\} \to \{1,2,\ldots,v\}$, and

$$S_{\Delta_1,\Delta_2,f} = \left\{T \mid T \in S_{m,n} \text{ and } T_{K_i}^{\mathcal{C}L_{f(i)}} = 0,\ \forall i \in \{1,2,\ldots,u\}\right\},$$

when $\emptyset \neq X \subseteq \{(i,j) \mid (i,j) \in \{1,2,\ldots,m\}^2$ and $\exists K \in \Delta_1$ such that $i,j \in K\}$. ($S_{\Delta_1,\Delta_2,f}$ is a generalization of $S_{\Delta,\sigma}$ in [7].)

In the next theorem (the its first part is a special case of Theorem 1.18 in [8]) we give a result similar to Theorem 1.13, but using $\overline{\gamma}_\Delta$ in place of $\overline{\alpha}$.

THEOREM 1.20 ([9]). *Let* $\Delta \in \operatorname{Par}(\{1,2,\ldots,n\})$. *If* $\xi \in \mathbf{R}^n$ *such that* $\xi^K (e')_K = 0$, $\forall K \in \Delta$, *and* $R \in S_{n,p}$, *then*

$$\|\xi R\|_1 \leq \|\xi\|_1 \overline{\gamma}_\Delta(R).$$

*More generally, if* $T \in R_{m,n}$ *such that* $T^K (e')_K = 0$, $\forall K \in \Delta$, *and* $R \in S_{n,p}$, *then*

$$|||TR|||_\infty \leq |||T|||_\infty \overline{\gamma}_\Delta(R).$$

*In particular, for* $T := C - D$ *and* $R := E$, *where* $C, D \in S_{m,n}$ *(more generally,* $C, D \in R_{m,n}$*),* $\sum_{j \in K} C_{ij} = \sum_{j \in K} D_{ij}$, $\forall i \in \{1, 2, \ldots, m\}$, $\forall K \in \Delta$, *and* $E \in S_{n,p}$, *we have*

$$\overline{N}_\infty (CE, DE) \leq \overline{N}_\infty (C, D) \overline{\gamma}_\Delta(E).$$

*Proof.* See [9, Proposition 1.7]. $\square$

Theorem 1.20 can be improved as follows.

THEOREM 1.21. *Let* $\Delta \in \operatorname{Par}(\{1, 2, \ldots, n\})$. *If* $\xi \in \mathbf{R}^n$ *such that* $\xi^K (e')_K = 0$, $\forall K \in \Delta$, *and* $R \in S_{n,p}$, *then*

$$\|\xi R\|_1 \leq \left\|\xi^W\right\|_1 \overline{\gamma}_{\Delta \cap W} (R_W), \ \forall W, \ W \neq \emptyset \textit{ and } \mathcal{C}Z (\xi) \subseteq W \subseteq \{1, 2, \ldots, n\}.$$

*More generally, if* $T \in R_{m,n}$ *such that* $T^K (e')_K = 0$, $\forall K \in \Delta$, *and* $R \in S_{n,p}$, *then*

$$|||TR|||_\infty \leq \left|\left|\left|T^W\right|\right|\right|_\infty \overline{\gamma}_{\Delta \cap W} (R_W),$$
$$\forall W, \ W \neq \emptyset \textit{ and } \mathcal{C}Z (T) \subseteq W \subseteq \{1, 2, \ldots, n\}.$$

*In particular, for* $T := C - D$ *and* $R := E$, *where* $C, D \in S_{m,n}$ *(more generally,* $C, D \in R_{m,n}$*),* $\sum_{j \in K} C_{ij} = \sum_{j \in K} D_{ij}$, $\forall i \in \{1, 2, \ldots, m\}$, $\forall K \in \Delta$, *and* $E \in S_{n,p}$, *we have*

$$\overline{N}_\infty (CE, DE) \leq \overline{N}_\infty \left(C^W, D^W\right) \overline{\gamma}_{\Delta \cap W} (E_W),$$
$$\forall W, \ W \neq \emptyset \textit{ and } \mathcal{C}Z (C - D) \subseteq W \subseteq \{1, 2, \ldots, n\}.$$

*Proof.* Let $W \neq \emptyset$ and $\mathcal{C}Z (T) \subseteq W \subseteq \{1, 2, \ldots, n\}$. By Theorem 1.20 ($T^K (e')_K = 0$, $\forall K \in \Delta \cap W$) we have

$$|||TR|||_\infty = \left|\left|\left|T^W R_W\right|\right|\right|_\infty \leq \left|\left|\left|T^W\right|\right|\right|_\infty \overline{\gamma}_{\Delta \cap W} (R_W). \quad \square$$

## 2. APPLICATIONS TO APPROXIMATE COMPUTATION

In this section we use the ergodicity coefficients to the approximate computation for products of stochastic matrices and probability distributions of finite Markov chains.

Consider a finite Markov chain $(X_n)_{n \geq 0}$ with state space $S = \{1, 2, \ldots, r\}$, initial distribution $p_0$, and transition matrices $(P_n)_{n \geq 1}$. We frequently shall refer to it as the (finite) Markov chain $(P_n)_{n\geq 1}$. For all integers $m \geq 0$, $n > m$, define

$$P_{m,n} = P_{m+1}P_{m+2}\dots P_n = ((P_{m,n})_{ij})_{i,j\in S}\,.$$

Let $\emptyset \neq B \subseteq \mathbf{N}$. We give below some definitions from $\Delta$-ergodic theory in a special case (for a more general framework, see [10] and [11]).

*Definition* 2.1 ([10]). Let $i, j \in S$. We say that $i$ and $j$ are in the same *weakly ergodic class on* (*time set*) $B$ if $\forall m \in B$, $\forall k \in S$ we have

$$\lim_{n\to\infty} [(P_{m,n})_{ik} - (P_{m,n})_{jk}] = 0.$$

Write $i \overset{B}{\sim} j$ when $i$ and $j$ are in the same weakly ergodic class on $B$. Then $\overset{B}{\sim}$ is an equivalence relation and determines a partition $\Delta = \Delta(B) = (C_1, C_2, \dots, C_s)$ of $S$. The sets $C_1, C_2, \dots, C_s$ are called *weakly ergodic classes on* $B$.

*Definition* 2.2 ([10]). Let $\Delta = (C_1, C_2, \dots, C_s)$ be the partition of weakly ergodic classes on $B$ of a Markov chain. We say that the chain is *weakly* $\Delta$-*ergodic on* $B$.

In connection with the above notions and notation we mention some special cases:

1. $\Delta = (S)$. In this case, a weakly $(S)$-ergodic chain on $B$ can be called *weakly ergodic on* $B$ for short.
2. $B = \{m\}$. In this case, a weakly $\Delta$-ergodic chain on $\{m\}$ can be called *weakly* $\Delta$-*ergodic at time* $m$. An important case is $m = 0$. (E.e., we need the asymptotic behaviour of $(P_{0,n})_{n\geq 1}$ to determine the limit distribution $\pi$, when it exists, of the Markov chain $(P_n)_{n\geq 1}$ because $\lim_{n\to\infty} p_0 P_{0,n} = \pi$ (see also [12]).)
3. $B = \mathbf{N}$. In this case, a weakly $\Delta$-ergodic chain on $\mathbf{N}$ can be called *weakly* $\Delta$-*ergodic* for short.
4. $\Delta = (S)$, $B = \{m\}$. In this case, a weakly $(S)$-ergodic chain at time $m$ can be called *weakly ergodic at time* $m$ for short.
5. $\Delta = (S)$, $B = \mathbf{N}$. In this case, a weakly $(S)$-ergodic chain on $\mathbf{N}$ can be called *weakly ergodic* for short.

*Definition* 2.3 ([10]). Let $i, j \in S$. We say that $i$ and $j$ are in the same *uniformly weakly ergodic class on* $B$ if $\forall k \in S$ we have

$$\lim_{n\to\infty} [(P_{m,n})_{ik} - (P_{m,n})_{jk}] = 0$$

uniformly with respect to $m \in B$.

Write $i \overset{u,B}{\sim} j$ when $i$ and $j$ are in the same uniformly weakly ergodic class on $B$. Then $\overset{u,B}{\sim}$ is an equivalence relation and determines a partition

$\Delta = \Delta(B) = (U_1, U_2, \ldots, U_t)$ of $S$. The sets $U_1, U_2, \ldots, U_t$ are called *uniformly weakly ergodic classes on* $B$.

*Definition* 2.4 ([10]). Let $\Delta = (U_1, U_2, \ldots, U_t)$ be the partition of uniformly weakly ergodic classes on $B$ of a Markov chain. We say that the chain is *uniformly weakly* $\Delta$*-ergodic on* $B$.

Since in Definitions 2.3 and 2.4 the case $B = \{m\}$ is nonimportant, we have three special cases for the simplification of language, namely, 1) $\Delta = (S)$, 2) $B = \mathbf{N}$, 3) $\Delta = (S)$, $B = \mathbf{N}$. These are left to the reader (see the corresponding cases from weak $\Delta$-ergodicity on $B$).

*Definition* 2.5 ([10]). Let $C$ be a weakly ergodic class on $B$. We say that $C$ is a *strongly ergodic class on* $B$ if $\forall i \in C$, $\forall m \in B$ the limit

$$\lim_{n\to\infty} (P_{m,n})_{ij} := \pi_{m,j} = \pi_{m,j}(C)$$

exists and does not depend on $i$.

*Definition* 2.6 ([10]). Let $C$ be a uniformly weakly ergodic class on $B$. We say that $C$ is a *uniformly strongly ergodic class on* $B$ if $\forall i \in C$ the limit

$$\lim_{n\to\infty} (P_{m,n})_{ij} := \pi_{m,j} = \pi_{m,j}(C)$$

exists uniformly with respect to $m \in B$ and does not depend on $i$.

*Definition* 2.7 ([10]). Consider a weakly (respectively, uniformly weakly) $\Delta$-ergodic chain on $B$. We say that the chain is *strongly* (respectively, *uniformly strongly*) $\Delta$*-ergodic on* $B$ if any $C \in \Delta$ is a strongly (respectively, uniformly strongly) ergodic class on $B$.

Also, in the last three definitions we can simplify the language when referring to $\Delta$ (Definition 2.7) and $B$ (Definitions 2.5, 2.6, and 2.7). These are left to the reader.

Let $P_1, P_2, \ldots, P_n \in S_r$. We can suppose, e.g., that these stochastic matrices are the first $n$ matrices of a Markov chain $(P_n)_{n\geq 1}$. First, we consider two problems (the finite case):

1) the approximate computation of product $P_{0,n}$ ($P_{0,n} = P_1P_2 \ldots P_n$);

2) the approximate computation of (probability) distribution $p_n$ (therefore of $p_0P_{0,n}$ because $p_n = p_0P_{0,n}$) at time $n$ of a Markov chain $(P_n)_{n\geq 1}$ with initial distribution $p_0$,

when $n$ is large in both cases.

In the first situation, we approximate $P_{0,n}$ by its tail $P_{k,n}$ ($0 < k < n$ and $n - k$ is its length) and, in the second one, $p_n$ by $p_{k,n}$, where $p_{k,n} := p_0P_{k,n}$. To get an approximation within an error $\varepsilon$ ($\varepsilon > 0$) for our problems, one way is to use the ergodicity coefficients of two matrices. This is made in the theorem below.

Let $(P_n)_{n\geq 1}$ be a Markov chain with state space $S=\{1,2,\ldots,r\}$. Set $P_{m,m}=I_r$, $\forall m\geq 0$.

THEOREM 2.8. *Let* $(P_n)_{n\geq 1}$ *be a Markov chain with state space* $S$ *and initial distribution* $p_0$. *Let* $\Delta\in\operatorname{Par}(S)$. *Then*

(i) $\overline{\zeta}\,(P_{0,n},P_{k,n})\leq\overline{\alpha}\left((P_{l,n})_W\right)$, $\forall l,\ k\leq l<n$, $\forall W=W\,(l)$, $W\neq\emptyset$ *and* $\mathcal{C}\Sigma_{S^2}\,(P_{0,l},P_{k,l})\subseteq W\subseteq S$;

(ii) $\overline{N}_\infty\,(P_{0,n},P_{k,n})\leq\overline{\alpha}\left((P_{l,n})_W\right)$, $\forall l,\ k\leq l<n$, $\forall W=W\,(l)$, $W\neq\emptyset$ *and* $\mathcal{C}Z\,(P_{0,l}-P_{k,l})\subseteq W\subseteq S$;

(iii) $\overline{\theta}_\Delta\,(P_{0,n},P_{k,n})\leq\overline{\gamma}_{\Delta\cap W}\left((P_{l,n})_W\right)$, $\forall l,\ k\leq l<n$, *for which the conditions of Theorem* 1.19 *hold with* $P:=P_{0,l}$, $Q:=P_{k,l}$, $R:=P_{l,n}$, $\Delta_1=\Delta_2:=\Delta$, *and* $X:=\left\{(i,j)\mid (i,j)\in S^2 \text{ and } \exists K\in\Delta \text{ such that } i,j\in K\right\}$, $\forall W=W\,(l)$, $W\neq\emptyset$ *and* $\mathcal{C}\Sigma_X\,(P_{0,l},P_{k,l})\subseteq W\subseteq S$;

(iv) $\overline{N}_\infty\,(P_{0,n},P_{k,n})\leq\overline{\gamma}_{\Delta\cap W}\left((P_{l,n})_W\right)$, $\forall l,\ k\leq l<n$, *for which the conditions of Theorem* 1.21 *hold with* $C:=P_{0,l}$, $D:=P_{k,l}$ *and* $E:=P_{l,n}$, $\forall W=W\,(l)$, $W\neq\emptyset$ *and* $\mathcal{C}Z\,(P_{0,l}-P_{k,l})\subseteq W\subseteq S$;

(v) $\|p_n-p_{k,n}\|_1\leq 2\overline{\alpha}\left((P_{l,n})_W\right)$, $\forall l,\ k\leq l<n$, $\forall W=W\,(l)$, $W\neq\emptyset$ *and* $\mathcal{C}Z\,(p_0\,(P_{0,l}-P_{k,l}))\subseteq W\subseteq S$;

(vi) $\|p_n-p_{k,n}\|_1\leq 2\overline{\gamma}_{\Delta\cap W}\left((P_{l,n})_W\right)$, $\forall l,\ k\leq l<n$, *for which the conditions of Theorem* 1.21 *hold with* $\xi:=p_0\,(P_{0,l}-P_{k,l})$, $\forall W=W\,(l)$, $W\neq\emptyset$ *and* $\mathcal{C}Z\,(p_0\,(P_{0,l}-P_{k,l}))\subseteq W\subseteq S$.

*Proof.* For all results we use the decompositions $P_{0,n}=P_{0,l}P_{l,n}$ and $P_{k,n}=P_{k,l}P_{l,n}$.

(i) See Theorem 1.12 ($X=S^2$ and, by Theorem 1.10(i), $0\leq\overline{\zeta}\,(E,F)\leq 1$, $\forall E,F\in S_{m,n}$).

(ii) See Theorem 1.14(iii).

(iii) See Theorem 1.19.

(iv) See Theorem 1.21.

(v) See Theorem 1.14(i).

(vi) See Theorem 1.21. $\square$

*Remark* 2.9. (a) We can make approximations in $|||\cdot|||_\infty, \overline{N}_\infty\,(\cdot), \overline{\zeta}\,(\cdot)$ etc. (See again Remark 1.15.) But (in the case when we have no errors)

$|||P-Q|||_\infty=0$ implies $P=Q$,

$\overline{N}_\infty(P,Q)=0$ implies $P=Q$,

$\overline{\zeta}(P,Q)=0$ implies that there exists a stable stochastic matrix $\Pi$ such that $P=Q=\Pi$ etc. Moreover, for approximation in $\overline{\zeta}\,(\cdot)$, we note that $\overline{\zeta}(P,Q)\leq\varepsilon$ (i.e., $P\simeq Q$ in $\overline{\zeta}\,(\cdot)$ within an error $\varepsilon$) implies that there exist two stable stochastic matrices $\Pi_1$ and $\Pi_2$, e.g., $\Pi_1=e'Q_{\{1\}}$ and $\Pi_2=e'P_{\{1\}}$, such that $\overline{\zeta}\,(P,\Pi_1)\leq\varepsilon$ and $\overline{\zeta}\,(Q,\Pi_2)\leq\varepsilon$. The proof is obvious. Also, we note that $\overline{\zeta}(P,Q)\leq\varepsilon$ implies that there exists a stable stochastic matrix $\Pi$, e.g., $\Pi = \frac{1}{2}\left(e'P_{\{1\}} + e'Q_{\{1\}}\right)$, such that $\overline{\zeta}\,(P,\Pi) \leq \frac{3}{2}\varepsilon$ and $\overline{\zeta}\,(Q,\Pi) \leq \frac{3}{2}\varepsilon$. The proof is as follows. By Remark 1.15 we have

$$\overline{\zeta}\,(P,\Pi) = \overline{N}_\infty\,(P,\Pi) = \frac{1}{2}\,|||P-\Pi|||_\infty = \frac{1}{2}\left|\left|\left|P - \frac{1}{2}\left(e'P_{\{1\}} + e'Q_{\{1\}}\right)\right|\right|\right|_\infty =$$

$$= \frac{1}{4}\,\left|\left|\left|2P - e'P_{\{1\}} - e'Q_{\{1\}}\right|\right|\right|_\infty \leq$$

$$\leq \frac{1}{2}\left(\frac{1}{2}\,\left|\left|\left|P - e'Q_{\{1\}}\right|\right|\right|_\infty + \frac{1}{2}\,|||P-Q|||_\infty + \frac{1}{2}\,\left|\left|\left|Q - e'P_{\{1\}}\right|\right|\right|_\infty\right) \leq \frac{3}{2}\,\varepsilon.$$

There arises a question: Is there a stable stochastic matrix $\Pi$ such that $\overline{\zeta}(P,Q) \leq \varepsilon$ implies $\overline{\zeta}\,(P,\Pi) \leq \varepsilon$ and $\overline{\zeta}\,(Q,\Pi) \leq \varepsilon$? The approximations in ergodicity coefficients we call the *ergodic approximations*. Finally, we note the next basic problems:

I) Given $\varepsilon > 0$ ($\varepsilon < 1$) and $P_1, P_2, \ldots, P_n \in S_r$, what is the greatest $k = k\,(n,\varepsilon)$, $0 \leq k < n$, such that $|||P_{0,n} - P_{k,n}|||_\infty \leq \varepsilon$? (This problem always has a solution because $k \geq 0$.) But such that $\overline{N}_\infty\,(P_{0,n}, P_{k,n}) \leq \varepsilon$? But such that $\overline{\zeta}\,(P_{0,n}, P_{k,n}) \leq \varepsilon$? Etc.

II) Given $\varepsilon > 0$, a probability distribution $p_0 \in \mathbf{R}^r$, and $P_1, P_2, \ldots, P_n \in S_r$, what is the greatest $k = k\,(p_0, n, \varepsilon)$, $0 \leq k < n$, such that $\|p_n - p_{k,n}\|_1 \leq \varepsilon$?

(b) In Theorem 2.8 we can use $P_{m,n}$ instead of $P_{0,n}$ (here we approximate $P_{m,n}$ by $P_{k,n}$, $m < k < n$). But this is an apparent generalization (because we can apply Theorem 2.8 to the chain $(P_n)_{n>m}$).

(c) To approximate $P_{0,n}$ by $P_{k,n}$ in $|||\cdot|||_\infty$ within an error $\varepsilon$, Theorem 2.8, e.g., (i), requires that $\exists l, k \leq l < n$, $\exists W = W\,(l)$, $W \neq \emptyset$ and $\mathcal{C}\Sigma_{S^2}\,(P_{0,l}, P_{k,l}) \subseteq W \subseteq S$, such that $\overline{\alpha}\left((P_{l,n})_W\right) \leq \frac{\varepsilon}{2}$ (recall that $\overline{N}_\infty \leq \overline{\zeta}$). On the other hand, Theorem 2.8, e.g., (i), says that $P_{0,n} \simeq P_{k,n}$ in $|||\cdot|||_\infty$ within an error $2\overline{\alpha}\left((P_{l,n})_W\right)$, $\forall l, k \leq l < n$, $\forall W = W\,(l)$, $W \neq \emptyset$ and $\mathcal{C}\Sigma_{S^2}\,(P_{0,l}, P_{k,l}) \subseteq W \subseteq S$.

(d) If $(P_n)_{n\geq 1}$ is weakly ergodic at time $l$ ($k \leq l < n$), then $\overline{\alpha}\,(P_{l,n}) \to 0$ as $n \to \infty$. Therefore, $\forall \varepsilon > 0$, $\exists n\,(\varepsilon, l) > l$ such that $\overline{\alpha}\,(P_{l,n}) \leq \varepsilon$, $\forall n \geq n\,(\varepsilon, l)$. This implies that, at least theoretically, we can approximate $P_{0,n}$ by $P_{k,n}$ in $\overline{\zeta}\,(\cdot)$ within an error $\varepsilon$, $\forall n \geq n\,(\varepsilon, l)$. A more restrictive condition than '$(P_n)_{n\geq 1}$ weakly ergodic at time $l$' is '$(P_n)_{n\geq 1}$ weakly ergodic'.

(e) Theorem 2.8, (iii), (iv), and (vi), can be applied, e.g., to the cyclic (homogeneous or nonhomogeneous) case. E.g., let

$$P_n = \begin{pmatrix} 0 & 0 & \frac{2}{4} & \frac{2}{4} \\ 0 & 0 & \frac{1}{4} & \frac{3}{4} \\ \frac{1}{4} & \frac{3}{4} & 0 & 0 \\ \frac{2}{4} & \frac{2}{4} & 0 & 0 \end{pmatrix} := P, \quad \forall n \geq 1,$$

and $p_0 = \left(\frac{1}{2}, \frac{1}{2}, 0, 0\right)$. (This is an example of a chain that does not have a limit distribution.) The period of this chain is $d = 2$. Let $\Delta = (\{1,2\}, \{3,4\})$ (the partition of cyclic subclasses of the chain). First, we can approximate $P^n$ by $P^{n-k}$ $(k < n)$ if $n \bmod 2 = (n-k) \bmod 2$. For $l = k$, since $\Sigma_X\left(P^k, I_4\right) = \emptyset$, where $X = \left\{(i,j) \mid (i,j) \in \{1,2,3,4\}^2 \text{ and } \exists K \in \Delta \text{ such that } i,j \in K\right\}$, we just take $W = S = \{1,2,3,4\}$ in Theorem 2.8, (iii) and (iv), and obtain

$$\overline{\theta}_\Delta\big(P^n, P^{n-k}\big) \le \overline{\gamma}_\Delta\big(P^{n-k}\big)$$

and

$$\overline{N}_\infty\big(P^n, P^{n-k}\big) \le \overline{\gamma}_\Delta\big(P^{n-k}\big),$$

respectively. Since the chain is weakly $\Delta$-ergodic, we have $\overline{\gamma}_\Delta\left(P^{n-k}\right) \to 0$ as $n \to \infty$. Second, we can approximate $p_n$ by $p_{k,n}$ if $n \bmod 2 = (n-k) \bmod 2$. For $l = k+1$, because $Z\left(p_0\left(P^{k+1} - P\right)\right) = \{1,2\}$ ($n \bmod 2 = (n-k) \bmod 2 \Rightarrow n = 2u + t$ and $n - k = 2v + t$, where $u, v \ge 0$ and $t \in \{0,1\} \Rightarrow k = 2(u-v)$, i.e., $k$ is an even number $\Rightarrow k+1$ is an odd number), we can take $W = W(l) = \{3,4\}$ in Theorem 2.8(vi). Therefore,

$$\|p_n - p_{k,n}\|_1 \le 2\overline{\gamma}_{(\{3,4\})}\left(\left(P^{n-(k+1)}\right)_{\{3,4\}}\right) = 2\overline{\alpha}\left(\left(P^{n-(k+1)}\right)_{\{3,4\}}\right),$$

where $p_{k,n} = p_0 P^{n-k}$. Since the chain is weakly $\Delta$-ergodic, we have $\overline{\gamma}_{(\{3,4\})}\left(P^{n-(k+1)}\right) \to 0$ as $n \to \infty$.

Second, we consider two other problems (the infinite case):

3) the approximate computation of limit of product $P_{0,n}$, if it exists (more generally, of $P_{m,n}$, $\forall m \in B$ $(\emptyset \ne B \subseteq \mathbf{N})$, if the limit of $P_{m,n}$ does exist, $\forall m \in B$),

4) the approximate computation of limit (probability) distribution of a Markov chain if it exists.

Related to these problems we shall approximate the limit matrix and the limit distribution of a uniformly strongly ergodic Markov chain ($\Delta = (S)$, $B = \mathbf{N}$).

Theorem 2.10 ([3]). *Let $(P_n)_{n\ge1}$ be a Markov chain. Then it is uniformly weakly ergodic if and only if $\exists a > 0$ $(a \le 1)$, $\exists n_0 \ge 1$ such that $\alpha(P_{m,m+n}) \ge a$, $\forall m \ge 0$, $\forall n \ge n_0$ (equivalently, $\exists a > 0$ $(a \le 1)$, $\exists n_0 \ge 1$ such that $\alpha(P_{m,m+n_0}) \ge a$, $\forall m \ge 0$).*

*Proof.* See [3] or [4, pp. 221–222]. □

This theorem can be modified, more precisely, we can isolate $a$ (with the difference that $0 < a \le 1$ above and $0 < a < 1$ below).

Theorem 2.11. *Let $(P_n)_{n\geq 1}$ be a Markov chain. Let $0 < a < 1$. Then it is uniformly weakly ergodic if and only if $\exists n_0 \geq 1$ such that $\alpha(P_{m,m+n}) \geq a$, $\forall m \geq 0$, $\forall n \geq n_0$.*

*Proof.* "$\Rightarrow$" By Theorem 2.10, $\exists b$, $0 < b \leq 1$, $\exists u_0 \geq 1$ such that $\alpha(P_{m,m+u}) \geq b$, $\forall m \geq 0$, $\forall u \geq u_0$. Let $k_0 = \min\left\{k \mid k \geq 1 \text{ and } (1-b)^k \leq 1-a\right\}$ and $n_0 = k_0 u_0$. Then $\forall m \geq 0$, $\forall n \geq n_0$ we have

$$\overline{\alpha}(P_{m,m+n}) \leq \overline{\alpha}(P_{m,m+n_0}) \leq$$

(by Theorem 1.4, taking $X = S^2$, where $S = \{1, 2, \ldots, r\}$ is the state space of $(P_n)_{n\geq 1}$)

$$\leq \overline{\alpha}(P_{m,m+u_0})\,\overline{\alpha}(P_{m+u_0,m+2u_0})\ldots\overline{\alpha}\left(P_{m+(k_0-1)u_0,m+n_0}\right) \leq (1-b)^{k_0} \leq 1-a.$$

Therefore, $\exists n_0 \geq 1$ such that $\alpha(P_{m,m+n}) \geq a$ (see Theorem 1.3(i)), $\forall m \geq 0$, $\forall n \geq n_0$.

"$\Leftarrow$" See Theorem 2.10. $\square$

*Remark* 2.12. The reader can try to find a proof of Theorem 2.11 without using Theorem 2.10 (for this see the proof of the latter).

*Remark* 2.13. There exist generalizations of Theorem 2.10 in [8, Theorems 3.14, (1)$\Leftrightarrow$(4), and 3.16, (1)$\Leftrightarrow$(4)] which use $\gamma_\Delta$ in place of $\alpha$. These lead to some generalizations of Theorem 2.11 which are left to the reader.

The next result is useful for problem 1) of this section in the uniformly weakly ergodic Markov chain case and is based on Theorem 2.11.

Theorem 2.14. *Let $(P_n)_{n\geq 1}$ be a uniformly weakly ergodic Markov chain. Let $0 < \varepsilon' < 1$. Let $n_0 \geq 1$ such that $\alpha(P_{m,m+n_0}) \geq \varepsilon'$, $\forall m \geq 0$ (see Theorem 2.11). Then*

(i) $\overline{\zeta}(P_{m,n}, P_{n-n_0,n}) \leq 1 - \varepsilon' := \varepsilon$, $\forall m \geq 0$, $\forall n \geq m + n_0$.

(ii) $\|p_n - p_{n-n_0,n}\|_1 \leq 2\varepsilon$, $\forall n \geq n_0$.

*Proof.* (i) By Theorem 1.11,

$$\overline{\zeta}(P_{m,n}, P_{n-n_0,n}) \leq \overline{\zeta}(P_{m,n-n_0}, I_r)\,\overline{\alpha}(P_{n-n_0,n}) \leq$$
$$\leq \overline{\alpha}(P_{n-n_0,n}) \leq 1 - \varepsilon' = \varepsilon, \quad \forall m \geq 0,\ \forall n \geq m + n_0.$$

(ii) By Theorem 1.13(i),

$$\|p_n - p_{n-n_0,n}\|_1 = \|p_0(P_{0,n-n_0} - I_r)P_{n-n_0,n}\|_1 \leq$$
$$\leq \|p_0(P_{0,n-n_0} - I_r)\|_1\,\overline{\alpha}(P_{n-n_0,n}) \leq 2\overline{\alpha}(P_{n-n_0,n}) \leq 2\varepsilon, \quad \forall n \geq n_0. \quad \square$$

*Remark* 2.15. By Theorem 2.14, $P_{m,n} \simeq P_{n-n_0,n}$ in $\overline{\zeta}(\cdot)$ within an error $\varepsilon$, $\forall m \geq 0$, $\forall n \geq m + n_0$, $P_{m,n} \simeq P_{n-n_0,n}$ in $\overline{N}_\infty(\cdot)$ within an error $\varepsilon$, $\forall m \geq 0$,

$\forall n \geq m + n_0$ (because $\overline{N}_\infty \leq \overline{\zeta}$), $P_{m,n} \simeq P_{n-n_0,n}$ in $|||\cdot|||_\infty$ within an error $2\varepsilon$, $\forall m \geq 0$, $\forall n \geq m + n_0$, and $p_n \simeq p_{n-n_0,n}$ in $\|\cdot\|_1$ within an error $2\varepsilon$, $\forall n \geq n_0$.

THEOREM 2.16 (see, e.g., [5, pp. 160–163]). *Let $(P_n)_{n\geq 1}$ be a Markov chain. If $(P_n)_{n\geq 1}$ is weakly ergodic and $\sum\limits_{n\geq 1} \|\psi_{n+1} - \psi_n\|_1 < \infty$, where $\psi_n$ is a probability vector satisfying $\psi_n P_n = \psi_n$, $\forall n \geq 1$, then the chain is strongly ergodic.*

*Proof.* See, e.g., [5, pp. 160−163]. □

THEOREM 2.17 ([6]). *Let $(P_n)_{n\geq 1}$ be a Markov chain. Then $(P_n)_{n\geq 1}$ is uniformly strongly ergodic if and only if it is uniformly weakly ergodic and strongly ergodic.*

*Proof.* See [6]. □

Combining Theorems 2.16 and 2.17 we have the next result.

THEOREM 2.18. *Let $(P_n)_{n\geq 1}$ be a Markov chain. If $(P_n)_{n\geq 1}$ is uniformly weakly ergodic and $\sum\limits_{n\geq 1} \|\psi_{n+1} - \psi_n\|_1 < \infty$, where $\psi_n$ is a probability vector satisfying $\psi_n P_n = \psi_n$, $\forall n \geq 1$, then the chain is uniformly strongly ergodic.*

*Proof.* See Theorems 2.16 and 2.17. □

*Remark* 2.19. Let $(P_n)_{n\geq 1}$ be a uniformly weakly ergodic Markov chain. Is $(P_{m,m+n})_{m\geq 0}$ convergent, $\forall n \geq 1$? The answer is 'no'. Indeed, let

$$P_{4n-3} = P_{4n-2} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} := P,$$

$$P_{4n-1} = P_{4n} = \begin{pmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix} := Q, \quad \forall n \geq 1.$$

By Theorem 2.10 (or Theorem 2.11), $(P_n)_{n\geq 1}$ is uniformly weakly ergodic. Further, by Theorem 2.18, $(P_n)_{n\geq 1}$ is uniformly strongly ergodic because it is uniformly weakly ergodic and $\pi P = P$ and $\pi Q = Q$, where $\pi = \left(\frac{1}{2}, \frac{1}{2}\right)$. But, e.g., $\nexists \lim\limits_{m\to\infty} P_m$ and $\nexists \lim\limits_{m\to\infty} P_{m,m+2}$.

The next theorem is our main result related to problem 3).

THEOREM 2.20. *Let $(P_n)_{n\geq 1}$ be a uniformly strongly ergodic Markov chain with limit* $\Pi$ (i.e., $\lim\limits_{n\to\infty} P_{m,n} = \Pi$, $\forall m \geq 0$). *Let $0 < \varepsilon < 1$. Let $n_0 \geq 1$ such that $\alpha\left(P_{m,m+n_0}\right) \geq 1 - \frac{\varepsilon}{2}$, $\forall m \geq 0$* (*see Theorem* 2.11). *Let $0 \leq m_1 \leq m_2 \leq \cdots$ such that $\lim\limits_{t\to\infty} P_{m_t,m_t+n_0}$ exists* ($\exists\, (m_t)_{t\geq 1}$ *with the above* *property because* $(P_{m,m+n_0})_{m\geq 0}$ *is bounded). Let* $\widetilde{\Pi} = \lim_{t\to\infty} P_{m_t,m_t+n_0}$*. Then* $\Pi \simeq \widetilde{\Pi}$ *in* $|||\cdot|||_\infty$ *within an error* $\varepsilon$*.*

*Proof.* Let $\delta > \varepsilon$. It follows from $\lim_{t\to\infty} P_{m_t,m_t+n_0} = \widetilde{\Pi}$ that $\exists t_1 \geq 1$ such that

$$\left|\left|\left|P_{m_t,m_t+n_0} - \widetilde{\Pi}\right|\right|\right|_\infty \leq \frac{\delta-\varepsilon}{2}, \quad \forall t \geq t_1.$$

Further,

$$\left|\left|\left|P_{0,m_t+n_0} - \widetilde{\Pi}\right|\right|\right|_\infty \leq \left|\left|\left|P_{0,m_t+n_0} - P_{m_t,m_t+n_0}\right|\right|\right|_\infty + \left|\left|\left|P_{m_t,m_t+n_0} - \widetilde{\Pi}\right|\right|\right|_\infty \leq$$

(by Theorem 1.13(i))

$$\leq 2\overline{\alpha}\left(P_{m_t,m_t+n_0}\right) + \left|\left|\left|P_{m_t,m_t+n_0} - \widetilde{\Pi}\right|\right|\right|_\infty \leq \varepsilon + \frac{\delta-\varepsilon}{2}, \quad \forall t \geq t_1.$$

Further, $\lim_{t\to\infty} P_{0,m_t+n_0} = \Pi$ implies that $\exists t_2 \geq 1$ such that

$$\left|\left|\left|P_{0,m_t+n_0} - \Pi\right|\right|\right|_\infty \leq \frac{\delta-\varepsilon}{2}, \quad \forall t \geq t_2.$$

Let $t_0 = \max\left(t_1, t_2\right)$. Then

$$\left|\left|\left|\Pi - \widetilde{\Pi}\right|\right|\right|_\infty \leq \left|\left|\left|\Pi - P_{0,m_t+n_0}\right|\right|\right|_\infty + \left|\left|\left|P_{0,m_t+n_0} - \widetilde{\Pi}\right|\right|\right|_\infty \leq \delta, \quad \forall t \geq t_0.$$

Therefore, $\Pi \simeq \widetilde{\Pi}$ in $|||\cdot|||_\infty$ within an error $\varepsilon$ because $\varepsilon = \inf_{\delta>\varepsilon} \delta$. □

*Remark* 2.21. One can give a proof of Theorem 2.20 using $\overline{N}_\infty$ or $\overline{\zeta}$ instead of $|||\cdot|||_\infty$. This is left to the reader.

Theorem 2.22 ([12]). *Let* $P, Q \in R_{m,n}$*. Then*
(i) $\|xP - xQ\|_1 \leq |||P - Q|||_\infty$, $\forall x \in R^m$ *with* $\|x\|_1 \leq 1$;
(ii) $\|xP - xQ\|_\infty \leq |||P - Q|||_\infty$, $\forall x \in R^m$ *with* $\|x\|_1 \leq 1$.

*Proof.* See [12]. □

The next theorem is our main result related to problem 4).

Theorem 2.23. *Let* $(P_n)_{n\geq 1}$ *be a uniformly strongly ergodic Markov chain with initial distribution* $p_0$ *and limit distribution* $\pi$ *(i.e.,* $\lim_{n\to\infty} p_0 P_{0,n} = \pi$*). Let* $0 < \varepsilon < 1$*. Let* $n_0 \geq 1$ *such that* $\alpha\left(P_{m,m+n_0}\right) \geq 1 - \frac{\varepsilon}{2}$, $\forall m \geq 0$ *(see Theorem* 2.11*). Let* $0 \leq m_1 \leq m_2 \leq \cdots$ *such that* $\lim_{t\to\infty} P_{m_t,m_t+n_0}$ *exists* $\left(\exists \left(m_t\right)_{t\geq 1}\right.$ *with the above property because* $\left.(P_{m,m+n_0})_{m\geq 0}\right.$ *is bounded). Let* $\widetilde{\Pi} = \lim_{t\to\infty} P_{m_t,m_t+n_0}$ *and* $\widetilde{\pi} = p_0\widetilde{\Pi}$*. Then* $\pi \simeq \widetilde{\pi}$ *in* $\|\cdot\|_1$ *within an error* $\varepsilon$*.*

*Proof.* We use notation and results from Theorem 2.20 and its proof. We have $\Pi = e'\pi$. Further,

$$\lim_{t\to\infty} p_0 P_{0,m_t+n_0} = p_0\Pi = p_0 e'\pi = \pi,$$

so that

$$\|p_0 P_{0,m_t+n_0} - \pi\|_1 = \|p_0 P_{0,m_t+n_0} - p_0\Pi\|_1 \le$$

(by Theorem 2.22(i))

$$\le \left|\left|\left|P_{0,m_t+n_0} - \Pi\right|\right|\right|_\infty \le \frac{\delta-\varepsilon}{2}, \quad \forall t \ge t_2.$$

Finally,

$$\|\pi - \widetilde{\pi}\|_1 = \left\|\pi - p_0\widetilde{\Pi}\right\|_1 \le \|\pi - p_0 P_{0,m_t+n_0}\|_1 + \left\|p_0 P_{0,m_t+n_0} - p_0\widetilde{\Pi}\right\|_1 \le$$

(by Theorem 2.22(i))

$$\le \|\pi - p_0 P_{0,m_t+n_0}\|_1 + \left|\left|\left|P_{0,m_t+n_0} - \widetilde{\Pi}\right|\right|\right|_\infty \le \delta, \quad \forall t \ge t_0.$$

Therefore, $\|\pi - \widetilde{\pi}\|_1 \le \varepsilon = \inf_{\delta>\varepsilon} \delta$, i.e., $\pi \simeq \widetilde{\pi}$ in $\|\cdot\|_1$ within an error $\varepsilon$. □

## 3. OTHER APPLICATIONS

In this section we use certain ergodicity coefficients of one or two matrices to give another proof of Theorem 4.3 in [4, p. 126] related to the convergence and the speed of convergence of homogeneous finite Markov chains. Other applications refer to new proofs of Theorem 2.8 in [9] and to the equivalence of weak and strong ergodicity for backward products.

THEOREM 3.1. *Let $P, Q \in S_{m,n}$. Then*

(i) $|P_{ik} - Q_{jk}| \le \frac{1}{2}\sum_{u=1}^{n} |P_{iu} - Q_{ju}|$, $\forall i,j \in \{1,2,\ldots,m\}$, $\forall k \in \{1,2,\ldots,n\}$;

(ii) $|P_{ik} - Q_{jk}| \le \overline{\zeta}\,(P,Q)$, $\forall i,j \in \{1,2,\ldots,m\}$, $\forall k \in \{1,2,\ldots,n\}$.

*Proof.* (i) Let $i,j \in \{1,2,\ldots,m\}$ and $k \in \{1,2,\ldots,n\}$. From

$$\sum_{u=1}^{n} P_{iu} = \sum_{u=1}^{n} Q_{ju} = 1,$$

we have

$$\sum_{u=1}^{n} (P_{iu} - Q_{ju}) = 0.$$

Further, setting

$$a^+ = \begin{cases} a & \text{if } a \ge 0, \\ 0 & \text{if } a < 0 \end{cases} \quad \text{and} \quad a^- = \begin{cases} 0 & \text{if } a \ge 0, \\ a & \text{if } a < 0, \end{cases}$$

where $a \in \mathbf{R}$, we have

$$\sum_{u=1}^{n} (P_{iu} - Q_{ju})^{+} = -\sum_{u=1}^{n} (P_{iu} - Q_{ju})^{-} = \frac{1}{2} \sum_{u=1}^{n} |P_{iu} - Q_{ju}| .$$

*Case* 1. $|P_{ik} - Q_{jk}| = P_{ik} - Q_{jk}$. Then

$$|P_{ik} - Q_{jk}| \leq \sum_{u=1}^{n} (P_{iu} - Q_{ju})^{+} = \frac{1}{2} \sum_{u=1}^{n} |P_{iu} - Q_{ju}| .$$

*Case* 2. $|P_{ik} - Q_{jk}| = -(P_{ik} - Q_{jk})$. Then

$$|P_{ik} - Q_{jk}| \leq -\sum_{u=1}^{n} (P_{iu} - Q_{ju})^{-} = \frac{1}{2} \sum_{u=1}^{n} |P_{iu} - Q_{ju}| .$$

Thus we proved (i).

(ii) Let $i, j \in \{1, 2, \ldots, m\}$ and $k \in \{1, 2, \ldots, n\}$. We have

$$\overline{\zeta}(P, Q) = \frac{1}{2} \max_{1 \leq v, w \leq m} \sum_{k=1}^{n} |P_{vk} - Q_{wk}| \geq \frac{1}{2} \sum_{u=1}^{n} |P_{iu} - Q_{ju}| \geq$$

(by (i))

$$\geq |P_{ik} - Q_{jk}| . \quad \square$$

*Definition* 3.2 (see, e.g., [4, p. 126]). Let $P \in S_r$. We say that $P$ is a *mixing matrix* if $\exists n \geq 1$ such that $\alpha(P^n) > 0$.

Theorem 3.3 ([4, p. 126]). *Consider a homogeneous* (*finite*) *Markov chain with mixing transition matrix $P$ and state space $S$. Let $R$ and $T$ be the set of recurrent states and the set of transient states, respectively. Then*

(i) $P^n \to \Pi = e'\pi$ *as* $n \to \infty$, *where* $\pi$ *is a probability vector,* $\pi P = \pi, \pi_i > 0$ *if* $i \in R$ *while* $\pi_i = 0$ *if* $i \in T$ *when* $T \neq \emptyset$;

(ii) $|(P^n)_{ij} - \pi_j| \leq (\overline{\alpha}(P^{n_0}))^{\left[\frac{n}{n_0}\right]}$, $\forall i, j \in S$, $\forall n \geq 1$, *where* $n_0 \geq 1$ *is taken such that* $\overline{\alpha}(P^{n_0}) > 0$.

*Proof.* See [4, pp. 123–126]. Another proof is as follows.

(i) First, we show that $(P^n)_{n \geq 1}$ is convergent. For this, we show that $(P^n)_{n \geq 1}$ is a Cauchy sequence. Let $0 < \varepsilon < 1$. By Theorem 1.4 (taking $X = S^2$), since $P$ is mixing, $\exists m = m_\varepsilon \geq 1$ such that

$$\overline{\alpha}(P^m) < \frac{\varepsilon}{2}.$$

Let $p \geq 1$ and $n \geq m$. Then (we can also prove that $(P^n)_{n \geq 1}$ is a Cauchy sequence using $\overline{N}_\infty$ or $\overline{\zeta}$)

$$|||P^{n+p} - P^n|||_\infty = ||| (P^{n+p-m} - P^{n-m}) P^m |||_\infty \leq$$

(by Theorem 1.13(ii))

$$\leq \left|\left|\left|P^{n+p-m} - P^{n-m}\right|\right|\right|_{\infty} \overline{\alpha}\left(P^{m}\right) \leq 2\overline{\alpha}\left(P^{m}\right) < \varepsilon.$$

Therefore, $\left(P^{n}\right)_{n\geq 1}$ is convergent. Further, this implies that $\exists \Pi \in S_r$ such that $P^n \to \Pi$ as $n \to \infty$.

Second, we show that $\Pi$ is a stable matrix. Since $P$ is mixing and $\overline{\alpha}$ is continuous, we have

$$\overline{\alpha}\left(\Pi\right) = \overline{\alpha}\left(\lim_{n\to\infty} P^{n}\right) = \lim_{n\to\infty} \overline{\alpha}\left(P^{n}\right) = 0.$$

Therefore, $\Pi$ is a stable matrix (see Remark 1.1).

Third, we show that there exists a probability vector $\pi$ such that $\Pi = e'\pi$, $\pi P = \pi$, $\pi_i > 0$ if $i \in R$ while $\pi_i = 0$ if $i \in T$ when $T \neq \emptyset$. Because $\Pi$ is a stable matrix, there exists a probability vector $\pi$ such that $\Pi = e'\pi$. Obviously, $\pi = \Pi_{\{i\}}$, $\forall i \in S$. Further, it follows from $P^{n+1} = P^n P$ that $\Pi = \Pi P$. Hence $\forall i \in S$ we have $\Pi_{\{i\}} = \Pi_{\{i\}} P$, i.e., $\pi P = \pi$. Further, the fact that $P$ is mixing implies that $P$ has only a recurrent class that is aperiodic and, perhaps, transient states. Therefore, $P$ can be written as

$$P = U \quad \text{or} \quad P = \begin{pmatrix} U & 0 \\ V & W \end{pmatrix},$$

where $U$ is a regular matrix (i.e., a matrix for which $\exists n \geq 1$ such that $U^n > 0$); recall that if $C \in N_l$ is a irreducible and aperiodic matrix, then $C$ is a regular matrix (see, e.g., Theorem 1.4 in [13, p. 21]). Now, $\pi_i > 0$ if $i \in R$ follows from the fact that $P$ is mixing and a theorem of O. Perron and G. Frobenius (see, e.g., Theorem 1.8 in [4, p. 51]) applied to the regular matrix $U$. If $T \neq \emptyset$, then by Theorem 2.10 (or 2.11), since $P$ is mixing, the chain is uniformly weakly ergodic (we need only to prove that the chain is weakly ergodic; note that in this case weak ergodicity and uniformly weak ergodicity are equivalent). This implies that $\Pi^T = 0$ since $P_R^T = 0$ ($P_R^T = 0 \Rightarrow \left(P^n\right)_R^T = 0$, $\forall n \geq 1$). Finally, from $\Pi^T = 0$ we have $\pi_i = 0$ if $i \in T$.

(ii) Let $i, j \in S$ and $n \geq 1$. By Theorem 3.1(ii),

$$\left|\left(P^n\right)_{ij} - \Pi_{ij}\right| \leq \overline{\zeta}\left(P^n, \Pi\right).$$

Because $\Pi_{kl} = \pi_l$, $\forall k, l \in S$ we must only prove that

$$\overline{\zeta}\left(P^n, \Pi\right) \leq \left(\overline{\alpha}\left(P^{n_0}\right)\right)^{\left[\frac{n}{n_0}\right]}.$$

We have

$$\overline{\zeta}\left(P^n, \Pi\right) = \overline{\zeta}\left(P^n, \lim_{m\to\infty} P^m\right) =$$

(using continuity of $\overline{\zeta}$)

$$= \lim_{m\to\infty} \overline{\zeta}\left(P^n, P^m\right) \leq$$

(for $m \geq n$ and using the fact that there exists the limit below)

$$\leq \lim_{m\to\infty} \overline{\zeta}\left(I, P^{m-n}\right)\overline{\alpha}\left(P^{n}\right) \leq \lim_{m\to\infty} \overline{\alpha}\left(P^{n}\right) = \overline{\alpha}\left(P^{n}\right).$$

Finally, setting $n = sn_0 + u$, where $s \geq 1$ and $0 \leq u < n_0$, we have

$$\overline{\alpha}\left(P^{n}\right) = \overline{\alpha}\left(P^{sn_0+u}\right) \leq$$

(by Theorem 1.4)

$$\leq \overline{\alpha}\left(P^{sn_0}\right) \leq \left(\overline{\alpha}\left(P^{n_0}\right)\right)^{s} = \left(\overline{\alpha}\left(P^{n_0}\right)\right)^{\left[\frac{n}{n_0}\right]}. \quad \square$$

Next, following the proof of Theorem 3.3 and using $\overline{\theta}$ instead of $\overline{\zeta}$ and, possibly, $\overline{\theta}$ instead of $|||\cdot|||_{\infty}$, the reader can try to give a new proof of Theorem 2.8 in [9] (see also Remark 2.9 there).

Another application of the ergodicity coefficients of one or two matrices refers to the asymptotic behaviour of backward products, i.e., of the products

$$U_{m,n} = P_n P_{n-1} \ldots P_{m+1}, \quad m \geq 0, \; n > m,$$

where $P_n \in S_r$, $\forall n \geq 1$ (see, e.g., [13, p. 153]). More precisely, we can prove the equivalence of weak and strong ergodicity for backward products (see [2] or [13, pp. 154–155]); obviously, strong ergodicity implies weak ergodicity and to prove the converse we need only to prove that $\left(U_{m,n}\right)_{n>m}$ is a Cauchy sequence, $\forall m \geq 0$ (the proof of this is similar to the proof of the fact that $\left(P^{n}\right)_{n\geq 1}$ is a Cauchy sequence, Theorem 3.3(i)). Moreover, taking as a guide the general $\Delta$-ergodic theory of finite Markov chains (see [10] and [11]), we can build a general $\Delta$-ergodic theory for backward products. The interested reader can develop this topic (for a starting point see also [2] and [13]).

## REFERENCES

[1] E. Behrends, *Introduction to Markov Chains with Special Emphasis on Rapid Mixing.* Vieweg, Wiesbaden, 2000.

[2] S. Chatterjee and E. Seneta, *Towards consensus: some convergence theorems on repeated averaging.* J. Appl. Probab. 14 (1977), 89−97.

[3] M. Iosifescu, *Conditions nécessaires et suffisantes pour l'ergodicité uniforme des chaînes de Markoff variables et multiples.* Rev. Roumaine Pures Appl. **11** (1966), 325–330.

[4] M. Iosifescu, *Finite Markov Processes and Their Applications.* Wiley, Chichester & Ed. Tehnică, Bucharest, 1980; corrected republication by Dover, Mineola, N.Y., 2007.

[5] D.L. Isaacson and R.W. Madsen, *Markov Chains: Theory and Applications.* Wiley, New York, 1976; republication by Krieger, 1985.

[6] U. Păun, *Ergodic theorems for finite Markov chains.* Math. Rep. (Bucur.) **3**(**53**) (2001), 383–390.

[7] U. Păun, *New classes of ergodicity coefficients, and applications.* Math. Rep. (Bucur.) **6**(**56**) (2004), 141–158.

[8] U. Păun, *Weak and uniform weak* $\Delta$*-ergodicity for* $[\Delta]$*-groupable finite Markov chains*. Math. Rep. (Bucur.) **6**(**56**) (2004), 275–293.
[9] U. Păun, *Ergodicity coefficients of several matrices*. Math. Rep. (Bucur.) **7**(**57**) (2005), 125–148.
[10] U. Păun, *General* $\Delta$*-ergodic theory of finite Markov chains*. Math. Rep. (Bucur.) **8**(**58**) (2006), 83–117.
[11] U. Păun, *General* $\Delta$*-ergodic theory*: *an extension*. Rev. Roumaine Math. Pures Appl. **53** (2008), 209–226.
[12] U. Păun, *What do we need for simulated annealing*? Math. Rep. (Bucur.) **11**(**61**) (2009), 231–247.
[13] E. Seneta, *Non-negative Matrices and Markov Chains*, 2nd Edition. Springer-Verlag, Berlin–Heidelberg–New York, 1981; revised printing, 2006.

*Received 15 February 2008*

*Romanian Academy*
*"Gheorghe Mihoc–Caius Iacob" Institute*
*of Mathematical Statistics and Applied Mathematics*
*Calea 13 Septembrie nr. 13*
*050711 Bucharest 5, Romania*
*paun@csm.ro*