Available online at www.sciencedirect.com

ScienceDirect

Mathematical Biosciences 206 (2007) 134–154

Mathematical Biosciences

www.elsevier.com/locate/mbs

Approximate reduction of multiregional models with environmental stochasticity

Luis Sanz [a,*], Rafael Bravo de la Parra [b]

[a] *Departamento de Matemáticas, E.T.S.I Industriales, Universidad Politécnica de Madrid, José Gutiérrez Abascal, 2, 28006 Madrid, Spain*
[b] *Departamento de Matemáticas, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, Spain*

Received 31 March 2004; received in revised form 9 December 2004; accepted 3 March 2005
Available online 7 November 2005

Abstract

In this work we extend previous results regarding the use of approximate aggregation techniques to simplify the study of discrete time models for populations that live in an environment that changes randomly with time. Approximate aggregation techniques allow one to transform a complex system involving many coupled variables and in which there are processes with different time scales, by a simpler reduced model with a fewer number of 'global' variables, in such a way that the dynamics of the former can be approximated by that of the latter.

We present the reduction of a stochastic multiregional model in which the population, structured by age and spatial location, lives in a random environment and in which migration is fast with respect to demography. However, the technique works in much more general settings as, for example, those of stage-structured populations living in a multipatch environment. By manipulating the original system and appropriately defining the global variables we obtain a simpler system.

The paper concentrates on establishing relationships between the original and the reduced systems for a given separation of time scales between the two processes. In particular, we relate the original state variables and the global variables and, in the case the pattern of temporal variation is Markovian, we relate the presence of strong stochastic ergodicity for the original and reduced systems. Moreover, we relate different measures of asymptotic population growth for these systems.

---

[*] Corresponding author. Fax: +34 913363001.
E-mail addresses: lsanz@etsii.upm.es (L. Sanz), rafael.bravo@uah.es (R. Bravo de la Parra).

0025-5564/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.mbs.2005.03.015

Keywords: Approximate aggregation; Environmental stochasticity; Time scales; Multiregional models; Ergodicity

1. Introduction

Nature offers many examples of systems with an inherent complexity whose study leads to mathematical models with a large number of state variables whose analytical study is, in most cases, not feasible. In order to be able to extract important information about the behavior of some of these complex models, one can resort to ‘approximate aggregation methods’. These are mathematical techniques, which are usually applied in systems governed by processes with different time scales, in which appropriate approximations are introduced in order to transform the system under consideration into a reduced system with a lesser number of variables, called ‘global variables’. In this way, the behavior of the original system can be approximated, but not known with exactitude, in terms of the knowledge of the behavior of the reduced system.

Approximate aggregation techniques have been widely studied in the context of time continuous systems with different time scales for both linear and density dependent models (see [1] for a list of references). The discrete time case, in which we will focus our attention in the sequel, has been thoroughly explored in linear, non-linear and non-autonomous deterministic contexts (see among others [2,14]).

Nowadays, a great part of the research in ecological modelling is devoted to models which incorporate stochasticity as a way of taking into account different factors which either have an intrinsic random nature or are too complex to be taken into account in a deterministic way. The literature contemplates two sources of stochasticity: demographic [3] and environmental. The reduction of models which incorporate demographic stochasticity has been addressed in Sanz et al. [17].

In this work we contemplate linear discrete models which incorporate ‘environmental stochasticity’, i.e., the randomness introduced when we consider random fluctuations in the environment and, consequently, in the vital rates which affect the population (see [3] and [21, Chapter 3] for an introduction, [20] for a full discussion and [9] for an application to actual fish populations). These models are analogous to the deterministic ones but in this case the matrix of vital rates in each projection interval is selected within a given set of matrices according to a certain (possibly time varying) probability distribution. The most relevant parameter in this kind of models is the so-called stochastic growth rate (s.g.r.); given certain hypotheses on the pattern of temporal variation and the vital rates in each environment, the s.g.r. is the stochastic analogue of the logarithm of the dominant eigenvalue for deterministic systems, which characterizes the asymptotic behavior of the total population.

In Sanz and Bravo [15], the authors deal with the reduction of discrete time systems for populations subjected to environmental stochasticity. A very general model with two time scales is presented and a technique for carrying out the reduction of the system is given. Regarding the relationships between the original and the reduced system, it is shown that the variables of the original system can be approximated in terms of those of the reduced system, and the approximation is exact when the separation of time scales tends to infinity. However, there are no results relating the s.g.r. of the two systems. Moreover, the relationships between the original and the aggregated systems in Ref. [15] are valid only in the limit when the separation of time scales tends to infinity.

The purpose of this paper is to extend the results of Ref. [15] in two directions. In the first place we deal with the general case in which there are no hypotheses imposed on the sequence of random variables that define the temporal variation. We give a way of approximating the variables of the original system in terms of those of the reduced system and give bounds of the error we make in terms of the value of parameter k that characterizes the separation of time scales between the two processes. In the second place, we relate several features of the two systems under the hypothesis that, as it is frequently assumed in stochastic modelling [3], the temporal variation for the original system is Markovian. Specifically, we begin relating the presence of strong stochastic ergodicity, a property by which the probability distribution for the population structure converges to a stationary probability distribution independent of initial conditions, for the original and reduced systems. Moreover, we relate two different measures of asymptotic population growth, among them the s.g.r., for these systems, both for finite k and in the limit when k tends to infinity. In all cases we give bounds for the error we make when we estimate the parameters of the original system in terms of their analogues in the aggregated system.

The results obtained, although valid in a very general context, have a special relevance in the particular case of stochastic multiregional models [8]. Indeed, in these models it is usually the case that migration among the spatial sites is fast with respect to reproduction or aging, and by using our aggregation technique we obtain a stochastic Leslie type age structured model whose study is much simpler. Now, the analytical study of general stochastic matrix models is very complex, but in the particular case of age structured models some aspects of that study simplify considerably [20] which makes our technique especially appropriate for this situation.

The structure of the paper is as follows; Section 2 briefly introduces the basic form of the matrix models that consider environmental stochasticity. Section 3, where ecologists interested mainly in applications may focus their attention, is devoted to presenting a stochastic multiregional model in which population is structured by age and patch and migration among the different patches is fast with respect to demography. By defining the global variables as the total population in each age class, we obtain a reduced stochastic Leslie model. The model, together with the aggregation procedure, is a particular case of a technique which was first presented by the authors in Ref. [15] and which is valid, for example, for any stage-structured population living in a multipatch environment.

Section 4, that constitutes the main contribution of this work, presents the above-mentioned relationships between the original system and the reduced system. This relationships allow one to approximate different features regarding the dynamics of the original stochastic system in terms of their analogues for the reduced stochastic system. Finally, Section 5 is a discussion on the application of the results obtained in this work and pointing out future lines of research.

2. Matrix models with environmental stochasticity

This section presents the basic form of the matrix models that consider environmental stochasticity. We restrict our attention to the case in which the number of environments is finite. We assume then that the population lives in an ambient in which there are s environmental states.

The vital rates corresponding to each one of these environments are given by the non-negative matrices $\mathbf{A}_\sigma \in \mathbb{R}^{N\times N}$, $\sigma = 1,\ldots,s$ in such a way that, for each σ, $\mathbf{A}_\sigma$ represents the vital rates of the population in environment σ. The environmental variation is characterized by a sequence of random variables τ_n, $n = 0,1,2,\ldots$ defined in a certain probability space $(\Omega,\mathscr{F},p)$ and with state space $\{1,\ldots,s\}$. For each realization $\omega \in \Omega$ of the process, the population is subjected to environmental conditions $\tau_{n+1}(\omega)$ during times n and $n+1$. In this way, the model reads

$$\mathbf{Z}_{n+1} = \mathbf{A}_{\tau_{n+1}}\mathbf{Z}_n, \tag{1}$$

where for each $n = 0,1,\ldots$, $\mathbf{Z}_n$ is a vector random variable in $\mathbb{R}^N$ which represents the population vector at time n. Throughout we will assume that $\mathbf{Z}_0$ is a fixed non-zero vector $\mathbf{Z}_0 \geqslant \mathbf{0}$. We will also consider the structure of the population vector

$$\mathbf{H}_n := \frac{\mathbf{Z}_n}{\|\mathbf{Z}_n\|}, \tag{2}$$

where $\|*\|$ denotes (unless specifically stated) the 1-norm in R^N, i.e. $\|\mathbf{z}\| = |z_1| + |z_2| + \cdots + |z_N|$.

3. Reduction of a stochastic multiregional model with fast migration

In this section we will deal with the reduction of a certain kind of stochastic multiregional models. The procedure is a particular case of that introduced by the authors in Ref. [15] to reduce general models which incorporate environmental stochasticity and in which there are two time scales. In this reference, a model and a reduction technique is presented both for the case in which demography is fast with respect to migration and the case in which the contrary happens. We now present the model and the aggregation procedure for the latter case, which is the more relevant for different reasons. In the first place, in most practical situations, migration is fast with respect to demography. In the second place, the aggregation procedure leads to a reduced system which is much simpler than the aggregated system obtained in the case in which demography is faster than migration, therefore simplifying the obtention of relationships between the original and the aggregated system that will be explored in the next section. And, finally, in Section 5 we will show why the general aggregation technique is specially useful in the context of this kind of models.

However, we have to stress that both the aggregation procedure and the results to relate the behavior of the original model and the reduced model are valid in a more general setting that we specify at the end of the section.

Multiregional models consider the dynamics of an age structured population distributed among different spatial patches among which they can migrate. These models have been used with profusion by Rogers [13] among others for the study of human populations. A list of ecological applications can be found in Refs. [3,12]. The usual approach has been deterministic, but the stochastic setting has also been used [8].

The above references do not explicitly consider the existence of different time scales in the multiregional system. The fact that in many situations migration and demography take place with different time scales (usually migration is fast with respect to demography although the contrary may also happen, see [11]) has been exploited in several works in order to reduce the complexity of the multiregional model in different contexts regarding the kind of model. Some examples in a linear deterministic setting are the works of the authors [14] (autonomous case) and [16] (non-autonomous case), while the non-linear case has been dealt with by Bravo et al. [2]. Regarding multiregional models with demographic stochasticity, we can point out the work of Sanz et al. [17]. The use of multiregional models to study specific instances of fish populations can be found in [4,5].

We think of a population living in an habitat in which there are s environmental conditions that influence the vital rates. The population is structured in q age classes (corresponding to groups) and spread out in different spatial patches (subgroups) among which they may migrate. We assume that individuals in each age class i may migrate among r spatial patches. Therefore, the total number of subgroups is $N := qr$ and the composition of the population is then given by vector $\mathbf{X}_n = (x_n^{11}, \ldots, x_n^{1r}, \ldots, x_n^{q1}, \ldots, x_n^{qr})^{\mathrm{T}} \in \mathbb{R}^N$ where x_n^{ij} is the number of individuals in age class i living in patch j at time n and T denotes transposition.

Demography and migration are responsible for the transference of individuals among the different stages, and we suppose that migration is a fast process in comparison with demography. Moreover, we choose as time step $\Delta_n = [n, n+1)$ for the model, the duration of each age class.

For each i and each environment $\sigma = 1, \ldots, s$, migration for individuals of age i is modeled by a matrix $\mathbf{P}_i(\sigma) \in \mathbb{R}^{r\times r}$ that, since migration is a conservative process for the total number of individuals, is stochastic. Besides we suppose that $\mathbf{P}_i(\sigma)$ is primitive for each i and each σ. This is the case, for example, if the fast process in each environment verifies: (a) transition from any patch to any other, in a sufficient number of steps, is allowed and (b) individuals of at least one patch are allowed to stay in that patch.

The demographic process in each environment σ is defined by the following generalized Leslie matrix:

$$\mathbf{M}_\sigma = \begin{bmatrix} \mathbf{F}_1(\sigma) & \mathbf{F}_2(\sigma) & \cdots & \mathbf{F}_{q-1}(\sigma) & \mathbf{F}_q(\sigma) \\ \mathbf{S}_1(\sigma) & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_2(\sigma) & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{S}_{q-1}(\sigma) & \mathbf{0} \end{bmatrix},$$

where

$$\begin{aligned} \mathbf{F}_i(\sigma) &:= \operatorname{diag}(F_i^1(\sigma), \ldots, F_i^r(\sigma)), \quad i = 1, \ldots, q, \\ \mathbf{S}_i(\sigma) &:= \operatorname{diag}(S_i^1(\sigma), \ldots, S_i^r(\sigma)), \quad i = 1, \ldots, q-1 \end{aligned} \tag{3}$$

and the coefficients $F_i^j(\sigma)$ and $S_i^j(\sigma)$ denote, respectively, the fertility and survival rates for individuals of age i in patch j in environment σ.

In order to approximate the effect of migration over the time step of the model, which is much longer than its corresponding projection interval, we assume that if the population is subjected to environment σ during Δ_n matrix $\mathbf{P}_\sigma$ operates a number k of times, where k (that we assume to be an integer) can be interpreted as the ratio between the projection intervals corresponding to demography and migration. Thus, the set of vital rates for our system in the different environments is

$$\mathscr{A} := \{\mathbf{M}_1\mathbf{P}_1^k, \mathbf{M}_2\mathbf{P}_2^k, \ldots, \mathbf{M}_s\mathbf{P}_s^k\}.$$

The pattern of environmental variation is defined by a sequence of random variables τ_n, $n = 1, 2, \ldots$ which take values in the set of environmental states $\{1, \ldots, s\}$. Therefore, the proposed model, to which we will refer as 'microsystem' or 'original system', consists in the following system of N random difference equations:

$$\mathbf{X}_{n+1} = \mathbf{M}_{\tau_{n+1}} \mathbf{P}^k_{\tau_{n+1}} \mathbf{X}_n. \tag{4}$$

Now, following Ref. [15], we make use of approximate aggregation to reduce the original system (4), consisting of N variables (microvariables) associated to the different subgroups, by an aggregated system of q variables (global variables), each of them associated to one group.

Let $\mathbf{v}_i(\sigma)$ be the right Perron eigenvector of $\mathbf{P}_i(\sigma)$, i.e., the eigenvector uniquely defined by the conditions $\mathbf{P}_i(\sigma)\mathbf{v}_i(\sigma) = \mathbf{v}_i(\sigma)$, $\mathbf{v}_i(\sigma) > 0$, $\mathbf{1}^{\mathrm{T}}\mathbf{v}_i(\sigma) = 1$ where $\mathbf{1} = (1, \overset{(r)}{\ldots}, 1)^{\mathrm{T}}$. $\mathbf{v}_i(\sigma)$ can be interpreted as that which would characterize the equilibrium structure of migration for individuals in age class i if the population were constantly subjected to environment σ.

In this way, the matrix that characterizes the conditions for migration in environment σ for group i is

$$\overline{\mathbf{P}}_i(\sigma) = \lim_{k\to\infty} \mathbf{P}^k_i(\sigma) = \mathbf{v}_i(\sigma)\mathbf{1}^{\mathrm{T}} > 0, \quad \sigma = 1, \ldots, s, \ \ i = 1, \ldots, q \tag{5}$$

and, for the total population we have matrix

$$\overline{\mathbf{P}}_\sigma := \operatorname{diag}(\overline{\mathbf{P}}_1(\sigma), \overline{\mathbf{P}}_2(\sigma), \ldots, \overline{\mathbf{P}}_q(\sigma)). \tag{6}$$

Now we define matrices

$$\begin{aligned} &\mathbf{V}_\sigma := \operatorname{diag}(\mathbf{v}_1(\sigma), \mathbf{v}_2(\sigma), \ldots, \mathbf{v}_q(\sigma)), \quad \sigma = 1, \ldots, s, \\ &\mathbf{U} := \operatorname{diag}(\mathbf{1}^{\mathrm{T}}, \mathbf{1}^{\mathrm{T}}, \ldots, \mathbf{1}^{\mathrm{T}}). \end{aligned} \tag{7}$$

Some of the properties of these matrices are gathered in the following lemma, whose proof is straightforward:

Lemma 1. *For all* $\sigma = 1, \ldots, s$, *matrices* $\mathbf{P}_\sigma, \overline{\mathbf{P}}_\sigma, \mathbf{V}_\sigma$ *and* $\mathbf{U}$ *verify*:

(*a*) $\mathbf{P}_\sigma\overline{\mathbf{P}}_\sigma = \overline{\mathbf{P}}_\sigma\mathbf{P}_\sigma = \overline{\mathbf{P}}_\sigma$,
(*b*) $\mathbf{P}_\sigma\mathbf{V}_\sigma = \mathbf{V}_\sigma$,
(*c*) $\mathbf{U}\overline{\mathbf{P}}_\sigma = \mathbf{U}$; $\mathbf{U}\mathbf{V}_\sigma = \mathbf{I}_q$; $\overline{\mathbf{P}}_\sigma = \mathbf{V}_\sigma\mathbf{U}$.

Now we introduce the 'auxiliary system' as the stochastic model defined by

$$\mathbf{X}'_{n+1} = \mathbf{M}_{\tau_{n+1}} \overline{\mathbf{P}}_{\tau_{n+1}} \mathbf{X}'_n. \tag{8}$$

Note that this system can be interpreted as the result of letting migration reach equilibrium in the original system.

For each age class i we define a global variable y^i_n corresponding to the total population of the auxiliary system with age i, i.e.,

$$y^i_n = x'^{i1}_n + x'^{i2}_n + \cdots + x'^{ir}_n, \quad i = 1, \ldots, q \tag{9}$$

and then we define the vector $\mathbf{Y}_n$ of global variables

$$\mathbf{Y}_n = (y_n^1, \dots, y_n^q)^{\mathrm{T}} := \mathbf{U}\mathbf{X}'_n \in \mathbb{R}^q \tag{10}$$

If we multiply both sides of (8) by $\mathbf{U}$ and make use of Lemma 1 we have

$$\mathbf{U}\mathbf{X}'_{n+1} = \mathbf{U}\mathbf{M}_{\tau_{n+1}}\overline{\mathbf{P}}_{\tau_{n+1}}\mathbf{X}'_n = \mathbf{U}\mathbf{M}_{\tau_{n+1}}\mathbf{V}_{\tau_{n+1}}\mathbf{U}\mathbf{X}'_n, \tag{11}$$

which can be written in terms of the global variables exclusively, i.e., we have an aggregated system defined by

$$\mathbf{Y}_{n+1} = \overline{\mathbf{M}}_{\tau_{n+1}}\mathbf{Y}_n, \tag{12}$$

where, for each n, $\overline{\mathbf{M}}_{\tau_n}$ is given by

$$\overline{\mathbf{M}}_{\tau_n} := \mathbf{U}\mathbf{M}_{\tau_n}\mathbf{V}_{\tau_n}.$$

The aggregated system (12) can be interpreted as a stochastic model in which the pattern of environmental variation coincides with that of the original system and in which the matrix of vital rates in each environment σ is

$$\overline{\mathbf{M}}_\sigma := \mathbf{U}\mathbf{M}_\sigma\mathbf{V}_\sigma \in \mathbb{R}^{q\times q}, \quad \sigma = 1, \dots, s \tag{13}$$

in such a way that the set of matrices for the different environmental conditions is

$$\mathscr{A}_{\mathrm{ag}} := \{\overline{\mathbf{M}}_1, \overline{\mathbf{M}}_2, \dots, \overline{\mathbf{M}}_s\}. \tag{14}$$

Note that, for each $\sigma = 1,\dots,s$, $\overline{\mathbf{M}}_\sigma$ is a classical Leslie matrix given by

$$\overline{\mathbf{M}}_\sigma = \begin{bmatrix} f_1(\sigma) & f_2(\sigma) & \cdots & f_{q-1}(\sigma) & f_q(\sigma) \\ s_1(\sigma) & 0 & \cdots & 0 & 0 \\ 0 & s_2(\sigma) & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & s_{q-1}(\sigma) & 0 \end{bmatrix}, \tag{15}$$

where the vital rates have the form

$$f_i(\sigma) = \mathbf{1}^{\mathrm{T}}\mathbf{F}_i(\sigma)\mathbf{v}_i(\sigma) = \sum_{j=1}^{r} v_i^j(\sigma)F_i^j(\sigma), \quad i = 1, \dots, q, \ \ \sigma = 1, \dots, s,$$

$$s_i(\sigma) = \mathbf{1}^{\mathrm{T}}\mathbf{S}_i(\sigma)\mathbf{v}_i(\sigma) = \sum_{j=1}^{r} v_i^j(\sigma)S_i^j(\sigma), \quad i = 1, \dots, q-1, \ \ \sigma = 1, \dots, s$$

i.e., each fertility rate $f_i(\sigma)$ in the aggregated system is a weighted linear combination of the fertility rates in the general system corresponding to individuals of age class i in environment σ, being the weights the coefficients of the equilibrium spatial distribution for migration in environment σ. Something analogous holds for the survival rates.

The original multiregional model has been transformed into a reduced system in which the spatial distribution has been averaged in a certain way and the population appears structured only by age.

As we mentioned above, both the reduction procedure we presented and the relationships between the original and the reduced systems, are valid for a more general kind of models subjected to the effects of environmental stochasticity. Specifically, the results are valid for the general kind of models governed with different time scales presented in Ref. [15], whenever hypothesis (D) in Section 4.1 holds. In particular, this includes the following setting:

(a) The population is divided into q groups (age classes in model (4)) attending to any characteristic of the life cycle, and each group $i = 1, 2, \ldots, q$ is itself split into N_i subgroups (spatial patches in model (4)). Then, x_n^{ij} is the number of individuals in group i and subgroup j at time n and the population vector is $\mathbf{X}_n = (x_n^{11}, \ldots, x_n^{1N_1}, \ldots, x_n^{q1}, \ldots, x_n^{qN_q})^{\mathrm{T}} \in \mathbb{R}^N$ where $N := N_1 + \cdots + N_q$.

(b) The projection interval of the model is that corresponding to the slow dynamics, on which we impose no special assumptions. Therefore, for any environmental condition $\sigma = 1, \ldots, s$ the slow process will be modeled by a non-negative projection matrix $\mathbf{M}_\sigma \in \mathbb{R}^{N \times N}$, which can be considered as divided into blocks $\mathbf{M}_{ij}(\sigma)$, $1 \leqslant i, j \leqslant q$ of dimensions $N_i \times N_j$ that characterize the rates of transference of individuals from the subgroups of group j to the subgroups of group i in environment σ.

(c) Regarding the fast process:
  (c.1) The fast dynamics is an internal process for each group, i.e., there is no transference of individuals from one group to a different one.
  (c.2) For each i and σ, matrix $\mathbf{P}_i(\sigma) \in \mathbb{R}^{N_i \times N_i}$ is a column stochastic primitive matrix. In particular, the fast process is conservative of the total number of individuals in each group.

This setting allows one to model a population structured by any characteristic of its life cycle (stage-structured models, see [3] for comments and references) and living in a multipatch environment with migration among the different patches. A particular case of this is the case of spatially heterogeneous size-structured models.

We have to stress that Lemma 1, and therefore the aggregation procedure described, remains valid under the assumptions above.

4. Relationships between the original system and the aggregated system

In Sanz and Bravo [15], the authors relate some features regarding the behavior of systems (4) and (12). Specifically, in the first place the paper relates the variables $\mathbf{X}_n$ and $\mathbf{Y}_n$ of the two systems for finite values of n when the parameter k tends to infinity. In the second place, and under the hypothesis that τ_n is a Markov chain, Ref. [15] deals with the relationship between the statistical moments of any order of the population vector of the two systems when k tends to infinity.

This section extends those results in two different situations. In the first place we deal with the general case in which there are no hypothesis imposed on the sequence τ_n and relate the variables of the original and the aggregated system for finite values of the parameter k. Secondly, we deal with the particular case in which τ_n is an homogeneous Markov chain, relating different asymptotic properties of the two systems as ergodicity and measures of population growth.

All the results presented are valid for the general setting described at the end of the previous section.

4.1. General temporal variation

The purpose of this section is to provide a bound for the error we make when we estimate the total population of the original system in terms of that of the aggregated system, for any values of n and k. To start with, let us partially state some features of a proposition (6.2) in [15] which allows one to relate the population vector of systems (4) and (12) in the limit when $k \to \infty$. In the first place the relationship between the variables corresponding to the aggregated system and to the auxiliary system is given by

$$\mathbf{X}'_n = \mathbf{M}_{\tau_n}\mathbf{V}_{\tau_n}\mathbf{Y}_{n-1}; \quad \mathbf{Y}_n = \mathbf{U}\mathbf{X}'_n, \tag{16}$$

meanwhile the variables $\mathbf{X}_n$ of the original system can be approximated by those of the aggregated system $\mathbf{Y}_n$ and reciprocally, in the following way:

$$\begin{aligned}&\mathbf{X}_n = \mathbf{M}_{\tau_n}\mathbf{V}_{\tau_n}\mathbf{Y}_{n-1} + \mathbf{o}(\delta^k); \quad k \to \infty,\\ &\mathbf{Y}_n = \mathbf{U}\mathbf{X}_n + \mathbf{o}(\delta^k); \quad k \to \infty,\end{aligned} \tag{17}$$

where $\delta < 1$ is a parameter which has to do with the 'subdominant' eigenvalues of matrices $\mathbf{P}_\sigma$ (see [15] for details).

The previous relationships show that we can obtain the population vector of the auxiliary system exactly in terms of that of the aggregated system and reciprocally. In addition, for each realization of the process, the variables $\mathbf{X}_n(\omega)$ of the original system can be approximated in terms of those corresponding to the aggregated system through $\mathbf{M}_{\tau_n(\omega)}\mathbf{V}_{\tau_n(\omega)}\mathbf{Y}_{n-1}(\omega)$, and for fixed n and big enough k, the discrepancy decays geometrically with k. We will refer to that discrepancy as the 'error' we make when carrying out the study of the original system in terms of that of the reduced system.

We are now interested in obtaining a bound for the error as a function of k and n. Specifically, we are interested in the total population size $\|\mathbf{X}_n\|$ and, since the systems under consideration are multiplicative, we will work in a logarithmic scale. Therefore, for each $\omega \in \Omega$ and each initial condition $\mathbf{X}_0 \geqslant 0$ we will consider the error

$$E_{n,k}(\omega) := \left|\log\|\mathbf{X}_n(\omega)\| - \log\|\mathbf{X}'_n(\omega)\|\right| = \left|\log\|\mathbf{X}_n(\omega)\| - \log\|\mathbf{M}_{\tau_n(\omega)}\mathbf{V}_{\tau_n(\omega)}\mathbf{Y}_{n-1}(\omega)\|\right|.$$

For notational convenience, let us define matrices

$$\mathbf{C}_\sigma = [c^\sigma_{\alpha\beta}] := \mathbf{M}_\sigma\overline{\mathbf{P}}_\sigma \geqslant 0; \quad \mathbf{D}_{\sigma,k} = [d^{\sigma,k}_{\alpha\beta}] := \mathbf{M}_\sigma\left(\mathbf{P}^k_\sigma - \overline{\mathbf{P}}_\sigma\right), \quad \sigma = 1,\ldots,s$$

and now let matrix $\mathbf{W}_{\sigma,k} := [w^{\sigma,k}_{\alpha\beta}],\ \sigma = 1,\ldots,s$ be given by

$$w^{\sigma,k}_{\alpha\beta} := \begin{cases} \dfrac{d^{\sigma,k}_{\alpha\beta}}{c^\sigma_{\alpha\beta}} & \text{if } c^\sigma_{\alpha\beta} \neq 0,\\ 0 & \text{if } c^\sigma_{\alpha\beta} = 0\end{cases} \tag{18}$$

and let numbers $\xi_M(k)$, $\xi_m(k)$ be defined as follows:

$$\xi_M(k) := \max_{\sigma\in\{1,\ldots,s\}}(\max \mathbf{W}_{\sigma,k}); \quad \xi_m(k) := -\min_{\sigma\in\{1,\ldots,s\}}(\min \mathbf{W}_{\sigma,k}), \tag{19}$$

where, given a matrix **A**, max **A** and min **A** denote, respectively, the maximum and minimum of the entries of **A**.

Lemma 10 (Appendix A) shows some properties of $\xi_M(k)$ and $\xi_m(k)$, among them the fact that they are positive and tend to zero when $k \to \infty$.

The next proposition provides a bound $C_{n,k}$ for the error $E_{n,k}(\omega)$:

Proposition 2. *For all k such that* $\xi_m(k) < 1$ *we have that for all n*

$$\forall w \in \Omega \ \forall \mathbf{X}_0 \geqslant 0, \quad E_{n,k}(\omega) \leqslant C_{n,k} := nB_k, \tag{20}$$

where

$$B_k := \max\left\{\log(1+\xi_M(k)), \log\left(\frac{1}{1-\xi_m(k)}\right)\right\} \tag{21}$$

Proof. See Appendix A. □

Note that the bound $C_{n,k}$ grows linearly with time and tends to zero when $k \to \infty$. In order to obtain $C_{n,k}$ we must compute matrices $\mathbf{W}_{\sigma,k}$ and therefore it is necessary to calculate the powers $\mathbf{P}_\sigma^k$ for all σ. The following result gives a coarser bound $C_{n,k}^a$ that can be obtained more easily, therefore allowing one to have a quick idea of the order of magnitude of the error:

Proposition 3. *Let* $v(k)$ *be defined by*

$$v(k) := \max_{\sigma \in \{1,\ldots,s\}} \left\{\frac{\max \mathbf{M}_\sigma}{\min^+(\mathbf{M}_\sigma \overline{\mathbf{P}}_\sigma)} \|\mathbf{P}_\sigma - \overline{\mathbf{P}}_\sigma\|^k\right\},$$

where $\min^+$ *denotes the minimum of the positive entries of the corresponding matrix and* $\|*\|$ *denotes the* 1*-matrix norm.*

Then if $v(k) < 1$ *we have*

$$\forall \omega \in \Omega \ \forall \mathbf{X}_0 \geqslant 0, \quad E_{n,k}(\omega) \leqslant C_{n,k}^a := n \max\left\{\log(1+v(k)), \ \log\left(\frac{1}{1-v(k)}\right)\right\}. \tag{22}$$

Proof. See Appendix A. □

Let us now contemplate a particular case of the general situation and for which we will obtain bounds that will be both sharper and easier to calculate.

Definition 4. We will say that hypothesis (H) holds when matrices $\mathbf{M}_{ij}(\sigma)$ are square and diagonal for all $i,j \in \{1,\ldots,q\}$ and all σ.

Note that condition (H) holds in the multiregional model of Section 3.

Now, given condition (H), we can obtain matrices $\mathbf{W}_{\sigma,k}$ and bounds $C_{n,k}$ and $C_{n,k}^a$ in a simpler way. Indeed we have:

Lemma 5. *Let us assume hypothesis* (*H*):

(*a*) *Let* $\sigma \in \{1,\ldots,s\}$ *and k be fixed. Matrix* $\mathbf{W}_{\sigma,k}$ *can be obtained as*

$$w_{\alpha\beta}^{\sigma,k} = \begin{cases} \dfrac{\left(\mathbf{P}_\sigma^k - \overline{\mathbf{P}}_\sigma\right)_{\alpha\beta}}{\left(\overline{\mathbf{P}}_\sigma\right)_{\alpha\beta}} & \textit{if } c_{\alpha\beta}^\sigma \neq 0, \\ 0 & \textit{if } c_{\alpha\beta}^\sigma = 0. \end{cases} \tag{23}$$

In particular, matrix $\mathbf{W}_{\sigma,k}$ *depends only on* $\mathbf{P}_\sigma$ *and on the pattern of non-zero elements of* $\mathbf{M}_\sigma$*, i.e.,* $\mathbf{W}_{\sigma,k}$ *is independent of the value of the non-zero elements of* $\mathbf{M}_\sigma$*.*

(b) Let $v_H(k)$ *be defined by*

$$v_H(k) := \max_{\sigma\in\{1,\ldots,s\}} \left\{ \frac{\left\|\mathbf{P}_\sigma - \overline{\mathbf{P}}_\sigma\right\|^k}{\min^+ \overline{\mathbf{P}}_\sigma} \right\}, \tag{24}$$

which verifies $v_H(k) \leqslant v(k)$*. Therefore, if* $v_H(k) < 1$ *bound* $C_{n,k}^a$ *can be improved by replacing* $v(k)$ *by* $v_H(k)$ *in (22).*

Proof. See Appendix A. □

Note that, given hypothesis (H), $C_{n,k}$ is independent of the value of the non-zero entries of the matrices defining the slow dynamics.

4.2. Markovian environment

So far, the pattern of environmental variation defined by τ_n did not have to meet any special requirements. Subsequently we will explore the relationships between the original system and the aggregated system in the case that the pattern of temporal variation for the former is an homogeneous Markov chain. Indeed, a great part of the models of the kind (1) found in the literature represent environmental change through Markov chains (see among others the works of [7,3,20]), for Markov chains can represent sequential dependence between environments and yet are simple enough to be analyzed with great detail. Besides, the systems under Markovian temporal variation usually have some desirable properties regarding their asymptotic behavior. Indeed, an appropriate Markovian temporal variation combined with certain conditions on the set of environmental matrices guarantee in one hand the presence of strong stochastic ergodicity, i.e., the probability distribution of the population structure converges to a stationary probability distribution independent of initial conditions [7] and, on the other hand, that almost all realization of the process ultimately grows or decays exponentially with the same growth rate. This approach contemplates, as a particular case, the possibility of the τ_n being independent and having the same probability distribution (i.i.d. case), which is an usual choice for the characterization of an uncorrelated temporal variation.

This section will relate some features regarding the asymptotic behavior of the total population for the two systems under the above-mentioned hypothesis that τ_n is a Markov chain.

In the first place let us recall some features regarding the asymptotic behavior of general matrix models with environmental stochasticity of the kind (1). In order to do so, we introduce the following concept:

Definition 6. A finite set $\mathcal{S} = \{\mathbf{A}_1, \ldots, \mathbf{A}_s\}$ of square non-negative matrices is said to be ergodic if there exists a positive integer r such that any product of r matrices (with repetitions allowed) drawn from $\mathcal{S}$ is a positive matrix.

In particular, any finite set of matrices which share a common primitive incidence matrix is an ergodic set. More generally, if the condition $\mathbf{A}_\sigma \geqslant \mathbf{B}$, $\sigma = 1, \ldots, s$ with $\mathbf{B}$ a primitive matrix holds, then $\mathcal{S}$ is an ergodic set.

The following theorem, which is a compilation of results extracted from Refs. [6,18,20], deals with some important features regarding the asymptotic behavior of systems of the kind (1) and motivates the kind of results which will be obtained in this section.

Theorem 7. *Let us consider system (1):*

(*a*) *Let τ_n be an homogenous Markov chain with initial and transition probabilities given by vector $\mathbf{q}$ and matrix $\mathbf{Q}$ respectively, where $q^i = p(\tau_1 = i)$ and $q_{ij} = p(\tau_{n+1} = i \,|\tau_n = j)$, $i,j = 1, \ldots, s$. Then, the logarithm of the mean population size for system (1) grows asymptotically with a rate $\log \mu$ where μ is the spectral radius of the non-negative matrix $\mathbf{D} = \mathrm{diag}(\mathbf{A}_1, \ldots, \mathbf{A}_s)(\mathbf{Q} \otimes \mathbf{I}_N)$ ($\otimes$ denotes the Kronecker matrix product).*

(*b*) *Moreover, assume that (b.1) the chain is irreducible and aperiodic, i.e., the matrix $\mathbf{Q}$ of transition probabilities is primitive, and (b.2) that the set $\{\mathbf{A}_1, \ldots, \mathbf{A}_s\}$ of vital rates is an ergodic set. Then we have:*

(1) *System (1) is stochastically strongly ergodic in the sense that there exists a stationary probability distribution F for the joint distribution of current age structure $\mathbf{H}_n$ and environment τ_n which is independent of the initial probabilities of the chain and of the environment and the population structure at any given time (see [6] for details).*

(2) *We can define the stochastic growth rate (s.g.r.) of system (1) through*

$$a := \lim_{n\to\infty} \frac{\log \|\mathbf{Z}_n\|}{n} \quad (a.s.), \tag{25}$$

where the limit is in the sense of almost sure convergence. Moreover, a is independent of the initial probabilities of the chain and of the initial (non-zero) population vector $\mathbf{Z}_0 \geqslant \mathbf{0}$. Parameter a can also be computed through

$$a = \lim_{n\to\infty} \mathbb{E}_F[\log \|\mathbf{A}_{\tau_{n+1}}\mathbf{H}_n\|], \tag{26}$$

where the average is taken with respect to the stationary distribution F.

(3) *Parameters a and μ verify $a \leqslant \log \mu$.*

From the above we have that, asymptotically, for almost every realization of the process the total population size grows exponentially with a rate e^a meanwhile the mean population size grows exponentially with a rate μ.

It is important to note that, both from a theoretical and from an applied point of view, the s.g.r. is the most important parameter to study the dynamics of matrix models with environmental stochasticity.

Our aim is, assuming that τ_n is an homogeneous Markov chain, to find conditions under which we can guarantee that the ergodicity of the reduced system (12) implies that of the original system (4). Therefore, in the sequel we will assume:

*H*1. *Let the process* τ_n *which characterizes the environmental change for the microsystem be an homogeneous Markov chain with transition and initial probabilities defined by matrix* $\mathbf{Q} := [q_{ij}] \in \mathbb{R}^{s\times s}$ *and the probability normed vector* $q := (q^1, \ldots, q^s)^{\mathrm{T}}$, *respectively.*

We have then

$$\begin{aligned} &q_{ij} = p(\tau_{n+1} = i | \tau_n = j), \quad i,j = 1,\ldots,s, \ \ n = 1,2,\ldots \\ &q^i = p(\tau_1 = i), \quad i = 1,\ldots,s. \end{aligned} \tag{27}$$

Let us introduce some concepts and notation which will be useful in the subsequent developments. A non-negative matrix **A** is said to be column allowable (row allowable) if it has at least a non-zero element in each one of its columns (rows). **A** is said to be allowable if it is both column and row allowable.

Proposition 8.

(*a*) *If the set* $\mathscr{A} = \{\mathbf{M}_1\mathbf{P}_1^k, \ldots, \mathbf{M}_s\mathbf{P}_s^k\}$ *is ergodic for big enough k then* $\mathscr{A}_{ag} = \{\overline{\mathbf{M}}_1, \ldots, \overline{\mathbf{M}}_s\}$ *is ergodic.*

(*b*) *If the set* $\mathscr{A}_{ag}$ *is ergodic, and the matrices* $\mathbf{M}_1, \ldots, \mathbf{M}_s$ *are row-allowable, then the set* $\mathscr{A}$ *is also ergodic for big enough k . Therefore, if the aggregated system meets the sufficient conditions of Theorem 7 for the existence of stochastic strong ergodicity and a s.g.r., and matrices* $\mathbf{M}_1, \ldots, \mathbf{M}_s$ *are row-allowable, then the original system will also meet those sufficient conditions for big enough k.*

Proof. See Appendix A. □

Let us now turn our attention to the relationships between the asymptotic behavior of the total population size for the original system and the aggregated system.

Theorem 9. *Let hypothesis H*1 *hold*:

(*a*) *Let* μ_k *and* μ *denote the asymptotic growth rate of the logarithm of the mean population size for the original and the aggregated system respectively, i.e.,* μ_k *and* μ *are the spectral radii of matrices* $\mathbf{D}_k := \mathrm{diag}(\mathbf{M}_1\mathbf{P}_1^k, \ldots, \mathbf{M}_s\mathbf{P}_s^k)(\mathbf{Q}\otimes\mathbf{I}_N)$ *and* $\overline{\mathbf{D}} := \mathrm{diag}(\overline{\mathbf{M}}_1, \ldots, \overline{\mathbf{M}}_s)(\mathbf{Q}\otimes\mathbf{I}_q)$ *respectively. Then*

$$\ln(1 - \xi_m(k)) \leqslant \log\mu_k - \log\mu \leqslant \ln(1 + \xi_M(k))$$

and so in particular $\lim_{k\to\infty}\log\mu_k = \log\mu$.

(*b*) *Let the chain* τ_n *be irreducible and aperiodic and let the sets* $\mathscr{A}$ *and* $\mathscr{A}_{ag}$ *be ergodic for big enough k so that both the original and the aggregated system verify conditions* (*b*.1) *and* (*b*.2) *of Theorem* 7. *Let* a_k *and a denote the respective s.g.r. of the original and aggregated systems. Then*

$$\ln(1 - \xi_m(k)) \leqslant a_k - a \leqslant \ln(1 + \xi_M(k)) \tag{28}$$

and so in particular $\lim_{k\to\infty} a_k = a$.

Proof. See Appendix A. □

4.3. An example of the application of the results

Now let us illustrate with an specific example the application of the previous results in a particular situation to deduce properties from the multiregional stochastic model (4) in terms of the reduced Leslie stochastic model (12). For a general kind of environmental variation τ_n, the variables of the multiregional model can be approximated in terms of those of the reduced model using (17), and the error we make can be bounded by Proposition 2 and Lemma 5. Now let us assume that the temporal variation is characterized by a homogeneous Markov chain with transition matrix $\mathbf{Q}$, in such that for any environmental condition σ there is a positive probability of staying in that environment and, moreover, there is a positive probability of going from environment σ to any other environment in a sufficiently high number of time steps. This condition ensures that matrix $\mathbf{Q}$ is primitive.

Moreover, and regarding the vital rates let us assume: (i) The fertilities and survival coefficients (3) can vary with the environment, but if they are non-zero in one environment they are non-zero in all other environments (i.e., the incidence matrices of the $\mathbf{M}_\sigma$ are independent of σ). (ii) $S_i^j(\sigma) > 0$ for all i and j and for all $j = 1, \ldots, r$ there exists i such that $F_i^j(\sigma) > 0$, i.e., all survival coefficients are non-zero and in every patch there is at least an age class which is reproductive.

It is immediate to check that the above conditions ensure that the environmental matrices $\overline{\mathbf{M}}_\sigma$ in the reduced system are all primitive and they all share the same incidence matrix. Therefore they constitute an ergodic set and so the hypothesis of Theorem 7 are met, so the reduced stochastic Leslie model is (stochastically) strongly ergodic and has a s.g.r. a. Since hypothesis (ii) ensures that all $\mathbf{M}_\sigma$ are row-allowable we have from Proposition 8 and Theorem 9 that the original multiregional model is strongly ergodic, that its s.g.r. a_k can be approximated in terms of a and that expression (28) gives a measure of the error we make. Analogous ideas apply to the parameter μ corresponding to the growth rate of the average population.

5. Discussion

This work provides new results that relate the dynamics of complex models with different time scales and subjected to the effects of environmental stochasticity, and their corresponding reduced models. In principle these models can be of a very general nature, including among them any spatially heterogeneous stage-structured matrix model. However, for reasons that will make clear in the next paragraphs, emphasis has been given to stochastic multiregional models in which migration is fast with respect to demography.

The results obtained include, in the first place, the derivation of explicit bounds for the error we make when we approximate the variables of the complex system in term of those of the reduced system. Moreover, in the case the environment evolves according to a Markov chain, we give sufficient conditions that guarantee that the original complex system is stochastically strongly ergodic and has a stochastic growth rate (s.g.r.) just by studying the reduced system.

Taking into account the importance of the s.g.r. of a stochastic system to characterize its dynamics [3,21], the following comments address the advantages of our results to obtain the s.g.r. of complex systems with different time scales both for the general kind of models considered and for the case of stochastic multiregional models with fast migration.

Except in very specific cases, it is very difficult to obtain analytically the s.g.r. of a system of the kind (1). Indeed, the s.g.r. can be in principle calculated through expression (26), but in general it is not possible to obtain analytically the stationary distribution F and so one must resort to computer simulations using the expression (25) [18]. Our results allow the modeler to reduce the dimension of the system and simplify the numerical simulation to approximate the s.g.r. of the system.

Moreover, in the case of stochastic multiregional models with fast migration, which for example have been used to model fish populations [4,5], our results are more relevant. Indeed, the aggregation technique presented allows one to reduce a complex multiregional model, difficult to analyze, to construct a stochastic Leslie model for which a number of special techniques are available. Indeed, following Tuljapurkar [20], we have that in the case of a stochastic Leslie model the general expression (26) to calculate the s.g.r. simplifies to

$$a = E_T \log s_1(\tau_n) - E_G \log \frac{y_n^2}{y_n^1}, \tag{29}$$

where the expectations are taken with respect to the distributions T and G that denote, respectively, the stationary equilibrium distribution for the environmental chain τ_n and for the age structure $\mathbf{H}_n$ derived from the stationary distribution F. Now, T can always be calculated from matrix $\mathbf{Q}$ and G can be explicitly calculated in some particular cases when there are only two age classes (see [20, Chapter 8] and [19]). Now, for example, if we consider a population with two age classes and distributed among different spatial patches, our aggregation procedure allows one to reduce the original complex model to obtain a reduced Leslie model with two age classes which can be treated with the techniques of Refs. [19,20] to obtain its s.g.r. a explicitly. Since our results guarantee that the s.g.r. of the complex model can be approximated by a, we can obtain analytical conclusions for the complex model, for example studying how changes in the migration rates affect the fate of the population, an area to which the authors will direct their attention in future research.

Even in the case of Leslie models with a general number of age classes, for which the stationary distribution F cannot be computed explicitly, (29) allows one to obtain a quick estimate of a by using some census information and some statistics on the sequence of environments [20]. Therefore, in the case of stochastic multiregional models with a general number of age classes, our aggregation technique allows one to obtain a reduced stochastic Leslie model, and then use (29) to estimate it.

Leaving aside the calculation of the s.g.r., the fact that for stochastic Leslie models it is possible to obtain, in a relatively easy way, bounds for the vector of age structure of the population (see [20, Chapter 9]), constitutes another advantage of our reduction method.

Therefore, our aggregation procedure, although valid in a very general context, has a special relevance in the particular case of stochastic multiregional models with fast migration, for it allows one to reduce a complex system for which analytical study is intractable in practical terms to a much simpler stochastic Leslie model for which there are a number of efficient tools at our disposal.

Acknowledgment

This work was supported by Proyecto de Investigación BFM2002-01615 (MCYT).

Appendix A

Lemma 10. *Let* $\sigma \in \{1,\ldots,s\}$.

(*a*) *If* $\mathbf{D}_{\sigma,k} \neq \mathbf{0}$ *then* $\mathbf{W}_{\sigma,k}$ *does not have a 'definite sign', i.e., neither* $\mathbf{W}_{\sigma,k}$ *nor* $-\mathbf{W}_{\sigma,k}$ *are non-negative matrices. Therefore, excluding the trivial case* $\mathbf{P}_\sigma^k = \overline{\mathbf{P}}_\sigma$ *for all* σ, *both* $\xi_M(k)$ *and* $\xi_m(k)$ *are positive numbers.*

(*b*) $\xi_M(k) = \mathrm{o}(\delta^k)$ *and* $\xi_m(k) = \mathrm{o}(\delta^k)$ *where* $\delta < 1$ *corresponds to the parameter* α *defined by Eq.* (6.1) *in* [15]. *Therefore* $\xi_M(k)$ *and* $\xi_m(k)$ *can be made arbitrarily close to zero by taking* k *big enough.*

(*c*) $\xi_m(k) \leqslant 1$ *for all* k, *and for* $k \geqslant k_0 := \max_{i=1,\ldots,q}\{N_i^2 - 2N_i + 2\}$, *we have* $\xi_m(k) < 1$.

Proof.

(a) According to the definition of $\mathbf{W}_{\sigma,k}$, and taking into account that matrix $\mathbf{C}_\sigma$ is non-negative, it suffices to prove that neither $\mathbf{D}_{\sigma,k}$ nor $-\mathbf{D}_{\sigma,k}$ are non-negative matrices. Let $\mathbf{v}_\sigma := \left(\mathbf{v}_1^{\mathrm{T}}(\sigma)|\cdots|\mathbf{v}_q^{\mathrm{T}}(\sigma)\right)^{\mathrm{T}}$ which by construction is a positive vector. Using Lemma 1 it is immediate to check that $\mathbf{P}_\sigma^k\mathbf{v}_\sigma = \mathbf{v}_\sigma$ and $\overline{\mathbf{P}}_\sigma\mathbf{v}_\sigma = \mathbf{v}_\sigma$ and so $\mathbf{M}_\sigma\mathbf{P}_\sigma^k\mathbf{v}_\sigma = \mathbf{M}_\sigma\overline{\mathbf{P}}_\sigma\mathbf{v}_\sigma = \mathbf{M}_\sigma\mathbf{v}_\sigma$. Now $\mathbf{D}_{\sigma,k}\mathbf{v}_\sigma = (\mathbf{M}_\sigma\mathbf{P}_\sigma^k - \mathbf{M}_\sigma\overline{\mathbf{P}}_\sigma)\mathbf{v}_\sigma = \mathbf{0}$. Since $\mathbf{D}_{\sigma,k} \neq \mathbf{0}$ and $\mathbf{v}_\sigma > 0$ if follows that it necessarily must be $\mathbf{D}_{\sigma,k} \ngeqslant 0$ and $\mathbf{D}_{\sigma,k} \nleqslant 0$ as we wanted to prove.

(b) Direct consequence of the fact that for all $\sigma \in \{1,\ldots,s\}$, $\mathbf{P}_\sigma^k = \overline{\mathbf{P}}_\sigma + \mathbf{o}(\delta^k)$ (Proposition 6.1 in Ref. [15]).

(c) $w_{\alpha\beta}^{\sigma,k} \geqslant -1$ and therefore $\xi_m(k) \leqslant 1$ follows immediately from the fact of $\mathbf{M}_\sigma\mathbf{P}_\sigma^k$ and $\mathbf{M}_\sigma\overline{\mathbf{P}}_\sigma$ being non-negative matrices. If $k \geqslant k_0$ then we have [10, p. 520] that $\mathbf{P}_i^k(\sigma) > 0$ for all i and σ. Since matrices $\overline{\mathbf{P}}_i(\sigma)$ are positive we have that $(\mathbf{P}_\sigma^k)_{\alpha\beta} > 0$ in every position (α,β) for which $(\overline{\mathbf{P}}_\sigma)_{\alpha\beta} > 0$. Therefore $(\mathbf{M}_\sigma\mathbf{P}_\sigma^k)_{\alpha\beta} > 0$ in every position (α,β) for which $c_{\alpha\beta}^\sigma > 0$, which implies $\frac{d_{\alpha\beta}^{\sigma,k}}{c_{\alpha\beta}^\sigma} = \frac{(\mathbf{M}_\sigma\mathbf{P}_\sigma^k)_{\alpha\beta}}{c_{\alpha\beta}^\sigma} - 1 > -1$ and so $\xi_m(k) < 1$ follows. □

Proof of Proposition 2. (a) We define, for all σ and k, $\mathbf{T}_{\sigma,k} = [t_{\alpha\beta}^{\sigma,k}] := \mathbf{M}_\sigma\mathbf{P}_\sigma^k$ and then we have directly from (18) that if $c_{\alpha\beta}^\sigma \neq 0$ then $t_{\alpha\beta}^{\sigma,k} = c_{\alpha\beta}^\sigma\left(1 + w_{\alpha\beta}^{\sigma,k}\right)$ $(*)$. Let us show that $c_{\alpha\beta}^\sigma = 0 \Rightarrow t_{\alpha\beta}^{\sigma,k} = 0$ and so the expression $(*)$ will be valid for all values of α and β. Indeed, matrix $\mathbf{C}_\sigma = \mathbf{M}_\sigma\overline{\mathbf{P}}_\sigma$ (resp. $\mathbf{T}_{\sigma,k} = \mathbf{M}_\sigma\mathbf{P}_\sigma^k$) is composed of q^2 blocks $\mathbf{M}_{ij}(\sigma)\overline{\mathbf{P}}_j(\sigma)$ (resp. $\mathbf{M}_{ij}(\sigma)\mathbf{P}^k(\sigma)$); $i,j = 1,\ldots,q$. Since matrices $\overline{\mathbf{P}}_j(\sigma)$ are positive for all j, that implies that if the element (r,l) of $\mathbf{M}_{ij}(\sigma)\overline{\mathbf{P}}_j(\sigma)$ is zero then the rth row of $\mathbf{M}_{ij}(\sigma)$ must be zero and therefore the element (r,l) of $\mathbf{M}_{ij}(\sigma)\mathbf{P}^k(\sigma)$ is also zero. Thus $c_{\alpha\beta}^\sigma = 0 \Rightarrow t_{\alpha\beta}^{\sigma,k} = 0$ as desired.

Let us define, for each n and k, the following random matrix products

$$\begin{aligned}
&\mathbf{\Pi}_{n,k} = [\Pi^{n,k}_{\alpha\beta}] := \mathbf{M}_{\tau_n}\mathbf{P}^k_{\tau_n}\cdots\mathbf{M}_{\tau_2}\mathbf{P}^k_{\tau_2}\mathbf{M}_{\tau_1}\mathbf{P}^k_{\tau_1},\\
&\mathbf{\Pi}'_n = [\Pi'^n_{\alpha\beta}] := \mathbf{M}_{\tau_n}\overline{\mathbf{P}}_{\tau_n}\cdots\mathbf{M}_{\tau_2}\overline{\mathbf{P}}_{\tau_2}\mathbf{M}_{\tau_1}\overline{\mathbf{P}}_{\tau_1},\\
&\overline{\mathbf{\Pi}}_n := \overline{\mathbf{M}}_{\tau_n}\cdots\overline{\mathbf{M}}_{\tau_1}\overline{\mathbf{M}}_{\tau_1}
\end{aligned}\tag{30}$$

so we have $\mathbf{X}_n = \mathbf{\Pi}_{n,k}\mathbf{X}_0$, $\mathbf{X}'_n = \mathbf{\Pi}'_n\mathbf{X}_0$ and $\mathbf{Y}_n = \overline{\mathbf{\Pi}}_n\mathbf{Y}_0$ where $\mathbf{Y}_0 = \mathbf{U}\mathbf{X}_0$.

Let us fix $\alpha, \beta \in \{1, 2, \ldots, N\}$. Then, for all n and k we have

$$\begin{aligned}
\Pi^{n,k}_{\alpha\beta} &= \sum_{I_{\alpha,\beta}} t^{\tau_n,k}_{\alpha h_{n-1}}\cdots t^{\tau_2,k}_{h_2h_1}t^{\tau_1,k}_{h_1\beta}\\
&= \sum_{I_{\alpha,\beta}} c^{\tau_n}_{\alpha h_{n-1}}\cdots c^{\tau_2}_{h_2h_1}c^{\tau_1}_{h_1\beta}\left(1+w^{\tau_n,k}_{\alpha h_{n-1}}\right)\cdots\left(1+w^{\tau_2,k}_{h_2h_1}\right)\left(1+w^{\tau_1,k}_{h_1\beta}\right),\\
\Pi'^n_{\alpha\beta} &= \sum_{I_{\alpha,\beta}} c^{\tau_n}_{\alpha h_{n-1}}\cdots c^{\tau_2}_{h_2h_1}c^{\tau_1}_{h_1\beta},
\end{aligned}$$

where $I_{\alpha,\beta}$ is the following set of indexes $I_{\alpha,\beta} := \{(h_{n-1}, \ldots, h_2, h_1)\colon h_i = 1, \ldots, N,\ i = 1, \ldots, n-1\}$. Therefore

$$\begin{aligned}
\Pi^{n,k}_{\alpha\beta} &\leqslant \sum_{I_{\alpha,\beta}} c^{\tau_n}_{\alpha h_{n-1}}\cdots c^{\tau_2}_{h_2h_1}c^{\tau_1}_{h_1\beta}(1+\xi_M(k))^n = (1+\xi_M(k))^n\Pi'^n_{\alpha\beta},\\
\Pi^{n,k}_{\alpha\beta} &\geqslant \sum_{I_{\alpha,\beta}} c^{\tau_n}_{\alpha h_{n-1}}\cdots c^{\tau_2}_{h_2h_1}c^{\tau_1}_{h_1\beta}(1-\xi_m(k))^n = (1-\xi_m(k))^n\Pi'^n_{\alpha\beta}.
\end{aligned}$$

Using these bounds we can write

$$\begin{aligned}
\log\|\mathbf{X}_n\| &= \log\|\mathbf{\Pi}_{n,k}\mathbf{X}_0\| = \log\sum_\alpha\left|\sum_\beta \Pi^{n,k}_{\alpha\beta}X^\beta_0\right| = \log\sum_\alpha\sum_\beta \Pi^{n,k}_{\alpha\beta}X^\beta_0\\
&\leqslant \log\left[(1+\xi_M(k))^n\sum_\alpha\sum_\beta \Pi'^n_{\alpha\beta}X^\beta_0\right] = n\log(1+\xi_M(k)) + \log\|\mathbf{\Pi}'_n\mathbf{X}_0\|\\
&= n\log(1+\xi_M(k)) + \log\|\mathbf{X}'_n\|,
\end{aligned}\tag{31}$$

where in the third and fourth equalities we have used that $\mathbf{X}_0$, $\mathbf{\Pi}_{n,k}$ and $\mathbf{\Pi}'_n$ have non-negative components. Analogously

$$\log\|\mathbf{X}_n\| \geqslant \log\left[(1-\xi_m(k))^n\sum_\alpha\sum_\beta \Pi'^n_{\alpha\beta}X^\beta_0\right] = n\log(1-\xi_m(k)) + \log\|\mathbf{X}'_n\|.\tag{32}$$

Now, from (31) and (32) we have

$$\begin{aligned}
&\forall\omega\in\Omega\ \forall\mathbf{X}_0 \geqslant 0,\quad E_{n,k}(\omega) = \left|\log\|\mathbf{X}_n(\omega)\| - \log\|\mathbf{X}'_n(\omega)\|\right|\\
&\qquad\leqslant n\max\left\{\log(1+\xi_M(k)), \log\left(\frac{1}{1-\xi_m(k)}\right)\right\}
\end{aligned}$$

as we wanted to show. □

Proof of Proposition 3. In the first place let us show that if **A** and **B** are square matrices of the same size then $\max|\mathbf{AB}| \leqslant \max|\mathbf{A}|\|\mathbf{B}\|$ (I). Indeed,

$$|(\mathbf{AB})_{ij}| = \left|\sum_l A_{il}B_{lj}\right| \leqslant \sum_l |A_{il}||B_{lj}| \leqslant \max|\mathbf{A}|\sum_l |B_{lj}|$$

and so $\max_{ij}|(\mathbf{AB})_{ij}| \leqslant \max|\mathbf{A}|\max_j\sum_l|B_{lj}| = \max|\mathbf{A}|\|\mathbf{B}\|$ as we wanted to show.

Let $\sigma \in \{1,\ldots,s\}$ and k be fixed. By definition (18) we have that if α and β are such that $c^{\sigma}_{\alpha\beta} = 0$ then $\left|w^{\sigma,k}_{\alpha\beta}\right| = 0$. In the case $c^{\sigma}_{\alpha\beta} \neq 0$ we can write

$$\left|w^{\sigma,k}_{\alpha\beta}\right| = \left|\frac{d^{\sigma,k}_{\alpha\beta}}{c^{\sigma}_{\alpha\beta}}\right| \leqslant \frac{\max\left|\mathbf{M}_\sigma\left(\mathbf{P}^k_\sigma - \overline{\mathbf{P}}_\sigma\right)\right|}{\min^+\left(\mathbf{M}_\sigma\overline{\mathbf{P}}_\sigma\right)} \leqslant \frac{\max\mathbf{M}_\sigma}{\min^+(\mathbf{M}_\sigma\overline{\mathbf{P}}_\sigma)}\left\|\mathbf{P}^k_\sigma - \overline{\mathbf{P}}_\sigma\right\| \leqslant \frac{\max\mathbf{M}_\sigma}{\min^+(\mathbf{M}_\sigma\overline{\mathbf{P}}_\sigma)}\left\|\mathbf{P}_\sigma - \overline{\mathbf{P}}_\sigma\right\|^k \leqslant v(k),$$

where in the second inequality we have used (I) and, in the third one the fact that

$$\mathbf{P}^k_\sigma - \overline{\mathbf{P}}_\sigma = \left(\mathbf{P}_\sigma - \overline{\mathbf{P}}_\sigma\right)^k \tag{33}$$

(which easily follows from Lemma 1) and the submultiplicativity of the 1-matrix norm. From the last expression we have $\max_{\sigma\in\{1,\ldots,s\}}\max|\mathbf{W}_{\sigma,k}| \leqslant v(k)$ and, since

$$\begin{aligned}\max_{\sigma\in\{1,\ldots,s\}}\max|\mathbf{W}_{\sigma,k}| &= \max_{\sigma\in\{1,\ldots,s\}}\max\{\max\mathbf{W}_{\sigma,k}, -\min\mathbf{W}_{\sigma,k}\}\\ &= \max\left\{\max_{\sigma\in\{1,\ldots,s\}}\max\mathbf{W}_{\sigma,k}, -\min_{\sigma\in\{1,\ldots,s\}}\min\mathbf{W}_{\sigma,k}\right\} = \max\{\xi_M(k), \xi_m(k)\},\end{aligned}$$

then $\max\{\xi_M(k), \xi_m(k)\} \leqslant v(k)$. Therefore, $1 + \xi_M(k) \leqslant 1 + v(k)$ and $(1 - \xi_m(k))^{-1} \leqslant (1 - v(k))^{-1}$ and so the bound (22) now follows directly from (20) and (21). □

Proof of Lemma 5. (a) The result easily follows taking into account the block structure of the matrices under consideration. Indeed, for each $j = 1,\ldots,q$ let us denote $\mathbf{S}_j(\sigma,k) = [S^{rl}_j(\sigma,k)] := \mathbf{P}^k_j(\sigma) - \overline{\mathbf{P}}_j(\sigma)$. Due to the block structure of matrices $\mathbf{M}_\sigma$, $\mathbf{P}^k_\sigma$ and $\overline{\mathbf{P}}_\sigma$, matrix $\mathbf{D}_{\sigma,k} = \mathbf{M}_\sigma(\mathbf{P}^k_\sigma - \overline{\mathbf{P}}_\sigma) = [d^{\sigma,k}_{\alpha\beta}] \in \mathbb{R}^{N\times N}$ (resp. $\mathbf{C}_\sigma = \mathbf{M}_\sigma\overline{\mathbf{P}}_\sigma = [c^{\sigma}_{\alpha\beta}] \in \mathbb{R}^{N\times N}$) can be though of as being composed of q^2 blocks $\mathbf{D}_{ij}(\sigma,k) = [d^{rl}_{ij}(\sigma,k)] := \mathbf{M}_{ij}(\sigma)\left(\mathbf{P}^k_j(\sigma) - \overline{\mathbf{P}}_j(\sigma)\right) = \mathbf{M}_{ij}(\sigma)\mathbf{S}_j(\sigma,k)$ (resp. $\mathbf{C}_{ij}(\sigma) = [c^{rl}_{ij}(\sigma)] = \mathbf{M}_{ij}(\sigma)\overline{\mathbf{P}}_j(\sigma)$); $i, j = 1,\ldots,q$. Analogously, matrix $\mathbf{W}_{\sigma,k} = [w^{\sigma,k}_{\alpha\beta}]$ can be thought of as composed of q^2 blocks $\mathbf{W}_{ij}(\sigma,k) = [w^{rl}_{ij}(\sigma,k)]$, i, $j = 1,\ldots,q$. Let i, $j \in \{1,\ldots,q\}$ and $r,l \in \{1,\ldots,\sigma_i\}$ be fixed. Since $\mathbf{M}_{ij}(\sigma) = [M^{rl}_{ij}(\sigma)] \in \mathbb{R}^{N_i\times N_i}$ is diagonal, then $c^{rl}_{ij}(\sigma) = M^{rr}_{ij}(\sigma)\overline{P}^{rl}_j(\sigma)$ and $d^{rl}_{ij}(\sigma,k) = M^{rr}_{ij}(\sigma)S^{rl}_j(\sigma,k)$. Note that the fact that the $\overline{\mathbf{P}}_j(\sigma)$ are positive matrices implies $c^{rl}_{ij}(\sigma) = 0$ if and only if $M^{rr}_{ij}(\sigma) = 0$. Then, by definition(18) we have that $c^{rl}_{ij}(\sigma) = 0$ implies $w^{rl}_{ij}(\sigma,k) = 0$. Moreover, if $c^{rl}_{ij}(\sigma) \neq 0$ then $M^{rr}_{ij}(\sigma) \neq 0$ and $w^{rl}_{ij}(\sigma,k) = \frac{M^{rr}_{ij}(\sigma)S^{rl}_j(\sigma,k)}{M^{rr}_{ij}(\sigma)\overline{P}^{rl}_j(\sigma)} = \frac{S^{rl}_j(\sigma,k)}{\overline{P}^{rl}_j(\sigma)}$. Then (23) is proved.

(b) Using (23) we have that if α and β are such that $c^{\sigma}_{\alpha\beta} \neq 0$, it follows:

$$\left|w^{\sigma,k}_{\alpha\beta}\right| = \left|\frac{(\mathbf{P}^k_\sigma - \overline{\mathbf{P}}_\sigma)_{\alpha\beta}}{(\overline{\mathbf{P}}_\sigma)_{\alpha\beta}}\right| \leqslant \frac{\max|\mathbf{P}^k_\sigma - \overline{\mathbf{P}}_\sigma|}{\min^+\overline{\mathbf{P}}_\sigma} \leqslant \frac{\|\mathbf{P}^k_\sigma - \overline{\mathbf{P}}_\sigma\|}{\min^+\overline{\mathbf{P}}_\sigma} \leqslant \frac{\|\mathbf{P}_\sigma - \overline{\mathbf{P}}_\sigma\|^k}{\min^+\overline{\mathbf{P}}_\sigma},$$

where in the second inequality we have used $\max \mathbf{A} \leqslant \|\mathbf{A}\|$ and, in the third one, (33) and the submultiplicativity of the 1-matrix norm. Now $\max_{\sigma\in\{1,\ldots,s\}}\max|\mathbf{W}_{\sigma,k}| \leqslant v_H(k)$ and, reasoning like in the proof of Proposition 3 we have $\max\{\xi_M(k), \xi_m(k)\} \leqslant v_H(k)$ from where (22) with $v(k)$ replaced by $v_H(k)$ follows immediately.

Last, let us prove $v_H(k) \leqslant v(k)$ for all k. Let $\sigma \in \{1,\ldots,s\}$ be fixed. Reasoning similarly to the proof of Lemma 5 on the block structure of matrix $\mathbf{C}_\sigma$ we have

$$c^{rl}_{ij}(\sigma) = M^{rr}_{ij}(\sigma)\overline{P}^{rl}_j(\sigma) \leqslant \max \mathbf{M}_{ij}(\sigma)\overline{P}^{rl}_j(\sigma)$$

for all $i, j \in \{1,\ldots,q\}$ and $r, l \in \{1,\ldots,N_i\}$. From here we have

$$\min{}^+\mathbf{C}_{ij}(\sigma) \leqslant \max \mathbf{M}_{ij}(\sigma) \min \overline{\mathbf{P}}_j(\sigma) \leqslant \max \mathbf{M}_\sigma \min \overline{\mathbf{P}}_j(\sigma).$$

Now, since $\min^+\mathbf{C}_\sigma = \min_{i,\ j} \min^+\mathbf{C}_{ij}(\sigma)$ and $\min^+\overline{\mathbf{P}}_\sigma = \min_j \min \overline{\mathbf{P}}_j(\sigma)$, it follows $\min^+\mathbf{C}_\sigma \leqslant \max \mathbf{M}_\sigma \min^+\overline{\mathbf{P}}_\sigma$, i.e.,

$$\frac{\max \mathbf{M}_\sigma}{\min^+(\mathbf{M}_\sigma\overline{\mathbf{P}}_\sigma)}\|\mathbf{P}_\sigma - \overline{\mathbf{P}}_\sigma\|^k \geqslant \frac{\|\mathbf{P}_\sigma - \overline{\mathbf{P}}_\sigma\|^k}{\min^+\overline{\mathbf{P}}_\sigma}$$

and the result follows. □

Proof of Proposition 8. In the first place, let us recall that the product of row (column) allowable matrices is row (column) allowable. Moreover, if $\mathbf{A}$ is row (column) allowable and $\mathbf{B}$ is a positive matrix then $\mathbf{AB} > 0$ (resp. $\mathbf{BA} > 0$) as long as the product is defined. We will write $\mathbf{A} \sim \mathbf{B}$ to denote that $\mathbf{A}$ and $\mathbf{B}$ have the same incidence matrix.

Since $\mathbf{P}_i(\sigma)$ is primitive for all σ and i, we have that for big enough k (specifically, for $k \geqslant k_0 := \max_{i=1,..,q}\{N_i^2 - 2N_i + 2\}$) it follows $\mathbf{P}^k_i(\sigma) > 0$. Now, since matrices $\overline{\mathbf{P}}_i(n)$ are positive, we have that for all $k \geqslant k_0$, $\overline{\mathbf{P}}_i(\sigma) \sim \mathbf{P}^k_i(n)$ and consequently $\mathbf{P}^k_\sigma \sim \overline{\mathbf{P}}_\sigma$ $(*)$ for all $\sigma = 1,\ldots,s$.

(a) Let t be such that any product of t matrices drawn from $\mathscr{A}$ is positive and let $\sigma_1,\ldots,\sigma_t \in \{1,\ldots,s\}$. Then, if $k \geqslant k_0$

$$\begin{aligned}\overline{\mathbf{M}}_{\sigma_1}\overline{\mathbf{M}}_{\sigma_2}\ldots\overline{\mathbf{M}}_{\sigma_t} &= \mathbf{U}\mathbf{M}_{\sigma_1}\mathbf{V}_{\sigma_1}\mathbf{U}\mathbf{M}_{\sigma_2}\mathbf{V}_{\sigma_2}\ldots\mathbf{M}_{\sigma_t}\mathbf{V}_{\sigma_t} = \mathbf{U}\mathbf{M}_{\sigma_1}\overline{\mathbf{P}}_{\sigma_1}\mathbf{M}_{\sigma_2}\overline{\mathbf{P}}_{\sigma_2}\ldots\mathbf{M}_{\sigma_t}\overline{\mathbf{P}}_{\sigma_t}\mathbf{V}_{\sigma_t}\\ &\sim \mathbf{U}\mathbf{M}_{\sigma_1}\mathbf{P}^k_{\sigma_1}\mathbf{M}_{\sigma_2}\mathbf{P}^k_{\sigma_2}\ldots\mathbf{M}_{\sigma_t}\mathbf{P}^k_{\sigma_t}\mathbf{V}_{\sigma_t},\end{aligned}\tag{34}$$

where we have used $(*)$ and Lemma 1. Since $\mathbf{M}_{\sigma_1}\mathbf{P}^k_{\sigma_1}\mathbf{M}_{\sigma_2}\mathbf{P}^k_{\sigma_2}\ldots\mathbf{M}_{\sigma_t}\mathbf{P}^k_{\sigma_t} > 0$ by hypothesis and $\mathbf{U}$ and $\mathbf{V}_\sigma$ are allowable matrices by construction, it follows that the product (34) is positive as we wanted to show.

(b) Let t such that any product of t matrices drawn from $\mathscr{A}_{\text{ag}}$ is positive and let $\sigma_1,\ldots,\sigma_t$, $\sigma_{t+1} \in \{1,\ldots,s\}$. Then, if $k \geqslant k_0$ using 1 we have

$$\begin{aligned}\mathbf{M}_{\sigma_1}\mathbf{P}^k_{\sigma_1}\mathbf{M}_{\sigma_2}\mathbf{P}^k_{\sigma_2}\dots\mathbf{M}_{\sigma_{t+1}}\mathbf{P}^k_{\sigma_{t+1}} &\sim \mathbf{M}_{\sigma_1}\overline{\mathbf{P}}_{\sigma_1}\mathbf{M}_{\sigma_2}\overline{\mathbf{P}}_{\sigma_2}\dots\mathbf{M}_{\sigma_{t+1}}\overline{\mathbf{P}}_{\sigma_{t+1}}\\ &= \mathbf{M}_{\sigma_1}\mathbf{V}_{\sigma_1}\mathbf{U}\mathbf{M}_{\sigma_2}\mathbf{V}_{\sigma_2}\dots\mathbf{U}\mathbf{M}_{\sigma_{t+1}}\mathbf{V}_{\sigma_{t+1}}\mathbf{U}\\ &= \mathbf{M}_{\sigma_1}\mathbf{V}_{\sigma_1}\overline{\mathbf{M}}_{\sigma_2}\dots\overline{\mathbf{M}}_{\sigma_t}\overline{\mathbf{M}}_{\sigma_{t+1}}\mathbf{U},\end{aligned}\tag{35}$$

where, again, we have used Lemma 1. Product $\overline{\mathbf{M}}_{\sigma_2}\dots\overline{\mathbf{M}}_{\sigma_t}\overline{\mathbf{M}}_{\sigma_{t+1}}$ is positive by hypothesis. Since matrices $\mathbf{M}_\sigma$ and $\mathbf{V}_\sigma$ are row allowable so is the product $\mathbf{M}_{\sigma_1}\mathbf{V}_{\sigma_1}$ and since $\mathbf{U}$ is allowable then product (35) is positive. □

Proof of Proposition 25. (a) In the first place, the spectral radius of matrices $\overline{\mathbf{D}}$ and $\overline{\mathbf{D}}' := \mathrm{diag}(\mathbf{M}_1\overline{\mathbf{P}}_1,\dots,\mathbf{M}_s\overline{\mathbf{P}}_s)(\mathbf{Q}\otimes\mathbf{I}_N)$ coincide (Proposition 6.4 in Ref. [15]) and so $\rho(\overline{\mathbf{D}}') = \mu$. Now, from the proof of Proposition 2 we have $(\mathbf{M}_\sigma\mathbf{P}^k_\sigma)_{\alpha\beta} = (\mathbf{M}_\sigma\overline{\mathbf{P}}_\sigma)_{\alpha\beta}\left(1+w^{\sigma,k}_{\alpha\beta}\right)$ for all values of α and β. Then it follows $0 \leqslant (\mathbf{M}_\sigma\mathbf{P}^k_\sigma)_{\alpha\beta} \leqslant (\mathbf{M}_\sigma\overline{\mathbf{P}}_\sigma)_{\alpha\beta}(1+\xi_M(k))$ from where we have

$$\mathbf{0} \leqslant \mathbf{D}_k \leqslant (1+\xi_M(k))\mathrm{diag}(\mathbf{M}_1\overline{\mathbf{P}}_1,\dots,\mathbf{M}_s\overline{\mathbf{P}}_s)(\mathbf{Q}\otimes\mathbf{I}_N) = (1+\xi_M(k))\overline{\mathbf{D}}'.$$

Since $|\mathbf{A}| \leqslant \mathbf{B} \Rightarrow \rho(\mathbf{A}) \leqslant \rho(\mathbf{B})$ [10, p. 491] we have $\mu_k \leqslant (1+\xi_M(k))\mu$ and so $\log\mu_k \leqslant \ln(1+\xi_M(k)) + \log\mu$ (*). On the other hand $(\mathbf{M}_\sigma\mathbf{P}^k_\sigma)_{\alpha\beta} \geqslant (\mathbf{M}_\sigma\overline{\mathbf{P}}_\sigma)_{\alpha\beta}(1-\xi_m(k)) \geqslant 0$ from where it follows:

$$\mathbf{D}_k \geqslant (1-\xi_m(k))\mathrm{diag}(\mathbf{M}_1\overline{\mathbf{P}}_1,\dots,\mathbf{M}_s\overline{\mathbf{P}}_s)(\mathbf{Q}\otimes\mathbf{I}_N) = (1-\xi_m(k))\overline{\mathbf{D}}' \geqslant 0$$

and reasoning analogously we have $\mu_k \geqslant (1-\sigma_m(k))\mu$. Now the result now follows by taking logarithms in both sides of the last expression and using (*).

(b) From (35) we have that if the set $\mathscr{A}$ is ergodic for big enough k, then so is the set $\{\mathbf{M}_1\overline{\mathbf{P}}_1,\dots,\mathbf{M}_s\overline{\mathbf{P}}_s\}$ and therefore the auxiliary system (8) has a s.g.r. that we will denote a'. By definition of global variables we have $\|\mathbf{Y}_n\| = \|\mathbf{X}'_n\|$ and so $a = a'$. Now let us relate the original and the auxiliary system. From (31) and (32) we have $\log\|\mathbf{X}_n\| \leqslant n\log(1+\xi_M(k)) + \log\|\mathbf{X}'_0\|$ and $\log\|\mathbf{X}_n\| \geqslant \log\left[(1-\xi_m(k))^n\sum_\alpha\sum_\beta \Pi'^n_{\alpha\beta}X^\beta_0\right] = n\log(1-\xi_m(k)) + \log\|\mathbf{X}'_n\|$. The result is now immediate from these expressions taking into account that, by definition, $a_k = \lim_{n\to\infty}\log\|\mathbf{X}_n\|/n$ (*a.s.*) and $a' = \lim_{n\to\infty}\log\|\mathbf{X}'_n\|/n$ (*a.s.*). □

References

[1] P. Auger, R. Bravo de la Parra, Methods of aggregation of variables in population dynamics, C.R. Acad. Sci. Paris, Life Sci. 323 (2000) 665.

[2] R. Bravo de la Parra, E. Sánchez, O. Arino, P. Auger, A discrete model with density dependent fast migration, Math. Biosci. 157 (1999) 91.

[3] H. Caswell, Matrix Population Models, second ed., Sinauer Associates Inc., Sunderland, 2001.

[4] S. Charles, R. Bravo de la Parra, J.P. Mallet, H. Persat, P. Auger, Population dynamics modelling in an hierarchical arborescent river network: an attempt with Salmo trutta, Acta Biot. 46 (3) (1998) 223.

[5] S. Charles, R. Bravo de la Parra, J.P. Mallet, H. Persat, P. Auger, Annual spawning migrations in modelling brown trout population dynamics inside an arborescent river network, Ecol. Modell. 133 (2000) 15.

[6] J.E. Cohen, Ergodicity of age structure in populations with Markovian vital rates. II. General states, Adv. Appl. Probab. 9 (1977) 18.

[7] J.E. Cohen, Ergodics theorems of demography, Bull. Amer. Math. Soc. N.S. 1 (1979) 275.

[8] J.E. Cohen, Multiregional age structured populations with changing vital rates: weak and strong stochastic ergodic theorems, in: K.C. Land, A. Rogers (Eds.), Multiregional Mathematical Demography, Academic Press, New York, 1982, p. 477.

[9] J.E. Cohen, S.W. Christensen, C.P. Goodyear, A stochastic age-structured model of Striped Bass (Morone saxatilis) in the Potomac River, Can. J. Fish. Aquat. Sci. 40 (1983) 2170.

[10] R.A. Horn, C.A. Johnson, Matrix Analysis, Cambridge University, Cambridge, 1985.

[11] K.L. Liaw, Multistate dynamics: the convergence of an age-by-region population system, Environ. Planning A 12 (1980) 589.

[12] D.O. Logofet, Matrices and graphsStability Problems in Mathematical Ecology, CRC, 1993.

[13] A. Rogers, Multiregional Demography, Chichester, New York, 1995.

[14] L. Sanz, R. Bravo de la Parra, Variables aggregation in a time discrete linear model, Math. Biosci. 157 (1999) 111.

[15] L. Sanz, R. Bravo de la Parra, Time scales in stochastic multiregional models, Nonlinear Anal.: Real World Applicat. 1 (2000) 89.

[16] L. Sanz, R. Bravo de la Parra, Time scales in a non autonomous linear discrete model, Math. Models Meth. Appl. Sci. 11 (7) (2001) 1203.

[17] L. Sanz, A. Blasco, R. Bravo de la Parra, Approximate reduction of Galton–Watson processes with two time scales, Math. Models Meth. Appl. Sci. 13 (2003) 491.

[18] S. Tuljapurkar, S. Orzack, Population dynamics in variable environments. I. Long-run growth rates and extinction, Theor. Popul. Biol. 18 (1980) 314.

[19] S. Tuljapurkar, Demography in stochastic environments. I. Exact distributions of age structure, J. Math. Biol. 19 (1984) 335.

[20] S. Tuljapurkar, Population Dynamics in Variable Environments, Springer, 1990.

[21] S. Tuljapurkar, H. Caswell (Eds.), Structured-Population Models in Marine, Terrestrial, and Freshwater Systems, Chapman and Hall, 1997.