21st International Symposium on
Mathematical Theory of Networks and Systems
July 7-11, 2014. Groningen, The Netherlands

Perturbation and Sensitivity of Inhomogeneous Markov Chains in Dynamic Environments

Manuela L. Bujorianu   Robert S. MacKay

Abstract— **In this paper, we study robustness problems of inhomogeneous Markov chains in dynamic environments. First, we define the natural distributions of nonstationary Markov chains as a sequence of probability distributions, which represents a non-autonomous dynamical system on a suitable state space. Then, we study the robustness and sensitivity analysis of these distributions under the assumption of geometric ergodicity of the underlying Markov chain.**

Keywords: time-inhomogeneous Markov chain, non-autonomous dynamical system, robustness, sensitivity analysis.

I. Introduction

Markov chains represent a class of probabilistic models widely used in practical applications. Many engineering, biological and social systems can be modelled by Markov chains. Different types of perturbation may affect the system because of the external environment and control. The purpose of perturbation analysis (PA) is to study the effects of small perturbations on system properties. Mainly, PA studies the robustness of the system and then estimates the deviations of its behaviours caused by different types of perturbation.

The literature regarding perturbation bounds for Markov chain is very rich. Some papers derive, using matrix analysis methods, bounds for the sensitivity of the stationary distributions of finite stationary Markov chains ([19], see also the survey [6] and the references therein). Others obtain perturbation bounds of the invariant measures of Markov chains with a general state space via ergodicity coefficients of the iterated transition kernel using functional analysis (operator theory) and probabilistic methods (see, e.g., [14]).

The purpose of this paper is to study robustness and sensitivity analysis of the probability distributions that may depend on some external parameters of an inhomogeneous Markov chain evolving in a dynamic environment under possible perturbations. A Markov chain is a very simple way of describing the probabilities that the system will move from state to state. When the transition matrix of a Markov chain changes at each step we deal with the concept of inhomogeneous (nonhomogeneous or nonstationary) Markov chain.

Our perspective is to view the sequence of probability distributions associated to an inhomogeneous Markov chain as a non-autonomous dynamical system on a suitable state space. Then, we study the robustness of these distributions under the assumption of geometric ergodicity of the underlying Markov chain. Furthermore, supposing that the perturbations to the dynamics depend on some parameter, we study the sensitivity of the resulting distributions with respect to this parameter.

This work has been supported by the project Management of Complex Systems funded by the Sloan Foundation.

Manuela Bujorianu is with the Center for Complexity, University of Warwick, UK, e-mail: {L-M.Bujorianu@warwick.ac.uk)

Robert Mackay is with the Centre for Complexity, University of Warwick, UK, e-mail: {R.S.MacKay@warwick.ac.uk)

II. Background

In this section, we present preliminary definitions regarding inhomogeneous Markov chains and their connections with other mathematical objects.

A. Inhomogeneous Markov Chains

An inhomogeneous Markov chain (IMC) is constructed using the following ingredients [13]:

- a finite set $X := \{1, ..., r\}$ whose elements are called states;
- a probability distribution $\mathbf{p} = (p(i))_{i\in X}$ over X, whose components are called initial probabilities; and
- a sequence of matrices $\mathbf{P}_t = (p_{ij}(t, t+1))_{i,j\in X}$, $t \geq 0$, whose entries are called transition probabilities.

A sequence of X-valued random variables $(X_t)_{t\geq 0}$ is called an inhomogeneous (nonstationary) finite Markov chain with state space X, initial probability distribution $\mathbf{p}$ and transition matrices $\mathbf{P}_t$, $t \geq 0$, if and only if

$$\mathbf{P}(X_0 = i) = p(i),\ i \in X,$$

and

$$\begin{aligned}\mathbf{P}(X_{t+1} &= i_{t+1}|X_t = i_t, ..., X_0 = i_0) \\ &= \mathbf{P}_t(X_{t+1} = i_{t+1}|X_t = i_t)\end{aligned}$$

whenever the states $i_0, ..., i_{t+1} \in X$ and $t \in \mathbb{Z}_+$ (discrete time inhomogeneous Markov chain, abbreviated DTIMC), or $t \in [0, \infty)$ (continuous time inhomogeneous Markov chain, abbreviated CTIMC).

We denote by $p_{ij}(r, s)$ the probability that the system is in state j at time s under the hypothesis that it was in state i at time r, where $r < s$. The $r\times r$ matrices $P(r, s)$ with elements $p_{ij}(r, s)$ are stochastic matrices (matrices with nonnegative elements and row sums equal to one) and they satisfy the Chapman-Kolmogorov equation, i.e.,

$$P(r, s)P(s, t) = P(r, t), \text{ for } r < s < t.$$

Note that the classification of states and the description of limiting behaviour from the case of homogeneous Markov chains, no longer apply.

1) Wasserstein space: Let

$$\mathcal{D}_r^1 \;:\; = \{(p_1, p_2, ..., p_r) \in \mathbb{R}_+^r | \sum_{i=1}^{r} p_i = 1, \\ 0 \;\leq\; p_i \leq 1,\, i = 1, .., r\}$$

the set of all stochastic vectors of dimension r.

The space $\mathcal{D}_r^1$ equipped with the transportation metric is known in the literature as the *Wasserstein space*. One way to define the transportation metric is as follows:

$$D_T(\nu, \mu) := \sup \left| \int_X f d\nu - \int_X f d\mu \right|$$

where the supremum is taken over all functions f in the space

$$lip_1(X) := \{f : X \to \mathbb{R} |\ |f(i) - f(j)| \leq d(i, j),\, \forall i, j \in X\}$$

where d is a metric on the space X. This metric is called also *Hutchinson distance* and it was introduced by Kantorovich and Rubinshteĭn in the 1950s. Here, d is an arbitrary discrete metric on X since the state space is finite.

2) Zero charge measure space: Recall that $\mathcal{D}_r^1$ is the simplex of probability distributions on the space $\{1, .., r\}$ and it is not a linear space. It can be 'embedded' in the space of zero-charge measures, denoted by $\mathcal{Z}(X)$. A Borel signed measure ν on X is with zero charge if its positive part ν^+ and its negative part ν^- have the same integral. Formally, for our discrete space X, we have:

$$\mathcal{Z}(X) := \{(a_1, a_2, ..., a_r) \in \mathbb{R}^r | \sum_{i=1}^{r} a_i = 0\}.$$

The Kantorovich-Rubinshtein norm $||\nu||_K$ of an element $\nu \in \mathcal{Z}(X)$ is defined as the Kantorovich distance between the positive and negative parts of ν, i.e.,

$$||\nu||_K := D_T(\nu^+, \nu^-).$$

The space of all Lipschitz (up to additive constants) functions with the Lipschitz norm is the conjugate normed space of the space $(\mathcal{Z}(X), || \cdot ||_K)$, so $\mathcal{Z}(X)$ is complete (Banach space). Note that the simplex of probability distributions $\mathcal{D}_r^1$ is homeomorphic with a closed set of the Banach space $\mathcal{Z}(X)$.

Denote

$$\mathcal{P}_r := \{M = (\mu_{ij})_{i,j=1,..,r} |\ \mu_{ij} \geq 0,\ \sum_{j=1}^{r} \mu_{ij} = 1\}$$

the set of r-dimensional stochastic matrices.

In the following, we will consider stochastic matrices from $\mathcal{P}_r$ with the operator norm derived from the '*transportation norm*' on the Banach space $\mathcal{Z}(X)$. The operator norm of a matrix $M \in \mathcal{P}_r$ will be

$$||M|| := \sup\{||\mu M||_K |\ \mu \in \mathcal{Z}(X);\ ||\mu||_K = 1\}. \quad (1)$$

This is a key hypothesis of our paper. It is worthy to note that this norm is not longer equivalent with the traditional matrix norms that can be defined also for stochastic matrices.

III. IMCs as Nonautonomous Dynamical Systems

In this section, we explain how an inhomogeneous Markov chain can be specified as nonautonomous dynamical system. The reason for doing this is the fact that the ergodic related issues treated in this paper are closely related to the analogous ones existing for dynamical systems.

An autonomous dynamical system is described via a group or semi-group of (homogeneous/autonomous) mappings (i.e., they depend only on the elapsed time $t - t_0$ since starting and not directly on the current time t or the starting time t_0).

For a *nonautonomous system* both the current time t and starting time t_0 are important rather than just their difference. The semi-group formalism existing for autonomous dynamical systems can be generalized to nonautonomous dynamical systems as the two-parameter semi-group or *process formalism* of a nonautonomous dynamical system, where both t and t_0 are the parameters. The process formulation is given below.

The process formulation of an abstract nonautonomous dynamical system on a metric state space (X, d) and time set T, where $T = \mathbb{R}$ (for a continuous time process) or $T = \mathbb{Z}$ (for a discrete time process) is motivated by the study of nonautonomous ordinary differential equations on $\mathbb{R}^d$.

Process formulation. A process is a continuous mapping

$$(t, t_0, x_0) \mapsto \varphi(t, t_0, x_0) \in X$$

for $t, t_0 \in T$ and $x_0 \in X$ with $t \geq t_0$, which satisfies the initial value and evolution properties:'

(i) $\varphi(t_0, t_0, x_0) = x_0$ for all $t_0 \in T$ and $x_0 \in X$,

(ii) $\varphi(t_2, t_0, x_0) = \varphi(t_2, t_1, \varphi(t_1, t_0, x_0))$ for all $t_0 \leq t_1 \leq t_2$ and $x_0 \in X$.

In order to present an IMC as a nonautonomous dynamical system, we need to identify its underlying evolution process.

Let us consider an inhomogeneous Markov chain defined on the state space $X = \{1, 2, ..., r\}$ with $r \times r$ transition probability matrices

$$P(t_0, t) = (p_{i,j}(t_0, t))_{i,j=1,..,r} \text{ for all } t, t_0 \in T \text{ with } t_0 \leq t.$$

Such matrices satisfy

$$P(t_0, t_0) = I_r \text{ for all } t_0 \in T$$

and the Chapman-Kolmogorov equation, i.e.,

$$P(t_0, s)P(s, t) = P(t_0, t) \text{ for all } t_0 \leq s \leq t.$$

If the states of the Markov chain at time t_0 satisfy the probability vector $p(t_0) \in \mathcal{D}_r^1$, then they are distributed according to a probability vector

$$p(t) = p(t_0)P(t_0, t) \text{ at } t \geq t_0.$$

Therefore, on $\mathcal{D}_r^1$ we can define a process

$$\varphi(t, t_0, p_0) := p_0 P(t_0, t).$$

Therefore, an inhomogeneous Markov chain induces a nonautonomous dynamical system defined on an appropriate space of probability distributions. At each step, a stochastic matrix (that depends on time) dictates the next 'state' of this dynamical system.

IV. Setting the Problem

In this section, we present the framework of inhomogeneous Markov chains in dynamic environments and the problem of perturbation analysis of its natural distribution.

A. DTIMC in dynamic environments

Let us consider a stochastic chain $\{X_k | k \in \mathbb{Z}_+\}$ with a finite state space $X := \{1, ..., r\}$ but with transition probabilities matrices $\{P_k | k \in \mathbb{Z}_+\}$. We see $\{P_k | k \in \mathbb{Z}_+\}$ a $\mathcal{P}_r$-valued semi-dynamical system. This semi-dynamical system is called dynamic environment and $\{X_k | k \in \mathbb{Z}_+\}$ is called a *Markov chain in a dynamic environment*. For the purpose of this work, it is convenient to extend the semi-dynamical system (P_k) from $\mathbb{Z}_+$ to $\mathbb{Z}$.

1) Geometric ergodicity: Now we define the concept of geometric ergodicity (exponentially mixing) for a DTIMC. This concept is essential for the results of this work, but it is well-defined only if we use the operator norm defined by (1).

Definition 1 (Geometric ergodicity): A Markov chain (X_k) is called exponentially mixing (or, geometrically ergodic) if there exist $C \geq 1$, $\lambda \in [0;1)$ such that

$$||P_k P_{k+1} ... P_{k+\tau-1}|| \leq C\lambda^\tau, \forall k \in \mathbb{Z}, \tau \in \mathbb{Z}_+^*. \quad (2)$$

The condition (2) can be written in a simplified way using the Chapman-Kolmogorov equation, as follows:

$$||P(s,t)|| \leq C\lambda^{t-s+1}, \forall s, t \in \mathbb{Z}, \ s \leq t$$

for some $C \geq 1$; $\lambda \in [0;1)$.

Note that in the exponential mixing condition can not be fulfilled if we use matrix norms. The novel idea is to use an operator norm defined on appropriate Banach space that is strictly embedded in $\mathbb{R}^r$.

Proposition 1: Exponential mixing is an open property with respect to supremum norm of the sequence $\mathcal{P}$. In fact

$$\begin{aligned} ||Q_k - P_k|| &\leq \delta < \frac{1-\lambda}{C}, \forall k \in \mathbb{Z} \Longrightarrow \\ ||Q_k Q_{k+1} ... Q_{k+\tau-1}|| &\leq C(\lambda + C\delta)^\tau, \forall k \in \mathbb{Z}, \tau \in \mathbb{Z}_+^*. \end{aligned}$$

Proof: Assume that $||Q_k - P_k|| \leq \delta < \frac{1-\lambda}{C}$, $\forall k \in \mathbb{Z}$.

The proof can be easily done by induction with respect to $\tau \in \mathbb{Z}_+^*$.

Step1. Take $\tau = 1$ and prove that $||Q_k|| \leq C(\lambda + C\delta)$, $\forall k \in \mathbb{Z}$. This comes as a consequence of the fact that $C \geq 1$ and then, we have:

$$\begin{aligned} ||Q_k|| &\leq ||Q_k - P_k|| + ||P_k|| \leq \\ &\leq \delta + C\lambda \leq C(\lambda + C\delta). \end{aligned}$$

Step 2. Suppose that $||Q_k Q_{k+1} ... Q_{k+l-1}|| \leq C(\lambda + C\delta)^l$, $\forall k \in \mathbb{Z}$, for all $l \in \mathbb{Z}_+^*$, $l \leq \tau \in \mathbb{Z}_+^*$. Then we have to prove that

$$||Q_k Q_{k+1} ... Q_{k+\tau}|| \leq C(\lambda + C\delta)^{\tau+1}, \forall k \in \mathbb{Z}.$$

This is an easy consequence of the following sequence of inequalities:

$||Q_k Q_{k+1} ... Q_{k+\tau}|| = ||(Q_k - P_k)Q_{k+1} ... Q_{k+\tau} + P_k Q_{k+1} ... Q_{k+\tau}|| = ||(Q_k - P_k)Q_{k+1} ... Q_{k+\tau} + P_k(Q_{k+1} - P_{k+1})Q_{k+2} ... Q_{k+\tau} + P_k P_{k+1} Q_{k+2} ... Q_{k+\tau}|| = ...$
$\leq ||Q_k - P_k|| \cdot ||Q_{k+1} ... Q_{k+\tau}|| + ||P_k|| \cdot ||Q_{k+1} - P_{k+1}|| \cdot ||Q_{k+2} ... Q_{k+\tau}|| + ||P_k P_{k+1}|| \cdot ||Q_{k+2} - P_{k+2}|| \cdot ||Q_{k+3} ... Q_{k+\tau}|| + ... + ||P_k P_{k+1} ... P_{k+\tau-1}|| \cdot ||Q_{k+\tau} - P_{k+\tau}|| \leq$
$\leq \delta[C(\lambda + C\delta)^\tau + C^2\lambda(\lambda + C\delta)^{\tau-1} + C^2\lambda^2(\lambda + C\delta)^{\tau-2} + ... + C\lambda^\tau] \leq \delta C^2 \frac{(\lambda+C\delta)^{\tau+1} - \lambda^{\tau+1}}{C\delta} \leq C(\lambda + C\delta)^{\tau+1}$. q.e.d. ■

2) The forward equation:

Definition 2: A sequence $(\mu^k)_{k \in \mathbb{Z}} \subset \mathcal{D}_r^1$ is called natural distribution (or, set of absolute probabilities in the terminology of [1], [16]) for $\mathcal{P}$ if

$$\mu^{k+1} = \mu^k P_k, \forall k \in \mathbb{Z}. \quad (3)$$

The absolute probabilities have been defined initially by Kolmogorov [16] as a set of $1 \times r$ Markov matrices that satisfy:

$$\mu^s P(s,t) = \mu^t, \ s < t.$$

Because (3) is defined recursively by going forward in time, this relation represent the *forward equations* for the Markov chain. In fact, the relation (3) describes the transfer (Perron Frobenius) operator associated to the underlying inhomogeneous Markov chain. The absolute probabilities always exist, but the uniqueness is true only under some certain conditions. The following characterization result was proved in [16]:

Theorem 2: There exists a unique set of absolute probabilities $(\mu^k)_{k \in \mathbb{Z}}$ if and only if

$$P(s,k) \to Q(k) \text{ as } s \to -\infty, \quad (4)$$

where $Q(k)$ has identical rows. Each row of $Q(k)$ is then identical with μ^k.

The condition in the Th.2 is called some times *weak ergodicity* for the Markov chain and the matrices $Q(k)$ are the *limit regimes*.

In the following, we suppose that $\mathcal{P} = (P_k)_{k \in \mathbb{Z}}$ is a discrete-time dynamical system.

Theorem 3 (Uniqueness of natural distribution): If Markov chain (X_k) is exponentially mixing then there exists a unique natural distribution $(\mu^k)_{k \in \mathbb{Z}}$ associated to our DTIMC.

Proof: According to the Th.2, we have to prove that when $s \to -\infty$ the matrix $P(s,k)$ converges to a matrix with identical rows. Using the exponential mixing condition, we have

$$||P(s,k)|| \leq C\lambda^{k-s+1}, \forall s, k \in \mathbb{Z}, \ s < k.$$

If we fix k, and make $s \to -\infty$, the quantity λ^{k-s+1} goes to zero since $\lambda \in [0,1)$. So, the norm of $P(s,k)$ goes to zero as $s \to -\infty$. Now, since this norm is defined with respect to the Kantorovich-Rubinshtein norm $|| \cdot ||_K$ on $\mathcal{Z}(X)$, we can take the supremum with respect to all zero charge measures μ with $||\mu||_K = 1$. In particular, if we take zero charge measures of the form $\mu := \delta_i - \delta_j$ $(i \neq j)$, with δ_i equal to the Dirac distribution, we obtain that the rows i and j in the matrix $Q(k) := \lim_{s\to-\infty} P(s,k)$ are identical, and so on. Note that since Kantorovich-Rubinshtein norm is compatible with the metric of the underlying space X, we have, indeed, that $||\mu||_K = ||\delta_i - \delta_j||_K = d(i,j) = 1$ $(i \neq j)$. ■

There is a unique probability over the σ-algebra generated by the Markov chain trajectories (paths consistent with the Markov chain realizations from time $-\infty$). We will denote it also by $\mathbf{P}$. It can be defined by its marginals on time intervals m to $n > m$ by starting in μ^m at time m and evolving forwards, i.e.,

$$\mathbf{P}(i_m, ..., i_n) = \mu^m(i_m)p_{i_m,i_{m+1}}(m, m{+}1)...p_{i_{n-1},i_n}(n{-}1, n).$$

Applications of Markov chains in dynamic environments can be found in robotics. An example at hand is the walking robot modelled as a DTIMC from the recent paper [22]. In the above reference, the authors develop a new control strategy for the dynamics of a walking robot in a real changing environment based on the properties of inhomogeneous Markov chains.

B. CTIMC in dynamic environments

Let us consider an inhomogeneous Markov chain $\{X_t | t \in \mathbb{R}_+\}$ with continuous time and the state space $X := \{1, ..., r\}$. Denote by $P(s,t)$ the transition matrix for the chain (X_t) and the intensity matrix $A(t) = (a_{ij}(t))_{i,j=0}^r$. The elements of the intensity matrix are defined as follows:

$$P(X_{t+h} = j | X_t = i) = \begin{cases} a_{ij}(t)/a_{ii}(t)h + o(h) & \text{if } j \neq i, \\ 1 - ha_{ii}(t) + o(h) & \text{if } j = i, \end{cases}$$

where all $o(h)$ are uniform in i and $a_{ii}(t) = -\sum_{k\neq i} a_{ik}(t)$.

To be consistent with the previous section, we extend the time interval from $[0, +\infty)$ to $(-\infty, +\infty)$.

1) Geometric ergodicity:

Definition 3 (Weak ergodicity): The Markov chain (X_t) is called weakly ergodic if for all t there exists a matrix $Q(t)$ with identical rows such that

$$\lim_{s\to-\infty} P(s,t) = Q(t).$$

We can remark that the weak ergodicity of the Markov chain is equivalent with

$$\lim_{s\to-\infty} \alpha P(s,t) = 0,$$

for any t and any $\alpha \in \mathcal{Z}(X)$.

Definition 4 (Geometric ergodicity): The Markov chain (X_t) is called geometrically ergodic (or, exponentially mixing) if there exist the constant $C > 1$ and $\lambda \in [0,1)$ such that

$$||P(s,t)|| \leq C\lambda^{t-s}, \forall s,t \in \mathbb{R},\ s < t. \tag{5}$$

It is clear that the geometric ergodicity implies weak ergodicity.

The 'computation' of the transition matrices using the intensity matrices can be realised using the *product integral* formula (see [7]):

$$P(s,t) = \prod_{(s,t]} (I + A(du)),\ s < t. \tag{6}$$

We expect that the geometric ergodicity for CTIMC to be an open property, but the proof is not straightforward as in the case DTIMC. Our conjecture is that we could use the formula (6) to prove the openness of the geometric ergodicity.

Using the properties of the product integral and the Duhamel equation, according to [8], it is possible to prove that if $\overline{A}$ is another intensity matrix such that $a(t) = A(t) - \overline{A}(t)$ is bounded with respect to the supremum norm, then

$$||\prod(I + A(du)) - \prod(I + \overline{A}(du))||_\infty \leq \overline{C}||a||_\infty$$

2) The forward equation:

Definition 5: A function $\mu : \mathbb{R} \to \mathcal{D}_r^1$ is called natural distribution (or, $\mu(t)$ represents the set of absolute probabilities) if

$$\mu(s)P(s,t) = \mu(t),\ s < t. \tag{7}$$

From the theory of Markov chains, it is well known that the natural distribution is solution of the forward Kolmogorov equation for the Markov chain, i.e.,

$$d\mu/dt = \mu(t)A(t),\ t \in \mathbb{R}, \tag{8}$$

which is a time varying linear system on the whole line.

As in the case of CTIMC, the uniqueness of the natural distribution is ensured by the weak ergodicity property.

V. Robustness of IMC Distributions

The aim of this section is to investigate those conditions that ensure the existence of the IMC natural distributions when the transition probabilities are perturbed at different moments of time. Note that due to the nonstationarity of the underlying Markov chain, we cannot use directly the operator semigroup and the potential kernel like in the homogeneous case. The mathematical tool to address this issue is the evolution semigroup of operators associated to the IMC. The exponentially mixing condition is equivalent with the exponential stability of this semigroup. Then the robustness of the natural distribution will be closely related with the existence of the Green operator under perturbations of the transition probabilities.

The purpose of this section is to study the robustness of the non-autonomous (time-varying) difference/differential equation that defines the natural distribution of an IMC (see (3) and (8)). The main assumption is that the underlying *IMC is exponentially mixing*.

Note that although $\mathcal{D}_r^1$ is not a Banach space, $\mathcal{D}_r^1$ is homeomorphic to a closed subset of the Banach space $\mathcal{Z}(X)$ through the following embedding:

$$\mathcal{D}_r^1 \ni \mu \mapsto \widehat{\mu} =: \mu - \mu_0 \in \mathcal{Z}(X),$$

where μ_0 is a fixed probability distribution.

For DTIMC, the non-autonomous difference equation on the space $\mathcal{Z}(X)$, which we need to study is

$$\widehat{\mu}^{k+1} = \widehat{\mu}^k P_k + \widehat{f}^{k+1}, \forall k \in \mathbb{Z}, \tag{9}$$

where $\widehat{f}^{k+1} := \mu_0(P_k - I) \in \mathcal{Z}(X)$ and I is the identity matrix.

For CTIMC, the non-autonomous differential equation on the space $\mathcal{Z}(X)$, which will be the studied is the following Cauchy equation:

$$d\widehat{\mu}/dt = \widehat{\mu}(t)A(t) + f(t),\ t \in \mathbb{R}, \tag{10}$$

where $\widehat{f}(t) := \mu_0 A(t)$, where $A(t)$ is the intensity matrix at time t. Note that $\widehat{f}(t) \in \mathcal{Z}(X)$ since any $A(t)$ is a zero row-sum matrix.

A. Discrete Evolution Semigroup

Let us consider the nonautonomous difference equation given by (9). In a standard way (see [2] and the references therein), we can associate an *evolution family* $\mathcal{U} = \{U_{n,m} | m \geq n \in \mathbb{Z}\}$ on the Banach space $\mathcal{Z}(X)$ by

$$U_{n,m} = \begin{cases} P(n,m) & \text{if } n < m; \\ I & \text{if } n = m \end{cases}. \tag{11}$$

Note that $U_{n,m} = U_{n,p}U_{p,m}$ for all $m \geq p \geq n \in \mathbb{Z}$ (due to the Chapman-Kolmogorov equation).

For each $j \in \mathbb{Z}_+$; let T_j be the linear operator given by

$$(T_j\widehat{f})(n) := \widehat{f}^{n-j}U_{n-j,n},\ \text{for all } n \in \mathbb{Z}, \tag{12}$$

for any $\widehat{f} = (\widehat{f}^k)_{k\in\mathbb{Z}} \in l_\infty(\mathbb{Z}, \mathcal{Z}(X))$. Here, $l_\infty(\mathbb{Z}, \mathcal{Z}(X))$ represents the Banach space of all $\mathcal{Z}(X)$-valued sequences which are bounded (w.r.t. the Kantorovich-Rubinshtein norm).

The family $\mathcal{T} = \{T_j\}_{j\in\mathbb{Z}_+}$ is called the *evolution semigroup* associated to $\mathcal{U}$ on $l_\infty(\mathbb{Z}, \mathcal{Z}(X))$.

The exponential mixing property ensures that the evolution family $\mathcal{U}$ and the evolution semigroup $\mathcal{T}$ are *uniformly exponentially stable*. As well, the image of natural distributions of the underlying IMC belong to the space $l_\infty(\mathbb{Z}, \mathcal{Z}(X))$.

Having in mind the notion of infinitesimal generator for a strongly continuous semigroup, we define the "infinitesimal generator" for a discrete semigroup as being

$$\mathcal{L} := T_1 - I. \tag{13}$$

The following proposition provides the expression of the kernel operator (Green kernel) associated to the semigroup $\mathcal{T}$. This is just a simple consequence of Lemma 4.1 from [2] for our evolution family $\mathcal{U}$.

Proposition 4: Let $\widehat{f} = (\widehat{f}^n)_{n\in\mathbb{Z}}$ and $\widehat{g} = (\widehat{\nu}^n)_{n\in\mathbb{Z}}$ be two elements in $l_\infty(\mathbb{Z}, \mathcal{Z}(X))$. The following two statements are equivalent:

(i) $\mathcal{L}\widehat{g} = \widehat{f}$;

(ii) For each $n \in \mathbb{Z}$, the following limit exists

$$\widehat{u}^n := \lim_{k\to-\infty} \sum_{l=k}^{n} \widehat{f}^l U_{l,n}$$

and $\widehat{g}^n = \widehat{u}^n$ for all $n \in \mathbb{Z}$.

As a consequence of Theorem 4.2 from [2] and the exponential mixing property, we obtain the following result.

Proposition 5: Given the evolution family $\mathcal{U}$ given by (11), which is uniformly exponential stable, and its associated semigroup $\mathcal{T}$ given by (12), the following two conditions are equivalent:

(i) The infinitesimal generator $\mathcal{L}$ of $\mathcal{T}$ given by (13) is invertible;

(ii) For each $\widehat{f} = (\widehat{f}^n)_{n\in\mathbb{Z}}$ in $l_\infty(\mathbb{Z}, \mathcal{Z}(X))$ and each $n \in \mathbb{Z}$ there exists

$$\widehat{\nu}^n := \lim_{k\to-\infty} \sum_{l=k}^{n} \widehat{f}^l U_{l,n}.$$

Moreover, $(\widehat{\nu}^n)$ belongs to $l_\infty(\mathbb{Z}, \mathcal{Z}(X))$.

Therefore, the kernel operator $G = \mathcal{L}^{-1}$ can be defined as

$$G\widehat{f}(n) := (T_1 - I)^{-1}\widehat{f}(n) = \lim_{k\to-\infty} \sum_{l=k}^{n} \widehat{f}^l U_{l,n}.$$

Considering the theory of non-autonomous difference equations, the solution of our difference equation (9) will be given by

$$\begin{aligned}
\widehat{\mu}^n &= (T_1 - I)^{-1}\widehat{f}(n) = \\
&= \lim_{k\to-\infty} \sum_{l=k}^{n} \widehat{f}^l U_{l,n} = \\
&= \lim_{k\to-\infty} \sum_{l=k}^{n-1} \mu_0(P_{l-1} - I)P(l,n) + \mu_0(P_{n-1} - I).
\end{aligned}$$

Doing the calculations, we can recover the solutions of (3) as the identical rows of the matrix $Q(k) := \lim_{s\to-\infty} P(s,k)$.

B. Continuous evolution semigroup

Let us consider the nonautonomous differential equation given by (10). It is known that to this equation (see [23] and the references therein), we can associate an *evolution family* (more precisely, an evolution system) $\mathcal{U} = \{U(s,t) | t \geq s \in \mathbb{R}\}$ on the Banach space $\mathcal{Z}(X)$ by

$$U(s,t) = \begin{cases} P(s,t) & \text{if } s < t; \\ I & \text{if } s = t \end{cases}$$

The evolution semigroup $\mathcal{T} = \{T_t\}_{t\geq 0}$ is defined in a similar way as in the discrete case, i.e.,

$$(T_t\widehat{f})(\tau) := \widehat{f}(\tau - t)U(\tau - t, \tau),\ \tau \in \mathbb{R},\ t \geq 0,$$

for any $\widehat{f} : \mathbb{R} \to \mathcal{Z}(X)$ that belongs to $L_\infty(\mathbb{R}, \mathcal{Z}(X))$. Here, $L_\infty(\mathbb{R}, \mathcal{Z}(X))$ represents the space of functions $\widehat{h} : \mathbb{R} \to \mathcal{Z}(X)$ such that

$$||\widehat{h}||_\infty = \sup_t ||\widehat{h}(t)||_K < \infty.$$

The generator $\mathcal{L}$ of the evolution semigroup $\mathcal{T}$ is given by

$$(\mathcal{L}\widehat{f})(\tau) := -\frac{d\widehat{f}}{d\tau} + \widehat{f}(\tau)A(\tau),\ \tau \in \mathbb{R}$$

with domain the set of differentiable functions $\widehat{f}$ such that $\mathcal{L}\widehat{f} \in L_\infty(\mathbb{R}, \mathcal{Z}(X))$ (see [5]).

Then the kernel operator G is defined

$$\begin{aligned}
G\widehat{f}(s) \quad : &= \int_0^\infty (T_t\widehat{f})(s)dt \\
&= \int_{-\infty}^{s} \widehat{f}(\tau)U(\tau, s)d\tau.
\end{aligned}$$

It is known that G is equal to $(-\mathcal{L})^{-1}$ provided that the semigroup $\mathcal{T}$ or the evolution family $\mathcal{U}$ is uniformly exponentially stable. The uniform exponential stability of $\mathcal{U}$ is nothing else but the exponential mixing of the underlying CTIMC, i.e. condition (5).

C. Perturbation of the IMC Distributions

Let us consider first a DTIMC and its associated forward equation (which is nonautonomous discrete linear difference equation):

$$\mu^{k+1} = \mu^k P_k, \, \forall k \in \mathbb{Z}.$$

A perturbed version of this equation can be written as follows:

$$\mu^{k+1} = \mu^k (P_k + \Delta_k), \, \forall k \in \mathbb{Z}. \tag{14}$$

where Δ_k is a zero row-sum perturbation r-dimensional matrix.
Then the corresponding perturbed equation defined on $\mathcal{Z}(X)$ can be written as follows:

$$\widehat{\mu}^{k+1} = \widehat{\mu}^k (P_k + \Delta_k) + \widehat{f}^{k+1}, \, \forall k \in \mathbb{Z}$$

where $\widehat{f}^{k+1} := \mu_0 (P_k - I) + \mu_0 \Delta_k \in \mathcal{Z}(X)$.

In a similar way, we can consider the forward equation of a CTIMC, and its perturbed version:

$$d\mu/dt = \mu(t)(A(t) + \Delta(t)), \, t \in \mathbb{R}.$$

Now, the corresponding perturbed equation defined on $\mathcal{Z}(X)$ is:

$$d\widehat{\mu}/dt = \widehat{\mu}(t)(A(t) + \Delta(t)) + \widehat{f}(t), \, t \in \mathbb{R}$$

where $\widehat{f}(t) := \mu_0(A(t) + \Delta(t))$.

Then the stability of the perturbed equation is determined by whether or not the perturbed generator $\mathcal{L} + \overline{\Delta}$ is invertible. Here, $\overline{\Delta}$ is the multiplication operator, which is defined:

- in the discrete case, as $\overline{\Delta}\widehat{f}(k) := \widehat{f}^k \Delta_k$ for all $\widehat{f} = (\widehat{f}^k)_{k\in\mathbb{Z}} \in l_\infty(\mathbb{Z}, \mathcal{Z}(X))$;
- in the continuous case, as $\overline{\Delta}\widehat{f}(t) := \widehat{f}(t)\Delta(t)$ for all $\widehat{f} \in L_\infty(\mathbb{R}, \mathcal{Z}(X))$.

The stability of such a linear system is usually measured using the concept of stability radius. This represents the "size" of the smallest operator under which the exponential stability no longer holds for the additively perturbed system. This concept was introduced by Hinrichsen and Pritchard for studying robustness of linear (autonomous and non-autonomous) systems [10], [11].

Therefore, the stability radius $\rho_{stab}(\mathcal{U})$ of $\mathcal{U}$ with respect to a matrix perturbation is the largest $||\Delta_{(*)}||_\infty$ such that $\mathcal{L} + \overline{\Delta}$ is invertible.

In the literature, there exist many results that provide upper and lower bounds for the stability radius. See, e.g., [4].

The geometric ergodicity allows us to apply the existing results that are usually based on algebraic arguments and to obtain the following proposition.

Proposition 6: If the IMC (X_k) (in discrete time) or (X_t) (in continuous time) is exponentially mixing, we have:

$$\frac{1}{||G||} \le \rho_{stab}(\mathcal{U}) \le \frac{1}{r(G)}$$

where $r(G)$ denotes the spectral radius of the kernel operator.

Recall that the spectral radius of an operator G is defined as

$$r(G) := \sup\{|\alpha| : \alpha \in \sigma(G)\},$$

where $\sigma(G)$ is the spectrum of G (i.e., the set of all scalars α for which $\alpha I - G$ is not invertible). It is relevant to mention the following known Beurling formula:

$$r(G) = \lim_{n\to\infty} ||G^n||^{1/n}.$$

D. Example: One Time Perturbation

As an example, suppose we perturb the probability transition matrix P_{k_0} at one time k_0. Specifically, we consider the perturbed stochastic matrix

$$Q_{k_0} = P_{k_0} + C_{k_0}$$

where C_{k_0} is a zero row-sum perturbation matrix. The zero row-sum condition is necessary to ensure that the perturbed matrix Q_{k_0} remains a stochastic matrix. The natural distribution associated to the perturbed Markov chain is denoted by $(\nu^k)_{k\in\mathbb{Z}}$.

Proposition 7: If the underlying DTIMC is exponentially mixing, then the change of the natural distribution for $k > k_0$ ($\nu^k = \mu^k + \Delta\mu^k$) is given by

$$\Delta\mu^k = \mu^{k_0} C_{k_0} U_{k_0+1,k-1}$$

and the effect decays at least exponentially:

$$||\Delta\mu^k|| \le ||C_{k_0}|| C \lambda^{k-k_0-1}. \tag{15}$$

Proof: The proof is an easy consequence of the exponentially mixing condition (2).

VI. Perturbation Control

Let us consider a DTIMC $\{X_k | k \in \mathbb{Z}_+\}$ with a finite state space $X := \{1, ..., r\}$ but with transition probabilities depending on some global control parameters θ. Suppose that $\theta \in \Theta$, where Θ is a compact convex set in $\mathbb{R}^l$. This could be a parameter of the system design, or of an input process, or one that is associated to a control policy.

Let us take $\{P_k(\theta) | k \in \mathbb{Z}_+\}$ a controlled $\mathcal{P}_r$-valued dynamical system. $P_k(\theta)$ will give the transition probabilities of the chain (X_k) at the time k depending on the control input θ.

In this case, we say that the chain $\{X_k | k \in \mathbb{Z}_+\}$ is Markov if for any $k, n \in \mathbb{Z}_+$ and $i_0, i_1, ..., i_n \in \{1, ..., r\}$ the following relation is satisfied

$$\begin{aligned} \mathbf{P}\{X_k &= i_0, X_{k+1} = i_1, ..., X_{k+n} = i_n | \theta\} \\ &= \mathbf{P}\{X_k = i_0 | \theta\} \prod_{l=0}^{n-1} p_{k+l}(\theta, i_l, i_{l+1}), \end{aligned}$$

where $p_k(\theta, i, j)$ are the elements of the matrix $P_k(\theta)$.

We extend the semi-dynamical system $(P_k(\theta))$ from $\mathbb{Z}_+$ to $\mathbb{Z}$.

For the Markov chain evolving in dynamic environment depending on the parameter θ, the following assumption will be in force for the most of our results.

Assumption 1 (Uniform exponential mixing): Suppose that the family of Markov chains $\{X_k(\theta)|\theta \in \Theta\}$ is a uniformly exponentially mixing w.r.t. θ, i.e. the sequence of transition matrices $\mathcal{P} = (P_k(\theta))_{k\in\mathbb{Z}}$ satisfies (2), where the constants $C \geq 1$, $\lambda \in [0;1)$ do not depend on $\theta \in \Theta$.

A natural control problem is to design an optimal zero row-sum perturbation matrix C_K such that the perturbed environment $(Q_k(\theta))_{k\in\mathbb{Z}}$ has a natural distribution that converges to a desired one, denoted by $(\mu_d^k)_{k\in\mathbb{Z}}$. In the light of [3], this problem can be called the *inverse perturbation problem at step* K. At the step K, we could have numerous perturbation matrices C_K (shocks) that can force the chain to make a transition from a nondesirable natural distribution to a desirable one. In principle, such environment perturbations constitute feasible control strategies and can therefore be used to drive the chain towards a targeted sequence of absolute probability distributions.

We assume that the system remains in the exponential mixing regime. The relation (15) shows that the perturbation effect at step K in the structure of the natural distribution acts mostly on the distribution ν^{K+1}, then it decays exponentially. Therefore, the control strategies we are looking for will be represented by continual changes to the environment and are applied step by step in a model predictive manner. Suppose we start with $(\mu^k)_{k\in\mathbb{Z}}$, apply a shock at step K obtaining $(\nu^k)_{k\in\mathbb{Z}}$, then apply another shock at step K' and getting another natural distribution $(\nu'^k)_{k\in\mathbb{Z}}$, and so on. We continue the process until we obtain a natural distribution that is close to the desired one (w.r.t. an appropriate norm).

Therefore, it is sufficient to describe the control strategy we apply only for one step. Let $\mathcal{D}_K$ be the set of *feasible control strategies* for which the perturbed Markov chain is forced to have as natural distribution the sequence $(\widetilde{\mu}^k)_{k\in\mathbb{Z}}$, i.e.,

$$\mathcal{D}_K := \{C_K(\theta) \mid \widetilde{\mu}^{K+1} = \widetilde{\mu}^K(P_K(\theta) + C_K(\theta))\}.$$

Mainly, finding a feasible control input θ means to design the perturbation matrix $C_K(\theta)$ such that $(\widetilde{\mu}^k)_{k\in\mathbb{Z}}$ is a natural distribution for the perturbed Markov chain. The definition of $\mathcal{D}_K$ is inspired by the similar concept of the reference [3]. $\mathcal{D}_K$ is nonempty since we can always find the at least one perturbation matrix as follows:

$$C_K^0(\theta) = \mathbf{1}\widetilde{\mu}^{K+1} - P_K(\theta)$$

where $\mathbf{1}$ is the column vector whose components are equal to 1.

As well, using the above mentioned reference we can also define the optimal perturbation matrix as solution of the following optimization problem:

Minimize $|||C_K|||$ subject to $C_K \in \mathcal{D}_K$, where $|||\cdot|||$ is a suitable matrix norm.

Since, $\mathcal{D}_K$ is a convex set, robust control techniques can be adapted to characterize solutions or optimal solutions for the desired perturbation matrices. See [3], and the references therein.

VII. Sensitivity Analysis

In this section we study the dependence of the probability distributions associated to an IMC w.r.t. the parameter $\theta \in \Theta$. When a cost is associated to each state, a performance index of the system can be defined. Optimal control of the system requires the estimation of the performance index for a given value of the parameter and, as well, sensitivity analysis, i.e., to modify the current value of the parameter for improving the performance.

First, we consider a DTIMC. Suppose that its transition matrices are described by the dynamical system $\mathcal{P} = (P_k(\theta))_{k\in\mathbb{Z}}$, which depends smoothly on the control parameter θ. In this paper, we consider the C^1-dependance on the parameter θ of the transition matrices and other probability distributions only with respect to the Kantorovich metric. More precisely, we consider the concept of strong differentiability with respect to this metric. We can extend $\Theta \subset \mathbb{R}^l$ to an open set or to the whole Euclidean space $\mathbb{R}^l$. Then the differentiability of the probability measures depending on θ will be given in terms of the Fréchet differentiability.

Let us recall some definitions.

Definition 6: (i) Let V and W be some Banach spaces, and $U \subset V$ be an open set of G. A function $F : U \to W$ is said to be Fréchet differentiable at a point $u \in U$ if there exists a bounded linear operator $A_x : V \to W$ such that

$$\lim_{h\to 0} \left\| \frac{F(u+h) - F(u) - A_u h}{||h||} \right\| = 0. \tag{16}$$

If the limit (16) exists then we denote $DF(u) := A_u$ and call it the Fréchet derivative of F at u.

(ii) A function F that is Fréchet differentiable for any point of U is called C^1 if the function $u \mapsto DF(u)$ is continuous.

Using the above definition, we can define differentiability of probability distributions (or of transition matrices) w.r.t. the parameter $\theta \in \Theta$.

Definition 7: (i) $\widehat{\mu}(\theta)$ is called strongly differentiable, with the derivative $\frac{d\widehat{\mu}}{d\theta}$, if the function $\theta \in \Theta \mapsto \widehat{\mu}(\theta) \in \mathcal{Z}(X)$ is Fréchet differentiable on Θ.

(ii) $\mu(\theta) \in \mathcal{D}_r^1$ is called strongly differentiable if the function $\theta \in \Theta \mapsto \widehat{\mu}(\theta) = \mu(\theta) - \mu_0 \in \mathcal{Z}(X)$ is Fréchet differentiable on Θ, in which case we write $\frac{d\mu}{d\theta} := \frac{d\widehat{\mu}}{d\theta}$.

Note that the Fréchet differentiability for our measures is making use of the transportation metric. Since the topology induced by this metric on a space of finite diameter coincides with the weak topology [15], the strong differentiability of our probability measures will coincide the weak differentiability [18].

Since the chain rule is also valid for the Fréchet derivative, we can differentiate the natural distribution w.r.t. θ as follows.

Proposition 8: $(\mu^k(\theta))_{k\in\mathbb{Z}}$ depends smoothly on $\mathcal{P}$ with derivatives

$$\frac{d\mu^k}{d\theta} = \sum_{\tau\in\mathbb{N}} \mu^{k-\tau}\frac{dP_{k-\tau}}{d\theta}P_{k-\tau+1}...P_{k-1},\ k\in\mathbb{Z}.$$

The proof of this product rule is an easy consequence of the Ass. 1.

This result can be used further for gradient estimations of a suitable performance index.

Let us consider now a CTIMC, whose transition matrices $\{P(s,t,\theta)|s<t\}$ depend also on a control parameter θ. Naturally, in this case, the intensity matrices depend on θ, so we have $A(t,\theta)$. The scope is to differentiate the natural distributions given by (7) and to obtain the dependance on the differential of $A(t,\theta)$.

By using the product integral differentiability [7] the Ass. 1, the following formula can be derived:

$$\frac{d\mu(t)}{d\theta} = \int_{-\infty}^{t}\mu(s)\int_{(s,t]}P(s,u-)\frac{dA}{d\theta}(du)P(u,t)ds,\ t\in\mathbb{R}.$$

VIII. Conclusions

In this paper, we have presented our first results regarding robustness of inhomogeneous Markov chains that evolve in dynamic environments. The results are still in their infancy, but we expect that a further study of sensitivity analysis of IMCs will be possible based on this work.

The main assumption that makes our results possible is the geometric ergodicity of the underlying IMC. This assumption is fulfilled for many other stochastic models, like homogeneous Markov families with unique and aperiodic communicating component, or “weakly dependent” probabilistic cellular automata (PCA). In order to extend our results to PCA we need to replace the Kantorovich norm with a more suitable norm that captures the spatiality of PCA. Such a norm is the Dobrushin norm [17]. This will constitute the topic of a follow-up paper.

References

[1] Blackwell, D.: Finite Non-Homogeneous Chains. *Annals of Mathematics* **46**(4) (1945): 594-599.

[2] Buse, C., Khan, A., Rahmat, G., Tabassum, A.: Uniform Exponential Stability for Discrete Non-Autonomous Systems via Discrete Evolution Semigroups.

[3] Bouaynaya, N, Shterenberg, R, Schonfeld, D.: Robustness of Inverse Perturbation for Discrete Event Control. Conf Proc IEEE Eng Med Biol Soc. (2011): 2422-2425.

[4] Clark, S., Latushkin, Y., Montgomery-Smith, S., Randolph, T.: Stability Radius and Internal versus External Stability in Banach Spaces: An Evolution Semigroup Approach.

[5] Chicone, C., Latushkin, Y.: *Evolution Semigroups in Dynamical Systems and Differential Equations*. Mathematical Surveys and Monographs, Volume 70, AMS (1999).

[6] Cho, G.E., Meyer, C.D.: Comparison of Perturbation Bounds for the Stationary Distribution of a Markov Chain. *Linear Algebra Appl.* **335** (2001): 137-150.

[7] Gill, R.D., Johansen, S.: A Survey of Product-Integration with a View Toward Application in Survival Analysis. *The Annals of Statistics* **18**(4) (1990): 1501-1555.

[8] Gill, R.D.: *Lectures on Survival Analysis*. Springer LNM **1581** (1994): 115-241.

[9] Heidergott, B., Leahu, H., Lopker, A., Pflug, G.: Perturbation Analysis of Inhomogeneous Markov Processes. Working Paper (2012).

[10] Hinrichsen, D., Pritchard, A.J.: Stability Radii of Linear Systems. *Systems & Control Letters* **7** (1986): 1-10.

[11] Hinrichsen, D., Ilchmann, A., Pritchard, A.J.: Robustness of Stability of Time-Varying Linear Systems. J. of Differential Equations **82** (1989): 219-250.

[12] Holling, C.S.: Understanding the complexity of economics, ecological, and social systems. Ecosystems **4**: 390-405.

[13] Iosifescu, M.: *Finite Markov Processes and their Applications*. John Wiley & Sons (1980).

[14] Kartashov, N.V.: *Strong Stable Markov Chains.* VSP, Utrecht (1996).

[15] Kravchenko, A.S.: Completness of the Space of Separable Measures in the Kantorovich-Rubinshtein Metric. *Siberian Math. J.* **47**(1) (2006): 68-76.

[16] Kolmogorov, A.: Zur Theorie der Markoffschen Ketten. *Mathematishe Annalen* **112**(6) (1935): 155-160.

[17] MacKay, R.S.: Robustness of Markov Processes on Large Networks. *J. Difference Eqns & AppIns* **17** (2011): 1155-67.

[18] Pflug, G.: Gradient Estimates for the Performance of Markov Chains and Discrete Event Processes. *Ann. Oper.Res.* **39** (1992): 173-194.

[19] Schweitzer, P.: Perturbation Theory and Finite Markov Chains. *J. of Applied Prob.* **5** (1968): 410-413.

[20] Skorokhod, A.V., Hoppensteadt, F.C., Salehi, H.: *Random Perturbation Methods with Applications in Science and Engineering*. Applied Mathematical Sciences Vol. **150** Springer-Verlag New York, Inc. (2002).

[21] Vershic, A.M.: Kantorovich Metric: Initial History and Little-Known Applications. *J. Math Sci* **133** (2006): 1410-7.

[22] Vladareanu, L., Tont, G., Yu, H., Gal, A.: Walking Robots using the Petri Nets and Markov Chains. Preprint (2012).

[23] Zeifman, A.I., Isaacson, D.L.: On Strong Ergodicity for Nonhomogeneous Continuous-Time Markov Chains. *Stochastic Processes and their Applications* **50** (1994): 263-273.

[24] Zeifman, A.I.: Quasi-Ergodicity for Nonhomogeneous Continuous-Time Markov Chains. *J. of Applied Probability* **26**(3) (1989): 643-648.

[25] Zeifman, A.I.: On the Weak Ergodicity of Nonhomogeneous Continuous-Time Markov Chains. *J. of Mathematical Sciences* **93**(4) (1999): 612-615.