WHAT DO WE NEED FOR SIMULATED ANNEALING?

UDREA PĂUN

Let $(X_n)_{n\geq 0}$ be a finite Markov chain with state space S. Let $0 \leq a \leq 1$ and $\emptyset \neq A \subseteq S$. We give necessary and/or sufficient conditions for $\lim_{n\to\infty} P(X_n \in A) = a$ and $\lim_{n\to\infty} P(X_n \in A) \geq a$ in the language of Δ-ergodic theory. These are applied, in particular, to the simulated annealing and lead to some basic matters in this field.

AMS 2000 Subject Classification: 60J10, 60J20, 90C15.

Key words: finite Markov chain, Δ-ergodic theory, simulated annealing, quantum annealing.

1. Δ-ERGODIC THEORY

In this section we recall elements of Δ-ergodic theory (see [14] and [16]) that we shall need. These are used to give necessary and/or sufficient conditions for $\lim_{n\to\infty} P(X_n \in A) = a$ and $\lim_{n\to\infty} P(X_n \in A) \geq a$, where $(X_n)_{n\geq 0}$ is a finite Markov chain with state space S, $\emptyset \neq A \subseteq S$, and $0 \leq a \leq 1$.

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

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

(The entries of a matrix Z will be denoted Z_{ij}.)

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.

Definition 1.1. Let $\Delta_1, \Delta_2 \in \mathrm{Par}(E)$. We say that Δ_1 *is finer than* Δ_2 if $\forall V \in \Delta_1$, $\exists W \in \Delta_2$ such that $V \subseteq W$.

Write $\Delta_1 \preceq \Delta_2$ when Δ_1 is finer than Δ_2.

MATH. REPORTS **11(61)**, *3* (2009), 231–247

In Δ-ergodic theory the natural space is $S \times \mathbf{N}$, called *state-time space.* Let $\emptyset \neq A \subseteq S$ and $\emptyset \neq B \subseteq \mathbf{N}$. Let $\Sigma \in \mathrm{Par}(A)$. Suppose that Σ is an ordered set. Frequently, when we only use a partition Σ of A or $\Sigma = (\{i\})_{i \in A}$ ([14] corresponds to the latter situation), we shall omit to precise this.

Definition 1.2 ([16]). Let $i, j \in S$. We say that i and j are in the same *weakly ergodic class on* $A \times B$ (or *on* $A \times B$ *with respect to* Σ, or *on* $(A \times B, \Sigma)$ when confusion can arise) if $\forall K \in \Sigma$, $\forall m \in B$ we have

$$\lim_{n \to \infty} \sum_{k \in K} \left[(P_{m,n})_{ik} - (P_{m,n})_{jk} \right] = 0.$$

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

Definition 1.3 ([16]). Let $\Delta = (C_1, C_2, \ldots, C_s)$ be the partition of weakly ergodic classes on $A \times B$ of a Markov chain. We say that the chain is *weakly* Δ*-ergodic on* $A \times B$. In particular, a weakly (S)-ergodic chain on $A \times B$ is called *weakly ergodic on* $A \times B$ for short.

Definition 1.4 ([16]). Let $(C_1, C_2, \ldots, C_s)$ be the partition of weakly ergodic classes on $A \times B$ of a Markov chain with state space S and $\Delta \in \mathrm{Par}(S)$. We say that the chain is *weakly* $[\Delta]$*-ergodic on* $A \times B$ if $\Delta \preceq (C_1, C_2, \ldots, C_s)$.

In connection with the above notions and notation we mention some special cases ($\Sigma \in \mathrm{Par}(A)$):

1. $A \times B = S \times \mathbf{N}$. In this case we can write $\sim$ instead of $\overset{S \times \mathbf{N}}{\sim}$ (or $\overset{\Sigma}{\sim}$ instead of $\overset{(S \times \mathbf{N}, \Sigma)}{\sim}$) and can omit '*on* $S \times \mathbf{N}$' in Definitions 1.2, 1.3, and 1.4.

2. $A = S$. In this case we can write $\overset{B}{\sim}$ instead of $\overset{S \times B}{\sim}$ (or $\overset{(B, \Sigma)}{\sim}$ instead of $\overset{(S \times B, \Sigma)}{\sim}$) and can replace '$S \times B$' by '(*time set*) B (*with respect to* Σ)' (or by '(B, Σ)') in Definitions 1.2, 1.3, and 1.4. A special subcase is $B = \{m\}$ ($m \geq 0$); in this case we can write $\overset{m}{\sim}$ (or $\overset{(m, \Sigma)}{\sim}$) and can replace '*on* (*time set*) $\{m\}$' by '*at time* m' in Definitions 1.2, 1.3, and 1.4.

3. $B = \mathbf{N}$. In this case we can set $\overset{A}{\sim}$ instead of $\overset{A \times \mathbf{N}}{\sim}$ (or $\overset{(A, \Sigma)}{\sim}$ instead of $\overset{(A \times \mathbf{N}, \Sigma)}{\sim}$) and can replace '$A \times \mathbf{N}$' by '(*state set*) A (*with respect to* Σ)' (or by '(A, Σ)') in Definitions 1.2, 1.3, and 1.4.

The following three definitions are special cases of Definitions 1.13, 1.16, and 1.17 in [16], respectively.

Definition 1.5. Let C be a weakly ergodic class on $A \times B$ (with respect to Σ). We say that C is a *strongly ergodic class on* $A \times B$ if $\forall i \in C$, $\forall K \in \Sigma$, $\forall m \in B$ the limit

$$\lim_{n \to \infty} \sum_{j \in K} (P_{m,n})_{ij} := \sigma_{m,K} = \sigma_{m,K}(C)$$

exists and does not depend on i.

Definition 1.6. Consider a weakly Δ-ergodic chain on $A \times B$. We say that the chain is *strongly* Δ*-ergodic on* $A \times B$ if any $C \in \Delta$ is a strongly ergodic class on $A \times B$. In particular, a strongly (S)-ergodic chain on $A \times B$ is called *strongly ergodic on* $A \times B$ for short.

Definition 1.7. Consider a weakly $[\Delta]$-ergodic chain on $A \times B$. We say that the chain is *strongly* $[\Delta]$*-ergodic on* $A \times B$ if any $C \in \Delta$ is included in a strongly ergodic class on $A \times B$.

In connection with the last three definitions we mention some special cases:

1. $A \times B = S \times \mathbf{N}$. In this case we can omit '*on* $S \times \mathbf{N}$'.
2. $A = S$. In this case we can replace '$S \times B$' by '(*time set*) B'. In the special subcase $B = \{m\}$ we can replace '*on* $S \times \{m\}$' by '*at time* m'.
3. $B = \mathbf{N}$. In this case we can replace '$A \times \mathbf{N}$' by '(*state set*) A'.

Set

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

Let $T = (T_{ij}) \in R_{m,n}$, $\emptyset \neq U \subseteq \{1, 2, \ldots, m\}$, $\emptyset \neq V \subseteq \{1, 2, \ldots, n\}$, and $\Sigma = (K_1, K_2, \ldots, K_p) \in \operatorname{Par}(V)$. Suppose that Σ is an ordered set. Define

$$T_U = (T_{ij})_{i \in U, j \in \{1,2,\ldots,n\}}, \quad T^V = (T_{ij})_{i \in \{1,2,\ldots,m\}, j \in V}, \quad T_U^V = (T_{ij})_{i \in U, j \in V},$$
$$T^+ = \left(T_{ij}^+\right), \quad T_{ij}^+ = \sum_{k \in K_j} T_{ik}, \quad \forall i \in \{1, 2, \ldots, m\}, \forall j \in \{1, 2, \ldots, p\}$$

(we call $T^+ = \left(T_{ij}^+\right)$ the *reduced matrix of* T *on* (V, Σ); $T^+ = T^+(V, \Sigma)$, i.e., it depends on (V, Σ) (if confusion can arise we write T^{+V} or $T^{+(V,\Sigma)}$ instead of T^+); see [16]),

$$\alpha(T) = \min_{1 \leq i,j \leq m} \sum_{k=1}^{n} \min(T_{ik}, T_{jk})$$

(if $T \in S_{m,n}$, then $\alpha(T)$ is called *Dobrushin's ergodicity coefficient of* T (see, e.g., [4] or [7, p. 56])),

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

$$\gamma_{\Delta}(T) = \min_{K \in \Delta} \alpha(T_K), \quad \overline{\gamma}_{\Delta}(T) = \max_{K \in \Delta} \overline{\alpha}(T_K)$$

(see [12] for γ_{Δ} and $\overline{\gamma}_{\Delta}$), and

$$|||T|||_{\infty} = \max_{1 \le i \le m} \sum_{j=1}^{n} |T_{ij}|$$

(the ∞-*norm of* T).

A vector $x \in \mathbf{R}^n$ will be understood as a row vector and x' is its transpose. Define $e = e(n) = (1, 1, \ldots, 1) \in \mathbf{R}^n$.

Definition 1.8. Let C be a strongly ergodic class on $A \times B$ (with respect to $\Sigma \in \operatorname{Par}(A)$). We say that C has *limits* Λ_m, $m \in B$, if

$$\lim_{n \to \infty} \left((P_{m,n})_C \right)^{+} = \Lambda_m, \quad \forall m \in B.$$

In particular, if there exists a matrix Λ such that $\Lambda_m = \Lambda$, $\forall m \in B$, then we say that C has *limit* Λ.

The following definition is a generalization of Definition 2.19 in [16].

Definition 1.9. Consider a strongly $[\Delta]$- or Δ-ergodic Markov chain on $A \times B$. We say that the chain has *limits* Π_m, $m \in B$, if

$$\lim_{n \to \infty} (P_{m,n})^{+} = \Pi_m, \quad \forall m \in B.$$

In particular, if there exists a matrix Π such that $\Pi_m = \Pi$, $\forall m \in B$, then we say that it has *limit* Π.

The following result is an improvement of Theorem 1.16 in [17].

THEOREM 1.10. *Let* $(P_n)_{n \ge 1}$ *be a Markov chain.*

(i) *If* $\exists \Delta \in \operatorname{Par}(S)$ *such that the chain is strongly* Δ*-ergodic on* $A \times B$ *with respect to* Σ *and has limits* Π_m, $m \in B$, *then* $\exists \Delta' \in \operatorname{Par}(S)$ *with* $\Delta \preceq \Delta'$ *such that it is strongly* Δ'*-ergodic on* $\mathcal{C}A \times B$ *with respect to* $(\mathcal{C}A)$, *where* $\mathcal{C}A$ *is the complement of* A, *and has limits* $e' - h'_m$, $m \in B$, *where* $h_m \in \mathbf{R}^r$, $(h_m)_i := \sum_{j=1}^{|\Sigma|} (\Pi_m)_{ij}$, $\forall m \in B$, $\forall i \in S$ *(*$|\Sigma|$ *is the cardinal of* Σ*).*

(ii) *The chain is strongly* Δ*-ergodic on* $A \times B$ *with respect to* (A) *and has limits* Π_m, $m \in B$, *if and only if it is strongly* Δ*-ergodic on* $\mathcal{C}A \times B$ *with respect to* $(\mathcal{C}A)$ *and has limits* $e' - \Pi_m$, $m \in B$.

Proof. Obvious. $\square$

Concerning Theorem 1.10(ii), if $\Pi_m = e'$, $\forall m \in B$, we can say more.

THEOREM 1.11. *Let $(P_n)_{n\geq 1}$ be a Markov chain. Then the following statements are equivalent.*

(i) *The chain is strongly ergodic on $A \times B$ with respect to (A) and has limit e'.*

(ii) *The chain is strongly ergodic on $\mathcal{C}A \times B$ with respect to $(\mathcal{C}A)$ and has limit* 0.

(iii) *The chain is strongly ergodic on $\mathcal{C}A \times B$ (with respect to $(\{i\})_{i\in\mathcal{C}A}$) and has limit* 0.

Proof. Obvious. □

THEOREM 1.12. *Let $(X_n)_{n\geq 0}$ be a Markov chain with state space S, initial distribution p_0, and transition matrices $(P_n)_{n\geq 1}$. Let $\emptyset \neq A \subseteq S$. Then*

$$P(X_n \in A) = p_0 (P_{0,n})^+ ,$$

where $(\cdot)^+ = (\cdot)^{+(A,(A))}$.

Proof. Let $A = \{i_1, i_2, \ldots, i_w\}$. Then

$$P(X_n \in A) = \sum_{t=1}^{w} P(X_n = i_t) = \sum_{t=1}^{w}\sum_{s=1}^{r} (p_0)_s (P_{0,n})_{si_t} =$$
$$= \sum_{s=1}^{r} (p_0)_s \sum_{t=1}^{w} (P_{0,n})_{si_t} = p_0 (P_{0,n})^+ . \quad \square$$

Let π be a probability distribution on S ($S = \{1, 2, \ldots, r\}$). Define

$$\operatorname{supp}\pi = \{i \mid i \in S \text{ and } \pi_i > 0\}$$

(the *support of* π). Note that $\operatorname{supp}\pi = S$ if and only if $\pi > 0$.

Below we give main results related to $\lim_{n\to\infty} P(X_n \in A) = a$ and $\lim_{n\to\infty} P(X_n \in A) \geq a$, $0 \leq a \leq 1$ (we build a bridge between these and Δ-ergodic theory).

THEOREM 1.13. *Let $(X_n)_{n\geq 0}$ be a Markov chain with state space S, initial distribution p_0, and transition matrices $(P_n)_{n\geq 1}$. Let $p_0 > 0$, $\emptyset \neq A \subseteq S$, and $0 \leq a \leq 1$. If the chain is strongly ergodic on A at time* 0 *with respect to (A) and has limit ae', then*

$$\lim_{n\to\infty} P(X_n \in A) = a.$$

Proof. By Theorem 1.12,

$$\lim_{n\to\infty} P(X_n \in A) = \lim_{n\to\infty} p_0 (P_{0,n})^+ = p_0 a e' = a p_0 e' = a. \quad \square$$

If $a = 1$, we can say more. (The converse of Theorem 1.13 is not true. Indeed, if $P_n = I_2$, $\forall n \geq 1$, $p_0 = \left(\frac{1}{2}, \frac{1}{2}\right)$, and $A = \{1\}$, then $\lim_{n\to\infty} P(X_n \in \{1\}) = \frac{1}{2}$, but $(X_n)_{n\geq 0}$ is not strongly ergodic on $\{1\}$ at time 0 with respect to $(\{1\})$ (it is strongly $(\{1\}, \{2\})$-ergodic on $\{1\}$ at time 0 with respect to $(\{1\})$ and has limit $\begin{pmatrix} 1 & 0 \end{pmatrix}'$).)

THEOREM 1.14. *Let $(X_n)_{n\geq 0}$ be a Markov chain with state space S, initial distribution p_0, and transition matrices $(P_n)_{n\geq 1}$. Let $p_0 > 0$ and $\emptyset \neq A \subseteq S$. Then the following statements are equivalent.*

(i) $\lim_{n\to\infty} P(X_n \in A) = 1$.

(ii) *The chain is strongly ergodic on A at time 0 with respect to (A) and has limit e'.*

(iii) *The chain is strongly ergodic on $\mathcal{C}A$ at time 0 with respect to $(\mathcal{C}A)$ and has limit 0.*

(iv) *The chain is strongly ergodic on $\mathcal{C}A$ at time 0 (with respect to $(\{i\})_{i\in\mathcal{C}A}$) and has limit 0.*

Proof. (i) $\Rightarrow$ (iii) By (i), $\lim_{n\to\infty} P(X_n \in \mathcal{C}A) = 0$. Let

$$b = \max_{s\in S} \limsup_{n\to\infty} \sum_{k\in\mathcal{C}A} (P_{0,n})_{sk}.$$

Suppose that $b > 0$. Then $\exists s_0 \in S$ such that

$$b = \limsup_{n\to\infty} \sum_{k\in\mathcal{C}A} (P_{0,n})_{s_0 k}.$$

It follows that there exists a sequence $1 \leq n_1 < n_2 < \cdots$ such that

$$\sum_{k\in\mathcal{C}A} (P_{0,n_t})_{s_0 k} \to b \text{ as } t \to \infty.$$

Further,

$$0 = \lim_{t\to\infty} P(X_{n_t} \in \mathcal{C}A) =$$

(by Theorem 1.12)

$$= \lim_{t\to\infty} p_0 (P_{0,n_t})^{+(\mathcal{C}A,(\mathcal{C}A))} \geq \lim_{t\to\infty} (p_0)_{s_0} \sum_{k\in\mathcal{C}A} (P_{0,n_t})_{s_0 k} = (p_0)_{s_0} b > 0,$$

and we have reached a contradiction. Therefore, $b = 0$ and this implies that $\lim_{n\to\infty} (P_{0,n})^{+(\mathcal{C}A,(\mathcal{C}A))}$ exists and is equal to 0, i.e., (iii).

(ii) $\Rightarrow$ (i) See Theorem 1.13.

(ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv) See Theorem 1.11. $\square$

Theorems 1.13 and 1.14 can be improved by removing the condition $p_0 > 0$.

THEOREM 1.15. *Let* $(X_n)_{n\geq 0}$ *be a Markov chain with state space* S*, initial distribution* p_0*, and transition matrices* $(P_n)_{n\geq 1}$*. Let* $0 \leq a \leq 1$ *and* $\emptyset \neq A \subseteq S$*. If* supp p_0 *is included in a strongly ergodic class on* A *at time* 0 *with respect to* (A) *and the class has limit* ae'*, then*

$$\lim_{n\to\infty} P(X_n \in A) = a.$$

Proof. Let C be a strongly ergodic class on A at time 0 with respect to (A) and with supp $p_0 \subseteq C$. By Theorem 1.12,

$$\lim_{n\to\infty} P(X_n \in A) = \lim_{n\to\infty} p_0 (P_{0,n})^+ = \lim_{n\to\infty} (p_0)^C \left((P_{0,n})^+\right)_C =$$
$$= \lim_{n\to\infty} (p_0)^C \left((P_{0,n})_C\right)^+ = (p_0)^C ae' = a (p_0)^C e' = a. \quad \square$$

If $a = 1$, we can say more. (The converse of Theorem 1.15 is not true. Indeed, if $P_n = I_3$, $\forall n \geq 1$, $p_0 = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$, and $A = \{1\}$, then $\lim_{n\to\infty} P(X_n \in \{1\}) = \frac{1}{2}$, but supp p_0 is not included in a strongly ergodic class on $\{1\}$ at time 0 with respect to $(\{1\})$.)

THEOREM 1.16. *Let* $(X_n)_{n\geq 0}$ *be a Markov chain with state space* S*, initial distribution* p_0*, and transition matrices* $(P_n)_{n\geq 1}$*. Let* $\emptyset \neq A \subseteq S$*. Then the following statements are equivalent.*

(i) $\lim_{n\to\infty} P(X_n \in A) = 1$.

(ii) supp p_0 *is included in a strongly ergodic class on* A *at time* 0 *with respect to* (A) *and the class has limit* e'*.*

(iii) supp p_0 *is included in a strongly ergodic class on* $\mathcal{C}A$ *at time* 0 *with respect to* $(\mathcal{C}A)$ *and the class has limit* 0*.*

(iv) supp p_0 *is included in a strongly ergodic class on* $\mathcal{C}A$ *at time* 0 (*with respect to* $(\{i\})_{i\in\mathcal{C}A}$) *and the class has limit* 0*.*

Proof. This is left to the reader. (It is easy to obtain the results analogous to Theorems 1.10 and 1.11 for classes.) $\square$

THEOREM 1.17. *Let* $(X_n)_{n\geq 0}$ *be a Markov chain with state space* S*, initial distribution* p_0*, and transition matrices* $(P_n)_{n\geq 1}$*. Let* $0 \leq a \leq 1$ *and* $\emptyset \neq A \subseteq S$*. If the chain is strongly* $[\Delta]$*-ergodic on* A *at time* 0 *with respect to* (A) *and has limit greater or equal to* ae'*, then*

$$\lim_{n\to\infty} P(X_n \in A) \geq a.$$

Proof. By Theorem 1.12,

$$\lim_{n\to\infty} P(X_n \in A) = \lim_{n\to\infty} p_0 (P_{0,n})^+ \geq p_0 ae' = ap_0 e' = a. \quad \square$$

Theorem 1.17 also can be improved.

THEOREM 1.18. *Let $(X_n)_{n\geq 0}$ be a Markov chain with state space S, initial distribution p_0, and transition matrices $(P_n)_{n\geq 1}$. Let $0 \leq a \leq 1$ and $\emptyset \neq A \subseteq S$. If* $\operatorname{supp} p_0$ *is included in a union of strongly ergodic classes on A at time* 0 *with respect to* (*A*) *and this union has limit* (*we join the limits of all classes belonging to union*) *greater or equal to ae', then*

$$\lim_{n\to\infty} P\left(X_n \in A\right) \geq a.$$

Proof. This is left to the reader. □

2. SIMULATED ANNEALING

In this section we show that some results in the previous section apply, in particular, to the simulated annealing. Then we consider some basic matters and make some considerations on them.

Set

$$a^+ = \begin{cases} a & \text{if } a > 0, \\ 0 & \text{if } a \leq 0, \end{cases}$$

where $a \in \mathbf{R}$.

Let $H : S \to \mathbf{R}$ be a nonconstant function. We want to find $\min\limits_{y\in S} H\left(y\right).$ A stochastic optimization technique for solving this problem approximately when S is very large is the simulated annealing (see, e.g., [1], [5], [9], [11], [15], [18], [20], and [21]). For this, consider a sequence $(\beta_n)_{n\geq 1}$ of positive real numbers with $\beta_n \to \infty$ as $n \to \infty$ ($(\beta_n)_{n\geq 1}$ is called the *cooling schedule*), an irreducible stochastic matrix $G = (G_{ij})_{i,j\in S}$ (G is called the *generation matrix*) and a Markov chain $(X_n)_{n\geq 0}$ with state space S and transition matrices $(P_n)_{n\geq 1}$, where

$$(P_n)_{ij} = \begin{cases} G_{ij}\mathrm{e}^{-\beta_n (H(j)-H(i))^+} & \text{if } i \neq j, \\ 1 - \sum\limits_{k\neq i} (P_n)_{ik} & \text{if } i = j, \end{cases}$$

$\forall i, j \in S$. $(X_n)_{n\geq 0}$ (or, by convention, $(P_n)_{n\geq 1}$) is called the (*classical*) *simulated annealing chain* (the (*classical*) *simulated annealing* for short). Note that the results below can be applied to all versions of simulated annealing (also, to all versions of quantum annealing (see, e.g., [3] and [19])), except for Theorem 2.26 and Remark 2.27.

Let

$$S^* = S^*\left(H\right) = \left\{i \mid i \in S \text{ and } \min_{y\in S} H\left(y\right) = H(i)\right\}$$

(*the set of global minima of* H; it only depends on H).

Taking $A = S^*$, Theorems 1.13, 1.14, 1.15, 1.16, 1.17, and 1.18 yield necessary and/or sufficient conditions for $\lim_{n\to\infty} P(X_n \in S^*) = a$ and $\lim_{n\to\infty} P(X_n \in S^*) \geq a$ $(0 \leq a \leq 1)$. For other approaches, see, e.g., [2], [5], and [20]. As for, e.g., Theorem 1.14, we need:

1) necessary and/or sufficient conditions for (ii), or (iii), or (iv);

2) ergodicity coefficients and other tools and methods for to solve 1);

3) speed of convergence for the types of chains which appear at 1).

Below we make some considerations related to the basic matters 1), 2), and 3).

First, we consider matter 1).

Remark 2.1. If a chain is strongly ergodic, then it is strongly ergodic at time 0. The converse is not true. E.g., let

$$P_1 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad P_{2n} = I_2, \quad P_{2n+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \forall n \geq 1.$$

Obviously, this chain is strongly ergodic at time 0 and has limit

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

but is not strongly ergodic. Note that the simulated annealing has the strong ergodicity property under some conditions (see, e.g., [9], [11], [18], and [21]), but nobody (as far as we know) not examined the conditions we need such that the strong ergodicity at time 0 holds. A problem appears here: Are there conditions for the simulated annealing which guarantee the equivalence between strong ergodicity and strong ergodicity at time 0?

Theorem 2.2 ([6]). *Let $(P_n)_{n\geq 1}$ be a Markov chain. Then it is weakly ergodic if and only if there exists a sequence $0 \leq n_1 < n_2 < \cdots$ such that*

$$\sum_{s\geq 1} \alpha\left(P_{n_s,n_{s+1}}\right) = \infty.$$

Proof. See, e.g., [6] or [7, p. 219]. □

Remark 2.3. A sufficient condition for weak ergodicity at time 0 follows from the above theorem, namely, if there exists a sequence $0 \leq n_1 < n_2 < \cdots$ such that

$$\sum_{s\geq 1} \alpha\left(P_{n_s,n_{s+1}}\right) = \infty,$$

then the chain is weakly ergodic at time 0. A proof of this result is as the proof of "$\Leftarrow$" of Theorem 2.2 and an other is as follows. By Theorem 2.2, the above hypothesis implies that the chain is weakly ergodic. But weak ergodicity implies weak ergodicity at time 0. Further, we mention, obviously, that the above condition is not necessary for weak ergodicity at time 0. An example is

$$P_1 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad P_n = I_2, \quad \forall n \geq 2,$$

since

$$P_{m,n} = \begin{cases} P_1 & \text{if } m = 0, \\ I_2 & \text{if } m > 0, \end{cases}$$

$\forall m, n,\ 0 \leq m < n$. A problem occurs here: Give necessary and sufficient conditions for weak ergodicity at time 0.

THEOREM 2.4 ([10]). *Let $(P_n)_{n\geq 1}$ be a Markov chain. If*

(i) *the chain is weakly ergodic*

and

(ii) $\sum_{n\geq 1} \|\psi_{n+1} - \psi_n\|_1 < \infty$, *where ψ_n is a probability vector and $\psi_n P_n = \psi_n$, $\forall n \geq 1$,*

then it is strongly ergodic.

Proof. See, e.g., [8, pp. 160−162] or [10]. □

Remark 2.5. As for strong ergodicity at time 0, if $(P_n)_{n\geq 1}$ is weakly ergodic at time 0 and (ii) from Theorem 2.4 holds, then it does not follow that it is strongly ergodic at time 0. Indeed, let

$$P_1 = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}, \quad P_{2n} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} := P, \quad P_{2n+1} = I_2, \quad \forall n \geq 1.$$

Then it is weakly ergodic at time 0 while (ii) in Theorem 2.4 holds since by taking $\psi_1 = \left(\frac{1}{3}, \frac{2}{3}\right)$, $\psi_n = \left(\frac{1}{2}, \frac{1}{2}\right)$, $\forall n \geq 2$, we have $\psi_n P_n = \psi_n$, $\forall n \geq 1$, and $\sum_{n\geq 1} \|\psi_{n+1} - \psi_n\|_1 = \|\psi_2 - \psi_1\|_1 < \infty$. But the chain is not strongly ergodic at time 0 since with

$$T_n := \{k \mid 1 \leq k \leq n \text{ and } P_k = P\}, \ \forall n \geq 1, \quad \text{and} \quad Q := \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

we have

$$P_{0,n} = \begin{cases} P_1 & \text{if } |T_n| \text{ is even}, \\ Q & \text{if } |T_n| \text{ is odd}, \end{cases}$$

$\forall n \geq 1$. Therefore, a result similar to Theorem 2.4, keeping (ii) from there, does not exist for strong ergodicity at time 0. Open problem: Find a result 'similar' to Theorem 2.4 for strong ergodicity at time 0.

Remark 2.6 (see again Remark 2.1). If a chain is strongly ergodic at time 0, then it is strongly ergodic on A at time 0 with respect to (A) ($\emptyset \neq A \subseteq S$).

More generally, if a chain is strongly ergodic at time 0, then it is strongly ergodic on A at time 0 with respect to Σ, $\forall \Sigma \in \mathrm{Par}(A)$. The converse is not true. Indeed, if

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

then the chain is strongly ergodic on $\{3\}$ at time 0 with respect to $(\{3\})$, but it is not strongly ergodic at time 0.

Remark 2.7. Related to (iv) in Theorem 1.14, a main objective is to obtain as many columns as possible belonging to the set

$$\left\{ j \mid j \in S \text{ and } \lim_{n \to \infty} (P_{0,n})^{\{j\}} = 0 \right\}.$$

For this matter there exist some results in [15].

Second, we consider matter 2).

THEOREM 2.8 ([16]). *Let $P \in R_{m,n}$, $Q \in R_{n,p}$, $\emptyset \neq V \subseteq \{1, 2, \ldots, p\}$, and $\Sigma \in \mathrm{Par}(V)$. Then*

$$(PQ)^{+} = PQ^{+}.$$

Proof. See [16]. □

Remark 2.9. The well-known equation $e' = Pe'$ when $P \in S_{m,n}$ follows from the above result taking $P \in S_{m,n}$, $Q \in S_{n,p}$, $V = \{1, 2, \ldots, p\}$, and $\Sigma = (V)$, since in this case $(PQ)^{+} = Q^{+} = e'$.

THEOREM 2.10 ([4]). *Let $P \in S_{m,n}$ and $Q \in S_{n,p}$. Then*

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

Proof. See, e.g., [4], or [7, pp. 58–59], or [8, pp. 145–146]. □

THEOREM 2.11. *Let $P \in S_{m,n}$, $Q \in S_{n,p}$, and $\Sigma \in \mathrm{Par}(\{1, 2, \ldots, p\})$. Then*

$$\overline{\alpha}\left((PQ)^{+}\right) \leq \overline{\alpha}(P)\, \overline{\alpha}\left(Q^{+}\right).$$

Proof. See Theorems 2.8 and 2.10. (We can apply Theorem 2.10 because $Q \in S_{n,p}$ and $\Sigma \in \mathrm{Par}(\{1, 2, \ldots, p\})$ implies that Q^{+} is a stochastic matrix.) □

By Theorems 2.8 and 2.11 and induction we have

$$\overline{\alpha}\left((Q_1 Q_2 \ldots Q_t)^{+}\right) \leq \overline{\alpha}(Q_1)\, \overline{\alpha}(Q_2) \ldots \overline{\alpha}\left((Q_t)^{+}\right),$$

$\forall Q_1 \in S_{m_1,m_2}, \forall Q_2 \in S_{m_2,m_3}, \ldots, \forall Q_t \in S_{m_t,m_{t+1}}, \forall \Sigma \in \mathrm{Par}(\{1, 2, \ldots, m_{t+1}\})$.

Remark 2.12. (a) A main drawback of the above inequality is the fact that the operator $(\cdot)^+$ only appears at the last position of the right term.

(b) In general, even when $\Sigma \in \text{Par}(\{1,2,\ldots,p\})$ we do not have

$$\overline{\alpha}\left((PQ)^+\right) \leq \overline{\alpha}\left(P^+\right)\overline{\alpha}\left(Q^+\right).$$

Indeed, let

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{and} \quad \Sigma = (\{1\},\{2,3\}).$$

Then

$$PQ = Q, \quad (PQ)^+ = Q^+ = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad P^+ = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix},$$

$$\overline{\alpha}\left((PQ)^+\right) = \overline{\alpha}\left(Q^+\right) = 1 \quad \text{and} \quad \overline{\alpha}\left(P^+\right) = 0.$$

Hence

$$1 = \overline{\alpha}\left((PQ)^+\right) > \overline{\alpha}\left(P^+\right)\overline{\alpha}\left(Q^+\right) = 0.$$

This negative result raises the following question: When do we have $\overline{\alpha}\left((PQ)^+\right) \leq \overline{\alpha}\left(P^+\right)\overline{\alpha}\left(Q^+\right)$?

(c) If $Q \in S_{n,p}$, $\emptyset \neq V \subset \{1,2,\ldots,p\}$, and $\Sigma \in \text{Par}(V)$, it is possible that Q^+ is not a stochastic matrix. This drawback can be removed if we replace Σ by $\Sigma' = \Sigma \cup \{\mathcal{C}V\}$.

Let $T \in R_{m,n}$, $\emptyset \neq W \subseteq \{1,2,\ldots,n\}$, and $\Sigma \in \text{Par}(W)$. Suppose that Σ is an ordered set. Define

$$\alpha^{\Sigma}(T) = \min_{1\leq i,j\leq m} \sum_{K\in\Sigma} \min\left(\sum_{k\in K} T_{ik}, \sum_{k\in K} T_{jk}\right)$$

and

$$\overline{\alpha}^{\Sigma}(T) = \frac{1}{2}\max_{1\leq i,j\leq m} \sum_{K\in\Sigma}\left|\sum_{k\in K} T_{ik} - \sum_{k\in K} T_{jk}\right|.$$

(Because $\Sigma \in \text{Par}(W)$, in fact α^{Σ} and $\overline{\alpha}^{\Sigma}$ only work on T^W; equivalently, we can work on T if instead of $\Sigma \in \text{Par}(W)$ we use $\Sigma \cup \{\mathcal{C}W\}$.)

THEOREM 2.13. *Let $T \in R_{m,n}$, $\emptyset \neq W \subseteq \{1,2,\ldots,n\}$, and $\Sigma \in \text{Par}(W)$. Then*

$$\alpha^{\Sigma}(T) = \alpha\left(T^+\right) \quad \textit{and} \quad \overline{\alpha}^{\Sigma}(T) = \overline{\alpha}\left(T^+\right).$$

Proof. Obvious. □

Remark 2.14. It follows from the above result that the coefficients α^{Σ} and $\overline{\alpha}^{\Sigma}$ bring no news for our problems.

Below, we generalize part of the above results (these generalizations are related to Theorems 1.17 and 1.18).

Definition 2.15. Let $T \in N_{m,n}$. We say that T is a *generalized stochastic matrix* if $\exists a \geq 0, \exists Z \in S_{m,n}$ such that $T = aZ$.

Let $\Delta_1 \in \mathrm{Par}(\{1,2,\ldots,m\})$ and $\Delta_2 \in \mathrm{Par}(\{1,2,\ldots,n\})$. Define (see [13])

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

In particular, if $m = n$ and $\Delta_1 = \Delta_2 := \Delta$, we set $G_\Delta = G_{\Delta,\Delta}$.

The following result is a generalization of Theorem 2.10.

THEOREM 2.16 ([13]). *Let* $\Delta_1 \in \mathrm{Par}(\{1,2,\ldots,m\})$ *and* $\Delta_2 \in \mathrm{Par}(\{1,2,\ldots,n\})$. *Let* $P \in G_{\Delta_1,\Delta_2}$ *and* $Q \in S_{n,p}$. *Then*

$$\overline{\gamma}_{\Delta_1}(PQ) \leq \overline{\gamma}_{\Delta_1}(P)\,\overline{\gamma}_{\Delta_2}(Q).$$

Proof. See [13]. □

The following result is a generalization of Theorem 2.11 (obviously, with the same main drawback).

THEOREM 2.17. *Let* $\Delta_1 \in \mathrm{Par}(\{1,2,\ldots,m\})$, $\Delta_2 \in \mathrm{Par}(\{1,2,\ldots,n\})$ *and* $\Sigma \in \mathrm{Par}(\{1,2,\ldots,p\})$. *Let* $P \in G_{\Delta_1,\Delta_2}$ *and* $Q \in S_{n,p}$. *Then*

$$\overline{\gamma}_{\Delta_1}\left((PQ)^+\right) \leq \overline{\gamma}_{\Delta_1}(P)\,\overline{\gamma}_{\Delta_2}\left(Q^+\right).$$

Proof. See Theorems 2.8 and 2.16. (We can apply Theorem 2.16 because $Q \in S_{n,p}$ and $\Sigma \in \mathrm{Par}(\{1,2,\ldots,p\})$ implies that Q^+ is a stochastic matrix.) □

Remark 2.18. As above, we can consider other two coefficients γ_Δ^Σ and $\overline{\gamma}_\Delta^\Sigma$ ($\gamma_\Delta^\Sigma(T) := \min_{K\in\Delta} \alpha^\Sigma(T_K)$ and $\overline{\gamma}_\Delta^\Sigma(T) := \max_{K\in\Delta} \overline{\alpha}^\Sigma(T_K)$), but neither of them brings some news (since $\gamma_\Delta^\Sigma(T) = \gamma_\Delta(T^+)$ and $\overline{\gamma}_\Delta^\Sigma(T) = \overline{\gamma}_\Delta(T^+)$, $\forall T \in R_{m,n}$).

THEOREM 2.19. *Let* $(P_n)_{n\geq1}$ *be a Markov chain. Then the chain is weakly ergodic on* $A \times B$ *with respect to* Σ *if and only if it is weakly ergodic on* $S \times B$ *with respect to* $\Sigma \cup \{\mathcal{C}A\}$.

Proof. Obvious. □

THEOREM 2.20. *Let* $(P_n)_{n\geq1}$ *be a Markov chain. Then the chain is weakly ergodic on* $A \times B$ *with respect to* Σ *if and only if*

$$\lim_{n\to\infty} \overline{\alpha}\left((P_{m,n})^+\right) = 0, \quad \forall m \in B.$$

Proof. Obvious. □

Theorem 2.11 can be used to prove

THEOREM 2.21. *Let* $(P_n)_{n\geq 1}$ *be a weakly ergodic Markov chain on* A *at time* $l \geq 0$ *with respect to* Σ $(\Sigma \in \text{Par}(A))$. *Then it is weakly ergodic on* A *at time* m *with respect to* Σ, $\forall m$, $0 \leq m \leq l$.

Proof. By Theorem 2.19, the chain is weakly ergodic on S at time l with respect to $\Sigma' := \Sigma \cup \{\mathcal{C}A\}$. Further, we show that the chain is weakly ergodic on S at time m with respect to Σ', $\forall m$, $0 \leq m \leq l$. Let $0 \leq m \leq l$.

Case 1. $l = 0$. Obvious.

Case 2. $l > 0$. The subcase $m = l$ is obvious. Now, let $0 \leq m < l$. Then, by Theorem 2.20, for $n > l$ we have

$$\overline{\alpha}\left((P_{m,n})^{+}\right) \leq \overline{\alpha}\left(P_{m,l}\right)\overline{\alpha}\left((P_{l,n})^{+}\right) \to 0 \text{ as } n \to \infty.$$

Therefore, the chain is weakly ergodic on S at time m with respect to Σ'. Now, by Cases 1 and 2 and using again Theorem 2.19, the chain is weakly ergodic on A at time m with respect to Σ, $\forall m$, $0 \leq m \leq l$. □

Remark 2.22. Using Theorem 2.17 and considering a chain $(P_n)_{n\geq 1}$ with $P_n \in G_{\Delta}$, $\forall n \geq 1$, the reader can try to prove a result similar to Theorem 2.21.

Third, we consider matter 3).

Definition 2.23 (see, e.g., [7, p. 222]). Let $T \in R_{m,n}$. The matrix obtained from T by replacing all its nonzero entries by ones is called the *incidence matrix of* T.

Let $T, Z \in R_{m,n}$. We write $T \sim Z$ if T and Z have the same incidence matrix.

THEOREM 2.24. *Let* $P, Q \in R_{m,n}$. *Then*
(i) $\|xP - xQ\|_1 \leq |||P - Q|||_\infty$, $\forall x \in \mathbf{R}^m$ *with* $\|x\|_1 \leq 1$;
(ii) $\|xP - xQ\|_\infty \leq |||P - Q|||_\infty$, $\forall x \in \mathbf{R}^m$ *with* $\|x\|_1 \leq 1$.

Proof. (i) Let $x \in \mathbf{R}^m$ with $\|x\|_1 \leq 1$. Then

$$\|xP - xQ\|_1 = \left\|(P-Q)'\,x'\right\|_1 \leq \left|\left|\left|(P-Q)'\right|\right|\right|_1 \left\|x'\right\|_1 \leq$$
$$\leq \left|\left|\left|(P-Q)'\right|\right|\right|_1 = |||P - Q|||_\infty\,.$$

(ii) This follows from (i) because $\|\cdot\|_\infty \leq \|\cdot\|_1$. □

THEOREM 2.25. *Let* $(P_n)_{n\geq 1}$ *be a strongly ergodic Markov chain at time* 0 *with limit* Π_0. *Let* p_n *be the distribution of Markov chain at time* n $(p_n = p_0 P_{0,n})$, $\forall n \geq 1$. *Setting* $\Pi_0 = e'\pi_0$, *we have*

$$\|p_n - \pi_0\|_1 \leq |||P_{0,n} - \Pi_0|||_\infty\,.$$

Proof. By Theorem 2.24(i), we have

$$\|p_n - \pi_0\|_1 = \|p_0 P_{0,n} - p_0 \Pi_0\|_1 \leq |||P_{0,n} - \Pi_0|||_\infty . \quad \square$$

Let (V, E) be a connected directed graph (V is the set of vertices and E is the set of edges). Let $i, j \in V$, $i \neq j$. Denote

$$d(i,j) = \min \{t \mid t \geq 1 \text{ and } \exists i_0, i_1, \ldots, i_t \in V \text{ such that } i_0 = i,\ i_t = j,\ \text{and } (i_0, i_1), (i_1, i_2), \ldots, (i_{t-1}, i_t) \in E\}.$$

THEOREM 2.26. *Let $(P_n)_{n\geq 1}$ be a strongly ergodic Markov chain at time* 0 *with limit Π_0. Let $0 \leq \varepsilon < 1$. Suppose that P_1 determines a connected directed graph, $P_s \sim P_t$, $\forall s, t \geq 1$, $\Pi_0 = e'\pi_0$, and $\operatorname{supp} \pi_0 \neq S$. If $|||P_{0,n} - \Pi_0|||_\infty \leq \varepsilon$, then*

$$n \geq n^* := \max_{i \in S - \operatorname{supp} \pi_0} \ \min_{j \in \operatorname{supp} \pi_0} d(i,j).$$

Proof. *Case* 1. $n^* = 1$. Obvious.
Case 2. $n^* \geq 2$. We have

$$\sum_{j \in \operatorname{supp} \pi_0} \left| (P_{0,n})_{ij} - (\pi_0)_j \right| = \sum_{j \in \operatorname{supp} \pi_0} \left| (P_{0,n})_{ij} - (\Pi_0)_{ij} \right| \leq \leq |||P_{0,n} - \Pi_0|||_\infty , \quad \forall i \in S.$$

Therefore,

$$\sum_{j \in \operatorname{supp} \pi_0} \left| (P_{0,n})_{ij} - (\pi_0)_j \right| \leq \varepsilon, \quad \forall i \in S.$$

Now, let $i_1 \in S - \operatorname{supp} \pi_0$ for which $\exists j_1 \in \operatorname{supp} \pi_0$ such that $n^* = d(i_1, j_1)$. Then

$$(P_{0,n})_{i_1 j} = 0, \quad \forall j \in \operatorname{supp} \pi_0,\ \forall n, 1 \leq n < n^*.$$

Suppose that $\exists n$, $1 \leq n < n^*$, such that

$$\sum_{j \in \operatorname{supp} \pi_0} \left| (P_{0,n})_{ij} - (\pi_0)_j \right| \leq \varepsilon, \quad \forall i \in S.$$

Then we have, for $i = i_1$,

$$1 > \varepsilon \geq \sum_{j \in \operatorname{supp} \pi_0} \left| (P_{0,n})_{i_1 j} - (\pi_0)_j \right| = \sum_{j \in \operatorname{supp} \pi_0} (\pi_0)_j = 1.$$

We have reached a contradiction. This means that $n \geq n^*$. $\square$

Remark 2.27. (a) From the proof of Theorem 2.26 we have

$$|||P_{0,n} - \Pi_0|||_\infty \geq 1, \quad \forall n, 1 \leq n < n^*, \quad \text{if } n^* \geq 2.$$

(b) In particular, Theorem 2.26 can be applied to the simulated annealing $(P_n)_{n\geq 1}$ because $P_s \sim P_t$, $\forall s, t \geq 1$, a.s.o.. Moreover, we have $n^* = n^*(H, G) \leq$

$d(\overline{G}) := \max_{i,j\in S} d(i,j)$, where $\overline{G}$ is directed graph associated with G. (We also must keep in mind the worst case $n^* = d\left(\overline{G}\right)$.)

We conclude with two remarks.

Remark 2.28. In [15] we defined the notion of breaking up. The reader might associate this notion with Theorems 1.15, 1.16, 1.17, and 1.18 here.

Remark 2.29. Section 2 refers to the study of limit behaviour of simulated annealing. In particular, this might contribute to the study of finite-time behaviour of simulated annealing.

REFERENCES

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

[2] D.P. Connors and P.R. Kumar, *Balance of recurrence orders in time-inhomogeneous Markov chains with applications to simulated annealing.* Probab. Engrg. Inform. Sci. **2** (1988), 157–184.

[3] A. Das and B.K. Chakrabarti (Eds.), *Quantum Annealing and Related Optimization Methods.* Lecture Notes in Phys. **679**. Springer, Berlin, 2005.

[4] R.L. Dobrushin, *Central limit theorem for nonstationary Markov chains,* I, II. Theory Probab. Appl. **1** (1956), 65–80, 329–383.

[5] B. Hajek, *Cooling schedules for optimal annealing.* Math. Oper. Res. **13** (1988), 311–329.

[6] J. Hajnal, *Weak ergodicity in non-homogeneous Markov chains.* Proc. Cambridge Philos. Soc. **54** (1958), 233–246.

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

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

[9] P.J.M. van Laarhoven and E.H.L. Aarts, *Simulated Annealing: Theory and Applications.* Reidel, Dordrecht, 1987.

[10] R.W. Madsen and D.L. Isaacson, *Strongly ergodic behaviour for non-stationary Markov processes.* Ann. Probab. **1** (1973), 329–335.

[11] D. Mitra, F. Romeo, and A. Sangiovanni-Vincentelli, *Convergence and finite-time behavior of simulated annealing.* Adv. in Appl. Probab. **18** (1986), 747–771.

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

[13] U. Păun, *Ergodicity coefficients of several matrices.* Math. Rep. (Bucur.) **7**(**57**) (2005), 125–148.

[14] U. Păun, *General* Δ*-ergodic theory of finite Markov chains.* Math. Rep. (Bucur.) **8**(**58**) (2006), 83–117.

[15] U. Păun, Δ*-ergodic theory and simulated annealing.* Math. Rep. (Bucur.) **9**(**59**) (2007), 279–303.

[16] U. Păun, *General* Δ*-ergodic theory: an extension.* Rev. Roumaine Math. Pures Appl. **53** (2008), 209–226.

[17] U. Păun, Δ-*ergodic theory and reliability theory*. Math. Rep. (Bucur.) **10**(**60**) (2008), 73–95.

[18] F. Romeo and A. Sangiovanni-Vincentelli, *A theoretical framework for simulated annealing*. Algorithmica **6** (1991), 302–345.

[19] G.E. Santoro and E. Tosatti, *Optimization using quantum mechanics*: *quantum annealing through adiabatic evolution*. J. Phys. A Math. Gen. **39** (2006), R393–R431.

[20] J.N. Tsitsiklis, *Markov chains with rare transitions and simulated annealing*. Math. Oper. Res. **14** (1989), 70–90.

[21] G. Winkler, *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*: *A Mathematical Introduction*, 2nd Edition. Springer-Verlag, Berlin, 2003.

Received 28 January 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