

Article

Rate of Convergence and Periodicity of the Expected Population Structure of Markov Systems that Live in a General State Space

P. -C. G. Vassiliou 

Department of Statistical Science, University College London, Gower st, London WC1E 6BT, UK;
vassiliou@math.auth.gr

Received: 25 May 2020; Accepted: 12 June 2020; Published: 22 June 2020



Abstract: In this article we study the asymptotic behaviour of the expected population structure of a Markov system that lives in a general state space (MSGS) and its rate of convergence. We continue with the study of the asymptotic periodicity of the expected population structure. We conclude with the study of total variability from the invariant measure in the periodic case for the expected population structure of an MSGS.

Keywords: Markov processes in general spaces; Markov systems; asymptotic periodicity

1. Introductory Notes

Ref [1] introduced the stochastic process non-homogeneous Markov system (NHMS) with countable state space. Ref [2] introduced the stochastic process non-homogeneous semi-Markov system (NHSMS) in countable spaces. The theory and applications of both processes have seen the realization of an interesting growth and have found many applications in fields with great diversity. The latest reviews on the theory are given in [3,4]. The motives of this theory go back in the homogeneous Markov chain models for manpower systems in [5,6] and the evolution of which, in a variety of populations, are well described in [7]. The stochastic process of an NHMS represents a general frame work for a large variety of applied probability models to be accommodated in. However, real motives were the works on non-homogeneous Markov models for manpower systems such as [8–10]. We could selectively only supply some applications of the theory due to their large numbers. We start with the evolution of the HIV virus within the human body of T-cells in [11–13]; asthma was studied in [14]; reliability applications exist in [15] for example; examples of biomedical studies exist in [16,17]; applications in gene sequences ([18]); in DNA and web navigation ([19]); in manpower systems in [20–25]; in Physical Chemistry ([26]); examples from ecological modelling ([27]). Finally, there is the work of the research school of Prof McClean in the health systems (see, for example, [28–32]). Note that health systems are large manpower systems and each member has many parameters that are used to categorize him/her into the different groups. This characteristic of the health systems is what makes it an area of potential application of the present results (see [33]).

The rigorous foundation of Markov systems in general spaces (MSGS) was introduced in [33]. Also, the problem of asymptotic behaviour or ergodicity of Markov systems was studied. Important theorems were proved on the ergodicity of the distribution of expected population structure of a Markov system, which lives in the general space $(X, \mathcal{B}(X))$. In addition, the total variability from the invariant measure was studied given that the Markov system is ergodic. It is shown that the total variation is finite.

Markov chains in general spaces have very important applications in the areas of Markov models in time series, models in control and systems theory, Markov models with regenerations times, Moran

dams, Markov chains Monte Carlo simulation algorithms, etc. For more details, see [34] Chapter 2. There is a belief that the introduction in the presence of a population which is modulated as a Markov chain with general state space will increase the dynamics for new applications considerably, as it was described above in the countable case. Hence, although the results in the present are purely theoretical, the prospects for interesting applications are present and promising.

In Sections 2 and 2.1, we start with the formal definition of a Markov system in a general state space, as it was introduced in [33]. In Section 2.2, we provide an abstract image for an MSGS in an attempt to help the reader in understanding the present paper. In Section 2.3, we introduce some known concepts and results useful in the foundation of the novel results that follow from this point to the end. In Section 3, we study the rate of convergence of the expected population structure in an MSGS. Two theorems are provided where in the first we provide conditions under which the MSGS is uniformly ergodic, and in the second we provide conditions under which the MSGS is V -uniformly ergodic. In Section 4, we proceed to study the asymptotic periodicity of the expected population structure, for a population that lives in a general state space. A basic theorem is proved, where it is shown that if d is the period of the inherent Markov chain, then the sequence of the expected population structures splits into d converging subsequences. The asymptotic periodicity of non-homogeneous Markov systems in countable spaces was studied by [35,36]. In the present study, the basic tools and methodology are different since, among other reasons, the equations are of completely different nature. That is, we move from difference equations to integral equations for the expected population structure. In Section 5, we study the total variability from the invariant measure in the periodic case for the expected population structure of an MSGS. In the form of two theorems, we provide conditions under which the total variability in the periodic case is finite.

2. The Markov System in a General State Space

We will start with a formal definition of a Markov system in a general state space ([33]) and then we provide an explanatory example. The reader could choose his own suitable order of reading the two subsections.

2.1. The Foundation of an MSGS

Let $(X, \mathcal{B}(X))$ be the state space where X is a general set and $\mathcal{B}(X)$ denote a countably generated σ -field on X . Assume that small letters are elements of X and A, B, C elements of $\mathcal{B}(X)$. Let us denote by $T(t)$ the population of the system at time t , which lives in $(X, \mathcal{B}(X))$. We assume that $\{T(t)\}_{t=0}^{\infty}$ is a known stochastic process with state space \mathbb{Z}^+ and in discrete time.

Let that $\lambda : X \rightarrow \mathbb{R}^+$ be a positive σ -finite measure $\lambda : X \rightarrow \mathbb{R}^+$, that is,

$$\lambda(A) \geq 0, \lambda(A) < \infty \text{ for every } A \in \mathcal{B}(X),$$

for which we have that $\lambda(X) = 1$. We assume that $\lambda(\cdot)$ represents the initial distribution of a member of the population in the space $(X, \mathcal{B}(X))$. We also assume that there is a set $W \in \mathcal{B}(X)$ for which $\lambda(W) = 0$. We will refer to W as the “gate” of the space $(X, \mathcal{B}(X))$.

In the space $(X, \mathcal{B}(X))$ where the population lives, it is assumed that each member has a membership, and that the leavers are leaving their memberships at the “gate”. Therefore, at any time t at the “gate” W , the number of memberships are those left by leavers, and the necessary $\Delta T(t) = T(t) - T(t - 1)$ memberships to complete the desired population $T(t)$. We will only work the cases for which $\Delta T(t) \geq 0$. New members of the population take their memberships at the gate “gate” W from which they then make a transition to any $A \in \mathcal{B}(X)$. Now let

$$\{P(x, A), x \in X, A \in \mathcal{B}(X)\}, \tag{1}$$

be the transition probability kernel from $x \in X$ to the set $A \in \mathcal{B}(X)$ in one time step for a member of the population. Denote by $\Phi_p = \left\{ \Phi_p^{(n)} \right\}_{n \in \mathbb{N}^+}$ the Markov chain with state space $(X, \mathcal{B}(X))$, which is defined uniquely by the probability kernel (2.1). We now make the following:

Assumption 1. For the Markov chain $\Phi_p = \left\{ \Phi_p^{(n)} \right\}_{n \in \mathbb{N}^+}$ the set W is an atom.

That is, there exists a measure $\mu(\cdot)$ on $\mathcal{B}(X)$ such that

$$P(x, A) = \mu(A) \text{ for every } x \in W \text{ and } A \in \mathcal{B}(X), \tag{2}$$

with $\mu(X) = 1$.

Now, define by

$$Q(x, A) = P(x, A) + P(x, W) \mu(A), \text{ for every } x \in X, A \in \mathcal{B}(X), x \notin W, \tag{3}$$

$$Q(x, A) = P(x, A) = \mu(A), \text{ for every } x \in W, A \in \mathcal{B}(X). \tag{4}$$

The transition probability kernel $Q(x, A)$ in (3) is the probability of a membership to make a transition from point x into the set A in one time step, either by direct transition of the member holding the membership with probability $P(x, A)$, or by the member leaving the system through the gate W with probability $P(x, W)$ and the entrance of a new member, who gets the membership of the member from gate W , into set A with probability $\mu(A)$. The transition probability kernel $Q(x, A)$ in (4) is the probability of new memberships entering the space $(X, \mathcal{B}(X))$ at each time interval $(t - 1, t]$. From the gate of the atom, W is then distributed according to the σ -finite probability measure $\mu(\cdot) : \mathcal{B}(X) \rightarrow [0, 1]$.

We will use the notation by $\Phi_Q = \left\{ \Phi_Q^{(n)} \right\}_{n \in \mathbb{N}^+}$ for the Markov chain with state space $(X, \mathcal{B}(X))$ defined uniquely by $Q(x, A)$ in (3) and (4). We also use the notation $Q^n(x, A)$ for probability that a membership in x will move in n time steps to the set $A \in \mathcal{B}(X)$.

Now, define:

$N(x, t, A)$: the number of memberships of the population that are in set $A \in \mathcal{B}(X)$ at time t given that were initially at $x \in X$.

Assume, as it is usual in many applications, a natural partition of the space $(X, \mathcal{B}(X))$, that is,

$$\{A_i\}_{i=1}^{k+1}, \text{ with } A_i \in \mathcal{B}(X) \text{ and } A_i \cap A_j = \emptyset \text{ for } i \neq j \text{ and such that} \tag{5}$$

$$\cup_{i=1}^{k+1} A_i = X \text{ and with no loss of generality } A_{k+1} = W. \tag{6}$$

It is of interest the *expected population structure* defined as:

$$\mathbb{E}_{\mathbb{P}}[\mathbf{N}(t, X)] = [\mathbb{E}_{\mathbb{P}}[N(x, t, A_1)], \mathbb{E}_{\mathbb{P}}[N(x, t, A_2)], \dots, \mathbb{E}_{\mathbb{P}}[N(x, t, A_k)]] , \tag{7}$$

where $\mathbf{N}(t, X)$ is a non-negative measurable function on X . We are interested also for the *relative expected population structure* defined by

$$\mathbb{E}_{\mathbb{P}}[\mathbf{q}(t, X)] = \mathbb{E}_{\mathbb{P}}[\mathbf{N}(t, X) / T(t)]. \tag{8}$$

It could be easily seen that $q : X \rightarrow \mathbb{R}^+$ is a positive σ -finite probability measure since

$$q(x, t, A_i) = \frac{N(x, t, A_i)}{T(t)} \geq 0; q(x, t, A_i) < \infty \forall i \in \mathbb{N}^+ \text{ and } q(x, t, X) = 1. \tag{9}$$

Of central importance is the evolution of the expected population structure $\mathbb{E}_{\mathbb{P}} [\mathbf{N} (t, X)]$ or the relative expected population structure $\mathbb{E}_{\mathbb{P}} [\mathbf{q} (t, X)]$ in the study of Markov systems in general state space. From Vassiliou (2014) we get that

$$\mathbb{E}_{\mathbb{P}} [N (x, t, A)] = T (0) \int_{\tilde{X}} \lambda (dx) Q^t (x, A) + \sum_{n=1}^{t-1} \Delta T (n) \int_{\tilde{X}} \mu (dx) Q^{t-n} (x, A). \tag{10}$$

Note that

$$T (0) \int_{\tilde{X}} \lambda (dx) Q^t (x, A),$$

is the expected number of memberships from the initial population $T (0)$ that survived up to time t and are in set A . The part

$$\sum_{n=1}^{t-1} \Delta T (n) \int_{\tilde{X}} \mu (dx) Q^{t-n} (x, A),$$

is the expected number of memberships that entered the system in the interval $[0, t]$ and survived up to time t and are in set A .

We call the random process described above with state space $(X, \mathcal{B} (X))$ a *Markov system in a general state space*. In addition, we will call the Markov chain $\Phi_Q = \{ \Phi_Q^{(n)} \}_{n \in \mathbb{N}^+}$ the *inherent Markov chain* of the Markov system.

2.2. An Abstract Image for MSGS

In the present subsection, in order to understand the various concepts and results, we will use the analogy of the example of the frog who jumps on lilies in a pond. This example of an abstract image was firstly used in [33] and the reader is referred to that study for a more extensive description and an explanation of the various concepts. In summary, imagine a lily pond that is covered by the leaves of the lilies. A population of frogs is living on the pond and a number of them are in every lily at each point in time and hence we have an expected population of frogs in the lilies. There are also leavers and newcomers in the pond at every point in time due to the antagonistic mating situation that exists between the males.

2.3. Introducing Some Important Concepts and Known Results

Throughout the paper, we will assume that the Markov chain Φ_Q is ψ -irreducible. That is, there is a σ -finite measure ψ on $(X, \mathcal{B} (X))$ such that, for any $A \in \mathcal{B} (X)$ with $\psi (A) > 0$, and any $x \in X$,

$$Q^n (x, A) > 0, \text{ for all } n \text{ sufficiently large.} \tag{11}$$

Assume that ψ is *maximal with respect to* ψ . We will also assume that the Markov chain is aperiodic, apart from the last section, where we will study periodicity. The concept of ψ -irreducibility for Markov chains on general state spaces was introduced by [37,38], and followed by many authors [39–44].

We will adopt the term ergodicity for the Markov chains for which the limit, $\lim_{n \rightarrow \infty} Q^n (x, A)$ exists and is equal with $\pi (A)$, where π is the invariant measure or stationary distribution of Φ_Q . As a mode of convergence in the space $(X, \mathcal{B} (X))$, the total variation norm will be used. For a signed measure μ on $(X, \mathcal{B} (X))$, the total variation norm $\|\mu\|$ is given by

$$\|\mu\| := \sup_{f:|f| \leq 1} |\mu (f)| = \sup_{A \in \mathcal{B}(X)} \mu (A) - \inf_{A \in \mathcal{B}(X)} \mu (A). \tag{12}$$

The key limit of interest in the present paper will be of the form

$$\lim_{n \rightarrow \infty} \|Q^n(x, \cdot) - \pi(\cdot)\| = 2 \lim_{n \rightarrow \infty} \sup_A |Q^n(x, \cdot) - \pi(\cdot)|. \tag{13}$$

For more on the spectral theory and rates of convergence for Markov chains in general spaces, see [34]. Research is motivated by elegant mathematics as well as a range of applications. Note that, for countable inhomogeneous Markov systems it was proved that the space of all possible expected population structures is a Hilbert space ([45]). Now, for a ψ -irreducible Markov chain who lives in $(X, \mathcal{B}(X))$, denoted by

$$\mathcal{B}^+(X) = \{A \in \mathcal{B}(X) : \psi(A) > 0\}. \tag{14}$$

Let Q_1 and Q_2 be transition kernels for Markov chains, then for a positive function $\infty > V \geq 1$, define the V -norm distance between Q_1 and Q_2 as

$$\|Q_1 - Q_2\|_V := \sup_{x \in X} \frac{\|Q_1(x, \cdot) - Q_2(x, \cdot)\|_V}{V(x)}. \tag{15}$$

We define the *outer product* of the function 1 and the invariant measure π the kernel

$$[1 \otimes \pi](x, A) = \pi(A), \quad x \in X, A \in \mathcal{B}(X). \tag{16}$$

In many applications, we consider the distance $\|Q^n - 1 \otimes \pi\|_V$ for large n . We provide the following definition from [34]:

Definition 1. (*V-uniform ergodicity*). An ergodic chain Φ with transition kernel Q is called V -uniformly ergodic if

$$\|Q^n - 1 \otimes \pi\|_V \rightarrow 0, \text{ as } n \rightarrow \infty. \tag{17}$$

Of great interest is the special case when $V = 1$. In this case, we provide the following definition:

Definition 2. (*Uniform ergodicity*). A chain Φ with transition kernel Q is called uniformly ergodic if it is V -uniformly ergodic with $V = 1$, that is, if

$$\sup_{x \in X} \|Q^n(x, \cdot) - \pi\| \rightarrow 0, \text{ as } n \rightarrow \infty. \tag{18}$$

The concept of a small set is very useful in order to have a finite break up into cyclic sets for ψ -irreducible chains.

Definition 3. For $C \in \mathcal{B}(X)$ we say that it is a small set if there exists an $m > 0$, and a non-trivial measure ν_m on $\mathcal{B}(X)$, such that for all $x \in C, B \in \mathcal{B}(X)$

$$P^m(x, B) \geq \nu_m(B). \tag{19}$$

In such a case we say that C is ν_m -small.

From [34] p. 105 we get the following Theorem:

Theorem 1. If Φ is ψ -irreducible, then for every $A \in \mathcal{B}^+(X)$, there exists $m \geq 1$ and a ν_m -small set $C \subseteq A$ such that $C \in \mathcal{B}^+(X)$ and $\nu_m(C) > 0$.

Given the existence of just one small set from the previous Theorem, the Proposition in [34] p. 106 states that it is further possible for set X to be covered by small sets in the ψ -irreducible case.

Assume that C is any ν_M -small set, and $\nu_M(C) > 0$, without loss of generality. With the use of the set C ν_M we define a cycle for a general irreducible Markov chain. We will suppress the subscript of ν for simplicity.

We have that

$$P^M(x, \cdot) \geq \nu(\cdot), x \in C \text{ and } \nu(C) > 0,$$

hence, when the chain starts in C , there is a positive probability that the chain will return to C at time M . Let

$$E_C = \{n \geq 1 : \text{the set } C \text{ is } \nu_n = \delta_n \nu \text{ small for some } \delta_n > 0\}.$$

Then the “period” for the set C , is given by the greatest common divisor of E_C . The following Theorem ([34] p. 113) will be useful in what follows:

Theorem 2. Assume that Φ is a ψ -irreducible Markov chain on X .

If $C \in \mathcal{B}^+(X)$ is ν_M -small and d the greatest common divisor of the set E_C . Then there exist disjoint sets $D_0, \dots, D_{d-1} \in \mathcal{B}(X)$ (a “ d -cycle”) such that

(i) for $x \in D_i, P(x, D_{i+1}) = 1, i = 0, \dots, d - 1 \pmod{d}$.

(ii) the set $N = \left[\bigcup_{i=1}^d D_i \right]^c$ is ψ -null.

The d -cycle $\{D_i\}$ is maximal in the sense that for any collection

$$\{d', D'_k, k = 0, \dots, d' - 1\}$$

satisfying (i)-(ii), we have d' dividing d ; whilst if $d = d'$, then, by reordering the indices if necessary $D'_i = D_i$ a.e. ψ .

It is obvious from the above that any small set must be essentially contained inside one specific member of the d -cycle $\{D_i\}$ cyclic class. From [34] p. 115, we need the following Proposition:

Proposition 1. Suppose Φ is a ψ -irreducible chain with period d and a D -cycle $\{D_i, i = 0, \dots, d - 1\}$. Then each of the sets D_i is an absorbing ψ -irreducible set for the chain Φ_d corresponding to the transition kernel P^d , and Φ_d on each D_i is aperiodic.

For any set $A \in \mathcal{B}(X)$, we denote by $oc(A)$ the number of visits by Φ to A after time zero and it is given by

$$oc(A) = \sum_{n=1}^{\infty} \mathbf{1} \{ \Phi^{(n)} \in A \}. \tag{20}$$

We define the kernel

$$U(x, A) := \sum_{n=1}^{\infty} Q^n(x, A) = \mathbb{E}_{\mathbb{P}} [oc_x(A)]. \tag{21}$$

The chain Φ is called recurrent if it is ψ -irreducible and $U(x, A) = \infty$ for every $x \in X$ and every $A \in \mathcal{B}^+(X)$. If the chain Φ is irreducible and admits an atom $a \in \mathcal{B}^+(X)$ then if a is recurrent, then every set in $\mathcal{B}^+(X)$ is recurrent.

Definition 4. The set $A \in \mathcal{B}^+(X)$ is called Harris recurrent if

$$P[oc_x(A) = \infty] = 1. \tag{22}$$

A chain Φ is called Harris recurrent if it is ψ -irreducible and every set in $\mathcal{B}^+(X)$ is Harris recurrent.

Definition 5. A σ -finite measure π on $\mathcal{B}(X)$ with the property

$$\pi(A) = \int_X \pi(dx) P(x, A), \quad A \in \mathcal{B}(X). \tag{23}$$

will be called *invariant*.

Suppose that Φ is irreducible, and admits an invariant probability measure π . Then Φ is called a *positive chain*. If Φ is Harris recurrent and positive, then Φ is called a *positive Harris chain*. From [34] p. 328 and p. 204, we get the following two results:

Theorem 3. If Φ is positive Harris and aperiodic, then for any initial distribution λ

$$\left\| \int_X \lambda(dx) P^n(x, \cdot) - \pi \right\| \rightarrow 0 \text{ as } n \rightarrow \infty. \tag{24}$$

For any Harris recurrent set H , we write $H^\infty = \{y : L(y, H) = 1\}$, where $L(y, H)$ is the probability of Φ of ever entering H from y , so that $H \subseteq H^\infty$ and H^∞ is absorbing. We will call H a *maximal absorbing set* if $H = H^\infty$.

Theorem 4. If Φ is recurrent, then we can write

$$X = H \cup N$$

where H is a non-empty maximal Harris set and N is transient.

3. Rate of Convergence of MSGS

In the present section, we will study the rate of convergence of Markov systems in general spaces. We will provide conditions under which the rate of convergence of the expected population structure of an ergodic Markov system is uniformly ergodic and V -uniformly ergodic. From [34] p. 393, we get the following two Theorems:

Theorem 5. Suppose that Φ is ψ -irreducible and aperiodic with transition kernel Q . Then the following are equivalent for $V \geq 1$:

- (i) Φ is V -uniformly ergodic.
- (ii) There exists $r > 1$ and $R < \infty$ such that for all $n \in \mathbb{N}^+$

$$\|Q^n - 1 \otimes \pi\|_V \leq Rr^{-n}.$$

- (iii) There exists some $n > 0$ such that $\|Q^i - 1 \otimes \pi\|_V < \infty$ for $i \leq n$ and

$$\|Q^i - 1 \otimes \pi\|_V < 1.$$

Theorem 6. For any Markov chain Φ with transition kernel Q , the following are equivalent:

- (i) Φ is uniformly ergodic.
- (ii) There exists $r > 1$ and $R < \infty$ such that for all x

$$\|Q^n(x, \cdot) - \pi(\cdot)\| \leq Rr^{-n}; \tag{25}$$

that is, the convergence takes place at a uniform geometric rate.

- (iii) For some $n \in \mathbb{N}^+$

$$\sup_{x \in X} \|Q^n(x, \cdot) - \pi(\cdot)\| < 1.$$

(iv) The chain is aperiodic and Doeblin condition holds: that is, there is a probability measure ϕ on $\mathcal{B}(X)$ and $\varepsilon < 1, \delta > 0, n \in \mathbb{N}^+$ such that when $\phi(A) > \varepsilon$

$$\inf_{x \in X} Q^n(x, A) > \delta.$$

(v) The set X for some m is ν_m -small.

(vi) For every set $A \in \mathcal{B}^+(X)$ and an aperiodic Markov chain there is a petite set C with

$$\sup_{x \in X} \mathbb{E}_x[\tau_C] < \infty,$$

and in this case $\sup_{x \in X} \mathbb{E}_x[\tau_A] < \infty$.

(vii) For every set $A \in \mathcal{B}^+(X)$ and an aperiodic Markov chain there is a petite set C and $\kappa > 1$ with

$$\sup_{x \in X} \mathbb{E}_x[\kappa^{\tau_C}] < \infty,$$

and we have for some $\kappa_A > 1$,

$$\sup_{x \in X} \mathbb{E}_x[\kappa_A^{\tau_A}] < \infty,$$

(viii) For an aperiodic Markov chain there is a bounded solution $V \geq 1$ to

$$\Delta V(x) \leq -\beta V(x) + b\mathbb{I}_C(x), \quad x \in X.$$

for some $\beta > 0, b < \infty$, and some petite set C ,

Under (v), we get that for any x ,

$$\|Q^n(x, \cdot) - \pi(\cdot)\| \leq 2\rho^{\lfloor n/m \rfloor},$$

where $\rho = 1 - \nu_m(X)$.

Now, from [46], we will borrow the following theorem adapted for an MSGS:

Theorem 7. Let a Markov system in a general state space $(X, \mathcal{B}(X))$, which is expanding ($\Delta T(t) = T(t) - T(t-1) \geq 0$). The following two statements are equivalent

- (i) The sequence $\{T(t)\}_{t=0}^\infty$ converges, that is, $\lim_{t \rightarrow \infty} T(t) = T$ geometrically.
- (ii) The non-negative sequence $\{\Delta T(t)\}_{t=0}^\infty$ tends to zero geometrically.

We will start by studying Uniform ergodicity for an MSGS for simplicity reasons and then proceed to study V -uniform ergodicity.

Definition 6. (Uniform ergodicity for an MSGS). Let a Markov system in a general state space $(X, \mathcal{B}(X))$. Also let $\{T(t)\}_{t=0}^\infty$ be the total population of memberships with $\lim_{t \rightarrow \infty} T(t) = T$; $Q(x, A)$ with $x \in X$ and $A \in \mathcal{B}(X)$ the transition kernel of the inherent Markov chain of memberships and $\mathbb{E}_Q[\mathbf{N}(x, A)]$ the expected population structure. We say that the MSGS is uniformly ergodic if and only if there exists a $C < \infty$ and an $0 < a < 1$ such that

$$\|\mathbb{E}_Q[N(x, t, A)] - T\pi(A)\| \leq Ca^t \text{ for every } x \in X, A \in \mathcal{B}(X). \tag{26}$$

We now provide the following theorem concerning uniform ergodicity for an MSGS.

Theorem 8. Let a Markov system that lives in $(X, \mathcal{B}(X))$. Assume in addition that

$$\lim_{t \rightarrow \infty} T(t) = T \text{ at a geometric rate, } \Delta T(t) \geq 0 \text{ for every } t \in \mathbb{N}_+,$$

and that the inherent Markov chain with transition kernel $Q(x, A)$ is uniformly ergodic. Then the MSGS is uniformly ergodic.

Proof. We have that

$$\begin{aligned} & \|\mathbb{E}_Q [N(x, t, A)] - T\pi(A)\| = \|\mathbb{E}_Q [N(x, t, A)] - T(0)\pi(A) - (T - T(0))\pi(A)\| \\ & = \text{(From Equation (10))} \\ & = \|T(0) \left[\int_X \lambda(dx) Q^t(x, A) - \pi(A) \right] + \sum_{n=1}^{t-1} \Delta T(n) \int_X \mu(dx) Q^{t-n}(x, A) - \sum_{n=1}^{\infty} \Delta T(n) \pi(A)\| \\ \leq & \|T(0) \left[\int_X \lambda(dx) Q^t(x, A) - \pi(A) \right] + \sum_{n=1}^{t-1} \Delta T(n) \left[\int_X \mu(dx) Q^{t-n}(x, A) - \pi(A) \right] - \sum_{n=t}^{\infty} \Delta T(n) \pi(A)\| \tag{27} \\ & \leq T(0) \left\| \int_X \lambda(dx) Q^t(x, A) - \pi(A) \right\| \\ & + \left\| \sum_{n=1}^{t-1} \Delta T(n) \left[\int_X \mu(dx) Q^{t-n}(x, A) - \pi(A) \right] \right\| + \left\| \sum_{n=t}^{\infty} \Delta T(n) \pi(A) \right\|. \end{aligned}$$

From (27) and since $\lim_{t \rightarrow \infty} T(t) = T$ is geometrically fast and $\Delta T(t) \geq 0$ for every $t \in \mathbb{N}_+$ from Theorem 6 we get that there exist a $c > 0$ and a $1 > b > 0$ such that

$$\left\| \sum_{n=t}^{\infty} \Delta T(n) \pi(A) \right\| \leq \|\pi(A)\| \sum_{n=t}^{\infty} \Delta T(n) \leq c(1-b)^{-1} b^t. \tag{28}$$

Also from (27), since $\int_X \lambda(dx) = 1$ and since the inherent Markov chain with kernel $Q(x, A)$ is uniformly ergodic, from Theorem 6 (ii) there exists $R < \infty$ and $r > 1$ such that

$$\begin{aligned} T(0) \left\| \int_X \lambda(dx) Q^t(x, A) - \pi(A) \right\| &= T(0) \left\| \int_X \lambda(dx) [Q^t(x, A) - \pi(A)] \right\| \\ &\leq T(0) \int_X \lambda(dx) \|Q^t(x, A) - \pi(A)\| \leq T(0) \int_X \lambda(dx) R a^t \leq T(0) R a^t, \end{aligned} \tag{29}$$

with $a = r^{-1}, 0 < a < 1$.

Now, again since $\lim_{t \rightarrow \infty} T(t) = T$ is geometrically fast and $\Delta T(t) \geq 0$ for every $t \in \mathbb{N}_+$ from Theorem 6 there exist a $c > 0$ and a $1 > b > 0$ with $\Delta T(t) \leq c b^t$; in addition since the inherent Markov chain with kernel $Q(x, A)$ is uniformly ergodic, from Theorem 6 (ii) there exists $R < \infty$ and $r > 1$, that is, $0 < a = r^{-1} < 1$ such that

$$\begin{aligned} & \left\| \sum_{n=1}^{t-1} \Delta T(n) \left[\int_X \mu(dx) Q^{t-n}(x, A) - \pi(A) \right] \right\| \\ & \leq \sum_{n=1}^{t-1} \Delta T(n) \left\| \int_X \mu(dx) Q^{t-n}(x, A) - \pi(A) \right\| \\ \leq & \sum_{n=1}^{t-1} \Delta T(n) \int_X \mu(dx) \|Q^{t-n}(x, A) - \pi(A)\| \leq \sum_{n=1}^{t-1} \Delta T(n) \int_X \mu(dx) R a^{t-n} \\ & \leq \sum_{n=1}^{t-1} c b^t R a^{t-n} \leq c R (1-b/a)^{-1} \left(1 - (b/a)^t\right) a^t = R^* a^t. \end{aligned} \tag{30}$$

Now, from (27)–(30) we have that

$$\|\mathbb{E}_Q [N(x, t, A)] - T\pi(A)\| \leq T(0) R a^t + R^* a^t + c(1-b)^{-1} b^t. \tag{31}$$

Let $a_1 = \max\{a, b\}$ then $0 < a_1 < 1$ and

$$\begin{aligned} \|\mathbb{E}_Q [N(x, t, A)] - T\pi(A)\| &\leq T(0) R a_1^t + R^* a_1^t + c(1-b)^{-1} a_1^t \\ &\leq \left[T(0) R + R^* + c(1-b)^{-1} \right] a_1^t \\ &\leq c^* a_1^t, \end{aligned}$$

which concludes the proof that MSGS is uniformly ergodic. \square

Generalizing the above results to V -uniform ergodicity is now straight forward. We start with the definition

Definition 7. (V -Uniform ergodicity for an MSGS). Let a Markov system in a general state space $(X, \mathcal{B}(X))$. Also let $\{T(t)\}_{t=0}^\infty$ be the total population of memberships with $\lim_{t \rightarrow \infty} T(t) = T$; $Q(x, A)$ with $x \in X$ and $A \in \mathcal{B}(X)$ the transition kernel of the inherent Markov chain of memberships and $\mathbb{E}_Q[\mathbf{N}(x, A)]$ the expected population structure. We say that the MSGS is V -uniformly ergodic if and only if there exists a $C < \infty$ and an $0 < a < 1$ such that

$$\| \mathbb{E}_Q [N(x, t, A)] - T\pi(A) \|_V \leq Ca^t \text{ for every } x \in X, A \in \mathcal{B}(X). \tag{32}$$

Following exactly the proof of Theorem 8, we arrive at the following result:

Theorem 9. Let a Markov system that leaves in $(X, \mathcal{B}(X))$. Assume in addition that

$$\lim_{t \rightarrow \infty} T(t) = T \text{ at a geometric rate, } \Delta T(t) \geq 0 \text{ for every } t \in \mathbb{N}_+,$$

and that the inherent Markov chain with transition kernel $Q(x, A)$ is V -uniformly ergodic. Then the MSGS is V -uniformly ergodic.

4. Asymptotic Periodicity of an MSGS

The study of the asymptotic behaviour of Markov chains has been very important for finite or countable spaces from the start of their history. The general state space results are due to [40,41,47]. In the non-homogeneous Markov chains when the state space is countable, the initial important results on asymptotic behaviour were achieved by [48–52]. For the semi Markov process analogue, a wealth of results exist in the books by [53,54]. For NHMS and NMSMS on countable spaces the asymptotic behaviour has also been a problem of central importance see for example [1,36,55–57]. In the present section, we will assume that the Markov chain is periodic and study the asymptotic behaviour of the expected population structure, which apparently is an important problem for a population that lives in $(X, \mathcal{B}(X))$. Let it be that the inherent Markov chain of the MSGS with kernel Q is positive Harris with period $d \geq 1$, and let the d -cycle described in Theorem 2 be the set $\{D_0, D_1, \dots, D_{d-1}\}$, where $D_i \in \mathcal{B}(X)$ for $i = 0, 1, \dots, d - 1$. We have assumed that $\lambda(x)$ is the initial distribution on X . Then the distribution on the set D_i is

$$\lambda_i(x) = \frac{\lambda(x)}{\lambda(D_i)} \text{ where } x \in D_i. \tag{33}$$

From Theorem 2, we have that if $x \in D_0$ then $Q(x, D_1) = 1$ and consequently $Q^2(x, D_2) = \int_{D_1} Q(x, dy) Q(y, D_2)$. However, for $y \in D_1$ we have $Q(y, D_2) = 1$, hence $Q^2(x, D_2) = \int_{D_1} Q(x, dy) = Q(x, D_1) = 1$. Similarly, since the period is d we have

$$\text{If } x \in D_i \text{ then } Q^{nd}(x, D_i) = 1, Q^{nd}(x, D_j) = 0 \text{ for } i \neq j. \tag{34}$$

Now, in general for $x \in D_i$ and for $d > i + v$ we have

$$\begin{aligned} Q^{nd+v}(x, D_{i+v}) &= \int_{D_i} Q^{nd}(x, dy) Q^v(y, D_{i+v}) \\ &= (\text{for } y \in D_i \text{ and } d > i + v, Q^v(y, D_{i+v}) = 1) \\ &= \int_{D_i} Q^{nd}(x, dy) = 1. \end{aligned} \tag{35}$$

It is apparent from (35) that

$$\text{If } x \in D_i \text{ and for } d > i + v \text{ we have } Q^{nd+v}(x, D_j) = 0 \text{ for } j \neq i + v. \tag{36}$$

If $x \in D_i$ and for $d \leq i + v$ we have

$$\begin{aligned} Q^{nd+v}(x, D_{i+v-d}) &= \int_{D_i} Q^{nd}(x, dy) Q^v(y, D_{i+v-d}) \\ &= (\text{for } y \in D_i \text{ and } d \leq i + v, Q^v(y, D_{i+v-d}) = 1) \\ &= \int_{D_i} Q^{nd}(x, dy) = 1 \end{aligned}$$

From the fact that the inherent Markov chain with kernel Q is assumed to be positive Harris with period d we know that for each cyclic set is positive Harris on the d -skeleton and aperiodic, and by Theorem 3 it is straightforward to check the following Theorem:

Theorem 10. *If the inherent Markov chain of the MSGS is positive Harris and periodic with period d , then for the sets $\{D_0, D_1, \dots, D_{d-1}\}$ with $D_i \in \mathcal{B}(X)$ we have that*

$$\lim_{n \rightarrow \infty} \left\| \int_{D_i} \lambda_i(dx) P^{nd+v}(x, A) - \pi_{i+v}(A) \right\| = 0 \text{ if } d > i + v \text{ for } i = 0, 1, \dots, d - 1,$$

for any set $A \subseteq D_{i+v}$, and

$$\lim_{n \rightarrow \infty} \left\| \int_{D_i} \lambda_i(x) P^{nd+v}(x, A) - \pi_{i+v-d}(A) \right\| = 0 \text{ if } d \leq i + v \text{ for } i = 0, 1, \dots, d - 1,$$

for any set $A \subseteq D_{i+v-d}$.

Now, from the above Theorem and the decomposition Theorem 4, we immediately get the following:

Theorem 11. *If the inherent Markov chain of the MSGS is positive recurrent and periodic with period d , then for the sets $\{D_0, D_1, \dots, D_{d-1}\}$ with $D_i \in \mathcal{B}(X)$, then for each i there exist a π_i -null set N_i which for every initial distribution λ_i with $\lambda_i(N_i) = 0$*

$$\lim_{n \rightarrow \infty} \left\| \int_{D_i} \lambda_i(dx) P^{nd+v}(x, A) - \pi_{i+v}(A) \right\| = 0 \text{ if } d > i + v \text{ for } i = 0, 1, \dots, d - 1,$$

for any set $A \subseteq D_{i+v}$, and

$$\lim_{n \rightarrow \infty} \left\| \int_{D_i} \lambda_i(dx) P^{nd+v}(x, A) - \pi_{i+v-d}(A) \right\| = 0 \text{ if } d \leq i + v \text{ for } i = 0, 1, \dots, d - 1,$$

for any set $A \subseteq D_{i+v-d}$.

We will now start building the proof for the periodicity Theorem for an MSGS and we will state the theorem in the end. The important question is the periodicity of the expected population structure

$$\mathbb{E}_{\mathbb{P}}[\mathbf{N}(t, X)] = [\mathbb{E}_{\mathbb{P}}[N(x, t, A_1)], \mathbb{E}_{\mathbb{P}}[N(x, t, A_2)], \dots, \mathbb{E}_{\mathbb{P}}[N(x, t, A_k)]] ,$$

when the inherent Markov chain with kernel Q is positive Harris and periodic with period d or is positive recurrent and periodic with period d . With no loss of generality and for simplicity reasons, in what follows we will use the set A instead of any of the sets $A_i, (i = 1, 2, \dots, k)$. From Equation (10) we have that

$$\mathbb{E}_{\mathbb{P}} [N(x, t, A)] = T(0) \int_X \lambda(dx) Q^t(x, A) + \sum_{n=1}^{t-1} \Delta T(n) \int_X \mu(dx) Q^{t-n}(x, A). \tag{37}$$

We will work with each part of the right hand side of Equation (37) separately. We have that

$$\begin{aligned} \int_X \lambda(dx) Q^t(x, A) &= \lambda(D_0) \int_{D_0} \lambda_0(dx) Q^t(x, A) + \lambda(D_1) \int_{D_1} \lambda_1(dx) Q^t(x, A) \\ &\quad + \dots + \lambda(D_{d-1}) \int_{D_{d-1}} \lambda_{d-1}(dx) Q^t(x, A) \end{aligned} \tag{38}$$

We could write the set A in general as follows

$$A = A \cap D_0 \cup A \cap D_1 \cup \dots \cup A \cap D_{d-1}, \text{ with } (A \cap D_i) \cap (A \cap D_j) = \emptyset, \text{ for } i \neq j. \tag{39}$$

Let without loss of generality that $t = \pi d + v$ and consider in what follows that for all $i = 0, 1, \dots, d - 1, d > i + v$. Then we have that

$$Q^t(x, A) = Q^{\pi d + v}(x, A) = Q^{\pi d + v}(x, A \cap D_0) + Q^{\pi d + v}(x, A \cap D_1) + \dots + Q^{\pi d + v}(x, A \cap D_{d-1}). \tag{40}$$

From (38) we have that

$$\begin{aligned} &\lim_{t \rightarrow \infty} \left\| \int_X \lambda(dx) Q^t(x, A) - \sum_{i=0}^{d-1} \lambda(D_i) \pi_{i+v}(A \cap D_{i+v}) \right\| = \\ &= \lim_{\pi \rightarrow \infty} \left\| \sum_{i=0}^{d-1} \lambda(D_i) \int_{D_i} \lambda_i(dx) Q^{\pi d + v}(x, A) - \sum_{i=0}^{d-1} \lambda(D_i) \pi_{i+v}(A \cap D_{i+v}) \right\| \\ &\leq \lim_{\pi \rightarrow \infty} \sum_{i=0}^{d-1} \lambda(D_i) \left\| \int_{D_i} \lambda_i(dx) Q^{\pi d + v}(x, A) - \pi_{i+v}(A \cap D_{i+v}) \right\| \end{aligned} \tag{41}$$

Now, from (40) and (41) we get that

$$\begin{aligned} &\lim_{t \rightarrow \infty} \left\| \int_X \lambda(dx) Q^t(x, A) - \sum_{i=0}^{d-1} \lambda(D_i) \pi_{i+v}(A \cap D_{i+v}) \right\| \\ &\leq \lim_{\pi \rightarrow \infty} \sum_{i=0}^{d-1} \lambda(D_i) \left\| \sum_{j=0}^{d-1} \int_{D_j} \lambda_j(dx) Q^{\pi d + v}(x, A \cap D_j) - \pi_{i+v}(A \cap D_{i+v}) \right\| \\ &\leq \lim_{\pi \rightarrow \infty} \sum_{i=0}^{d-1} \lambda(D_i) \left\| \int_{D_i} \lambda_i(dx) Q^{\pi d + v}(x, A \cap D_{i+v}) - \pi_{i+v}(A \cap D_{i+v}) \right\| \\ &\quad + \lim_{\pi \rightarrow \infty} \sum_{i=0}^{d-1} \lambda(D_i) \sum_{j=0, j \neq i+v}^{d-1} \left\| \int_{D_i} \lambda_i(dx) Q^{\pi d + v}(x, A \cap D_j) \right\| \\ &\quad \text{(from Theorem 10 and relation (36) we get that)} \\ &\leq \in. \end{aligned} \tag{42}$$

Hence, we proved that when $t = \pi d + v$ and $d > i + v$ then the limit of $\int_X \lambda(dx) Q^t(x, A)$ as $t \rightarrow \infty$ is

$$\sum_{i=0}^{d-1} \lambda(D_i) \pi_{i+v}(A \cap D_{i+v}). \tag{43}$$

Similarly, following the same steps we could prove that for $t = \pi d + v$ and $d \leq i + v$ then

$$\lim_{t \rightarrow \infty} \left\| \int_X \lambda(dx) Q^t(x, A) - \sum_{i=0}^{d-1} \lambda(D_i) \pi_{i+v-d}(A \cap D_{i+v-d}) \right\| = 0. \tag{44}$$

From (42) and (43) it is not difficult to check that

$$\begin{aligned} \lim_{t \rightarrow \infty} \left\| \int_X \lambda(dx) Q^t(x, A) - \sum_{i=0}^{d-v-1} \lambda(D_i) \pi_{i+v}(A \cap D_{i+v}) \right. \\ \left. - \sum_{i=d-v}^{d-1} \lambda(D_i) \pi_{i+v-d}(A \cap D_{i+v-d}) \right\| = 0. \end{aligned} \tag{45}$$

Assume now that $\lim_{t \rightarrow \infty} T(t) = T$ and $\Delta T(t) \geq 0$ for every t . Let again that $t = \pi d + v$ then

$$\begin{aligned} U(t) &= \sum_{n=1}^{t-1} \Delta T(n) \int_X \mu(dx) Q^{t-n}(x, A) = \sum_{n=1}^{\pi d+v-1} \Delta T(n) \int_X \mu(dx) Q^{\pi d+v-n}(x, A) \\ &= \sum_{m=0}^{d-1} \sum_{\tau=0}^{\pi-1} \Delta T(\tau d + m) \int_X \mu(dx) Q^{\pi d+v-(\tau d+m)}(x, A) \\ &\quad + \sum_{m=0}^{v-1} \Delta T(\pi d + m) \int_X \mu(dx) Q^{\pi d+v-(\pi d+m)}(x, A). \end{aligned} \tag{46}$$

Now, since $\lim_{t \rightarrow \infty} \Delta T(t) = 0$ we have

$$\begin{aligned} \lim_{\pi \rightarrow \infty} \left\| \sum_{m=0}^{v-1} \Delta T(\pi d + m) \int_X \mu(dx) Q^{\pi d+v-(\pi d+m)}(x, A) \right\| \\ \leq \sum_{m=0}^{v-1} \lim_{\pi \rightarrow \infty} \Delta T(\pi d + m) \lim_{\pi \rightarrow \infty} \left\| \int_X \mu(dx) Q^{\pi d+v-(\pi d+m)}(x, A) \right\| = 0. \end{aligned} \tag{47}$$

Hence, from (46) and (47) we get that

$$\begin{aligned} \lim_{t \rightarrow \infty} U(t) &= \lim_{\pi \rightarrow \infty} \sum_{m=0}^{d-1} \sum_{\tau=0}^{\pi-1} \Delta T(\tau d + m) \int_X \mu(dx) Q^{\pi d+v-(\tau d+m)}(x, A) \\ &= \lim_{\pi \rightarrow \infty} \sum_{m=0}^v \sum_{\tau=0}^{\pi-1} \Delta T(\tau d + m) \int_X \mu(dx) Q^{(\pi-\tau)d+v-m}(x, A) \\ &\quad + \lim_{\pi \rightarrow \infty} \sum_{m=v+1}^{d-1} \sum_{\tau=0}^{\pi-1} \Delta T(\tau d + m) \int_X \mu(dx) Q^{(\pi-\tau-1)d+(d+v-m)}(x, A). \end{aligned} \tag{48}$$

It is apparent that since $\Delta T(t) \geq 0$

$$\sum_{\tau=0}^{\pi-1} \Delta T(\tau d + m) \leq \sum_{t=0}^{\pi d+r} \Delta T(t) = T(\pi d + r) - T(0),$$

and thus the series

$$\sum_{\tau=0}^{\pi-1} \Delta T(\tau d + m),$$

is bounded by $T - T(0)$, its elements are non-negative and thus it converges; so let

$$w_m = \frac{\lim_{\pi \rightarrow \infty} \sum_{\tau=0}^{\pi-1} \Delta T(\tau d + m)}{\lim_{t \rightarrow \infty} \sum_{\tau=0}^t \Delta T(t)}. \tag{49}$$

Now, let us define by

$$\lim_{\pi \rightarrow \infty} U_1(t) = \sum_{m=0}^v \lim_{\pi \rightarrow \infty} \sum_{\tau=0}^{\pi-1} \Delta T(\tau d + m) \int_X \mu(dx) Q^{(\pi-\tau)d+v-m}(x, A). \tag{50}$$

Now from (49) we get that

$$\lim_{\pi \rightarrow \infty} \sum_{\tau=0}^{\pi-1} \Delta T(\tau d + m) = w_m \lim_{t \rightarrow \infty} \sum_{\tau=0}^t \Delta T(t) = w_m (T - T(0)). \tag{51}$$

Therefore (50) becomes

$$\lim_{\pi \rightarrow \infty} U_1(t) = (T - T(0)) \sum_{m=0}^v w_m \tag{52}$$

$$\begin{aligned} & \lim_{\pi \rightarrow \infty} \sum_{\tau=0}^{\pi-1} \frac{\Delta T(\tau d + m)}{\sum_{\tau=0}^{\pi-1} \Delta T(\tau d + m)} \int_X \mu(dx) Q^{(\pi-\tau)d+v-m}(x, A) \times \\ & \lim_{\pi \rightarrow \infty} \sum_{\tau=0}^{\pi-1} \frac{\Delta T(\tau d + m)}{\sum_{\tau=0}^{\pi-1} \Delta T(\tau d + m)} \int_X \mu(dx) Q^{(\pi-\tau)d+v-m}(x, A). \end{aligned}$$

We now provide the following useful Lemma:

Lemma 1. *If $a_n / \sum_{k=0}^n a_k \rightarrow 0$ as $n \rightarrow \infty$ and $\{a_n\}$ non-negative, and $\{b_n\} \rightarrow b$ then*

$$\lim_{n \rightarrow \infty} \frac{b_0 a_n + b_1 a_{n-1} + \dots + b_n a_0}{\sum_{k=0}^n a_k} = b. \tag{53}$$

From Lemma 1, (52) and the result in (45) where we may replace $\lambda(dx)$ by $\mu(dx)$ it is not difficult to check that the following holds

$$\begin{aligned} & \lim_{\pi \rightarrow \infty} \| U_1(t) - (T - T(0)) \sum_{m=0}^v \sum_{i=0}^{d-(v-m)-1} w_m \mu(D_i) \pi_{i+(v-m)}(A \cap D_{i+(v-m)}) \\ & - (T - T(0)) \sum_{m=0}^v \sum_{\substack{i=d-(v-m) \\ i \leq d-1}}^{d-1} w_m \mu(D_i) \pi_{i-d+(v-m)}(A \cap D_{i-d+(v-m)}) \| = 0. \end{aligned} \tag{54}$$

Define by

$$\lim_{t \rightarrow \infty} U_2(t) = \lim_{\pi \rightarrow \infty} \sum_{m=v+1}^{d-1} \sum_{\tau=0}^{\pi-1} \Delta T(\tau d + m) \int_X \mu(dx) Q^{(\pi-\tau-1)d+(d+v-m)}(x, A). \tag{55}$$

Then similarly we could prove that

$$\begin{aligned} & \lim_{\pi \rightarrow \infty} \| U_2(t) - (T - T(0)) \times \\ & \sum_{m=v+1}^{d-1} \sum_{i=0}^{d-(d+v-m)-1} w_m \mu(D_i) \pi_{i+d+(v-m)} \left(A \cap D_{i+d+(v-m)} \right) \\ & - (T - T(0)) \sum_{m=v+1}^{d-1} \sum_{\substack{i=d-(d+m-v) \\ i \leq d-1}}^{d-1} w_m \mu(D_i) \times \\ & \pi_{i-d+(d+v-m)} \left(A \cap D_{i-d+(d+v-m)} \right) \| = 0. \end{aligned} \tag{56}$$

Therefore from (37), (45), (48), (54) and (56) for $t = \pi d + v$ we get that

$$\begin{aligned} & \lim_{t \rightarrow \infty} \| \mathbb{E}_{\mathbb{P}} [N(x, t, A)] - T(0) \sum_{i=0}^{d-v-1} \lambda(D_i) \pi_{i+v} (A \cap D_{i+v}) \\ & - T(0) \sum_{i=d-v}^{d-1} \lambda(D_i) \pi_{i+v-d} (A \cap D_{i+v-d}) \\ & - (T - T(0)) \sum_{m=0}^v \sum_{i=0}^{d-(v-m)-1} w_m \mu(D_i) \pi_{i+(v-m)} \left(A \cap D_{i+(v-m)} \right) \\ & - (T - T(0)) \sum_{m=0}^v \sum_{\substack{i=d-(v-m) \\ i \leq d-1}}^{d-1} w_m \mu(D_i) \pi_{i-d+(v-m)} \left(A \cap D_{i-d+(v-m)} \right) \\ & - (T - T(0)) \sum_{m=v+1}^{d-1} \sum_{i=0}^{(m-v)-1} w_m \mu(D_i) \pi_{i+d+(v-m)} \left(A \cap D_{i+d+(v-m)} \right) \\ & - (T - T(0)) \sum_{m=v+1}^{d-1} \sum_{\substack{i=(m-v) \\ i \leq d-1}}^{d-1} w_m \mu(D_i) \pi_{i+(v-m)} \left(A \cap D_{i+(v-m)} \right) \| = 0, \end{aligned} \tag{57}$$

for $v = 0, 1, \dots, d - 1$. Hence, we have proved the following Theorem:

Theorem 12. *Let an MSGS and let that the inherent Markov chain of the MSGS be positive Harris and periodic with period d .*

OR let it be that the inherent Markov chain of the MSGS is positive recurrent and periodic with period d , and let the sets $\{D_0, D_1, \dots, D_{d-1}\}$ with $D_i \in \mathcal{B}(X)$. Then the expected population structure splits into d converging subsequences, that is,

$$\begin{aligned} & \lim_{t \rightarrow \infty} \| \mathbb{E}_{\mathbb{P}} [N(x, t, A)] - T(0) \sum_{i=0}^{d-v-1} \lambda(D_i) \pi_{i+v} (A \cap D_{i+v}) \\ & - T(0) \sum_{i=d-v}^{d-1} \lambda(D_i) \pi_{i+v-d} (A \cap D_{i+v-d}) \\ & - (T - T(0)) \sum_{m=0}^v \sum_{i=0}^{d-(v-m)-1} w_m \mu(D_i) \pi_{i+(v-m)} \left(A \cap D_{i+(v-m)} \right) \\ & - (T - T(0)) \sum_{m=0}^v \sum_{\substack{i=d-(v-m) \\ i \leq d-1}}^{d-1} w_m \mu(D_i) \pi_{i-d+(v-m)} \left(A \cap D_{i-d+(v-m)} \right) \\ & - (T - T(0)) \sum_{m=v+1}^{d-1} \sum_{i=0}^{(m-v)-1} w_m \mu(D_i) \pi_{i+d+(v-m)} \left(A \cap D_{i+d+(v-m)} \right) \end{aligned}$$

$$-(T - T(0)) \sum_{m=v+1}^{d-1} \sum_{\substack{i=(m-v) \\ i \leq d-1}}^{d-1} w_m \mu(D_i) \pi_{i+(v-m)}(A \cap D_{i+(v-m)}) = 0,$$

for $v = 0, 1, 2, \dots, d - 1$.

5. Total Variability from the Invariant Measures in the Periodic Case. Coupling Theorems

In this section, we study the total variability from the invariant measures in the periodic case for an MSGS. We show that the total variation in the periodic case is finite. This is also known as the coupling problem (see [58]). From [34] (p. 332) we get the following:

Theorem 13. *Suppose that Φ is positive Harris and aperiodic chain and assume that the chain has an atom $a \in \mathcal{B}^+(X)$. Then for any λ, μ regular initial distributions*

$$\sum_{n=1}^{\infty} \int_X \int_X \lambda(dx) \mu(dy) \|P^n(x, \cdot) - P^n(y, \cdot)\| < \infty,$$

and in the case that Φ is regular, then for any $x, y \in X$

$$\sum_{n=1}^{\infty} \|P^n(x, \cdot) - P^n(y, \cdot)\| < \infty.$$

Theorem 14. *Suppose Φ is a positive Harris and aperiodic chain. For any λ, μ regular initial distributions*

$$\sum_{n=1}^{\infty} \int_X \int_X \lambda(dx) \mu(dy) \|P^n(x, \cdot) - P^n(y, \cdot)\| < \infty.$$

Theorem 15. *Suppose that Φ is a positive Harris and aperiodic chain and then let α be an accessible atom. If $\mathbb{E}_\alpha[\tau_\alpha^2] < \infty$, then for any regular initial distribution*

$$\sum_{n=1}^{\infty} \|\lambda P^n - \pi\| < \infty.$$

Now, if we assume that the inherent Markov chain Φ_Q with kernel Q of the MSGS is positive Harris with period d then by Proposition 1 when restricted to each cyclic set is aperiodic and positive Harris on the d -skeleton, and by Theorems 13–15 it is straightforward to check the following theorem:

Theorem 16. *If the inherent Markov chain Φ_Q with kernel Q of the MSGS is positive Harris with period d , then for the sets $\{D_0, D_1, \dots, D_{d-1}\}$ with $D_i \in \mathcal{B}(X)$ we have that*

(a) *For any regular initial distributions λ, μ and for $d > i + v$,*

$$\sum_{n=1}^{\infty} \int_{D_i} \int_{D_i} \lambda(dx) \mu(dy) \left\| Q^{nd+v}(x, A \cap D_{i+v}) - Q^{nd+v}(y, A \cap D_{i+v}) \right\| < \infty, \quad x, y \in D_i, \text{ for } i, v = 0, 1, \dots, d - 1.$$

Also, for $d \leq i + v$

$$\sum_{n=1}^{\infty} \int_{D_i} \int_{D_i} \lambda(dx) \mu(dy) \left\| Q^{nd+v}(x, A \cap D_{i+v-d}) - Q^{nd+v}(y, A \cap D_{i+v-d}) \right\| < \infty, \quad x, y \in D_i, \text{ for } i, v = 0, 1, \dots, d - 1.$$

for $i, v = 0, 1, \dots, d - 1$.

(b) If for the accessible atom W we have that $\mathbb{E}_w [\tau_w^2] < \infty$, then for any regular initial distribution λ we have for $d > i + v, x \in D_i$ and $i, v = 0, 1, \dots, d - 1$

$$\sum_{n=1}^{\infty} \left\| \lambda(dx) Q^{nd+v}(x, A \cap D_{i+v}) - \pi_{i+v}(A \cap D_{i+v}) \right\| < \infty,$$

and for $d \leq i + v, x \in D_i$ and $i, v = 0, 1, \dots, d - 1$

$$\sum_{n=1}^{\infty} \left\| \lambda(dx) Q^{nd+v}(x, A \cap D_{i+v-d}) - \pi_{i+v-d}(A \cap D_{i+v-d}) \right\| < \infty.$$

We will now prove the following coupling theorem for an MSGS with inherent Markov chain Φ_Q , which is assumed to be positive Harris and periodic.

Theorem 17. Assume that the inherent Markov chain Φ_Q of a Markov system that lives in $(X, \mathcal{B}(X))$ with kernel $Q(x, \cdot)$ is positive Harris with period d . Also, let it be that

$$\lim_{t \rightarrow \infty} T(t) = T, \Delta T(t) \geq 0 \text{ for every } t \in \mathbb{Z}^+.$$

Then for any regular distributions $\lambda_1, \mu_1, \lambda_2, \mu_2$ we have

$$\sum_{t=1}^{\infty} \left\| \mathbb{E}_P N(x, t, A) - \mathbb{E}_P [N(y, t, A)] \right\| < \infty.$$

Proof. We have that

$$\begin{aligned} \sum_{t=1}^{\infty} \left\| \mathbb{E}_P N(x, t, A) - \mathbb{E}_P [N(y, t, A)] \right\| &= \sum_{t=1}^{\infty} \left\| T(0) \int_X \lambda_1(dx) Q^t(x, A) + \right. \\ &\quad \sum_{n=1}^{t-1} \Delta T(n) \int_X \mu_1(dx) Q^{t-n}(x, A) - T(0) \int_X \lambda_2(dy) Q^t(y, A) - \\ &\quad \left. - \sum_{n=1}^{t-1} \Delta T(n) \int_X \mu_2(dy) Q^{t-n}(y, A) \right\| \\ &\leq \sum_{t=1}^{\infty} T(0) \left\| \int_X \lambda_1(dx) Q^t(x, A) - \int_X \lambda_2(dy) Q^t(y, A) \right\| + \\ &\quad \sum_{t=1}^{\infty} \sum_{n=1}^{t-1} \Delta T(n) \left\| \int_X \mu_1(dx) Q^{t-n}(x, A) - \int_X \mu_2(dy) Q^{t-n}(y, A) \right\| = U_\lambda(t) + U_\mu(t) \end{aligned} \tag{58}$$

With no loss of generality assume that $t = \pi d + v$. Denoted by

$$\lambda_{1i}(dx) = \frac{\lambda_1(dx)}{\lambda_1(D_i)} \text{ and } \lambda_{2i}(dx) = \frac{\lambda_2(dx)}{\lambda_2(D_i)}. \tag{59}$$

From (58) and (59) we have that

$$U_\lambda(t) \leq \sum_{\pi=0}^{\infty} \sum_{i=0}^{d-1} T(0) \left\| \lambda_1(D_i) \int_{D_i} \lambda_{1i}(dx) Q^{\pi d+v}(x, A) - \lambda_2(D_i) \int_{D_i} \lambda_{2i}(dy) Q^{\pi d+v}(y, A) \right\|. \tag{60}$$

We have that

$$\sum_{i=0}^{d-1} \lambda_1(D_i) = 1 \text{ and } \sum_{i=0}^{d-1} \lambda_2(D_i) = 1. \tag{61}$$

From (61) it is apparent that with no loss of generality we may assume that

$$\lambda_1(D_i) \leq \lambda_2(D_i) \text{ for } i = 0, 1, \dots, j \text{ and } \lambda_1(D_i) \geq \lambda_2(D_i) \text{ for } j = j + 1, \dots, d - 1. \tag{62}$$

Also, without loss of generality again we may assume that $d > j + v$, then from (60) and (62)

$$\begin{aligned} U_\lambda(t) &\leq \sum_{\pi=0}^{\infty} \sum_{i=0}^j T(0) \lambda_2(D_i) \left\| \int_{D_i} \lambda_{1i}(dx) Q^{\pi d+v}(x, A \cap D_{i+v}) \right. \\ &\quad \left. - \int_{D_i} \lambda_{2i}(dy) Q^{\pi d+v}(y, A \cap D_{i+v}) \right\| \\ &+ \sum_{\pi=0}^{\infty} \sum_{i=j+1}^{d-v-1} T(0) \lambda_1(D_i) \left\| \int_{D_i} \lambda_{1i}(dx) Q^{\pi d+v}(x, A \cap D_{i+v}) \right. \\ &\quad \left. - \int_{D_i} \lambda_{2i}(dy) Q^{\pi d+v}(y, A \cap D_{i+v}) \right\| \\ &+ \sum_{\pi=0}^{\infty} \sum_{i=d-v}^{d-1} T(0) \lambda_1(D_i) \left\| \int_{D_i} \lambda_{1i}(dx) Q^{\pi d+v}(x, A \cap D_{i+v-d}) \right. \\ &\quad \left. - \int_{D_i} \lambda_{2i}(dy) Q^{\pi d+v}(y, A \cap D_{i+v-d}) \right\| = U_\lambda^{(1)}(t) + U_\lambda^{(2)}(t) + U_\lambda^{(3)}(t). \end{aligned} \tag{63}$$

Now, it is not difficult to check that

$$\int_{D_i} \lambda_{1i}(dx) Q^{\pi d+v}(x, A \cap D_{i+v}) = \int_{D_i} \int_{D_i} \lambda_{1i}(dx) \lambda_{2i}(dy) Q^{\pi d+v}(x, A \cap D_{i+v}), \tag{64}$$

and

$$\int_{D_i} \lambda_{2i}(dy) Q^{\pi d+v}(y, A \cap D_{i+v}) = \int_{D_i} \int_{D_i} \lambda_{1i}(dx) \lambda_{2i}(dy) Q^{\pi d+v}(y, A \cap D_{i+v}). \tag{65}$$

Therefore from (63)–(65) we get that

$$\begin{aligned} U_\lambda^{(1)}(t) &\leq \sum_{i=0}^j T(0) \lambda_2(D_i) \sum_{\pi=0}^{\infty} \int_{D_i} \int_{D_i} \lambda_{1i}(dx) \lambda_{2i}(dy) \\ &\quad \times \left\| Q^{\pi d+v}(x, A \cap D_{i+v}) - Q^{\pi d+v}(y, A \cap D_{i+v}) \right\|. \end{aligned} \tag{66}$$

From Theorem 16 and (66) we get that

$$U_\lambda^{(1)}(t) \leq \sum_{i=0}^j T(0) \lambda_2(D_i) K_1 < \infty, \tag{67}$$

In a similar way we get that $U_\lambda^{(2)}(t) < \infty$ and $U_\lambda^{(3)}(t) < \infty$ and consequently from (63) we get that

$$U_\lambda(t) < \infty. \tag{68}$$

The case for $d \leq j + v$ is proved in an analogous way.

We have that

$$\begin{aligned}
 U_\mu(t) &= \sum_{t=1}^{\infty} \left[\sum_{n=0}^{t-1} \Delta T(n) \int_X \mu_1(dx) Q^{t-n}(x, A) \right. \\
 &\quad \left. - \sum_{n=0}^{t-1} \Delta T(n) \int_X \mu_2(dy) Q^{t-n}(y, A) \right] \\
 &= \sum_{\pi=0}^{\infty} \sum_{v=0}^{d-1} \left[\sum_{n=0}^{\pi d+v-1} \Delta T(n) \left\{ \int_X \mu_1(dx) Q^{t-n}(x, A) - \int_X \mu_2(dy) Q^{t-n}(y, A) \right\} \right] \\
 &= \sum_{\pi=0}^{\infty} \sum_{v=0}^{d-1} \left[\sum_{m=0}^{d-1} \sum_{\tau=0}^{\pi-1} \Delta T(\tau d + m) \times \right. \\
 &\quad \left. \left\{ \int_X \mu_1(dx) Q^{(\pi d+v)-(\tau d+m)}(x, A) - \int_X \mu_2(dy) Q^{(\pi d+v)-(\tau d+m)}(y, A) \right\} \right] \\
 &\quad + \sum_{\pi=0}^{\infty} \left[\sum_{m=0}^{v-1} \Delta T(\tau d + m) \left\{ \int_X \mu_1(dx) Q^{(\pi d+v)-(\pi d+m)}(x, A) \right. \right. \\
 &\quad \left. \left. - \int_X \mu_2(dy) Q^{(\pi d+v)-(\pi d+m)}(y, A) \right\} \right] = U_\mu^{(\pi)}(t) + U_\mu^{(v)}(t).
 \end{aligned} \tag{69}$$

We start with $U_\mu^{(v)}(t)$ and we have that

$$\begin{aligned}
 \|U_\mu^{(v)}(t)\| &\leq \sum_{\pi=0}^{\infty} \left[\sum_{m=0}^{v-1} \Delta T(\tau d + m) \right. \\
 &\quad \left. \times \left\| \int_X \mu_1(dx) Q^{v-m}(x, A) - \int_X \mu_2(dy) Q^{v-m}(y, A) \right\| \right] \\
 &\leq \sum_{\pi=0}^{\infty} \left[\sum_{m=0}^{v-1} \Delta T(\tau d + m) \sum_{i=0}^{d-1} \left\| \int_{D_i} \mu_1(dx) Q^{v-m}(x, A \cap D_{i+v-m}) \right. \right. \\
 &\quad \left. \left. - \int_{D_i} \mu_2(dy) Q^{v-m}(y, A \cap D_{i+v-m}) \right\| \right].
 \end{aligned} \tag{70}$$

Denote by

$$\mu_{1i}(dx) = \frac{\mu_1(dx)}{\mu_1(D_i)} \text{ and } \mu_{2i}(dx) = \frac{\mu_2(dx)}{\mu_2(D_i)}. \tag{71}$$

Then with no loss of generality we may assume that there is an r such that

$$\mu_1(D_i) \leq \mu_2(D_i) \text{ for } i = 0, 1, \dots, r \text{ and } \mu_1(D_i) > \mu_2(D_i) \text{ for } i = r + 1, \dots, d - 1. \tag{72}$$

From (70)–(72) and using Theorem 16 we get that

$$\begin{aligned}
 \|U_\mu^{(v)}(t)\| &\leq \sum_{\pi=0}^{\infty} \left[\sum_{m=0}^{v-1} \Delta T(\tau d + m) \sum_{i=0}^r \mu_2(D_i) \times \right. \\
 &\quad \left. \left\| \int_{D_i} \int_{D_i} \mu_{1i}(dx) \mu_{2i}(dy) Q^{v-m}(x, A \cap D_{i+v-m}) \right. \right. \\
 &\quad \left. \left. - \int_{D_i} \int_{D_i} \mu_{1i}(dx) \mu_{2i}(dy) Q^{v-m}(y, A \cap D_{i+v-m}) \right\| \right] \\
 &\quad + \sum_{\pi=0}^{\infty} \left[\sum_{m=0}^{v-1} \Delta T(\tau d + m) \sum_{j=r+1}^{d-1} \mu_1(D_j) \times \right. \\
 &\quad \left. \left\| \int_{D_j} \int_{D_j} \mu_{1j}(dx) \mu_{2j}(dy) Q^{v-m}(x, A \cap D_{j+v-m}) \right. \right. \\
 &\quad \left. \left. - \int_{D_j} \int_{D_j} \mu_{1j}(dx) \mu_{2j}(dy) Q^{v-m}(y, A \cap D_{j+v-m}) \right\| \right] \\
 &\leq \sum_{\pi=0}^{\infty} \sum_{m=0}^{v-1} \Delta T(\tau d + m) \left[\sum_{j=0}^r \mu_2(D_j) K_1(v-m) + \sum_{j=r+1}^{d-1} \mu_1(D_j) K_2(v-m) \right] < \infty.
 \end{aligned} \tag{73}$$

We now move to $U_\mu^{(\pi)}(t)$ and we have that

$$\begin{aligned} \|U_\mu^{(\pi)}(t)\| &\leq \sum_{\pi=0}^{\infty} \sum_{v=0}^{d-1} \left[\sum_{m=0}^v \sum_{\tau=0}^{\pi-1} \Delta T(\tau d + m) \left\| \int_X \mu_1(dx) Q^{(\pi-\tau)d+v-m}(x, A) \right. \right. \\ &\quad \left. \left. - \int_X \mu_2(dy) Q^{(\pi-\tau)d+v-m}(y, A) \right\| \right. \\ &\quad \left. \sum_{\pi=0}^{\infty} \sum_{v=0}^{d-1} \left[\sum_{m=v+1}^{d-1} \sum_{\tau=0}^{\pi-1} \Delta T(\tau d + m) \left\| \int_X \mu_1(dx) Q^{(\pi-\tau-1)d+d+v-m}(x, A) \right. \right. \right. \\ &\quad \left. \left. - \int_X \mu_2(dy) Q^{(\pi-\tau-1)d+d+v-m}(y, A) \right\| \right] = \|U_\mu^{(\pi,1)}(t)\| + \|U_\mu^{(\pi,2)}(t)\|. \end{aligned} \tag{74}$$

We start first with $U_\mu^{(\pi,1)}(t)$ and we get that

$$\begin{aligned} \|U_\mu^{(\pi,1)}(t)\| &\leq \sum_{\tau=0}^{\infty} \sum_{v=0}^{d-1} \sum_{m=0}^v \Delta T(\tau d + m) \sum_{\pi-\tau=1}^{\infty} \sum_{i=0}^{d-1} \left\| \int_{D_i} \mu_1(dx) Q^{(\pi-\tau)d+v-m}(x, A) \right. \\ &\quad \left. - \int_{D_i} \mu_2(dy) Q^{(\pi-\tau)d+v-m}(y, A) \right\| \\ &\leq \sum_{\tau=0}^{\infty} \sum_{v=0}^{d-1} \sum_{m=0}^v \Delta T(\tau d + m) \sum_{\pi-\tau=1}^{\infty} \sum_{i=0}^r \mu_2(D_i) \times \\ &\quad \left\| \int_{D_i} \mu_{1i}(dx) \mu_{2i}(dy) Q^{(\pi-\tau)d+v-m}(x, A) \right. \\ &\quad \left. - \int_{D_i} \mu_{1i}(dx) \mu_{2i}(dy) Q^{(\pi-\tau)d+v-m}(x, A) \right\| \\ &\quad + \sum_{\tau=0}^{\infty} \sum_{v=0}^{d-1} \sum_{m=0}^v \Delta T(\tau d + m) \sum_{\pi-\tau=1}^{\infty} \sum_{i=r+1}^{d-1} \mu_1(D_i) \times \\ &\quad \left\| \int_{D_i} \mu_{1i}(dx) \mu_{2i}(dy) Q^{(\pi-\tau)d+v-m}(x, A) \right. \\ &\quad \left. - \int_{D_i} \mu_{1i}(dx) \mu_{2i}(dy) Q^{(\pi-\tau)d+v-m}(x, A) \right\| = \|U_\mu^{(\pi,1,1)}(t)\| + \|U_\mu^{(\pi,1,2)}(t)\|. \end{aligned} \tag{75}$$

Now, using Theorem 17 and from (75) we have that

$$\|U_\mu^{(\pi,1,1)}(t)\| \leq \sum_{\tau=0}^{\infty} \sum_{v=0}^{d-1} \sum_{m=0}^v \Delta T(\tau d + m) \sum_{\pi-\tau=1}^{\infty} \sum_{i=0}^r \mu_2(D_i) C(v-m) < \infty. \tag{76}$$

Similarly, we get also that

$$\|U_\mu^{(\pi,1,2)}(t)\| < \infty \text{ and } \|U_\mu^{(\pi,2)}(t)\| < \infty. \tag{77}$$

From (68), (69), (76), and (77), we conclude the proof of the Theorem. \square

Similarly, and with the use of Theorem 16.b we could prove the following Theorem

Theorem 18. Assume that the inherent Markov chain Φ_Q of a Markov system that lives in $(X, \mathcal{B}(X))$ with kernel $Q(x, \cdot)$ is positive Harris with period d . Assume, in addition, that for the accessible atom W , we have $\mathbb{E}_\omega[\tau_\omega^2] < \infty$ and that

$$\lim_{t \rightarrow \infty} T(t) = T, \Delta T(t) \geq 0 \text{ for every } t \in \mathbb{Z}^+.$$

Then for any regular distribution λ we have

$$\sum_{t=1}^{\infty} \|\mathbb{E}[N(x, t, A)] - T(0) \sum_{v=0}^{d-1} \sum_{i=0}^{d-v-1} \lambda(D_i) \pi_{i+v}(A \cap D_{i+v})\|$$

$$\begin{aligned}
 & -T(0) \sum_{v=0}^{d-1} \sum_{i=d-v}^{d-1} \lambda(D_i) \pi_{i+v-d} (A \cap D_{i+v-d}) \\
 & - (T - T(0)) \sum_{v=0}^{d-1} \sum_{m=0}^v \sum_{i=0}^{d-(v-m)-1} w_m \mu(D_i) \pi_{i+(v-m)} (A \cap D_{i+(v-m)}) \\
 & - (T - T(0)) \sum_{v=0}^{d-1} \sum_{m=0}^v \sum_{\substack{i=d-(v-m) \\ i \leq d-1}}^{d-1} w_m \mu(D_i) \pi_{i-d+(v-m)} (A \cap D_{i-d+(v-m)}) \\
 & - (T - T(0)) \sum_{v=0}^{d-1} \sum_{m=v+1}^{d-1} \sum_{i=0}^{(m-v)-1} w_m \mu(D_i) \pi_{i+d+(v-m)} (A \cap D_{i+d+(v-m)}) \\
 & - (T - T(0)) \sum_{v=0}^{d-1} \sum_{m=v+1}^{d-1} \sum_{\substack{i=(m-v) \\ i \leq d-1}}^{d-1} w_m \mu(D_i) \pi_{i+(v-m)} (A \cap D_{i+(v-m)}) \parallel = 0.
 \end{aligned}$$

6. Conclusions and Further Research

We studied a population that lives in a general state space that at every instant has leavers and new entrants, which evolves under the Markov property. We proved that under certain conditions the expected population structure converges with geometrical rate to an invariant population structure with loss of memory. We also provided the analogous result for the V-uniformly ergodicity of MSGS. Moreover, under the assumption that the inherent Markov chain is periodic with period d , we proved that the sequence of the expected population structure splits into d convergent subsequences and provided their limits in closed forms. Finally, we proved that the total variability from the invariant measure in the periodic case is finite. We concluded the novel results by proving two coupling theorems under the assumption that the inherent Markov chain is positive Harris with period d using different additional conditions in each case. The above theoretical results have important applications both in the classical areas where Markov chains in general spaces have already been applied but especially to the important area of the health systems that, nowadays, are in crisis all over the world.

The results in the present paper together with the ones introduced in [33] provide the nucleus for a new path for extensive research in the area. Immediate further research of interest is to establish Laws of Large numbers for MSGS as a possible extension of the Laws of Large numbers for populations that live in countable spaces [59].

Funding: This research received no external funding.

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

References

1. Vassiliou, P.-C.G. Asymptotic behaviour of Markov systems. *J. Appl. Probab.* **1982**, *19*, 851–857. [CrossRef]
2. Vassiliou, P.-C.G. and Papadopoulou, A.A. Nonhomogeneous semi-Markov systems and maintainability of the state sizes. *J. Appl. Probab.* **1992**, *29*, 519–534. [CrossRef]
3. Vassiliou, P.-C.G. The evolution of the theory of non-homogeneous Markov systems. *Appl. Stoch. Models Data Anal.* **1997**, *13*, 159–176. [CrossRef]
4. Ugwuogo, F.I.; Mc Clean, S.I. Modelling heterogeneity in manpower systems: A review. *Appl. Stoch. Models Bus. Ind.* **2000**, *2*, 99–110. [CrossRef]
5. Young, A.; Almond, G. Predicting distributions of staff. *Comput. J.* **1961**, *3*, 246–250. [CrossRef]
6. Bartholomew, D.J. A multistage renewal process. *J. R. Stat. Soc. B* **1963**, *25*, 150–168.
7. Bartholomew, D.J. *Stochastic Models for Social Processes*; Wiley: New York, NY, USA, 1982.
8. 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]

9. Vassiliou, P.-C.G. A markov model for wastage in manpower systems. *Oper. Res. Quart.* **1976**, *27*, 57–70. [[CrossRef](#)]
10. 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](#)]
11. Mathew, E.; Foucher, Y.; Dellamonika, P. Daures, J.-P. *Parametric and Non-Homogeneous Semi-Markov Process for HIV Control*; Working paper; Archer Hospital: Nice, France, 2006.
12. Foucher, Y.; Mathew, E.; Saint Pierre P.; Durant, J.-F.; Daures, J.P. A semi-Markov model based on generalized Weibull distribution with an illustration for HIV disease. *Biom. J.* **2005**, *47*, 825–833. [[CrossRef](#)]
13. Dessie, Z.G. Modeling of HIV/AIDS dynamic evolution using non-homogeneous semi-Markov process. *Springerplus* **2014**, *3*, 537. [[CrossRef](#)] [[PubMed](#)]
14. Saint Pierre, P. Modelles multi-etas de type Markovien et Application a la Astme. Ph.D. Thesis, University of Montpellier I, Montpellier, France, 2005.
15. Barbu, V.; Boussemment, M.; Limnios, N. Discrete-time semi Markov model for reliability and survival analysis. *Commun. Stat. Theory Methods* **2004**, *33*, 2833–2886. [[CrossRef](#)]
16. Perez-Ocon, R.; Castro, J.E.R. A semi-Markov model in biomedical studies. *Commun. Stat. Theory Methods* **2004**, *33*, 437–455.
17. Ocana-Riola, R. Non-homogeneous Markov process for biomedical data analysis. *Biom. J.* **2005**, *47*, 369–376. [[CrossRef](#)] [[PubMed](#)]
18. McClean, S.I.; Scotney, B.; Robinson, S. Conceptual clustering of heterogeneous gene expression sequences. *Artif. Intell. Rev.* **2003**, *20*, 53–73. [[CrossRef](#)]
19. Papadopoulou, A.A. Some results on modeling biological sequences and web navigation with a semi-Markov chain. *Commun. Stat. Theory Methods* **2013**, *41*, 2853–2871. [[CrossRef](#)]
20. De Freyter, T. Modelling heterogeneity in manpower planning: Dividing the personel system in more homogeneous subgroups. *Appl. Stoch. Model. Bus. Ind.* **2006**, *22*, 321–334. [[CrossRef](#)]
21. Nilakantan, K.; Raghavendra, B.G. Control aspects in proportionality Markov manpower systems. *Appl. Stoch. Models Data Anal.* **2005**, *7*, 27–341. [[CrossRef](#)]
22. Yadavalli, V.S.S.; Natarajan, R.; Udayabhaskaram, S. Optimal training policy for promotion-stochastic models of manpower systems. *Electron. Publ.* **2002**, *13*, 13–23. [[CrossRef](#)]
23. Sen Gupta, A.; Ugwuogo, F.L. Modelling multi-stage processes through multivariate distributions. *J. Appl. Stat.* **2006**, *33*, 175–187. [[CrossRef](#)]
24. De Freyter, T.; Guerry, M. Markov manpower models: A review. In *Handbook of Optimization Theory: Decision Analysis and Applications*; Varela, J., Acuidja, S., Eds.; Nova Science Publishers: Hauppauge, NY, USA, 2011; pp. 67–88.
25. Guerry, M.A. Some results on the Embeddable problem for discrete time Markov models. *Commun. Stat. Theory Methods* **2014**, *43*, 1575–1584. [[CrossRef](#)]
26. Crooks, G.E. Path ensemble averages in system driven far from equilibrium. *Phys. Rev. E* **2000**, *61*, 2361–2366. [[CrossRef](#)]
27. Patoucheas, P.D.; Stamou, G. Non-homogeneous Markovian models in ecological modelling: A study of the zoobenthos dynamics in Thermaikos Gulf, Greece. *Ecol. Model.* **1993**, *66*, 197–215. [[CrossRef](#)]
28. Faddy, M.J.; McClean, S.I. Markov chain for geriatric patient care. *Methods Inf. Med.* **2005**, *44*, 369–373. [[CrossRef](#)] [[PubMed](#)]
29. Gournescu, F.; McClean, S.I.; Millard, P.H. A queueing model for bed-occupancy management and planning of hospitals. *J. Oper. Res. Soc.* **2002**, *53*, 19–24. [[CrossRef](#)]
30. Gournescu, F.; McClean, S.I.; Millard, P.H. Using a queueing model to help plan bed allocation in a department of geriatric medicine. *Health Care Manag. Sci.* **2002**, *5*, 307–312. [[CrossRef](#)]
31. Marshall, A.H.; McClean, S.I. Using Coxian phase-type distributions to identify patient characteristics for duration of stay in hospital. *Health Care Manag. Sci.* **2004**, *7*, 285–289. [[CrossRef](#)]
32. McClean, S.I.; Millard, P. Where to treat the older patient? Can Markov models help us better understand the relationship between hospital and community care. *J. Oper. Res. Soc.* **2007**, *58*, 255–261. [[CrossRef](#)]
33. Vassiliou, P.-C.G. Markov systems in a General State Space. *Commun. Stat.-Theory Methods* **2014**, *43*, 1322–1339. [[CrossRef](#)]
34. Meyn, S.; Tweedie, R.L. *Markov Chains and Stochastic Stability*; Cambridge University Press: Cambridge, UK, 2009.

35. Tsaklidis, G.M.; Vassiliou, P.-C.G. Asymptotic periodicity of the variances and covariances of the state sizes in non-homogeneous Markov systems. *J. Appl. Probab.* **1988**, *25*, 21–33. [[CrossRef](#)]
36. Georgiou, A.C.; Vassiliou, P.-C.G. Periodicity of asymptotically attainable structures in non-homogeneous Markov systems. *Linear Algebra Its Appl.* **1992**, *176*, 137–174. [[CrossRef](#)]
37. Doeblin, W. Sur les proprietes asymptotiques de mouvement regis par certain types de chaines simples. *Bull. Math. Soc. Roum. Sci.* **1937**, *39*, 57–115.
38. Doeblin, W. Elements d'une theorie general des chaines simple constantes de Markov. *Ann. Sci. l'Ecole. Norm. Super.* **1940**, *57*, 61–111. [[CrossRef](#)]
39. Doob, J.L. *Stochastic Processes*; John Wiley: New York, NY, USA, 1953.
40. Harris, T.E. The existence of stationary measures for certain Markov processes. In *Proceedings of the 3rd Berkeley Symposium on Mathematical Statistics and Probability*; University of California Press: Berkeley, CA, USA, 1956; Volume 2, pp. 113–124.
41. Orey, S. Recurrent Markov chains. *Pacific J. Math.* **1959**, *9*, 805–827. [[CrossRef](#)]
42. Tweedie, R.L. R-theory for Markov chains on a general state space I: Solidarity properties and R-recurrent chains. *Adv. Appl. Probab.* **1974**, *2*, 840–864. [[CrossRef](#)]
43. Tweedie, R.L. R-theory for Markov chains on a general state space II: R-subinvariant measures for R-transient chains. *Adv. Appl. Probab.* **1974**, *2*, 865–878. [[CrossRef](#)]
44. 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](#)]
45. Vassiliou, P.-C.G. Exotic properties of non-homogeneous Markov and semi-Markov systems. *Commun. Stat.-Theory Methods* **2013**, *42*, 2971–2990. [[CrossRef](#)]
46. Vassiliou, P.-C.G.; Tsaklidis, G.M. On the rate of convergence of the vector of variances and covariances in non-homogeneous Markov systems. *J. Appl. Probab.* **1989**, *26*, 776–783. [[CrossRef](#)]
47. Orey, S. *Limit Theorems for Markov Chain Transition Probabilities*; Van Nostrand Reinhold: London, UK, 1971.
48. Hajnal, J. The ergodic properties of non-homogeneous finite Markov chains. *Proc. Camb. Philos. Soci.* **1956**, *52*, 67–77. [[CrossRef](#)]
49. Hajnal, J. Weak ergodicity in non-homogeneous Markov chains. *Proc. Camb. Philos. Soc.* **1958**, *54*, 233–246. [[CrossRef](#)]
50. Hajnal, J. Shuffling with two matrices. *Contemp. Math.* **1993**, *149*, 271–287.
51. Dobrushin, R.L. Central limit theorem for non-stationary Markov chains, I, II. *Theory Probab. Its Appl.* **1956**, *1*, 65–80. [[CrossRef](#)]
52. Nummelin, E. *General Irreducible Markov chains and Nonnegative Operators*; Cambridge University Press: Cambridge, UK, 1984.
53. Howard, R.A. *Dynamic Probabilistic Systems*; Wiley: Chichester, UK, 1971; Volumes I and II.
54. Hunter, J.J. *Mathematical Techniques of Applied Probability*; Academic Press: Cambridge, MA, USA, 1983.
55. Vassiliou, P.-C.G. On the limiting behavior of a non-homogeneous Markov chain model in manpower systems. *Biometrika* **1981**, *68*, 557–561.
56. Papadopoulou, A.A.; Vassiliou, P.-C.G. Asymptotic behaviour of non-homogeneous Markov systems. *Linear Algebra Its Appl.* **1994**, *210*, 153–198. [[CrossRef](#)]
57. Vassiliou, P.-C.G. Onn the periodicity of non-homogeneous Markov chains and systems. *Linear Algebra Its Appl.* **2015**, *471*, 654–684. [[CrossRef](#)]
58. Lindvall, T. *Lectures on the Coupling Method*; John Wiley: Hoboken, NJ, USA, 1992.
59. Vassiliou, P.-C.G. Laws of Large numbers for non-homogeneous Markov systems. *Methodol. Comput. Appl. Probab.* **2018**, *1*–28. [[CrossRef](#)]

