mathematics

Article

Strong Ergodicity in Nonhomogeneous Markov Systems with Chronological Order

P.-C.G. Vassiliou

Department of Statistical Sciences, University College London, Gower St., London WC1E 6BT, UK; vasiliou@math.auth.gr

Abstract: In the present, we study the problem of strong ergodicity in nonhomogeneous Markov systems. In the first basic theorem, we relax the fundamental assumption present in all studies of asymptotic behavior. That is, the assumption that the inherent inhomogeneous Markov chain converges to a homogeneous Markov chain with a regular transition probability matrix. In addition, we study the practically important problem of the rate of convergence to strong ergodicity for a nonhomogeneous Markov system (NHMS). In a second basic theorem, we provide conditions under which the rate of convergence to strong ergodicity is geometric. With these conditions, we in fact relax the basic assumption present in all previous studies, that is, that the inherent inhomogeneous Markov chain converges to a homogeneous Markov chain with a regular transition probability matrix geometrically fast. Finally, we provide an illustrative application from the area of manpower planning.

Keywords: strong ergodicity; nonhomogeneous Markov systems; rate of convergence

MSC: 60J10; 60J20

1. Introductory Notes

Consider a stochastic system that has a population of members categorized in different states. Three types of movements are possible in the system. Firstly, movements of members are probable among the states of the system; secondly, members are leaving the system from the various states; and thirdly, new members are entering the system to replace leavers and to expand the population. When the various movements of the system are modeled by a nonhomogeneous Markov chain, we call such a system a nonhomogeneous Markov system (NHMS).

An NHMS is actually a generalization of the classical Markov chain where we have one particle moving among the states without the possibility of leaving the system and probably being replaced by another with possibly different characteristics. Of great importance is the vector of absolute probabilities which consists of the probabilities of the particle to be in any state of the Markov chain. On the other hand, in an NHMS, we have a population of particles categorized according to their characteristics in the various states. Particles are leaving the population from all the states, and new particles are entering the population to replace them and to expand the population. Of great importance is the vector of the expected relative population structure. Hence, the problems to be solved are a lot harder, and new strategies and tools are used other than the simple Markov chain. The roots of the motive for the development of the theory of NHMS, which was first introduced in Vassiliou [1], could be summarized in the use of Markov models in manpower systems. This started with the work of Young and Almond [2], Young [3,4], and Bartholomew [5,6], and it was extended in the works of Young and Vassiliou [7], Vassiliou [8,9], and McClean [10,11]. In the book by Vassiliou [12], one can find the evolution of the theory of NHMS and the large diversity of its developments in various directions, that is, NHMS in discrete and continuous time, stochastic control in NHMS, Laws of Large Numbers for NHMS,

check for updates

Citation: Vassiliou, P.-C.G. Strong Ergodicity in Nonhomogeneous Markov Systems with Chronological Order. *Mathematics* **2024**, *12*, 660. https://doi.org/10.3390/math12050660

Academic Editor: Michael Voskoglou

Received: 11 January 2024
Revised: 5 February 2024
Accepted: 18 February 2024
Published: 23 February 2024

Perturbations theory, NHMS in a stochastic environment, Markov systems, and others. In Section 5.4 of [12], there is a synopsis of real and potential applications of NHMS which illustrates the breadth of applications and some of the reasons why the entire theory is central to these processes. Work on manpower planning using results from NHMS and new areas has continued throughout the years up to nowadays, for example, Garg, et al. [13–15], Ugwugo and McClean [16], Vassiliadis [17,18], Georgiou, et al. [19], Guerry [20,21], Pollard [22], Esquivel, et al. [23–25].

In Section 2 of the present study, we define and describe the NHMS in discrete time and space in a compact but hopefully readable way for the reader who comes in contact for the first time with these processes. We also state the expected relative population structure in the various states as a function of the parameters of the population that could be estimated from the available data. In Section 3, we start with some basic definitions of concepts and mathematical tools, as well as useful known results that will be used in what follows. The novel part of this section is Theorem 1, where we study strong ergodicity for NHMS by relaxing the basic assumption present in all studies of strong ergodicity for NHMS. That is, we will not assume that the inherent inhomogeneous Markov chain converges as time goes to infinity to a homogeneous Markov chain with a regular (it consists of one communicating class of states, which is aperiodic) transition probability matrix or, equivalently, that the inhomogeneous Markov chain is strongly ergodic. In Section 3, we start with some basic definitions of concepts and mathematical tools, as well as useful known results that will be used in what follows. The novel part of this section is Theorem 4, where we prove under what conditions the rate of convergence of strong ergodicity in a NHMS is geometrically fast. This is an important question in NHMS due to its large practical value. In Theorem 4, we relax the basic assumption present in all studies of the rate of convergence to its asymptotic behavior for a NHMS. That is, we will not assume that the inherent inhomogeneous Markov chain converge as time goes to infinity to a homogeneous Markov chain with a regular transition probability matrix geometrically fast. Finally, in Section 5, we provide an illustrative application from the area of manpower planning.

2. The NHMS in Discrete Time and Space

Let a population consist of any kind of entities and let us denote by $T(t)$ for $t = 1, 2, \ldots$, the total number of memberships at time t, that is, at the end of the interval $(t-1, t]$, which are being held by its members. At every point of time that a member leaves the population, the membership is being transferred to a new member. For example, members could be patients in a hospital and memberships the beds they occupy. It is assumed that the total number of memberships are known or the sequence $\{T(t)\}_{t=0}^{\infty}$ is a realization of a known stochastic process depending on the application. The memberships are distributed in a finite number of states and let $S = \{1, 2, \ldots, k\}$ be the state space. Important aspect of an NHMS (see Vassiliou [12], Section 5.2) is the *population structure*, that is, the vector of random variables

$$\mathbf{N}(t) = [N_1(t), N_2(t), \ldots, N_k(t)],$$

where $N_i(t)$ is the number of memberships in state i at time t. Also, very important is the *relative population structure,* which is the vector of random variables $\mathbf{q}(t) = \mathbf{N}(t)/T(t)$. We denote by $\mathbf{P}(t)$ the transition probability matrix of the internal transitions of the members of the population during the interval $(t-1, t]$, that is, the t-th interval. Also, we have probable leavers from the states in S in every time interval t and let us denote by

$$\mathbf{p}_{k+1}(t) = [p_{1,k+1}(t), p_{2,k+1}(t), \ldots, p_{k,k+1}(t)],$$

where the state $k+1$ represents the external environment. Finally, we have new entrants of memberships to the population in order to replace leavers and to expand the population.

Let us collect the probabilities of allocation of the memberships to the various states in the t-th interval in the following stochastic vector

$$\mathbf{p}_0(t) = [p_{01}(t), p_{02}(t), \ldots, p_{0k}(t)],$$

where the state 0 represents the new entrants with their memberships waiting to be allocated in the various states. Note that in what follows we assume that $\Delta T(t) = T(t) - T(t-1) \geq 0$.

The transition probability matrix of the memberships $\mathbf{Q}(t)$ in the t-th interval can be shown (Vassiliou [12], p. 193), and it is given by

$$\mathbf{Q}(t) = \mathbf{P}(t) + \mathbf{p}_{k+1}(t)\mathbf{p}_0(t). \tag{1}$$

We call the inhomogeneous Markov chain defined uniquely by the sequence of transition probability matrices $\{\mathbf{Q}(t)\}_{t=0}^{\infty}$ the *embedded or inherent nonhomogeneous Markov chain of* NHMS. A population or any physical phenomenon that could be modeled in the above described way is defined to be a *nonhomogeneous Markov system.*

Now, as previously, we define the relative population structure for a population which started at time s and is at time t to be $\mathbf{q}(s,t) = \mathbf{N}(s,t)/T(t)$, where $\mathbf{N}(s,t)$ is the population structure for the population. It could be proved that (Georgiou and Vassiliou [26] and Vassiliou [12], p. 195)

$$\mathbb{E}[\mathbf{q}(s,t)] = \mathbb{E}[\mathbf{q}(s,t-1)]\alpha(t-1)\mathbf{Q}(t) + b(t-1)\mathbf{p}_0(t), \tag{2}$$

where s is the initial time and, therefore, $\mathbf{q}(s)$ is the initial relative population structure which is known and

$$\alpha(t-1) = \frac{T(t-1)}{T(t)} \text{ and } b(t-1) = \frac{T(t) - T(t-1)}{T(t)}, \tag{3}$$

and where $\mathbf{q}(s,t)$ is the relative population structure at time t for the system that started with initial relative population structure $\mathbf{q}(s)$, which apparently is a random variable, and we denote by $\mathbb{E}[\mathbf{q}(s,t)]$ its expected value.

From (2), recursively, we obtain

$$\mathbb{E}[\mathbf{q}(s,t)] = \mathbf{q}(s)\frac{T(s)}{T(t)}\mathbf{Q}(s,t) + \frac{1}{T(t)}\sum_{\tau=1}^{t}\Delta T(s+t)\mathbf{p}_0(s+t)\mathbf{Q}(s+\tau,t), \tag{4}$$

where

$$\mathbf{Q}(s,t) = \mathbf{Q}(s+1)\mathbf{Q}(s+2)\ldots\mathbf{Q}(t).$$

3. Strong Ergodicity in NHMS with Chronological Order

The asymptotic behavior of NHMS and of nonhomogeneous Markov chains, as well as of homogeneous Markov chains, has been one of the central problems for many years, as can be seen in Refs. [12,27–31]. The asymptotic behavior of NHMS started with Vassiliou [1,32], and an updated evolution of these theorems and their variants could be found in Vassiliou [12]. In the present section, we provide and prove a basic theorem for strong ergodicity in NHMS when the transition probabilities matrices of the inherent Markov chain $\{\mathbf{Q}(t)\}_{t=0}^{\infty}$ are given in chronological order; that is, it is assumed that the time order of the elements of the sequence $\{\mathbf{Q}(t)\}_{t=0}^{\infty}$ is given and will not be changed. In Theorem 1 we relax the basic assumption present in all studies of asymptotic behavior for NHMS. That is, we will not assume that the inherent inhomogeneous Markov chain converge as time goes to infinity to a homogeneous Markov chain with a regular (it consists of one communicating class of states which is aperiodic) transition probability matrix or, equivalently, that the inhomogeneous Markov chain is strongly ergodic. We start with some basic definitions and results, which will be useful in what follows.

For what follows, we assume a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and consider an NHMS in discrete time and space. Therefore, we will not repeat this in every Definition, Lemma, Proposition, or Theorem.

Definition 1. *We say that the NHMS is strongly ergodic if and only if there exists a stochastic vector $\boldsymbol{\psi}$ such that*

$$\lim_{\nu \to \infty} \|\mathbb{E}[\mathbf{q}(t, t+\nu)] - \boldsymbol{\psi}\| = 0. \text{ for } t = 0, 1, 2, \ldots, \tag{5}$$

where $\|.\|$ from now on is any vector norm, except if it is otherwise stated.

Definition 2. *If $\mathbf{Q}$ is any finite stochastic matrix with state space $S = \{1, 2, \ldots, k\}$, then the Dobrushin ergodicity coefficient that is known is given by*

$$\tau_1(\mathbf{Q}) = \frac{1}{2} \max_{i,r} \sum_{j=1}^{k} \left| q_{ij} - q_{kj} \right|. \tag{6}$$

We now define another class of ergodicity coefficients which will be generated by different norms.

Definition 3 (Seneta [27])**.** *Let the set*

$$D^n = \left\{ \mathbf{x} : \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \geq \mathbf{0}, \mathbf{x}\mathbf{1}^\top = 1 \right\},$$

and by $d(\ldots)$ any metric on this set. Then, the quantity

$$\tau(\mathbf{P}) = \sup_{\mathbf{x}, \mathbf{y} \in D^n} \frac{d(\mathbf{x}\mathbf{P}, \mathbf{y}\mathbf{P})}{d(\mathbf{x}, \mathbf{y})} \text{ with } \mathbf{x} \neq \mathbf{y}, \tag{7}$$

for any stochastic matrix, $\mathbf{P}$ *is called a coefficient of ergodicity.*

Remark 1. *All matrix norms on $\mathbb{R}^n$ provide an appropriate metric on D^n via $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$. Then for any stochastic matrix* $\mathbf{P}$ *we get*

$$\tau_{\|.\|}(\mathbf{P}) = \sup_{\mathbf{x}, \mathbf{y} \in D^n} \frac{\|(\mathbf{x} - \mathbf{y})\mathbf{P}\|}{\|\mathbf{x} - \mathbf{y}\|} \text{ with } \mathbf{x} \neq \mathbf{y}. \tag{8}$$

We call $\tau_{\|.\|}(\mathbf{P})$ the coefficient of ergodicity induced by the norm $\|.\|$. It is proved that (8) can be written equivalently (Vassiliou [12], p. 118) as follows:

$$\tau_{\|.\|}(\mathbf{P}) = \sup_{\substack{\|z\|=1 \\ \mathbf{z}\mathbf{1}^\top = 1}} \|\mathbf{z}\mathbf{P}\| \text{ over all } \mathbf{z} \in \mathbb{R}^n. \tag{9}$$

When the L_1 norm is used, that is, for an $n \times n$ matrix $\mathbf{A}$ with elements from $\mathbb{C}$, we obtain that

$$\|\mathbf{A}\|_1 = \sum_{i,j=1}^{n} |a_{ij}|, \tag{10}$$

then

$$\tau_1(\mathbf{P}) = \sup_{\substack{\|z\|_1=1 \\ \mathbf{z}\mathbf{1}^\top = 1}} \|\mathbf{z}\mathbf{P}\|_1 \text{ over all } \mathbf{z} \in \mathbb{R}^n. \tag{11}$$

We now give the definitions of strong and weak ergodicity for a nonhomogeneous Markov chain:

Definition 4. *Consider an inhomogeneous Markov chain* $\{X_t\}_{t=0}^{\infty}$ *in discrete time and space. We say that* $\{X_t\}_{t=0}^{\infty}$ *with a sequence of transition probability matrices* $\{\mathbf{Q}(t)\}_{t=0}^{\infty}$ *is strongly ergodic if there exists a stable stochastic matrix* $\mathbf{Q}$ *such that for every* t

$$\lim_{\nu\to\infty} \|\mathbf{Q}(t, t+\nu) - \mathbf{Q}\| = 0. \tag{12}$$

If the limit is zero uniformly in t *we say that* $\{X_t\}_{t=0}^{\infty}$ *or equivalently* $\{\mathbf{Q}(t)\}_{t=0}^{\infty}$ *is uniformly strongly ergodic.*

Definition 5. *Consider an inhomogeneous Markov chain* $\{X_t\}_{t=0}^{\infty}$ *in discrete time and space. We say that* $\{X_t\}_{t=0}^{\infty}$ *with sequence of transition probability matrices* $\{\mathbf{Q}(t)\}_{t=0}^{\infty}$ *is weakly ergodic if for all states* i, j, r, t

$$q_{ir}(t, t+\nu) - q_{jr}(t, t+\nu) \to 0 \text{ as } \nu \to \infty. \tag{13}$$

Note that in (13), $\lim_{\nu\to\infty} q_{ij}(t, t+\nu)$ *is not actually necessary to exist.*

Remark 2. *Note that equivalently a non-homogeneous Markov chain is weakly ergodic if* $\tau_1(\mathbf{Q}(t, t+\nu)) < 1$ *for every* t*.*

We now state the following Lemma the proof of which exists in Vassiliou ([12], p. 119).

Lemma 1. *The coefficient of ergodicity generated by any metric as in (7) or induced by any vector norm on* $\mathbb{R}^n$ *has the following properties:*

(1) $\tau(\mathbf{P}_1\mathbf{P}_2) \leq \tau(\mathbf{P}_1)\tau(\mathbf{P}_2)$ *for any* $\mathbf{P}_1, \mathbf{P}_2$ *stochastic matrices.*
(2) *For any stochastic matrix* $\mathbf{P}$, $\tau(\mathbf{P}) = 0$ *if and only if* $rank(\mathbf{P}) = 1$.

The following Lemma proof, which can be found in Paz [33], is useful in what follows:

Lemma 2. *If* $\mathbf{P}$ *is a stochastic matrix and if* $\mathbf{R}$ *is any real matrix such that* $\mathbf{R}\mathbf{1}^\top = \mathbf{0}$ *and in addition* $\|\mathbf{R}\| < \infty$*, then*

$$\|\mathbf{R}\mathbf{P}\| \leq \|\mathbf{R}\|\tau_1(\mathbf{P}).$$

We now state and prove one of the basic theorems of the present paper.

Theorem 1. *Let there be the complete probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and consider an NHMS in discrete time and space. We assume that*

$$\lim_{t\to\infty} T(t) = T \text{ with } \Delta T(t) \geq 0. \tag{14}$$

Let $\{Y(t)\}_{t=0}^{\infty}$ *be the inherent nonhomogeneous Markov chain of the movement of memberships. If* $\{Y(t)\}_{t=0}^{\infty}$ *is weakly ergodic with* $\{\mathbf{Q}(t)\}_{t=0}^{\infty}$*, the sequence of transition probabilities, and in addition, there exists a stochastic vector*

$$\boldsymbol{\psi}(t) = [\psi_1(t), \psi_2(t), \ldots, \psi_k(t)], \tag{15}$$

which is the left eigenvector of $\mathbf{Q}(t)$ *for* $t = 0, 1, 2, \ldots$*, that is*

$$\boldsymbol{\psi}(t) = \boldsymbol{\psi}(t)\mathbf{Q}(t) \text{ for } t = 0, 1, 2, \ldots, \tag{16}$$

and

$$\sum_{t=0}^{\infty} \|\boldsymbol{\psi}(t+1) - \boldsymbol{\psi}(t)\| < \infty, \tag{17}$$

then the NHMS is strongly ergodic.

Proof. From (17) we obtain that there exists a stochastic vector $\boldsymbol{\psi}$ such that

$$\lim_{t\to\infty}\|\boldsymbol{\psi}(t)-\boldsymbol{\psi}\|=0. \tag{18}$$

From (16) and (18), we obtain that there exists a stochastic matrix $\mathbf{Q}$ such that

$$\lim_{t\to\infty}\|\mathbf{Q}(t)-\mathbf{Q}\|=0 \text{ with } \mathbf{Q}\mathbf{1}^\top=\mathbf{1}^\top. \tag{19}$$

Denote by $\boldsymbol{\Psi}(t)$ to be the stable matrix with row the stochastic vector $\boldsymbol{\psi}(t)$ and $\boldsymbol{\Psi}$ the stable matrix with row $\boldsymbol{\psi}$. Then, we have

$$\|\boldsymbol{\Psi}(t+1)-\boldsymbol{\Psi}(t)\|=\|\boldsymbol{\psi}(t+1)-\boldsymbol{\psi}(t)\| \text{ and } \|\boldsymbol{\Psi}(t)-\boldsymbol{\Psi}\|=\|\boldsymbol{\psi}(t)-\boldsymbol{\psi}\|. \tag{20}$$

We now show that

$$\lim_{\nu\to\infty}\|\mathbf{Q}(t,t+\nu)-\boldsymbol{\Psi}\|=0. \tag{21}$$

We have that

$$\begin{aligned}\mathbf{Q}(t,t+\nu)-\boldsymbol{\Psi}&=\mathbf{Q}(t,t+r)\mathbf{Q}(t+r,t+\nu)-\boldsymbol{\Psi}(t+r)\mathbf{Q}(t+r,t+\nu)+\\ &\boldsymbol{\Psi}(t+r)\mathbf{Q}(t+r,t+\nu)-\boldsymbol{\Psi}(t+\nu-1)+\boldsymbol{\Psi}(t+\nu-1)-\boldsymbol{\Psi}.\end{aligned} \tag{22}$$

Taking norms on (22) we obtain that

$$\begin{aligned}\|\mathbf{Q}(t,t+\nu)-\boldsymbol{\Psi}\|&\le\|\mathbf{Q}(t,t+r)\mathbf{Q}(t+r,t+\nu)-\boldsymbol{\Psi}(t+r)\mathbf{Q}(t+r,t+\nu)\|\\ &+\|\boldsymbol{\Psi}(t+r)\mathbf{Q}(t+r,t+\nu)-\boldsymbol{\Psi}(t+\nu-1)\|+\|\boldsymbol{\Psi}(t+\nu-1)-\boldsymbol{\Psi}\|.\end{aligned} \tag{23}$$

We now have

$$\begin{aligned}&N_1(t,\nu)=\|\mathbf{Q}(t,t+r)\mathbf{Q}(t+r,t+\nu)-\boldsymbol{\Psi}(t+r)\mathbf{Q}(t+r,t+\nu)\|\le\\ &\|[\mathbf{Q}(t,t+r)-\boldsymbol{\Psi}(t+r)]\mathbf{Q}(t+r,t+\nu)\|\le(\text{Due to Lemma 2})\le\\ &\|\mathbf{Q}(t,t+r)-\boldsymbol{\Psi}(t+r)\|\tau_1(\mathbf{Q}(t+r,t+\nu))\le 2\tau_1(\mathbf{Q}(t+r,t+\nu)),\end{aligned} \tag{24}$$

where $\tau_1(\mathbf{Q}(t+r,t+\nu))$ is less than one due to weak ergodicity of $\{Y(t)\}_{t=0}^{\infty}$ the inherent nonhomogeneous Markov chain. Also for fixed r, we can always choose ν such that

$$N_1(t,\nu)\le\frac{\varepsilon}{3}, \text{ with } \varepsilon>0 \text{ a small number.} \tag{25}$$

We now have that

$$\begin{aligned}&\boldsymbol{\Psi}(t+r)\mathbf{Q}(t+r,t+\nu)=(\text{due to Equation (16)})=\\ &[\boldsymbol{\Psi}(t+r)-\boldsymbol{\Psi}(t+r+1)]\mathbf{Q}(t+r+1,t+\nu)+\\ &+\boldsymbol{\Psi}(t+r+1)\mathbf{Q}(t+r+1,t+\nu).\end{aligned} \tag{26}$$

Now, similarly, we obtain

$$\begin{aligned}&\boldsymbol{\Psi}(t+r+1)\mathbf{Q}(t+r+1,t+\nu)=\\ &[\boldsymbol{\Psi}(t+r+1)-\boldsymbol{\Psi}(t+r+2)]\mathbf{Q}(t+r+2,t+\nu)\\ &+\boldsymbol{\Psi}(t+r+2)\mathbf{Q}(t+r+2,t+\nu),\end{aligned} \tag{27}$$

and using this equation recursively, we obtain

$$\mathbf{\Psi}(t+r)\mathbf{Q}(t+r,t+\nu) = \sum_{j=t+r+1}^{t+\nu-1} [\mathbf{\Psi}(j-1) - \mathbf{\Psi}(j)]\mathbf{Q}(j,t+\nu)$$

$$+\mathbf{\Psi}(t+\nu-1)\mathbf{Q}(t+\nu-1,t+\nu). \tag{28}$$

From condition (16), we obtain that

$$\mathbf{\Psi}(t+\nu-1) = \mathbf{\Psi}(t+\nu-1)\mathbf{Q}(t+\nu-1), \tag{29}$$

and

$$\boldsymbol{\psi}(t+\nu-1) = \boldsymbol{\psi}\mathbf{\Psi}(t+\nu-1). \tag{30}$$

Therefore, from (28)–(30), as well as, Lemma 2, we have that

$$N_2(t,\nu) = \|\mathbf{\Psi}(t+r)\mathbf{Q}(t+r,t+\nu) - \mathbf{\Psi}\mathbf{\Psi}(t+\nu-1)\| \le$$

$$\left\| \sum_{j=t+r+1}^{t+\nu-1} [\mathbf{\Psi}(j-1) - \mathbf{\Psi}(j)]\tau_1(\mathbf{Q}(j,t+\nu)) \right\| \le$$

$$\sum_{j=t+r+1}^{t+\nu-1} \|\boldsymbol{\psi}(j-1) - \boldsymbol{\psi}(j)\|\tau_1(\mathbf{Q}(j,t+\nu)) \le$$

$\le$ $(\tau_1(\mathbf{Q}(t+r,t+\nu)) < 1$ due to weak ergodicity of $\{Y(t)\}_{t=0}^{\infty}$; see also Remark 2).

$$\le \sum_{j=t+r+1}^{t+\nu-1} \|\boldsymbol{\psi}(j-1) - \boldsymbol{\psi}(j)\| < \infty, \tag{31}$$

due to condition (17). Since $N_2(t,\nu) < \infty$ for every r and ν and it is a sum of positive numbers, we have that its tail goes to zero. Hence, we could, for every $\varepsilon > 0$, fix $t+r$ such that $N_2(t,\nu) < \varepsilon/3$ for $\nu - 1 \ge r$ and we can always take a ν large enough so that $N_1(t,\nu) < \varepsilon/3$ and

$$N_3(t,\nu) = \|\mathbf{\Psi}(t+\nu-1) - \mathbf{\Psi}\| \le \frac{\varepsilon}{3}, \tag{32}$$

therefore, we obtain that

$$\lim_{\nu\to\infty} \|\mathbf{Q}(t,t+\nu) - \mathbf{\Psi}\| = 0. \tag{33}$$

Now, from (4), we have that

$$\|\mathbb{E}[\mathbf{q}(t,t+\nu)] - \boldsymbol{\psi}\| =$$

$$\left\| \mathbf{q}(t)\frac{T(t)}{T(t+\nu)}\mathbf{Q}(t,t+\nu) + \frac{1}{T(t+\nu)}\sum_{\tau=t}^{t+\nu} \Delta T(\tau)\mathbf{Q}(\tau,t+\nu) - \boldsymbol{\psi} \right\| \le$$

$$\frac{1}{T(t+\nu)}\left\| \mathbf{q}(t)T(t)\mathbf{Q}(t,t+\nu) + \sum_{\tau=t}^{t+\nu} \Delta T(\tau)\mathbf{Q}(\tau,t+\nu) - T(t+\nu)\boldsymbol{\psi} \right\| \le$$

$$\|\mathbf{q}(t)T(t)\mathbf{Q}(t,t+\nu) - T(t)\boldsymbol{\psi}\| +$$

$$\left\| \sum_{\tau=t}^{t+\nu} \Delta T(\tau)\mathbf{p}_0(\tau)\mathbf{Q}(\tau,t+\nu) - [T(t+\nu) - T(t)]\boldsymbol{\psi} \right\|. \tag{34}$$

Now, we have

$$\mathcal{A}(t,t+\nu) = \|\mathbf{q}(t)T(t)\mathbf{Q}(t,t+\nu) - T(t)\boldsymbol{\psi}\| \le$$

$$T(t)\|\mathbf{q}(t)\|\|\mathbf{Q}(t,t+\nu)-\boldsymbol{\psi}\|. \tag{35}$$

By fixing $\varepsilon > 0$, we can always find ν_0 such that

$$\mathcal{A}(t,t+\nu) \leq \frac{\varepsilon}{2} \text{ for } \nu \geq \nu_0. \tag{36}$$

$$\mathcal{B}(t,t+\nu) = \left\| \sum_{\tau=t}^{t+\nu} \Delta T(\tau)\mathbf{p}_0(\tau)\mathbf{Q}(\tau,t+\nu) - [T(t+\nu)-T(t)]\boldsymbol{\psi} \right\| =$$

$$= (\text{since } \mathbf{p}_0(\tau)\boldsymbol{\Psi} = \mathbf{p}_0\boldsymbol{\Psi} = \boldsymbol{\psi}) =$$

$$\left\| \sum_{\tau=t}^{t+\nu} \Delta T(\tau)\mathbf{p}_0(\tau)\mathbf{Q}(\tau,t+\nu) - \sum_{\tau=t}^{t+\nu} \Delta T(\tau)\mathbf{p}_0\boldsymbol{\Psi} \right\| =$$

$$\left\| \sum_{\tau=t}^{t+\nu} \Delta T(\tau)[\mathbf{p}_0(\tau)\mathbf{Q}(\tau,t+\nu) - \mathbf{p}_0\boldsymbol{\Psi}] \right\| \leq$$

$$\sum_{\tau=t}^{t+\nu} \Delta T(\tau)\|\mathbf{p}_0(\tau)\mathbf{Q}(\tau,t+\nu) - \mathbf{p}_0\boldsymbol{\Psi}\| \leq$$

$$\sum_{\tau=t}^{t+\nu} \Delta T(\tau)\|\mathbf{p}_0(\tau)[\mathbf{Q}(\tau,t+\nu) - \boldsymbol{\Psi}] + \mathbf{p}_0(\tau)\boldsymbol{\Psi} - \mathbf{p}_0\boldsymbol{\Psi}\|$$

$$\leq \sum_{\tau=t}^{t+\nu} \Delta T(\tau)\|\mathbf{Q}(\tau,t+\nu) - \boldsymbol{\Psi}\| \tag{37}$$

Now, for $\nu > \nu_0$ from (37), we have that

$$\mathcal{B}(t,t+\nu) \leq \sum_{\tau=t}^{t+\nu-\nu_0} \Delta T(\tau)\|\mathbf{Q}(\tau,t+\nu) - \boldsymbol{\Psi}\| + \sum_{\tau=t+\nu-\nu_0}^{t+\nu} \Delta T(\tau)\|\mathbf{Q}(\tau,t+\nu) - \boldsymbol{\Psi}\|$$

$$\leq \sum_{\tau=t}^{t+\nu-\nu_0} \Delta T(\tau)\frac{\varepsilon}{2} + 2|T(t+\nu) - T(t+\nu-\nu_0+1)|, \tag{38}$$

where the second part of (38) for $\nu >> \nu_0$ is less than $\varepsilon/2$ due to condition (14). The first part is also for, $\nu >> \nu_0$, less than $\varepsilon/2$ since $\Delta T(t) \to_{t\to\infty} 0$ due to condition (14). From (34) and (36), and a $\nu > \nu_0$ large enough ($\nu >> \nu_0$), we easily see that

$$\|\mathbb{E}[\mathbf{q}(t,t+\nu)] - \boldsymbol{\psi}\| \leq \varepsilon \text{ for every } \nu >> \nu_0,$$

hence, the NHMS is strongly ergodic. □

4. Rate of Convergence in NHMS with Chronological Order

An important question in nonhomogeneous Markov chains and NHMS, due to its very large practical value, is the rate of convergence to their asymptotic structure. In fact, it is important to find the necessary conditions under which the rate of convergence is geometric because then the value of the asymptotic result is greater. The roots of the study of finding conditions under which the rate of convergence is geometric for nonhomogeneous Markov chains are in Huang, et al. [34,35] and Seneta [31]. The study of the geometric rate of convergence in NHMS started in Vassiliou and Tsaklidis [36] and Georgiou and Vassiliou [26], and an updated evolution of these theorems and their variants could be found in Vassiliou [12]. The importance of answering this problem for practical purposes is apparent in Bartholomew [37] for the homogeneous Markov system, which is a very special case of an NHMS. In the present section, we provide and prove a basic theorem for the rate of convergence to strong ergodicity in NHMS when the transition probabilities matrices of the inherent Markov chain $\{\mathbf{Q}(t)\}_{t=0}^{\infty}$ are given in chronological order. In

Theorem 4, we relax the basic assumption present in all studies of the rate of convergence to its asymptotic behavior for an NHMS. That is, we will not assume that the inherent inhomogeneous Markov chain converges as time goes to infinity to a homogeneous Markov chain with a regular transition probability matrix geometrically fast. We start with some basic definitions and results which will be useful in what follows.

Definition 6. *We say that a sequence of matrices* $\{\mathbf{A}_n\}_{n=0}^{\infty}$ *converges with geometrical rate to a matrix* $\mathbf{A}$ *if there exists constants* $c > 0$ *and* $0 < b < 1$ *such that*

$$\|\mathbf{A}_n - \mathbf{A}\| \leq cb^n \text{ for } n = 1, 2, \ldots. \tag{39}$$

Lemma 3 (Vassiliou and Tsaklidis [36])**.** *The following statements are equivalent:*

(i) *The sequence* $\{\Delta T(t)\}_{t=0}^{\infty}$ *converges to zero geometrically fast.*
(ii) *The sequence* $\{T(t)\}_{t=0}^{\infty}$ *converges to* T *geometrically fast.*

Definition 7. (i) *A stochastic matrix* $\mathbf{P}$ *is called Markov if at least one column of* $\mathbf{P}$ *is entirely positive. Let* $\mathcal{M}$ *be the set of all Markov matrices.* (ii) *We say that the stochastic matrix* $\mathbf{P} \in \mathcal{G}_2$ *if* (a) $\mathbf{P} \in \mathcal{G}_1$ *the set of all regular matrices;* (b) $\mathbf{QP} \in \mathcal{G}_1$ *for any* $\mathbf{Q} \in \mathcal{G}_1$. (iii) *We say that the stochastic matrix* $\mathbf{P} \in \mathcal{G}_3$ *the set of all scrambling matrices if* $\tau_{\|\cdot\|_1}(\mathbf{P}) < 1$.

Remark 3. *The distinction of the set* $\mathcal{G}_2$ *from all stochastic regular matrices is due to the fact that the product of two regular matrices is not always regular. In addition, the product of two nonregular stochastic matrices could be regular. A practical way to check if a stochastic matrix of small dimension is scrambling is the following: given any two rows* i, j*, there is at least one column* k *such that* $p_{ik} > 0$ *and* $p_{jk} > 0$.

Definition 8. *The incidence matrix of a stochastic matrix* $\mathbf{P}$ *is a matrix where in the positions of positive elements we put the number 1. Therefore, two stochastic matrices* $\mathbf{P}$ *and* $\mathbf{Q}$ *of the same dimension have the same incidence matrix if they have the positive elements in the same positions. Then, we write* $\mathbf{P} \sim \mathbf{Q}$.

We now state some known Lemmas and Theorems, the proofs of which can be found in Vassiliou ([12], p. 143).

Theorem 2. *For all stochastic matrices, we have* $\mathcal{M} \subset \mathcal{G}_3 \subset \mathcal{G}_2 \subset \mathcal{G}_1$.

Lemma 4. *If* $\mathbf{P}(t, t+\nu) \in \mathcal{G}_1$ *with* $t \geq 0, n \geq 1$*, then* $\mathbf{P}(t, t+\nu) \in \mathcal{M}$ *for* $t + \nu \geq \mu$ *the number of distinct incidence matrices corresponding to* $\mathcal{G}_1$ *with the same dimension as* $\mathbf{P}(t, t+\nu)$.

Theorem 3. *Let there be a complete probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, and consider a nonhomogeneous Markov chain* $\{X_t\}_{t=0}^{\infty}$ *in discrete time and space with a sequence of transition probabilities matrices* $\{\mathbf{P}(t)\}_{t=0}^{\infty}$. *If* $\mathbf{P}(t) \in \mathcal{G}_2$ *for every* $t = 1, 2, \ldots$ *and*

$$\min_{i,j}\left(p_{ij}(t), 0\right)^{+} \geq \gamma > 0, \tag{40}$$

uniformly for all $t \geq 1$*, then weak ergodicity obtains at a uniform geometric rate.*

We now define the geometrically strongly ergodic NHMS, which is a central concept in the present section.

Definition 9. *Let there be a complete probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, and consider an NHMS in discrete time and space. We say that the NHMS is strongly ergodic if there exists a stochastic vector* $\boldsymbol{\psi}$ *and constants* $c > 0$ *and* $0 < b < 1$ *such that*

$$\|\mathbb{E}[\mathbf{q}(t, t+\nu) - \boldsymbol{\psi}]\| \leq cb^{\nu} \text{ for } c > 0 \text{ and } 0 < b < 1.$$

We now state and prove the basic theorem of this section.

Theorem 4. *Let there be a complete probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, and consider an NHMS in discrete time and space. We assume that the total number of memberships is increasing* ($\Delta T(t) \geq 0$)*, and it converges geometrically fast to* T*. That is,*

$$\lim_{t\to\infty} T(t) = T \text{ geometrically fast with } \Delta T(t) \geq 0. \tag{41}$$

Let $\{Y(t)\}_{t=0}^{\infty}$ *be the inherent nonhomogeneous Markov chain of the movement of memberships. If* $\{Y(t)\}_{t=0}^{\infty}$ *is weakly ergodic with* $\{\mathbf{Q}(t)\}_{t=0}^{\infty}$ *the sequence of transition probabilities to be such that*

$$(i)\ \mathbf{Q}(t) \in \mathcal{G}_2 \text{ for every } t = 1, 2, \ldots, \tag{42}$$

$$(ii)\ \min_{i,j}\left(q_{ij}(t), 0\right)^{+} \geq \gamma > 0, \tag{43}$$

and if in addition there exists a stochastic vector

$$\boldsymbol{\psi}(t) = [\psi_1(t), \psi_2(t), \ldots, \psi_k(t)], \tag{44}$$

which is the left eigenvector for $t = 1, 2, \ldots$ *of* $\mathbf{Q}(t)$*, that is,*

$$\boldsymbol{\psi}(t) = \boldsymbol{\psi}(t)\mathbf{Q}(t) \text{ for every } t, \tag{45}$$

and in addition

$$\lim_{\nu\to\infty} \|\boldsymbol{\psi}(t+\nu) - \boldsymbol{\psi}(t+\nu-1)\| = 0 \text{ geometrically fast}, \tag{46}$$

then the NHMS is geometrically strongly ergodic.

Proof. In order to prove that the NHMS is geometrically strongly ergodic, we must show that the expected relative structure satisfies Definition 9.

From (46), we have that there exists a vector $\boldsymbol{\psi}$ and constants $c_1 > 0$ and $0 < b_1 < 1$ such that

$$\|\boldsymbol{\psi}(t+\nu) - \boldsymbol{\psi}\| \leq c_1 b_1^{t+\nu}. \tag{47}$$

We define $\boldsymbol{\Psi}(t)$ and $\boldsymbol{\Psi}$ as the stable matrices with rows $\boldsymbol{\psi}(t)$ and $\boldsymbol{\psi}$, respectively.

We now show that the inherent nonhomogeneous Markov chain $\{Y(t)\}_{t=0}^{\infty}$ is geometrically strongly ergodic. That is, we need to show that there exists constants $c_2 > 0$ and $0 < b_2 < 1$ such that

$$\|\mathbf{Q}(t, t+\nu) - \boldsymbol{\Psi}\| \leq c_2 b_2^{\nu} \text{ for every } t. \tag{48}$$

Let us denote by μ the number of distinct incidence matrices corresponding to $\mathcal{G}_1$ with the same dimension as $\mathbf{P}(t, t+\nu)$. Then, for $\nu > \mu$, we have that

$$\begin{aligned} \|\mathbf{Q}(t, t+\nu) - \boldsymbol{\Psi}\| &\leq \|\mathbf{Q}(t, t+\mu)\mathbf{Q}(t+\mu, t+\nu) - \boldsymbol{\Psi}(t+\mu)\mathbf{Q}(t+\mu, t+\nu)\| \\ &+ \|\boldsymbol{\Psi}(t+\mu)\mathbf{Q}(t+\mu, t+\nu) - \boldsymbol{\Psi}(t+\nu-1)\| + \|\boldsymbol{\Psi}(t+\nu-1) - \boldsymbol{\Psi}\|. \end{aligned} \tag{49}$$

Now, we have that

$$\begin{aligned} \mathcal{D}_1(t, \nu, \mu) &= \|\mathbf{Q}(t, t+\mu)\mathbf{Q}(t+\mu, t+\nu) - \boldsymbol{\Psi}(t+\mu)\mathbf{Q}(t+\mu, t+\nu)\| \\ &\leq \|[\mathbf{Q}(t, t+\mu) - \boldsymbol{\Psi}(t+\mu)]\mathbf{Q}(t+\mu, t+\nu)\| \\ &\leq \|\mathbf{Q}(t, t+\mu) - \boldsymbol{\Psi}(t+\mu)\|\tau_1(\mathbf{Q}(t+\mu, t+\nu)) \\ &\leq 2\tau_1(\mathbf{Q}(t+\mu, t+\nu)). \end{aligned} \tag{50}$$

For arbitrary but fixed t and ν with $m\mu \leq \nu$ with m, the largest such integer from (50), we have that

$$\mathcal{D}_1(t,\nu,\mu) \leq 2\tau_1(\mathbf{Q}(t+\mu,t+2\mu))\tau_1(\mathbf{Q}(t+\mu,t+2\mu))\ldots$$
$$\tau_1(\mathbf{Q}(t+(m-1)\mu,t+m\mu))\tau_1(\mathbf{Q}(t+m\mu,t+\nu)). \tag{51}$$

Since $\mathbf{Q}(t) \in \mathcal{G}_2$ and by Lemma 4 we have that $\tau_1(\mathbf{Q}(t+(i-1)\mu,t+i\mu)) \in \mathcal{M}$ and

$$\tau_1(\mathbf{Q}(t+(i-1)\mu,t+i\mu)) \leq 1-\gamma^{\mu} \text{ with } 0<\gamma<1 \text{ for } i=2,3,\ldots,m. \tag{52}$$

From the weak ergodicity of the inherent nonhomogeneous Markov chain $\{Y(t)\}_{t=0}^{\infty}$, we have that

$$\tau_1(\mathbf{Q}(t+m\mu,t+\nu)) \leq 1. \tag{53}$$

From (51)–(53), we arrive at

$$\mathcal{D}_1(t,\nu,\mu) \leq (1-\gamma^{\mu})^m, \tag{54}$$

which as $\nu \to \infty$ goes to zero at a uniform geometric rate.

Following the steps of arriving at relation (31), we straightforwardly obtain that

$$\begin{aligned}\mathcal{D}_2(t,\nu,\mu) &= \|\mathbf{\Psi}(t+\mu)\mathbf{Q}(t+\mu,t+\nu)-\mathbf{\Psi}(t+\nu-1)\| \\ &\leq \|\mathbf{\Psi}(t+\mu)\mathbf{Q}(t+\mu,t+\nu)-\mathbf{\Psi}\mathbf{\Psi}(t+\nu-1)\| \\ &\leq \sum_{j=t+\mu+1}^{t+\nu-1}\|\boldsymbol{\psi}(j-1)-\boldsymbol{\psi}(j)\|\tau_1(\mathbf{Q}(j,t+\nu)).\end{aligned} \tag{55}$$

We now have the largest integer for m, such that $m\mu \leq \nu$

$$\begin{aligned}\tau_1(\mathbf{Q}(t+\mu,t+\nu)) &= \tau_1(\mathbf{Q}(t+\mu,t+2\mu)\mathbf{Q}(t+2\mu,t+3\mu)\ldots \\ &\quad \mathbf{Q}(t+(m-1)\mu,t+m\mu)\mathbf{Q}(t+m\mu,t+\nu\mu)) \\ &= (\text{by Lemma 1}) \\ &\leq \tau_1(\mathbf{Q}(t+\mu,t+2\mu))\tau_1(\mathbf{Q}(t+2\mu,t+3\mu))\ldots \\ &\quad \tau_1(t+(m-1)\mu,t+m\mu)\tau_1(t+m\mu,t+\nu) \leq \\ &\leq (\text{using (52)}) \\ &\leq (1-\gamma^{\mu})(1-\gamma^{\mu})\ldots(1-\gamma^{\mu}) = (1-\gamma^{\mu})^m.\end{aligned} \tag{56}$$

From (56), we have that

$$\tau_1(\mathbf{Q}(t+\mu,t+\nu)) \to_{\nu\to\infty} 0 \text{ uniformly geometrically fast.} \tag{57}$$

From (55), we have that

$$\begin{aligned}\mathcal{D}_2(t,\nu,\mu) &\leq \|\boldsymbol{\psi}(t+\mu-1)-\boldsymbol{\psi}(t+\mu)\|(1-\gamma^{\mu})^m + \\ &\|\boldsymbol{\psi}(t+\mu)-\boldsymbol{\psi}(t+\mu+1)\|(1-\gamma^{\mu})^{m-1} + \|\boldsymbol{\psi}(t+\mu)-\boldsymbol{\psi}(t+\mu+1)\|(1-\gamma^{\mu})^{m-1} \\ &+\ldots+\|\boldsymbol{\psi}(t+\mu(m-1)-1)-\boldsymbol{\psi}(t+\mu(m-1))\|(1-\gamma^{\mu}) \\ &+\|\boldsymbol{\psi}(t+\mu m-1)-\boldsymbol{\psi}(t+\mu m)\|.\end{aligned} \tag{58}$$

Now, we have that

$$(1-\gamma^{\mu})^m \to_{m\to\infty} 0 \text{ geometrically fast },$$

since $(1-\gamma^{\mu}) < 1$. Also, from condition (46), we have that

$$\|\boldsymbol{\psi}(t+\mu m-1)-\boldsymbol{\psi}(t+\mu m)\| \to_{m\to\infty} 0 \text{ geometrically fast,}$$

Therefore,

$$\mathcal{D}_2(t, \nu, \mu) \rightarrow_{\nu \to \infty} 0 \text{ geometrically fast.} \tag{59}$$

Now, from (49), (54) and (59), and condition (46), we easily have that

$$\|\mathbf{Q}(t, t+\nu) - \mathbf{\Psi}\| \rightarrow_{\nu \to \infty} 0 \;\text{ geometrically fast;} \tag{60}$$

hence, there exist constants $c > 0$ and $0 < b < 1$ such that

$$\|\mathbf{Q}(t, t+\nu) - \mathbf{\Psi}\| \leq cb^{\nu} \text{ for every } t, \text{ and } c > 0, 0 < b < 1. \tag{61}$$

What remains, according to Definition 9, is to show that the expected relative population structure $\mathbb{E}[\mathbf{q}(t, t+\nu)]$ converges for every t to the vector $\boldsymbol{\psi}$ as ν goes to infinity geometrically fast.

From (35), we have that

$$\begin{aligned} \mathcal{A}(t, t+\nu) = \|\mathbf{q}(t)T(t)\mathbf{Q}(t, t+\nu) - T(t)\boldsymbol{\psi}\| \leq \\ T(t)\|\mathbf{q}(t)\|\|\mathbf{Q}(t, t+\nu) - \boldsymbol{\psi}\|. \end{aligned} \tag{62}$$

From (61), we have that

$$\mathcal{A}(t, t+\nu) \leq T(t)\|\mathbf{q}(t)\|\|\mathbf{Q}(t, t+\nu) - \boldsymbol{\psi}\| \leq c_1 b^{\nu}, \tag{63}$$

with $c_1 > 0$ and $0 < b < 1$.

From Lemma 3, we have that since $T(t)$ converges geometrically fast to T, then $\Delta T(t)$ converges geometrically fast to zero. Hence,

$$\Delta T(t) \leq c_2 b_2^t \;\text{ with } c_2 > 0 \text{ and } 0 < b_2 < 1. \tag{64}$$

Now, from (37), we have that

$$\mathcal{B}(t, t+\nu) \leq \sum_{\tau=t}^{t+\nu} \Delta T(\tau)\|\mathbf{Q}(\tau, t+\nu) - \mathbf{\Psi}\|. \tag{65}$$

With no loss of generality, we may assume that $b > b_2$, and then, from (63) and (64), we have that

$$\begin{aligned} \mathcal{B}(t, t+\nu) &\leq c_1 c_2 \sum_{\tau=t}^{t+\nu} \left(\frac{b_2}{b}\right)^{\tau} b^{t+\nu} = c_3 b^{t+\nu} \sum_{\tau=t}^{t+\nu} \left(\frac{b_2}{b}\right)^{\tau} \\ &\leq c_3 b^{t+\nu} \left(1 - \frac{b_2}{b}\right)^{\nu} \left(1 - \frac{b_2}{b}\right)^{-1} \leq c_4 b_3^{\nu}. \end{aligned} \tag{66}$$

From (63) and (66), we easily arrive at the proof of the theorem. □

5. An Illustrative Application

We will illustrate the results in the previous section with an example from a population of manpower. To possibly better visualize, the reader may have in mind a University system with three grades. That is, grade one is those with the level of Professors, grade two belongs to the Associate Professors, and finally, in grade three, there are the Assistant Professors. The University has a plan for funding $T(t)$ for $t = 0, 1, 2, \ldots$ memberships for the next few years. When a member of staff is leaving, their membership remains with the University, that is, the funding of their position is not lost but remains, and the University could go on to appoint someone at any grade. The external environment to which the leavers go and their membership is retained by the University and it is the population of members of academic staff from Universities almost all over the world, as practices have shown. Hence, it is from this external environment that the new members will obtain the memberships available from the organization of the University.

A fundamental question for the practitioner is the estimation of the transition probabilities from the historical records. The way to estimate the transition probabilities of the memberships is a small extension of the way it is performed in manpower planning, which is well documented in Bartholomew [38]. Another quite similar problem is that of the competing risk model in the medical literature, as it was presented by Kalbfleisch and Prentice ([39], Chapter 8), Lee [40], and Cox and Oakes [41]. Also, similar problems exist in the study of reliability models and various actuarial studies as discussed in Elandt–Johnson and Johnson ([42], Chapter 7).

Now define by

$N_{ij}(t)$: the number of memberships moving from grade i to grade j in the t-th interval.

$N_i(t-1)$: the number of memberships in state i at the beginning of the t-th interval.

Assume that the number of years available in the historical data of the University is n. Then, the maximum likelihood estimate of the probability $q_{ij}(t)$ is the following:

$$\hat{q}_{ij}(t) = \frac{N_{ij}(t)}{N_i(t-1)} \text{ for any } t. \tag{67}$$

It is an apparent advantage that the probabilities $\hat{q}_{ij}(t)$ are separately estimated for every (i,j). In this way, the number of years of historical records necessary are significantly reduced. At this point, it is useful to test the hypothesis that the probabilities $\hat{q}_{ij}(t)$ are indeed functions of time. That is,

$$H_0\colon \hat{q}_{ij}(t) = \hat{q}_{ij} \text{ for every } t. \tag{68}$$

Considering the flow of memberships which move from grade i to grade j as a multinomial random variable, then (see Andersen and Goodman [43]) hypothesis (68) is tested by the statistic:

$$\chi^2(i,j) = \sum_{t=1}^{n} N_i(t-1)\frac{\left(\hat{q}_{ij}(t) - \hat{q}_{ij}\right)}{\hat{q}_{ij}}, \tag{69}$$

where

$$\hat{q}_{ij} = \frac{\sum_{t=0}^{n} N_{ij}(t)}{\sum_{t=1}^{n} N_i(t-1)}, \tag{70}$$

is the maximum likelihood estimate under the null hypothesis and is chi-square distributed with $n-1$ degrees of freedom.

Now, let that the $\chi^2(i,j)$ showed that the probabilities $\hat{q}_{ij}(t)$ are functions of time. Then, there is a need to predict their values as functions of time. For a specific pair (i,j), let that

$$x_{1ij}(t), x_{2ij}(t), \ldots, x_{mij}(t), \tag{71}$$

are probable covariates for the specific application. Then, logistic stepwise regression is an appropriate model for these probabilities. Let us define by

$$\log it(x) = \log\left(\frac{x}{1-x}\right), \tag{72}$$

then we obtain

$$\log it\left(\hat{q}_{ij}(t)\right) = a_0 + a_1 x_{1ij}(t) + a_2 x_{2ij}(t) + \ldots + a_m x_{mij}(t). \tag{73}$$

Now, it is obvious that stepwise regression will show what are the important covariates to predict $\hat{q}_{ij}(t)$ (see also Vassiliou [44]).

Let that following the above-described steps, we ended with the matrix:

$$\mathbf{Q}(t) = \begin{pmatrix} 0.2 + \frac{1}{4+t^3} & 0.8 - \frac{1}{4+t^3} & 0 \\ 0.3 & 0.5 - \frac{1}{8+t^4} & 0.2 + \frac{1}{8+t^4} \\ 0 & 0.2 + \frac{2}{10+t^8} & 0.8 - \frac{2}{10+t^8} \end{pmatrix} \text{ for t = 1,2,}\dots \tag{74}$$

The total population of memberships was planned according to the following sequence

$$T(0) = 400,\ T(1) = 430,\ T(2) = 450,$$

$$T(3) = 475,\ T(4) = 500 \text{ and } T(t) = 500 \text{ for } t = 5, 6, \ldots, \tag{75}$$

that is, the total number of memberships converge geometrically fast with $\Delta T(t) \geq 0$, satisfying the condition (41) in Theorem 4.

It is not difficult to check that the sequence (74) satisfies condition (42), that is, $\mathbf{Q}(t) \in \mathcal{G}_2$ for $t = 1, 2, \ldots$. Also, it satisfies condition (43) since

$$\min_{i,j}\left(q_{ij}(t), 0\right)^+ \geq \gamma = 0.2 > 0. \tag{76}$$

For the sequence of transition probability matrices for the memberships (74), we find for the condition (45) that the vectors $\boldsymbol{\psi}(t)$ that satisfy it were the following:

$$\text{For } t = 1 \ \ \boldsymbol{\psi}(1) = [0.216, 0.432, 0.352], \tag{77}$$

$$\text{for } t = 2 \ \ \boldsymbol{\psi}(2) = [0.166, 0.390, 0.444], \tag{78}$$

$$\text{for } t = 3 \ \ \boldsymbol{\psi}(3) = [0.164, 0.416, 0.420], \tag{79}$$

$$\text{for } t = 4 \ \ \boldsymbol{\psi}(4) = [0.164, 0.424, 0.412], \tag{80}$$

$$\text{for } t = 5 \ \ \boldsymbol{\psi}(5) = [0.163, 0.426, 0.411], \tag{81}$$

$$\text{for } \ t = 6 \ \ \boldsymbol{\psi}(6) = [0.163, 0.427, 0.410], \tag{82}$$

$$\text{for } \ t = 7 \ \ \boldsymbol{\psi}(7) = [0.163, 0.427, 0.410]. \tag{83}$$

We observe that already for t = 6 in (82) and (83) we have convergence of $\boldsymbol{\psi}(t)$ which satisfies condition (45) of Theorem 4, that is the convergence is geometrically fast. Hence, we conclude that

$$\boldsymbol{\psi} = [0.16, 0.43, 0.41].$$

Calculating using the transition probability matrices given in (74) the matrix product $\mathbf{Q}(t, t+\nu)$ we find that for $\nu \geq 9$ it converges to $\boldsymbol{\psi} = [0.16, 0.43, 0.41]$. Hence, we conclude that $\mathbf{Q}(t, t+\nu)$ as $\nu \to \infty$ converges geometrically fast as was expected from Theorem 4 relation (48).

Now, given the convergence of $\mathbf{Q}(t, t+\nu)$ as $\nu \to \infty$ it is straight forward to find from Equation (2) applying it recursively or equivalently from Theorem 1 that $\mathbb{E}[\mathbf{q}(s,t)]$ converges geometrically fast (in fact in 9 time steps) to

$$\boldsymbol{\psi} = [0.16, 0.43, 0.41].$$

Hence, the NHMS is geometrically strongly ergodic.

6. Conclusions and Further Research

Two fundamental theorems have been founded that relax previous assumptions and provide conditions for ergodicity and for the convergence rate of the relative population structure in an NHMS. More specifically, in Theorem 1 the strong ergodicity of an NHMS is studied without assuming the convergence of the inherent inhomogeneous Markov chain to infinity to a homogeneous Markov chain with a regular transition probability matrix. In Theorem 4, it is proved under what conditions the rate of convergence of strong ergodicity in a NHMS is geometrically fast. This is done by departing from the basic assumption in which the inherent inhomogeneous Markov chain converges after a large time to a homogeneous Markov chain with a regular transition probability matrix. The proved theorems are expanding the understanding of the dynamics and behavior of an NHMS. The paper concludes with an illustrative application from the field of Manpower Planning, showcasing the vital practical relevance of the discussed concepts. Further research paths may include the relaxation of relative assumptions in the many variant models of the NHMS in diverse populations. Of particular interest may be the theorems of Laws of Large numbers in an NHMS.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The author declare no conflict of interest.

References

1. Vassiliou, P.-C.G. Asymptotic behavior of Markov systems. *J. Appl. Prob.* **1982**, *19*, 433–438. [CrossRef]
2. Young, A.; Almond, G. Predicting distributions of staff. *Comput. J.* **1961**, *3*, 144–153. [CrossRef]
3. Young, A. Models for planning recruitment and promotion of staff. *Brit. J. Indust. Rel.* **1965**, *3*, 301–310. [CrossRef]
4. Young, A. Demographic and ecological models for manpower planning. In *Aspects of Manpower Planning*; Bartholomew, D.J., Morris, B.R., Eds.; English University Press: London, UK, 1971.
5. Bartholomew, D.J. A multistage renewal processes. *J. R. Stat. Soc. B* **1963**, *25*, 150–168.
6. Bartholomew, D.J. *Stochastic Models for Social Processes*, 1st ed.; Wiley: New York, NY, USA, 1967.
7. Young, A.; Vassiliou, P.-C.G. A non-linear model on the promotion of staff. *J. R. Stat. Soc. A* **1974**, *138*, 584–595. [CrossRef]
8. Vassiliou, P.-C.G. A Markov model for wastage in manpower systems. *Oper. Res. Quart.* **1976**, *27*, 57–70. [CrossRef]
9. Vassiliou, P.-C.G. A high order non-linear Markovian model for promotion in manpower systems. *J. R. Stat. Soc. A* **1978**, *141*, 86–94. [CrossRef]
10. McClean, S.I. A continuous-time population model with Poisson recruitment. *J. Appl. Prob.* **1976**, *13*, 348–354. [CrossRef]
11. McClean, S.I. Continuous-time stochastic models for multigrade population. *J. Appl. Prob.* **1978**, *15*, 26–37. [CrossRef]
12. Vassiliou, P.-C.G. *Non-Homogeneous Markov Chains and Systems, Theory and Applications*; Chapman and Hall: London, UK; CRC Press: Boca Raton, FL, USA, 2023.
13. Garg, L.; McClean, S.I.; Meenan, B.; Millard, P. A non-homogeneous discrete time Markov model for admission scheduling and resource planning in a cost capacity constaint healthcare system. *Health Care Manag. Sci.* **2010**, *13*, 155–169. [CrossRef]
14. Garg, L.; McClean, S.I.; Meenan, B.; Millard, P. Non-homogeneous Markov models for sequential pattern mining of healthcare data. *Ima J. Manag. Math.* **2009**, *20*, 327–344. [CrossRef]
15. Garg, L.; McClean, S.I.; Meenan, B.; Millard, P. Phase-Type survival trees and mixed distribution syrvival trees for clustering patient's hospital length of stay. *Informatika* **2011**, *22*, 57–72.
16. Ugwuogo, F.I.; McClean, S.I. Modelling heterogeneity in manpower systems: A review. *Appl. Stoch. Models Bus. Ind.* **2000**, *2*, 99–110. [CrossRef]
17. Vassiliadis, G. Transient analysis of the M/M/k/N/N queue using a continuous time homogeneous Markov chain system with finite state capacity. *Commun. Stat. Theory Methods* **2014**, *43*, 1548–1562. [CrossRef]
18. Vassiliadis, G. Transient analysis of a finite source discrete-time queueing system using homogeneous Markov system with state size capacities. *Commun. Stat. Theory Methods* **2014**, *45*, 1403–1423. [CrossRef]
19. Georgiou, A.C.; Thanassoulis, E.; Papdopoulou, A. Using data envelopment analysis in markovian decision making. *Eur. J. Oper. Res.* **2022**, *298*, 276–292. [CrossRef]
20. Guerry, M.A. On the evolution of stock vectors in a deterministic integer-valued Markov system. *Linear Algebra Its Appl.* **2008**, *429*, 1944–1953. [CrossRef]
21. Guerry, M.A. Some results on the embeddable problem for discrete time Markov models in manpower planning. *Commun. Stat. Theory Methods* **2014**, *43*, 1575–1584. [CrossRef]
22. Pollard, B.S. Open Markov processes: A compositional perspective on a non-equilibrium steady state in biology. *Entropy* **2016**, *18*, 140. [CrossRef]
23. Esquivel, M.L.; Fernandes, J.M.; Guerriero, G.R. On the evolution and asymptotic analysis of open Markov populations: Application to consumption credit. *Stoch. Model.* **2014**, *30*, 365–389. [CrossRef]
24. Esquivel, M.L.; Guerriero, G.R.; Fernandes, J.M. Open Markov chain scheme models fed by second order stationary and non stationary processes. *Revstat-Stat. J.* **2017**, *15*, 277.

25. Esquivel, M.L.; Krasil, N.P.; Guerriero, G.R. Open type population models: From discrete to continuous time. *Mathematics* **2021**, *9*, 1496. [CrossRef]
26. Georgiou, A.C.; Vassiliou, P.-C.G. Periodicity of asymptotically attainable stuctures in non-homogeneous Markov systems. *Linear Algebra Its Appl.* **1992**, *176*, 137–174. [CrossRef]
27. Seneta, E. *Non-Negative Matrices and Markov Chains*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 1981.
28. Brémaud, P. Markov Chains. In *Gibbs Fields, Monte Carlo Simulation and Queues*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2020.
29. Tweedie, R.L. Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space. *Stoch. Proc. Appl.* **1975**, *3*, 385–403. [CrossRef]
30. Mitrophanov, A.Y. Sensitivity and convergence of uniformly ergodic Markov chains. *J. Appl. Prob.* **2005**, *42*, 1003–1014. [CrossRef]
31. Seneta, E. Inhomogeneous Markov chains and Ergodicity Coefficients: John Hajnal (1924–2008). *Commun. Stat. Theory Methods* **2014**, *43*, 1575–1584. [CrossRef]
32. Vassiliou, P.-C.G. On the limiting behavior of a non-homogeneous Markov model in manpower systems. *Biometrika* **1981**, *68*, 557–561.
33. Paz, A. *Introduction to Probabilistic Automata*; Academic Press: Cambridge, MA, USA, 1971.
34. Huang, C.; Isaacson, D.; Vinograde, B. The rate of convergence of certain non-homogeneous Markov chains. *Zeitsh. Wahrsch. Geb.* **1976**, *35*, 141–146. [CrossRef]
35. Huang, C.; Isaacson, D. Ergodicity using mean visit times. *J. London Math. Soc.* **1976**, *14*, 570–576. [CrossRef]
36. Vassiliou, P.-C.G.; Tsaklidis, G. The rate of convergence of the vector of variances and covariances in non-homogeneous Markov systems. *J. Appl. Prob.* **1989**, *27*, 776–783. [CrossRef]
37. *Bartholomew, D.J. Stochastic Models for Social Processes*, 3rd ed.; Wiley: New York, NY, USA, 1981.
38. Bartholomew, D.J.; Forbes, A.F.; McClean, S. *Statistical Techniques for Manpower Planning*; John Wiley: Chichester, UK, 1991.
39. Kalbfleisch, J.D.; Prentice, R.L. *The Statistical Analysis of Failure Time Data*, 2nd ed.; John Wiley: New York, NY, USA, 2002.
40. Lee, E.T. *Statistical Methods for Survival Data Analysis*, 2nd ed.; John Wiley: New York, NY, USA, 1992.
41. Cox, D.R.; Oakes, D. *Analysis of Survival Data*; Chapman and Hall: London, UK, 1984.
42. Elandt-Johnson, R.C.; Johnson, N.L. *Survival Models and Data Analysis*; John Wiley: New York, NY, USA, 1980.
43. Anderson, T.W.; Goodman, L.A. Statistical Inference about Markov chains. *Ann. Math. Statist.* **1957**, *28*, 89–110. [CrossRef]
44. Vasileiou, A.; Vassiliou, P.-C.G. An inhomogeneous semi-Markov model for the term structure of credit risk spreads. *Adv. Appl. Prob.* **2006**, *38*, 171–198. [CrossRef]