Time scales in stochastic multiregional models

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

[a] *Depto. Matemáticas, E.T.S.I Industriales, c/o José Gutierrez Abascal, 2, 28006 Madrid, Spain*
[b] *Depto. Matemáticas, Univ. de Alcalá, 28871 Alcalá de Henares, Madrid, Spain*

1. Introduction

In order to study ecological systems, one has to decide the level of complexity and detail one should incorporate into the model so as to optimize the study. Indeed, any model is a compromise between generality and simplicity in one hand and realism on the other. Models which include many biological details in its specification become complicated, and their analytical study is often non-viable. On the other hand, very simple models, although analytically tractable, often do not justify the assumptions needed to obtain such simplicity.

Nature offers many examples of systems with an inherent complexity. For example, communities are sets of interacting populations. Populations themselves have an internal structure, for individuals may have different ages or be in different stages. These stages may correspond to size, spatial patches, genotypes, individual activities as rest or search for food, etc.

Aggregation methods study the relationships between a large class of complex systems, in which many variables are involved, and their corresponding reduced or aggregated systems, governed by a few variables. The aim of aggregation techniques is twofold: on the one hand they allow one to rigorously construct, starting from a complex system, a simpler model which summarizes some characteristics of its dynamics, therefore simplifying their analytical study. On the other hand, the complex systems are explanations of the simpler ones at a finer level of detail. The essential

---

* Corresponding author. Fax: +91-336-3001.

E-mail addresses: lsanz@math.etsii.upm.es (L. Sanz), mtbravo@alcala.es (R.B. de la Parra).

property of complex systems that allows their aggregation is the existence of two time scales. This allows one to think of a hierarchically organized system with a division into subsystems, in such a way that these subsystems are weakly coupled through the slow process and however exhibit strong internal dynamics corresponding to the fast process.

In many of the complex systems found in nature, it is possible to distinguish several processes which take place with different time scales. For example, it is usually the case that processes that take place at the individual behavioral level, as migrations, are fast with respect to those regarding the population level, as the growth of the population [1]. In the majority of models found in the literature, it is implicitly assumed that the fast process reaches equilibrium very fast in comparison to the time scale corresponding to the slow process, and therefore the fast dynamics is supposed to have a negligible impact in the dynamics of the system. However, by using aggregation techniques we may consider the dynamics of both the fast and the slow processes without paying a high cost in terms of the complexity of the models we have to analyze. The idea of aggregation is to choose some (usually one) global variable for each subsystem and to build a reduced system for those global variables. The aggregated system reflects in a certain way both the slow and the fast dynamics of the original system. The dynamics of the aggregated system usually corresponds to the dynamics of the original system, while the fast dynamics of the original system is reflected in the coefficients of the reduced one in such a way that it is possible to study the influences between the different hierarchical levels.

Aggregation techniques have been widely studied in the context of time continuous systems with different time scales for both linear and density-dependent models [2–5]. Aggregation techniques have also been extended to deal with autonomous time discrete systems in both linear and nonlinear cases [38], under the hypothesis that the fast process is conservative of the total number of individuals as, for example, migration. The complex models proposed can be classified in two groups; a first one in which the time step of the model corresponds to the characteristic time of the fast process [6–8], and a second group in which the slow dynamics is the time unit for the model [31,9]. In [33], the authors have extended the second group to deal with linear systems in which the fast process can have a general nature, while [11] applies the theoretical results to a practical case regarding the study of an arborescent river network. Aggregation in the non-autonomous linear case has also been considered by the authors in [32].

The aim of this work is to extend variables aggregation techniques to deal with time discrete models that contemplate environmental stochasticity, developing a general procedure that is used to reduce the complexity of some multiregional models in which two time scales are involved. Environmental stochasticity refers to the incorporation of randomness in a model by considering that the population under study lives in an habitat which presents different environmental conditions. The sequence of environments to which the population is subjected at different times is not deterministic but can only be characterized in a probabilistic way.

Section 2 briefly goes through the different kinds of stochastic matrix models, focusing in the essential aspects of models with environmental stochasticity. In Section 3

we present a very general stochastic matrix model with two time scales (microsystem) for a population living in a random environment. Both the fast and the slow dynamics may correspond to any two biological processes, with the only restriction that the fast process in each environment tends to an equilibrium in a certain sense that it will be specified. In Section 4 we develop a procedure to transform the microsystem into an aggregated system in two different cases. In the first place, we address the aggregation of the microsystem under an additional hypothesis which has to do with the characteristics of the fast process. We present a criterium to define the global variables as linear combinations of the state variables and transform the original system into a stochastic aggregated system for which the pattern of environmental variation coincides with that of the microsystem. The entries of the matrices that represent the vital rates for the aggregated system in each environment are obtained as linear combinations of the ones corresponding to the slow process in that environment, being the coefficients of the combination functions of the equilibrium characteristics of the fast process. In second place, we aggregate the microsystem in the general case. The procedure is similar to the one mentioned above, but in this case we end up with an aggregated system in which the environmental variation is more complex than that corresponding to the microsystem.

Section 5, where ecologists interested mainly in applications may focus their attention, is devoted to applying the techniques above mentioned to simplify the study of two kinds of stochastic multiregional models, with population structured by age and patch, in which there are different time scales involved. In the first case, in which the simplifying assumption of Section 4 is met, we aggregate a stochastic multiregional model in which the migratory process is fast in comparison with demography. The averaging procedure results in an averaging over the spatial component of the population, obtaining a stochastic Leslie model. In the second case, in which the above mentioned simplifying assumption is not verified, we address the aggregation of a multiregional model in which demography is fast in comparison with migration. In this case, we define the global variables by averaging over the age structure of the population in each patch, therefore ending up with an aggregated system in which population is structured attending only to its spatial location.

Section 6 studies some relationships between the microsystem and the aggregated system in the general case. To start with, we obtain relationships between the state variables and the global variables in terms of the separation of time scales between the slow and fast processes. In this way we can describe, in an approximate way, the dynamics of the microsystem if we know that of the reduced one. Next, we particularize our study to the case in which environmental change is modeled by a Markov chain, and obtain relationships between the statistical moments of the population vector for original and aggregated systems. Section 7 briefly goes over the results of the work and points to some open problems.

2. Matrix models and environmental stochasticity

The literature offers several approaches to include stochasticity in matrix deterministic models for the study of structured populations. Roughly we can distinguish three

kinds of models. The first type consists in time-series models [12] that consider a deterministic model in which a random error is added (see for example [30]). These models present some inconveniences to study population dynamics [35] and have not been treated with much profusion in this field. The second kind of models consider "demographic stochasticity". In them (see among others [25–27]), individuals transfer among different states according to certain probabilities (instead of in fixed proportions as it is implied in the deterministic models), which introduces the random character in the model. They are particular cases of the multitype branching processes [18] and their interest seems to be restricted to the study of small populations, for they cannot account for any substantial part of the observed variability in the dynamics of moderate to large populations [35]. The third type of models takes into account "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. 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 greatest part of the literature in stochastic matrix models for the study of populations is devoted to this kind of models (see [38, Chapter 3] for a basic introduction and [37] for a full discussion) for they are conceptually adequate and lead to projections that agree with the variability observed in populations.

Next, we will illustrate a little deeper the characteristics of the models that consider environmental stochasticity. We will restrict our attention to the case in which the number of environments is finite. Therefore, we assume 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 kind of environmental variation is characterized by a sequence of random variables τ_n; $n = 1, 2, \ldots$ 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

$$\mathbf{Z}_{n+1} = \mathbf{A}_{\tau_{n+1}} \mathbf{Z}_n, \tag{2.1}$$

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.

3. A model with two time scales

The model we propose is an adaptation of models previously presented by the authors in [33] to deal with environmental stochasticity in systems in which there are two time scales. We assume that the population under study lives in an environment in which there are s environmental conditions $1, 2, \ldots, s$. For a given time, the environment to

which the population is subjected is chosen randomly from among the s possible states according to a certain probability distribution.

We suppose a stage-structured population in which the population is classified into stages or groups attending to any characteristic of the life cycle. Moreover, each of these groups is divided into several subgroups that may correspond to different spatial patches, different individual activities or any other characteristic that could change the life cycle parameters. The model is therefore general in the sense that we do not state in detail the nature of the population or the subpopulations, although in Section 5 we will particularize it to deal with multiregional models.

We consider the population being subdivided in q populations (groups). Each group is subdivided in subpopulations (subgroups) in such a way that for each $i = 1, 2, \ldots, q$, group i has N_i subgroups. Therefore, the total number of subgroups is $N = N_1 + N_2 + \cdots + N_q$.

We will denote by x_n^{ij} the random variable corresponding to the density of subpopulation j of population i at time n, with $i = 1, 2, \ldots, q$ and $j = 1, 2, \ldots, N_i$. In order to describe the population of group i we will use the random vector $\bar{\mathbf{x}}_n^i = (x_n^{i1}, x_n^{i2}, \ldots, x_n^{iN_i}) \in \mathbb{R}^{N_i}$, $i = 1, 2, \ldots, q$. The composition of the total population is then given by vector $\mathbf{X}_n = (\bar{\mathbf{x}}_n^1, \bar{\mathbf{x}}_n^2, \ldots, \bar{\mathbf{x}}_n^q)^{\mathrm{T}} \in \mathbb{R}^N$, where T denotes transposition.

In the evolution of the population we will consider two processes whose corresponding characteristic time scales, and consequently their projection intervals, are very different from each other. In order to include both time scales in our model we will model these two processes, to which we will refer as the fast and the slow dynamics, by two different matrices. Both the slow and the fast processes may be functions of the environmental conditions.

We will choose as the projection interval of our model, that corresponding to the slow dynamics, i.e., the time elapsed between times n and $n + 1$ is the projection interval of the slow process. For simplicity, we will denote the time span $[n, n+1)$ as Δ_n.

We will make no special assumptions regarding the characteristics of the slow dynamics. Therefore, for a certain fixed projection interval, the slow dynamics will be represented by a non-negative projection matrix $\mathbf{M}_\sigma \in \mathbb{R}^{N \times N}$, which we consider divided into blocks $\mathbf{M}_{ij}(\sigma)$, $1 \leq i, j \leq q$. We have then

$$
\mathbf{M}_\sigma = \begin{bmatrix} \mathbf{M}_{11}(\sigma) & \mathbf{M}_{12}(\sigma) & \cdots & \mathbf{M}_{1q}(\sigma) \\ \mathbf{M}_{21}(\sigma) & \mathbf{M}_{22}(\sigma) & \cdots & \mathbf{M}_{2q}(\sigma) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{M}_{q1}(\sigma) & \mathbf{M}_{q2}(\sigma) & \cdots & \mathbf{M}_{qq}(\sigma) \end{bmatrix},
\tag{3.1}
$$

where each block $\mathbf{M}_{ij}(\sigma) = [M_{ij}^{ml}(\sigma)]$ has dimensions $N_i \times N_j$ and characterizes the rates of transference of individuals from the subgroups of group j to the subgroups of group i in environment σ. More specifically, for each $m = 1, 2, \ldots, N_i$ and each $l = 1, 2, \ldots, N_j$, $M_{ij}^{ml}(\sigma)$ represents the rate of transference due to the slow process, of individuals from subgroup l of group j to subgroup m of group i, in environment σ.

Regarding the fast process we will make the following assumptions:

(A) 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. Therefore, if we consider a fixed projection interval, the fast dynamics of group i in environmental conditions σ will be represented by a non-negative matrix $\mathbf{P}_i(\sigma)$ of dimensions $N_i \times N_i$ for each $i = 1, \ldots, q$ and $\sigma = 1, 2, \ldots, s$. The matrix that characterizes the fast dynamics for the whole population in each environment σ will be then:

$$\mathbf{P}_\sigma = diag(\mathbf{P}_1(\sigma), \mathbf{P}_2(\sigma), \ldots, \mathbf{P}_q(\sigma)). \tag{3.2}$$

(B) For each i and σ, matrix $\mathbf{P}_i(\sigma)$ is irreducible and primitive, which is usually assumed in the modeling of biological systems.

(C) For each i and σ, matrix $\mathbf{P}_i(\sigma)$ has spectral radius equal to one.

As we stated before, we choose the projection interval of the slow process as the time step of the model. Therefore, we need to approximate the effect of the fast dynamics over a period much longer than its corresponding projection interval. Assume that, for fixed n, the population is subjected to environment σ during Δ_n. We will suppose that, during Δ_n, matrix $\mathbf{P}_\sigma$ operates a number k of times, where k can be interpreted as the ratio between the projection intervals corresponding to the slow and fast dynamics. Since the projection intervals of both processes are supposed to be very different from each other, k is a big number and, moreover, we assume it is an integer. Thus, the matrix that characterizes the dynamics of the population during Δ_n will be $\mathbf{M}_\sigma \mathbf{P}_\sigma^k$. In this way, the set of vital rates for our system in the different environments will be $\{\mathbf{M}_1 \mathbf{P}_1^k, \mathbf{M}_2 \mathbf{P}_2^k, \ldots, \mathbf{M}_s \mathbf{P}_s^k\}$.

The pattern of environmental variation will be 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

$$\mathbf{X}_{n+1} = \mathbf{M}_{\tau_{n+1}} \mathbf{P}_{\tau_{n+1}}^k \mathbf{X}_n. \tag{3.3}$$

Hypothesis (C), by which the spectral radius of the $\mathbf{P}_i(\sigma)$ is one, has a clear biological justification. Indeed, if the spectral radius of any $\mathbf{P}_i(\sigma)$ were smaller or greater than one then, for a big enough separation between the two time scales, the fast process in group i during environmental conditions σ would lead the total population of this group to zero or infinity, respectively, before the slow process has time to act. Therefore, the dynamics of the system would be controlled by the fast process and the distinction of two processes in the modeling of the system would be unnecessary.

The primitivity of the $\mathbf{P}_i(\sigma)$, together with assumption (C), guarantees that for a separation of the two time scales sufficiently high, the fast process in each environment approaches an equilibrium distribution.

Let us denote by $\| * \|_1$ the 1-norm in the appropriate $\mathbb{R}^N$ space, that is, if $\mathbf{z} = (z_1, z_2, \ldots, z_N)^T$ we have $\|\mathbf{z}\|_1 = |z_1| + |z_2| + \cdots + |z_N|$. For each $i = 1, \ldots q$ and $\sigma = 1, \ldots, s$ let the positive vectors $\mathbf{v}_i(\sigma)$ and $\mathbf{u}_i(\sigma)$ be defined by

$$\mathbf{P}_i(\sigma)\mathbf{v}_i(\sigma) = \mathbf{v}_i(\sigma); \quad \mathbf{u}_i^T(\sigma)\mathbf{P}_i(\sigma) = \mathbf{u}_i^T(\sigma),$$

$$\|\mathbf{u}_i(\sigma)\|_1 = 1; \quad \mathbf{u}_i^T(\sigma)\mathbf{v}_i(\sigma) = 1, \tag{3.4}$$

i.e., $\mathbf{v}_i(\sigma)$ and $\mathbf{u}_i(\sigma)$ are, respectively, the right and left positive eigenvectors of $\mathbf{P}_i(\sigma)$ associated to eigenvalue 1 and normalized in such a way that the sum of the components of $\mathbf{u}_i(\sigma)$ is one and $\mathbf{u}_i^{\mathrm{T}}(\sigma)\mathbf{v}_i(\sigma) = 1$. Hypothesis (B) and (C) guarantee that $\mathbf{v}_i(\sigma)$ and $\mathbf{u}_i(\sigma)$ are defined in a unique way by conditions (3.4).

Vectors $\mathbf{v}_i(\sigma)$ and $\mathbf{u}_i(\sigma)$ may be interpreted in terms of the behavior of the fast process of group i in environment σ. Consider the hypothetical situation in which the system was governed by the fast process exclusively. Assume that for one sample path of the stochastic process, to interval Δ_n there corresponds environment σ. Suppose besides that Δ_n is long enough with respect to the projection interval corresponding to the fast process for this to reach its equilibrium conditions during Δ_n. Then, for any non-zero "initial condition" of the system at time n, we would have that at the end of Δ_n, the structure of the population of group i would be characterized by the direction of vector $\mathbf{v}_i(\sigma)$, meanwhile the reproductive value of the individuals of that group would be defined by $\mathbf{u}_i(\sigma)$ [10].

In this way, we have that for each $\sigma = 1, \ldots, s$ and each $i = 1, \ldots, q$, the matrix that characterizes the equilibrium fast process in environment σ for group i is

$$\bar{\mathbf{P}}_i(\sigma) = \lim_{k \to \infty} \mathbf{P}_i^k(\sigma) = \mathbf{v}_i(\sigma)\mathbf{u}_i^{\mathrm{T}}(\sigma) > 0 \tag{3.5}$$

and, therefore, the equilibrium conditions for the fast dynamics in environment σ is defined by

$$\bar{\mathbf{P}}_\sigma = diag(\bar{\mathbf{P}}_1(\sigma), \bar{\mathbf{P}}_2(\sigma), \ldots, \bar{\mathbf{P}}_q(\sigma)). \tag{3.6}$$

We define matrices

$$\mathbf{V}_\sigma = diag(\mathbf{v}_1(\sigma), \mathbf{v}_2(\sigma), \ldots, \mathbf{v}_q(\sigma)),$$

$$\mathbf{U}_\sigma = diag(\mathbf{u}_1^{\mathrm{T}}(\sigma), \mathbf{u}_2^{\mathrm{T}}(\sigma), \ldots, \mathbf{u}_q^{\mathrm{T}}(\sigma)), \quad \sigma = 1, \ldots, s, \tag{3.7}$$

whose interpretation is immediate bearing in mind what we pointed out about $\mathbf{v}_i(\sigma)$ and $\mathbf{u}_i(\sigma)$.

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

Lemma 3.1. *For all $\sigma = 1, \ldots, s$, matrices $\mathbf{P}_\sigma, \bar{\mathbf{P}}_\sigma, \mathbf{V}_\sigma$ and $\mathbf{U}_\sigma$ verify*:
  (a) $\mathbf{P}_\sigma \bar{\mathbf{P}}_\sigma = \bar{\mathbf{P}}_\sigma \mathbf{P}_\sigma = \bar{\mathbf{P}}_\sigma$
  (b) $\mathbf{P}_\sigma \mathbf{V}_\sigma = \mathbf{V}_\sigma$
  (c) $\mathbf{U}_\sigma \bar{\mathbf{P}}_\sigma = \mathbf{U}_\sigma$; $\mathbf{U}_\sigma \mathbf{V}_\sigma = \mathbf{I}_q$; $\bar{\mathbf{P}}_\sigma = \mathbf{V}_\sigma \mathbf{U}_\sigma$

4. The aggregated system

In this section we will approximate microsystem (3.3), consisting of N variables (microvariables) associated to the different subgroups, by an aggregated system (or macrosystem) of q variables (global variables), each of them associated to one group.

In the general case it is not possible to perfectly aggregate the microsystem, i.e., it is not possible to define a set of global variables as functions of the microvariables

$\mathbf{Y}_n = f_n(\mathbf{X}_n)$ in such a way that the microsystem depends solely on these global variables. Biological systems can be perfectly aggregated only in some cases and for very particular values of the parameters involved [20], so perfect aggregation has only a theoretical interest. In [17], conditions for perfect aggregation in a certain kind of time continuous stochastic models are explored.

As a general technique, in order to aggregate the microsystem we will make use of approximate aggregation. For this we will consider an auxiliary system which approximates the dynamics of the general system and that can be perfectly aggregated.

We introduce the "auxiliary system" as the stochastic model defined by

$$\mathbf{X}_{n+1} = \mathbf{M}_{\tau_{n+1}} \bar{\mathbf{P}}_{\tau_{n+1}} \mathbf{X}_n, \tag{4.1}$$

i.e., a system in which environmental variability is defined in the same way as in the original system and where, for each realization ω, $\mathbf{M}_{\tau_{n+1}(\omega)} \bar{\mathbf{P}}_{\tau_{n+1}(\omega)}$ is the matrix of vital rates for the auxiliary system corresponding to interval Δ_n. Each matrix $\mathbf{M}_\sigma \bar{\mathbf{P}}_\sigma$ can be interpreted as the matrix of vital rates corresponding to environment σ in the original system, when we substitute the fast process in σ by the equilibrium characteristics of the fast dynamics in σ. In this way the auxiliary system can be considered, if k is big enough, as a perturbation of the original system. Note that we employ the same notation for the variables corresponding to the original and auxiliary systems, and every time they appear, the context will determine which of them we are referring to.

4.1. Particular case

In the first place, we will address the aggregation of system (3.3) in the case the fast dynamics, besides meeting the three hypothesis introduced in the last section, verifies an additional restriction. As it will be seen later on, this is a simplifying assumption that holds in many practical situations and that leads to a relatively simple aggregated system. This hypothesis will be dropped later on to deal with the general case.

Hypothesis (D). *for each $i = 1, \ldots, q$, the reproductive value of individuals of group i does not depend on the environment, i.e., $\mathbf{u}_i(\sigma) = \mathbf{u}_i$ for all $\sigma = 1, \ldots, s$.*

Some situations in which this holds are:

(1) Fast process independent of the environmental conditions, i.e., $\mathbf{P}_\sigma = \mathbf{P}$ for each σ.

(2) Fast dynamics corresponding to processes which are conservative of the total number of individuals, as it is the case with migration. In this situation, matrices $\mathbf{P}_i(\sigma)$ are stochastic and $\mathbf{u}_i(\sigma) = (1/N_i)(1, 1, \ldots, 1)^{\mathrm{T}}$ for all σ and i. (see Section 5.1).

Hypothesis (D) is satisfied also in some other particular cases. In Section 5.2 we will see an example of a situation in which a fast dynamics corresponding to demography meets this condition.

Under hypothesis (D) we have $\mathbf{U}_\sigma = \mathbf{U}$ for all σ. Let us now consider the auxiliary system (4.1). In order to show that this system can be perfectly aggregated, let us multiply both sides by $\mathbf{U}$. Then we have

$$\mathbf{U}\mathbf{X}_{n+1} = \mathbf{U}\mathbf{M}_{\tau_{n+1}} \bar{\mathbf{P}}_{\tau_{n+1}} \mathbf{X}_n = \mathbf{U}\mathbf{M}_{\tau_{n+1}} \mathbf{V}_{\tau_{n+1}} \mathbf{U}\mathbf{X}_n, \tag{4.2}$$

where we have used Lemma 3.1. Then, if we define the global variables by

$$\mathbf{Y}_n = \mathbf{U}\mathbf{X}_n \in \mathbb{R}^q, \tag{4.3}$$

we can express (4.2) as a function of the global variables exclusively. In this way, the aggregated system is defined by

$$\mathbf{Y}_{n+1} = \bar{\mathbf{M}}_{\tau_{n+1}} \mathbf{Y}_n, \tag{4.4}$$

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

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

Note that in our aggregation procedure we have used that, as a consequence of hypothesis (D), $\bar{\mathbf{P}}_\sigma = \mathbf{V}_\sigma \mathbf{U}$ for all σ.

The aggregated system (4.4) can be interpreted as a stochastic model in which, as in the original system (3.3), there are s environmental conditions. Besides, the pattern of environmental variation coincides with that of the original system and, for each σ, the matrix of vital rates in environment σ for the aggregated system is

$$\bar{\mathbf{M}}_\sigma = \mathbf{U}\mathbf{M}_\sigma\mathbf{V}_\sigma \in \mathbb{R}^{q \times q}, \quad \sigma = 1, \ldots, s, \tag{4.5}$$

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

$$\{\bar{\mathbf{M}}_1, \bar{\mathbf{M}}_2, \ldots, \bar{\mathbf{M}}_s\}. \tag{4.6}$$

Note that each $\bar{\mathbf{M}}_\sigma$ only depends on the characteristics of the fast and slow processes in environment σ and has the explicit form

$$\bar{\mathbf{M}}_\sigma = \begin{bmatrix} \mathbf{u}_1^T\mathbf{M}_{11}(\sigma)\mathbf{v}_1(\sigma) & \mathbf{u}_1^T\mathbf{M}_{12}(\sigma)\mathbf{v}_2(\sigma) & \cdots & \mathbf{u}_1^T\mathbf{M}_{1q}(\sigma)\mathbf{v}_q(\sigma) \\ \mathbf{u}_2^T\mathbf{M}_{21}(\sigma)\mathbf{v}_1(\sigma) & \mathbf{u}_2^T\mathbf{M}_{22}(\sigma)\mathbf{v}_2(\sigma) & \cdots & \mathbf{u}_2^T\mathbf{M}_{2q}(\sigma)\mathbf{v}_q(\sigma) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_q^T\mathbf{M}_{q1}(\sigma)\mathbf{v}_1(\sigma) & \mathbf{u}_q^T\mathbf{M}_{q2}(\sigma)\mathbf{v}_2(\sigma) & \cdots & \mathbf{u}_q^T\mathbf{M}_{qq}(\sigma)\mathbf{v}_q(\sigma) \end{bmatrix}.$$

In this way, the element of row i and column j of matrix $\bar{\mathbf{M}}_\sigma$ is $\mathbf{u}_i^T\mathbf{M}_{ij}(\sigma)\mathbf{v}_j(\sigma) = \sum_{m,l} M_{ij}^{ml}(\sigma)u_j^l v_j^l(\sigma)$, which is a linear combination of the coefficients of the slow process relative to the transference of individuals from group j to group i. The coefficients of the linear combination are defined by the equilibrium characteristics of the fast process in environment σ.

4.2. The general case

Now, we will aggregate the auxiliary system (4.1) in the general case that results when we drop hypothesis (D). The methodology is similar to the one considered above but, as we will see, the resulting aggregated system is more complex.

Multiplying both sides of (4.1) by $\mathbf{U}_{\tau_{n+2}}$ we have

$$\mathbf{U}_{\tau_{n+2}}\mathbf{X}_{n+1} = \mathbf{U}_{\tau_{n+2}}\mathbf{M}_{\tau_{n+1}}\bar{\mathbf{P}}_{\tau_{n+1}}\mathbf{X}_n = \mathbf{U}_{\tau_{n+2}}\mathbf{M}_{\tau_{n+1}}\mathbf{V}_{\tau_{n+1}}\mathbf{U}_{\tau_{n+1}}\mathbf{X}_n. \tag{4.7}$$

Then, if we define the global variables by

$$\mathbf{Y}_n = \mathbf{U}_{\tau_{n+1}}\mathbf{X}_n \in \mathbb{R}^q, \tag{4.8}$$

we can express (4.2) as a function of the global variables exclusively. In this way, the aggregated system is defined by

$$\mathbf{Y}_{n+1} = \mathbf{U}_{\tau_{n+2}}\mathbf{M}_{\tau_{n+1}}\mathbf{V}_{\tau_{n+1}}\mathbf{Y}_n. \tag{4.9}$$

Note that if we define the random variable $\bar{\tau}_n = (\tau_{n+1}, \tau_n)$ with state space $\{1,\dots,s\} \times \{1,\dots,s\}$ we can write (4.9) as

$$\mathbf{Y}_{n+1} = \bar{\mathbf{M}}_{\bar{\tau}_{n+1}} \mathbf{Y}_n, \tag{4.10}$$

where matrices $\bar{\mathbf{M}}_{\gamma,\sigma}$ are given by

$$\bar{\mathbf{M}}_{\gamma,\sigma} = \mathbf{U}_\gamma \mathbf{M}_\sigma \mathbf{V}_\sigma \in \mathbb{R}^{q \times q}; \quad \gamma, \sigma = 1,\dots,s. \tag{4.11}$$

In this case, the aggregated system (4.10) can be interpreted as a stochastic model with q variables which has the following characteristics:

(a) Unlike the original system, in this case there are s^2 environmental conditions that we can consider numbered (γ,σ), $\gamma,\sigma = 1,\dots,s$. Each environment in the aggregated system is defined by a pair of environments corresponding to the original system. The matrix of vital rates for system (4.10) in each environment (γ,σ) is given by

$$\bar{\mathbf{M}}_{\gamma,\sigma} = \begin{bmatrix} \mathbf{u}_1^{\mathrm{T}}(\gamma)\mathbf{M}_{11}(\sigma)\mathbf{v}_1(\sigma) & \mathbf{u}_1^{\mathrm{T}}(\gamma)\mathbf{M}_{12}(\sigma)\mathbf{v}_2(\sigma) & \cdots & \mathbf{u}_1^{\mathrm{T}}(\gamma)\mathbf{M}_{1q}(\sigma)\mathbf{v}_q(\sigma) \\ \mathbf{u}_2^{\mathrm{T}}(\gamma)\mathbf{M}_{21}(\sigma)\mathbf{v}_1(\sigma) & \mathbf{u}_2^{\mathrm{T}}(\gamma)\mathbf{M}_{22}(\sigma)\mathbf{v}_2(\sigma) & \cdots & \mathbf{u}_2^{\mathrm{T}}(\gamma)\mathbf{M}_{2q}(\sigma)\mathbf{v}_q(\sigma) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_q^{\mathrm{T}}(\gamma)\mathbf{M}_{q1}(\sigma)\mathbf{v}_1(\sigma) & \mathbf{u}_q^{\mathrm{T}}(\gamma)\mathbf{M}_{q2}(\sigma)\mathbf{v}_2(\sigma) & \cdots & \mathbf{u}_q^{\mathrm{T}}(\gamma)\mathbf{M}_{qq}(\sigma)\mathbf{v}_q(\sigma) \end{bmatrix}. \tag{4.12}$$

The vital rates in this environment are therefore given by $\mathbf{u}_i^{\mathrm{T}}(\gamma)\mathbf{M}_{ij}(\sigma)\mathbf{v}_j(\sigma) = \sum_{m,l} M_{ij}^{ml}(\sigma)u_j^l(\gamma)v_j^l(\sigma)$ and so they are functions of (1) the slow process for the original system in environment σ and (2) the equilibrium characteristics of the fast process in environments γ and σ.

(b) The pattern of environmental variation for (4.10) is defined by the sequence of random variables $\bar{\tau}_n = (\tau_{n+1}, \tau_n)$. Therefore, for each realization ω the environment for the aggregated system during Δ_n is $(\tau_{n+2}(\omega), \tau_{n+1}(\omega))$, which depends on the environment selected for the original system during intervals Δ_n and Δ_{n+1}. Thus, we can say that the state of the environment for the aggregated system at a given time n depends on the state of the environment for the original system at times n and $n+1$.

As we have stated before, one of the main purposes of aggregation techniques is that of reducing the complexity of systems. In the case of stochastic models of kind (2.1), this complexity is determined by both the number of the variables of the system and the number of possible environmental states. The general system has N variables and s environmental states. In the case in which hypothesis (D) holds, the aggregated system is governed by q variables and the number of possible environmental states is s, so aggregation always results in a simplification of the system. In the general case, however, even though the aggregation procedure reduces the number of variables from N to q, the number of environmental states increases from s to s^2. Therefore, aggregation will be an useful technique when N is much greater than q and s is small.

The set of matrices of vital rates for the different environments in the aggregated system is then

$$\mathscr{A}_{ag} = \{\bar{\mathbf{M}}_{1,1}, \bar{\mathbf{M}}_{2,1}, \dots, \bar{\mathbf{M}}_{s,1}, \bar{\mathbf{M}}_{1,2}, \bar{\mathbf{M}}_{2,2}, \dots, \bar{\mathbf{M}}_{s,2}, \dots, \bar{\mathbf{M}}_{1,s}, \bar{\mathbf{M}}_{2,s}, \dots, \bar{\mathbf{M}}_{s,s}\}. \tag{4.13}$$

The following lemma characterizes the incidence matrix of the matrices $\bar{\mathbf{M}}_{\gamma,\sigma}$ in terms of that of matrices $\mathbf{M}_\sigma$.

Lemma 4.1. *For* $\gamma, \sigma = 1, \ldots, s$, $\bar{\mathbf{M}}_{\gamma, \sigma}$ *is a nonnegative matrix in which the element of row i and column j of* $\bar{\mathbf{M}}_{\gamma, \sigma}$ *is non zero if and only if matrix* $\mathbf{M}_{ij}(\sigma)$ *is not zero.*

Note from this last result that the pattern of non-zero elements in $\bar{\mathbf{M}}_{\gamma, \sigma}$ coincides with the pattern of non zero blocks $\mathbf{M}_{ij}(\sigma)$ for the slow dynamics in environment σ.

4.2.1. Global variables

The global variables $\mathbf{Y}_n = (y_n^1, \ldots, y_n^q)^{\mathrm{T}}$, defined by (4.8), have the following expression in terms of the variables $\mathbf{X}_n$ of the auxiliary system:

$$y_n^i = \mathbf{u}_i^{\mathrm{T}}(\tau_{n+1})\bar{\mathbf{x}}_n^{iT} = u_i^1(\tau_{n+1})x_n^{i1} + u_i^2(\tau_{n+1})x_n^{i2} + \cdots + u_i^{N_i}(\tau_{n+1})x_n^{iN_i}, \quad i = 1, \ldots, q, \tag{4.14}$$

in such a way that if for a realization ω the environment during Δ_{n+1} is σ, i.e., $\tau_{n+1}(\omega) = \sigma$, then

$$y_n^i(\omega) = \mathbf{u}_i^{\mathrm{T}}(\sigma)\bar{\mathbf{x}}_n^{iT}(\omega) = u_i^1(\sigma)x_n^{i1}(\omega) + u_i^2(\sigma)x_n^{i2}(\omega) + \cdots + u_i^{N_i}(\sigma)x_n^{iN_i}(\omega),$$
$$i = 1, \ldots, q.$$

As a consequence we have:

$y_n^i(\omega)$ is a linear combination of the microvariables of the auxiliary system corresponding to group i, being the coefficients of the combination the components of vector $\mathbf{u}_i(\sigma)$. Recall that, as it was stated earlier, $\mathbf{u}_i(\sigma)$ is a vector of reproductive values for the fast process in group i and environment σ. Therefore we have that, for each $j = 1, \ldots, N_i$, the microvariable $x_n^{ij}(\omega)$ corresponding to subgroup j has a relative weight which is proportional to $u_i^j(\sigma)$. Recall that $u_i^j(\sigma)$ is proportional to the contribution to the total population at time $n + 1$ that an individual present in group i and subgroup j at time n would have in the case that the system were subjected to environmental condition σ, were governed by the fast process exclusively and this process reached equilibrium within Δ_n. In this way, the higher the relative value of $u_i^j(\sigma)$ in vector $\mathbf{u}_i(\sigma)$ is, the higher the relative contribution of $x_n^{ij}(\omega)$ to $y_n^i(\omega)$.

As we stated before, in the case the fast process is conservative of the total number of individuals in each one of the groups, we have that matrices $\mathbf{P}_i(\sigma)$ are stochastic and $\mathbf{u}_i(\sigma) = (1/N_i)(1, 1, \ldots, 1)^{\mathrm{T}}$ for all σ and i. Therefore, the global variable corresponding to each group is proportional to the total population in that group, i.e.,

$$y_n^i = \frac{1}{N_i}(x_n^{i1} + x_n^{i2} + \cdots + x_n^{iN_i}), \quad i = 1, \ldots, q.$$

Note that in the case hypothesis (D) holds, we have $\bar{\mathbf{M}}_{\gamma, \sigma} = \mathbf{U}\mathbf{M}_\sigma\mathbf{V}_\sigma = \bar{\mathbf{M}}_\sigma$ and $\bar{\mathbf{M}}_{\bar{\tau}_n} = \bar{\mathbf{M}}_{\tau_n}$, so the aggregated model (4.10) simplifies to system (4.4) which we considered in Section 4.1. In particular, set (4.13) reduces to (4.6) and the global variables simplify to

$$y_n^i(\omega) = \mathbf{u}_i^{\mathrm{T}}\bar{\mathbf{x}}_n^{iT}(\omega) = u_i^1 x_n^{i1}(\omega) + u_i^2 x_n^{i2}(\omega) + \cdots + u_i^{N_i} x_n^{iN_i}(\omega)$$
$$\text{for all } \omega \in \Omega, \ i = 1, \ldots, q.$$

Unless otherwise stated, in the sequel we will always consider the aggregated system in the general case (4.10).

5. Stochastic multiregional models

In this section we use the aggregation procedure considered in the preceding sections to simplify the treatment of stochastic multiregional models. We will show how, in some practical cases found in the literature, we can distinguish two time scales that make possible the use of aggregation in order to simplify the corresponding model.

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 [29] among others for the study of human populations. A list of ecological applications can be found in [10,24]. The usual approach has been deterministic, but the stochastic setting has also been used (see for example [15]).

The above references do not explicitly consider the existence of different time scales in the multiregional system. The approaches of Bravo et al. [6,7,9], Sánchez et al. [31] and the authors [32], have considered migration to be fast in comparison with the demographic process. In our first application we aggregate a multiregional model under these conditions. We will see that hypothesis (D) of Section 4 is met and, therefore, we end up with a very simple aggregated system.

5.1. A multiregional model with fast migration

We suppose that the population under study is divided into q age classes (corresponding to groups) and spread out in r spatial patches (subgroups) among which they may migrate. In this way, the population is structured in qr stages, each of them corresponding to an age class and a spatial location. Besides, the population lives in an habitat in which there are s environmental conditions that influence the vital rates.

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

We will denote by x_n^{ij} the number of individuals of age i in the jth spatial patch at time n, with $i = 1, 2, \ldots, q$ and $j = 1, 2, \ldots, r$. We use vectors $\bar{\mathbf{x}}_n^i = (x_n^{i1}, x_n^{i2}, \ldots, x_n^{ir}) \in \mathbb{R}^r$; $i = 1, 2, \ldots, q$, to describe the spatial allocation of individuals in age class i. The population will be described by vector $\mathbf{X}_n = (\bar{\mathbf{x}}_n^1, \bar{\mathbf{x}}_n^2, \ldots, \bar{\mathbf{x}}_n^q)^{\mathrm{T}} \in \mathbb{R}^{rq}$. 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 (1) transition from any patch to any other, in a sufficient number of steps, is allowed and (2) individuals of at least one patch are allowed to stay in that patch). Since the $\mathbf{P}_i(\sigma)$ are stochastic, the left eigenvectors associated to eigenvalue one and verifying (3.4) are independent of σ and have the form $\mathbf{u}_i = (1/r)(1, 1, \ldots, 1)^{\mathrm{T}} \in \mathbb{R}^r$. We have then that hypothesis (A)–(D) of Sections 3 and 4 are met.

Since matrices $\mathbf{P}_i(\sigma)$ are primitive, if the duration of each age class is sufficiently long with respect to the projection interval of migration, the migratory process in each

environment σ will approximately reach equilibrium within each Δ_n. Let $\mathbf{v}_i(\sigma)$ be a (positive) vector whose direction characterizes the equilibrium structure of the population for group i in environment σ. As normalization condition we impose, according to (3.4), that the sum of its components is r. So we have $\mathbf{V}_\sigma = diag(\mathbf{v}_1(\sigma), \ldots, \mathbf{v}_q(\sigma))$; $\mathbf{U} = \mathbf{U}_\sigma = diag(\mathbf{u}_1^{\mathrm{T}}, \ldots, \mathbf{u}_q^{\mathrm{T}})$. Migration for the whole of the population is given, for each environment σ, by matrix $\mathbf{P}_\sigma = diag(\mathbf{P}_1(\sigma), \ldots, \mathbf{P}_q(\sigma))$.

The demographic process is defined through the following coefficients:

Fertility coefficients: $F_i^j(\sigma) =$ fertility rate for individuals of age i in patch j under environment σ; $i = 1, \ldots, q$; $j = 1, \ldots, r$; $\sigma = 1, \ldots, s$.

Survival coefficients: $S_i^j(\sigma) =$ survival rate for individuals of age i in patch j under environment σ; $i = 1, \ldots, q-1$; $j = 1, \ldots, r$; $\sigma = 1, \ldots, s$.

We define matrices $\mathbf{F}_i(\sigma) = diag(F_i^1(\sigma), \ldots, F_i^r(\sigma))$, $i = 1, \ldots, q$ and $\mathbf{S}_i(\sigma) = diag(S_i^1(\sigma), \ldots, S_i^r(\sigma))$, $i = 1, \ldots, q-1$, and then demography for the whole of the population in each environment σ is defined by the following generalized Leslie matrix:

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

We will assume that the pattern of environmental variation is characterized by a sequence τ_n; $n = 1, 2, \ldots$ of random variables with values in the set $\{1, \ldots, s\}$ in such a way that for each realization ω, the environment during Δ_n is given by $\tau_{n+1}(\omega)$.

The general system or microsystem consists then in the following system of rq random difference equations:

$$
\mathbf{X}_{n+1} = \mathbf{M}_{\tau_{n+1}} \mathbf{P}_{\tau_{n+1}}^k \mathbf{X}_n, \tag{5.1}
$$

where k can be interpreted as an integer representing the ratio between the projection intervals corresponding to the demographic and migratory processes.

Following the aggregation procedure developed in Section 4.1, we collapse the rq state variables of the original multiregional model into q global varibles, each one of them corresponding to one age class, given by

$$
y_n^i = \mathbf{u}_i^{\mathrm{T}} \bar{\mathbf{x}}_n^{i\mathrm{T}} = \frac{1}{r}(x_n^{i1} + x_n^{i2} + \cdots + x_n^{ir}), \quad i = 1, \ldots, q,
$$

i.e., for a given instant n, the global variable corresponding to each class i is proportional to the total population with that age.

The aggregated system has the form

$$
\mathbf{Y}_{n+1} = \bar{\mathbf{M}}_{\tau_{n+1}} \mathbf{Y}_n, \tag{5.2}
$$

where, for each $\sigma = 1,\ldots,s$, $\bar{\mathbf{M}}_\sigma$ is a classical Leslie matrix given by

$$\bar{\mathbf{M}}_\sigma = \begin{bmatrix} f_1(\sigma) & f_2(\sigma) & \cdots & f_{q-1}(\sigma) & f_q(\sigma) \\ s_1(\sigma) & 0 & \cdots & 0 & 0 \\ 0 & s_2(\sigma) & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & s_{q-1}(\sigma) & 0 \end{bmatrix}. \tag{5.3}$$

The vital rates in (5.3) have the form

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

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

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 the migratory process in environment σ. Something analogous holds for the survival rates. Note that $f_i(\sigma)$ is non-zero if and only if $\mathbf{F}_i(\sigma)$ is. In the same way, $s_i(\sigma)$ is non-zero if and only if $\mathbf{S}_i(\sigma)$ is.

Note that the aggregated system is a stochastic Leslie model with q variables in which, as it is the case for the original multiregional system, there are s possible environmental conditions. 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.

In order to illustrate the above discussions we will particularize them to a very simple case in which there are only two patches, two age classes and two environmental conditions, i.e., $q=r=s=2$. Then the matrices $\mathbf{M}_\sigma$ and $\mathbf{P}_\sigma$ that characterize demography and migration are given, for $\sigma = 1, 2$, by

$$\mathbf{M}_\sigma = \begin{bmatrix} F_1^1(\sigma) & 0 & F_2^1(\sigma) & 0 \\ 0 & F_1^2(\sigma) & 0 & F_2^2(\sigma) \\ S^1(\sigma) & 0 & 0 & 0 \\ 0 & S^2(\sigma) & 0 & 0 \end{bmatrix},$$

$$\mathbf{P}_\sigma = diag(\mathbf{P}_1(\sigma), \mathbf{P}_2(\sigma)) = \begin{bmatrix} 1-p_1^\sigma & t_1^\sigma & 0 & 0 \\ p_1^\sigma & 1-t_1^\sigma & 0 & 0 \\ 0 & 0 & 1-p_2^\sigma & t_2^\sigma \\ 0 & 0 & p_2^\sigma & 1-t_2^\sigma \end{bmatrix},$$

where p_1^σ and p_2^σ denote the migration rates in environment σ of individuals of age 1 and 2, respectively, from patches 1 to 2, and t_1^σ and t_2^σ have an analogous meaning for migration from patches 2 to 1. $S^1(\sigma)$ and $S^2(\sigma)$, denote the survival rates, in

environmental condition σ, of individuals of age 1 in patches 1 and 2, respectively. If we assume that, for $\sigma = 1, 2$, coefficients p_1^σ, p_2^σ, t_1^σ and t_2^σ are different from 0 or 1, then matrices $\mathbf{P}_1(\sigma)$ and $\mathbf{P}_2(\sigma)$ are primitive. Vectors $\mathbf{v}_i(\sigma)$ and $\mathbf{u}_i$ have the form

$$\mathbf{v}_1(\sigma) = 2 \left(\frac{t_1^\sigma}{p_1^\sigma + t_1^\sigma}, \frac{p_1^\sigma}{p_1^\sigma + t_1^\sigma} \right)^{\mathrm{T}}, \quad \mathbf{v}_2(\sigma) = 2 \left(\frac{t_2^\sigma}{p_2^\sigma + t_2^\sigma}, \frac{p_2^\sigma}{p_2^\sigma + t_2^\sigma} \right)^{\mathrm{T}}, \quad \sigma = 1, 2$$

$$\mathbf{u}_1 = \mathbf{u}_2 = \tfrac{1}{2}(1, 1)^{\mathrm{T}}.$$

In this way, the general system has the following expression:

$$\begin{bmatrix} x_{n+1}^{11} \\ x_{n+1}^{12} \\ x_{n+1}^{21} \\ x_{n+1}^{22} \end{bmatrix} = \mathbf{M}_{\tau_{n+1}} \mathbf{P}_{\tau_{n+1}}^k \begin{bmatrix} x_n^{11} \\ x_n^{12} \\ x_n^{21} \\ x_n^{22} \end{bmatrix}.$$

The global variables y_n^1 and y_n^2 corresponding, respectively, to age classes 1 and 2 are

$$y_n^1 = \mathbf{u}_1^{\mathrm{T}}(x_n^{11}, x_n^{12})^{\mathrm{T}} = \tfrac{1}{2}(x_n^{11} + x_n^{12}), \quad y_n^2 = \mathbf{u}_2^{\mathrm{T}}(x_n^{21}, x_n^{22})^{\mathrm{T}} = \tfrac{1}{2}(x_n^{21} + x_n^{22}),$$

so that they are proportional to the total population with ages 1 and 2, respectively.

The aggregated system takes the form

$$\begin{bmatrix} y_{n+1}^1 \\ y_{n+1}^2 \end{bmatrix} = \bar{\mathbf{M}}_{\tau_{n+1}} \begin{bmatrix} y_n^1 \\ y_n^2 \end{bmatrix},$$

where the vital rates in environments 1 and 2 are given by the following Leslie matrices:

$$\bar{\mathbf{M}}_\sigma = \mathbf{U} \mathbf{M}_\sigma \mathbf{V}_\sigma = \begin{bmatrix} \dfrac{t_1^\sigma F_1^1(\sigma) + p_1^\sigma F_1^2(\sigma)}{p_1^\sigma + t_1^\sigma} & \dfrac{t_2^\sigma F_2^1(\sigma) + p_2^\sigma F_{12}^2(\sigma)}{p_2^\sigma + t_2^\sigma} \\ \dfrac{t_1^\sigma S^1(\sigma) + p_1^\sigma S^2(\sigma)}{p_1^\sigma + t_1^\sigma} & 0 \end{bmatrix}, \quad \sigma = 1, 2.$$

5.2. A multiregional model with fast demography

In some practical cases found in the literature on multiregional models, demography can be considered a fast process in relation to migration. Investigating the dynamics of the population of Canada and Great Britain, Liaw has shown, both empirically and analytically, that the tendency of a multiregional population projection to a fixed asymptotic distribution takes place in two stages: first, a relatively rapid convergence to the stable age compositions in each region and, second, a gradual convergence towards a stable interregional allocation of the national population [22,23]. This suggests [28,29], the use of aggregation techniques, with the role of the fast process being played by demography, in order to simplify the multiregional system.

Our next application consists in the formulation and aggregation of a stochastic multiregional model in which, unlike the model of Section 5.1, demography is fast

with respect to migration. Again, we will consider a population living in an habitat in which there are several possible environmental conditions, and we will allow the demographic and migratory coefficients to be dependent on the environment. We will first set up the general situation and then we will particularize it to a very simple case that will illustrate the methodology used.

We suppose a population living in an environment composed of q patches (corresponding to groups) and divided in r discrete age classes (which in this case correspond to the subgroups). We will denote by x_n^{ij} the number of individuals of age j living in patch i at time n, with $i = 1, \ldots, q$ and $j = 1, \ldots, r$. The variables corresponding to patch i are given by vector $\bar{\mathbf{x}}_n^i = (x_n^{i1}, x_n^{i2}, \ldots, x_n^{ir})$ and the whole population is described by vector $\mathbf{X}_n = (\bar{\mathbf{x}}_n^1, \bar{\mathbf{x}}_n^2, \ldots, \bar{\mathbf{x}}_n^q)^{\mathrm{T}}$ which therefore has rq components. Note that the notation in this case has a different meaning with respect to the one employed in Section 5.1. This is a consequence of the fact that now the fast and slow processes are interchanged with respect to those corresponding to the first application.

Now, we choose the projection interval of the migratory process as time step $\Delta_n = [n, n+1)$ for the model.

The characteristics of migration between the different patches are given, for each $\sigma = 1, \ldots, s$, by the following non-negative coefficients:

$m_{ij}^t(\sigma)=$ rate of transference of individuals of age t from patch j to patch i, in environment σ; $i, j = 1, \ldots, q$; $t = 1, \ldots, r$.

These coefficients must satisfy the obvious restriction $\sum_{i=1}^q m_{ij}^t(\sigma) = 1$ for all j, t and σ. In this way, migration for the whole population in environment σ is defined by the following stochastic matrix of dimensions $qr \times qr$:

$$
\mathbf{M}_\sigma =
\begin{bmatrix}
\mathbf{M}_{11}(\sigma) & \mathbf{M}_{12}(\sigma) & \cdots & \mathbf{M}_{1q}(\sigma) \\
\mathbf{M}_{21}(\sigma) & \mathbf{M}_{22}(\sigma) & \cdots & \mathbf{M}_{2q}(\sigma) \\
\vdots & \vdots & \ddots & \vdots \\
\mathbf{M}_{q1}(\sigma) & \mathbf{M}_{q2}(\sigma) & \cdots & \mathbf{M}_{qq}(\sigma)
\end{bmatrix},
$$

where matrix $\mathbf{M}_{ij}(\sigma) \in \mathbb{R}^{r \times r}$ characterizes the transference of individuals from patch j to patch i in environment σ, and is given by

$$
\mathbf{M}_{ij}(\sigma) = diag(m_{ij}^1(\sigma), m_{ij}^2(\sigma), \ldots, m_{ij}^r(\sigma)) \quad i, j = 1, \ldots, q, \quad \sigma = 1, \ldots, s.
$$

Demography is an internal process for each patch. For each environmental condition σ, the demographic process in patch i is defined by a Leslie matrix $\mathbf{P}_i(\sigma)$ in the following way:

$$
\mathbf{P}_i(\sigma) =
\begin{bmatrix}
F_i^1(\sigma) & F_i^2(\sigma) & \cdots & F_i^{r-1}(\sigma) & F_i^r(\sigma) \\
S_i^1(\sigma) & 0 & \cdots & 0 & 0 \\
0 & S_i^2(\sigma) & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & S_i^{r-1}(\sigma) & 0
\end{bmatrix} \in \mathbb{R}^{r \times r},
$$

where the vital rates have the classical interpretation, i.e.:

$F_i^j(\sigma) =$ fertility rate of individuals of age j in patch i in environment σ.

$S_i^j(\sigma) =$ survival rate of individuals of age j in patch i in environment σ.

We will suppose that: (a) All the survival rates are non-zero, i.e., $S_i^j(\sigma) \neq 0$ for all i, j and σ. (b) For all patches and environments, the fertility coefficient of the last age class is non-zero, i.e., $F_i^r(\sigma) \neq 0$ for all i and σ. (c) For all σ we have that given patch i there is an age j such that $F_i^j(\sigma) \neq 0$ and $g.c.d.(r, j) = 1$. These three conditions, which are usually met in most practical situations, guarantee that for all i and σ matrices $\mathbf{P}_i(\sigma)$ are primitive [24]. Therefore, if Δ_n is long enough with respect to the projection interval corresponding to migration, the demographic process in all patches and environments tends to an equilibrium. Besides we will assume that, for all i and σ, the dominant eigenvalue of $\mathbf{P}_i(\sigma)$ is one, i.e., the demography in each patch and each environment makes the individuals change age classes but does not make the total population of the patch grow indefinitely or decay to zero. This last assumption is reflected in the requirement that the vital rates of $\mathbf{P}_i(\sigma)$ satisfy the restriction (see [24])

$$F_i^1(\sigma) + F_i^2(\sigma)S_i^1(\sigma) + \cdots + F_i^r(\sigma)S_i^1(\sigma)S_i^2(\sigma)...S_i^{r-1}(\sigma) = 1,$$
$$i = 1,\ldots,q, \ \sigma = 1,\ldots,s.$$

We define $\mathbf{u}_i(\sigma)$ as the positive left eigenvector of $\mathbf{P}_i(\sigma)$ associated to 1 and normalized so the sum of its components is 1. This vector characterizes the reproductive values of the individuals for the demographic process in patch i and environment σ. For each j, $u_i^j(\sigma)$ measures the contribution that one individual of patch i initially present in age j would have in the total equilibrium population of this patch i if the system were subjected to environment σ and there were no migration. Note that, in general, vectors $\mathbf{u}_i(\sigma)$ depend on σ and therefore hypothesis (D) of Section 4 is not verified.

We denote by $\mathbf{v}_i(\sigma)$ the positive eigenvector of $\mathbf{P}_i(\sigma)$ associated to eigenvalue 1 and verifying the normalization condition $\mathbf{u}_i^T(\sigma)\mathbf{v}_i(\sigma) = 1$. This vector has the direction of the equilibrium population structure corresponding to the demography of group i in environment σ. The explicit expressions for vectors $\mathbf{u}_i(\sigma)$ and $\mathbf{v}_i(\sigma)$ in terms of the vital rates in patch i can be found in [10].

Demography for the whole population will be characterized by matrix $\mathbf{P}_\sigma = diag(\mathbf{P}_1(\sigma), \mathbf{P}_2(\sigma),\ldots,\mathbf{P}_q(\sigma))$.

In order to approximate the effect of demography over the time interval of the model, which is much longer than its own projection interval, we suppose that during each interval Δ_n, matrix $\mathbf{P}_\sigma$ operates a number k of times, where k is a big number, that we assume to be an integer, that in this case can be interpreted as the ratio between the projection intervals of the migratory and demographic processes.

We will assume that the pattern of environmental variation is characterized by a sequence τ_n; $n = 1, 2,\ldots$ of random variables with values in the set $\{1,\ldots,s\}$ in such a way that for each realization ω, the environment during Δ_n is given by $\tau_{n+1}(\omega)$.

Therefore, the proposed multiregional model will be the following system of qr stochastic difference equations:

$$\mathbf{X}_{n+1} = \mathbf{M}_{\tau_{n+1}} \mathbf{P}_{\tau_{n+1}}^k \mathbf{X}_n. \tag{5.4}$$

Following the aggregation procedure proposed in Section 4.2, we will end up with q random global variables, one for each one of the patches, governing the aggregated system. These random variables are given by

$$y_n^i = \mathbf{u}_i^T(\tau_{n+1})\bar{\mathbf{x}}_n^{iT} = u_i^1(\tau_{n+1})x_n^{i1} + u_i^2(\tau_{n+1})x_n^{i2} + \cdots + u_i^r(\tau_{n+1})x_n^{ir}, \quad i = 1, \ldots, q.$$

Let us assume that for a realization ω of the stochastic process, the environment during Δ_{n+1} is σ, i.e., $\tau_{n+1}(\omega) = \sigma$. Then

$$y_n^i(\omega) = \mathbf{u}_i^T(\sigma)\bar{\mathbf{x}}_n^{iT}(\omega) = u_i^1(\sigma)x_n^{i1}(\omega) + u_i^2(\sigma)x_n^{i2}(\omega) + \cdots + u_i^r(\sigma)x_n^{ir}(\omega),$$
$$i = 1, \ldots, q,$$

so $y_n^i(\omega)$ is a linear combination of the microvariables corresponding to patch i, being the coefficients of each microvariable $x_n^{ij}(\omega)$ the reproductive value $u_i^j(\sigma)$ for the demographic process of individuals of age j in that patch and environment σ.

The aggregated system has the form

$$\mathbf{Y}_{n+1} = \bar{\mathbf{M}}_{\tau_{n+2}, \tau_{n+1}} \mathbf{Y}_n \tag{5.5}$$

and can be interpreted as a stochastic system with s^2 possible environmental conditions that we can consider numbered (γ, σ), $\gamma, \sigma = 1, \ldots, s$. The matrix $\bar{\mathbf{M}}_{\gamma, \sigma} \in \mathbb{R}^{q \times q}$ of vital rates in each environment (γ, σ) is given by

$$\bar{\mathbf{M}}_{\gamma, \sigma} = \begin{bmatrix} \sum_{t=1}^{r} u_1^t(\gamma)v_1^t(\sigma)m_{11}^t(\sigma) & \sum_{t=1}^{r} u_1^t(\gamma)v_2^t(\sigma)m_{12}^t(\sigma) & \cdots & \sum_{t=1}^{r} u_1^t(\gamma)v_q^t(\sigma)m_{1q}^t(\sigma) \\ \sum_{t=1}^{r} u_2^t(\gamma)v_1^t(\sigma)m_{21}^t(\sigma) & \sum_{t=1}^{r} u_2^t(\gamma)v_2^t(\sigma)m_{22}^t(\sigma) & \cdots & \sum_{t=1}^{r} u_2^t(\gamma)v_q^t(\sigma)m_{2q}^t(\sigma) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{t=1}^{r} u_q^t(\gamma)v_1^t(\sigma)m_{q1}^t(\sigma) & \sum_{t=1}^{r} u_q^t(\gamma)v_2^t(\sigma)m_{q2}^t(\sigma) & \cdots & \sum_{t=1}^{r} u_q^t(\gamma)v_q^t(\sigma)m_{qq}^t(\sigma) \end{bmatrix}.$$

As we can see, the entries of the matrix of the aggregated system for environment (γ, σ) are functions of the migratory rates in environment σ as well as of the equilibrium characteristics of the demographic process in environments σ and γ. We have collapsed the original system with rq variables and s environmental conditions into an aggregated system with q variables, each of them corresponding to one patch, in which there are s^2 environments.

Note that, unlike the model studied in Section 5.1, in this case the original multi-regional model is collapsed into a reduced multiregional system in which the age structure has been averaged and, as a consequence, the population retains only its spatial structure.

Let us illustrate the above discussions by particularizing them to a simple case in which there are only two spatial patches, two age classes and two environmental conditions. Under these assumptions we have four microvariables for our multiregional

system x_n^{11}, x_n^{12}, x_n^{21}, x_n^{22} and matrices $\mathbf{M}_\sigma$ and $\mathbf{P}_\sigma$ are given by

$$\mathbf{M}_\sigma = \begin{bmatrix} 1 - p_1^\sigma & 0 & t_1^\sigma & 0 \\ 0 & 1 - p_2^\sigma & 0 & t_2^\sigma \\ p_1^\sigma & 0 & 1 - t_1^\sigma & 0 \\ 0 & p_2^\sigma & 0 & 1 - t_2^\sigma \end{bmatrix}, \mathbf{P}_\sigma = \begin{bmatrix} F_1^1(\sigma) & F_1^2(\sigma) & 0 & 0 \\ S_1(\sigma) & 0 & 0 & 0 \\ 0 & 0 & F_2^1(\sigma) & F_2^2(\sigma) \\ 0 & 0 & S_2(\sigma) & 0 \end{bmatrix},$$

where the migratory coefficients have the same meaning than in Section 5.1 and $S_1(\sigma)$ and $S_2(\sigma)$ denote the survival rates, in environmental condition σ, of individuals of age 1 in patches 1 and 2, respectively.

The restrictions over the vital rates so matrices $\mathbf{P}_1(\sigma)$ and $\mathbf{P}_2(\sigma)$ have dominant eigenvalue 1 are

$$F_1^1(\sigma) + F_1^2(\sigma)S_1(\sigma) = 1; \quad F_2^1(\sigma) + F_2^2(\sigma)S_2(\sigma) = 1. \tag{5.6}$$

Besides we will suppose that all the fertility and survival rates are positive in all environments, so matrices $\mathbf{P}_1(\sigma)$ and $\mathbf{P}_2(\sigma)$ are primitive. Vectors $\mathbf{v}_i(\sigma)$ and $\mathbf{u}_i(\sigma)$ are given by

$$\mathbf{v}_1(\sigma) = \frac{1 + F_1^2(\sigma)}{1 + F_1^2(\sigma)S_1(\sigma)} \begin{pmatrix} 1 \\ S_1(\sigma) \end{pmatrix}, \qquad \mathbf{v}_2(\sigma) = \frac{1 + F_2^2(\sigma)}{1 + F_2^2(\sigma)S_2(\sigma)} \begin{pmatrix} 1 \\ S_2(\sigma) \end{pmatrix},$$

$$\mathbf{u}_1(\sigma) = \frac{1}{1 + F_1^2(\sigma)} \begin{pmatrix} 1 \\ F_1^2(\sigma) \end{pmatrix}, \qquad \mathbf{u}_2(\sigma) = \frac{1}{1 + F_2^2(\sigma)} \begin{pmatrix} 1 \\ F_2^2(\sigma) \end{pmatrix}. \tag{5.7}$$

Therefore, the microsystem has the form

$$\begin{bmatrix} x_{n+1}^{11} \\ x_{n+1}^{12} \\ x_{n+1}^{21} \\ x_{n+1}^{22} \end{bmatrix} = \mathbf{M}_{\tau_{n+1}} \mathbf{P}_{\tau_{n+1}}^k \begin{bmatrix} x_n^{11} \\ x_n^{12} \\ x_n^{21} \\ x_n^{22} \end{bmatrix}$$

and the global variables y_n^1 and y_n^2 corresponding, respectively, to patches 1 and 2 are defined by

$$y_n^1 = \mathbf{u}_1^T(\tau_{n+1})(x_n^{11}, x_n^{12})^T = \frac{1}{1 + F_1^2(\tau_{n+1})}(x_n^{11} + F_1^2(\tau_{n+1})x_n^{12}),$$

$$y_n^2 = \mathbf{u}_2^T(\tau_{n+1})(x_n^{21}, x_n^{22})^T = \frac{1}{1 + F_2^2(\tau_{n+1})}(x_n^{21} + F_2^2(\tau_{n+1})x_n^{22}).$$

The aggregated system is a stochastic model with four environmental conditions (γ, σ); $\gamma, \sigma = 1, 2$ with the form

$$\begin{bmatrix} y_{n+1}^1 \\ y_{n+1}^2 \end{bmatrix} = \bar{\mathbf{M}}_{\tau_{n+2}, \tau_{n+1}} \begin{bmatrix} y_n^1 \\ y_n^2 \end{bmatrix}.$$

Matrix $\bar{\mathbf{M}}_{\gamma,\sigma}$ corresponding to environment (γ,σ) in this system is given by

$$\bar{\mathbf{M}}_{\gamma,\sigma} = \begin{bmatrix} \dfrac{1+F_1^2(\sigma)}{1+F_1^2(\gamma)} \dfrac{(1-p_1^\sigma)+F_1^2(\gamma)S_1(\sigma)(1-p_2^\sigma)}{1+F_1^2(\sigma)S_1(\sigma)} & \dfrac{1+F_2^2(\sigma)}{1+F_1^2(\gamma)} \dfrac{t_1^\sigma+F_1^2(\gamma)S_1(\sigma)t_2^\sigma}{1+F_2^2(\sigma)S_2(\sigma)} \\[2em] \dfrac{1+F_1^2(\sigma)}{1+F_2^2(\gamma)} \dfrac{p_1^\sigma+F_2^2(\gamma)S_2(\sigma)p_2^\sigma}{1+F_1^2(\sigma)S_1(\sigma)} & \dfrac{1+F_2^2(\sigma)}{1+F_2^2(\gamma)} \dfrac{(1-t_1^\sigma)+F_2^2(\gamma)S_2(\sigma)(1-t_2^\sigma)}{1+F_2^2(\sigma)S_2(\sigma)} \end{bmatrix}.$$

Note from (5.7) that, if we assumed $F_1^2(\sigma) = F_1^2$ and $F_2^2(\sigma) = F_2^2$ for $\sigma = 1,2$, we would have that vectors $\mathbf{u}_1(\sigma)$ and $\mathbf{u}_2(\sigma)$ would be independent of the environment, and so hypothesis (D) would hold. So, by keeping the fertilities in the second age class independent of the environment and allowing the rest of the parameters to vary randomly (always meeting (5.6)), we end up with a simpler aggregated model with two environments $\sigma = 1,2$

$$\begin{bmatrix} y_{n+1}^1 \\ y_{n+1}^2 \end{bmatrix} = \bar{\mathbf{M}}_{\tau_{n+1}} \begin{bmatrix} y_n^1 \\ y_n^2 \end{bmatrix},$$

in such a way that

$$\bar{\mathbf{M}}_\sigma = \begin{bmatrix} \dfrac{(1-p_1^\sigma)+F_1^2 S_1(\sigma)(1-p_2^\sigma)}{1+F_1^2 S_1(\sigma)} & \dfrac{1+F_2^2}{1+F_1^2} \dfrac{t_1^\sigma+F_1^2 S_1(\sigma)t_2^\sigma}{1+F_2^2 S_2(\sigma)} \\[2em] \dfrac{1+F_1^2}{1+F_2^2} \dfrac{(p_1^\sigma+F_2^2 S_2(\sigma)p_2^\sigma)}{1+F_1^2 S_1(\sigma)} & \dfrac{(1-t_1^\sigma)+F_2^2 S_2(\sigma)(1-t_2^\sigma)}{1+F_2^2 S_2(\sigma)} \end{bmatrix}, \quad \sigma = 1,2.$$

6. Relationships between micro and macro systems

In this section we will relate some important features of the microsystem (3.3), the auxiliary system (4.1) and the aggregated system (4.10). As we will see, from the behavior of the aggregated system we can exactly know the behavior of the auxiliary system and we can approximate, in terms of the separation of time scales, the behavior of the microsystem.

In the following, we will say that a given property holds "for big enough k", when there exists an integer k_0 such that for all $k \geq k_0$ that property holds.

6.1. Global variables and microvariables

In this section we will relate the variables corresponding to microsystem (3.3), auxiliary system (4.1) and aggregated system (4.10).

For each $\sigma = 1,\ldots,s$, let the eigenvalues $\lambda_i(\sigma)$ of $\mathbf{P}_\sigma$ ordered by decreasing modulus be

$$1 = \lambda_1(\sigma) = \cdots = \lambda_q(\sigma) > |\lambda_{q+1}(\sigma)| \geq \cdots \geq |\lambda_N(\sigma)| \tag{6.1}$$

and let

$$\alpha > \max\{|\lambda_{q+1}(1)|, |\lambda_{q+1}(2)|, \ldots, |\lambda_{q+1}(s)|\}, \tag{6.2}$$

i.e., α is any number greater that the modulus of the subdominant eigenvalue of matrices $\mathbf{P}_\sigma$. Note we can always take α to be less than one.

Let w_k be a sequence of random variables defined in a probability space $(\Omega, \mathbb{F}, p)$. If a_k is a sequence of non-zero real numbers we will write $w_k = o(a_k)$ to denote $w_k(\omega) = o(a_k)$ for all $\omega \in \Omega$, i.e., $\lim_{k \to \infty} \dfrac{w_k(\omega)}{a_k} = 0$ for all $\omega \in \Omega$.

The following result establishes that every matrix $\mathbf{P}_\sigma^k$ is a perturbation of the corresponding $\bar{\mathbf{P}}_\sigma$.

Proposition 6.1. $\mathbf{P}_\sigma^k = \bar{\mathbf{P}}_\sigma + o(\alpha^k); \ k \to \infty$ *for all* $\sigma = 1, \ldots, s$ *with* α *given by* (6.2). *Besides* $\mathbf{P}_{\tau_n}^k = \bar{\mathbf{P}}_{\tau_n} + o(\alpha^k); \ k \to \infty$ *for all* n.

Proof. See the appendix.

The next result shows we can approximately obtain the variables corresponding to the microsystem from the knowledge of the global variables, and conversely. Those relationships turn exact if we consider the connections between the variables corresponding to the auxiliary and aggregated systems.

Proposition 6.2. *For all* $n = 1, 2, \ldots$ *we have*

(a) *variables* $\mathbf{X}_n$ *and* $\mathbf{Y}_n$ *corresponding to the auxiliary system* (4.1) *and aggregated system* (4.10), *respectively, verify*

$$\mathbf{X}_n = \mathbf{M}_{\tau_n} \mathbf{V}_{\tau_n} \mathbf{Y}_{n-1}, \quad \mathbf{Y}_n = \mathbf{U}_{\tau_{n+1}} \mathbf{X}_n,$$

(b) *variables* $\mathbf{X}_n$ *and* $\mathbf{Y}_n$ *corresponding to the microsystem* (3.3) *and aggregated system* (4.10) *respectively, are related in the following way*:

$$\begin{aligned} \mathbf{X}_n &= \mathbf{M}_{\tau_n} \mathbf{V}_{\tau_n} \mathbf{Y}_{n-1} + o(\alpha^k), \quad k \to \infty, \\ \mathbf{Y}_n &= \mathbf{U}_{\tau_{n+1}} \mathbf{X}_n + o(\alpha^k), \quad k \to \infty, \end{aligned} \tag{6.3}$$

where α *is given by* (6.2).

Proof. (a) the equality $\mathbf{Y}_n = \mathbf{U}_{\tau_{n+1}} \mathbf{X}_n$ holds by definition. Besides, we know that for the auxiliary system $\mathbf{X}_n = \mathbf{M}_{\tau_n} \bar{\mathbf{P}}_{\tau_n} \mathbf{X}_{n-1} = \mathbf{M}_{\tau_n} \mathbf{V}_{\tau_n} \mathbf{U}_{\tau_n} \mathbf{X}_{n-1} = \mathbf{M}_{\tau_n} \mathbf{V}_{\tau_n} \mathbf{Y}_{n-1}$.

(b) For the original system we have, using Proposition 6.1

$$\begin{aligned} \mathbf{X}_n &= \mathbf{M}_{\tau_n} \mathbf{P}_{\tau_n}^k \ldots \mathbf{M}_{\tau_2} \mathbf{P}_{\tau_2}^k \mathbf{M}_{\tau_1} \mathbf{P}_{\tau_1}^k \mathbf{X}_0 \\ &= \mathbf{M}_{\tau_n} (\bar{\mathbf{P}}_{\tau_n} + o(\alpha^k)) \ldots \mathbf{M}_{\tau_2} (\bar{\mathbf{P}}_{\tau_2} + o(\alpha^k)) \mathbf{M}_{\tau_1} (\bar{\mathbf{P}}_{\tau_1} + o(\alpha^k)) \mathbf{X}_0 \\ &= \mathbf{M}_{\tau_n} \bar{\mathbf{P}}_{\tau_n} \ldots \mathbf{M}_{\tau_2} \bar{\mathbf{P}}_{\tau_2} \mathbf{M}_{\tau_1} \bar{\mathbf{P}}_{\tau_1} \mathbf{X}_0 + o(\alpha^k) = \mathbf{M}_{\tau_n} \mathbf{V}_{\tau_n} \mathbf{Y}_{n-1} + o(\alpha^k). \end{aligned}$$

where in the last equality we have used a). Similarly, for the aggregated system we have

$$\begin{aligned} \mathbf{Y}_n &= \mathbf{U}_{\tau_{n+1}} \mathbf{M}_{\tau_n} \bar{\mathbf{P}}_{\tau_n} \ldots \mathbf{M}_{\tau_2} \bar{\mathbf{P}}_{\tau_2} \mathbf{M}_{\tau_1} \bar{\mathbf{P}}_{\tau_1} \mathbf{X}_0 \\ &= \mathbf{U}_{\tau_{n+1}} \mathbf{M}_{\tau_n} (\mathbf{P}_{\tau_n}^k + o(\alpha^k)) \ldots \mathbf{M}_{\tau_2} (\mathbf{P}_{\tau_2}^k + o(\alpha^k)) \mathbf{M}_{\tau_1} (\mathbf{P}_{\tau_1}^k + o(\alpha^k)) \mathbf{X}_0 \\ &= \mathbf{U}_{\tau_{n+1}} \mathbf{M}_{\tau_n} \mathbf{P}_{\tau_n}^k \ldots \mathbf{M}_{\tau_2} \mathbf{P}_{\tau_2}^k \mathbf{M}_{\tau_1} \mathbf{P}_{\tau_1}^k \mathbf{X}_0 + o(\alpha^k) = \mathbf{U}_{\tau_{n+1}} \mathbf{X}_n + o(\alpha^k). \end{aligned}$$

6.2. Markovian and i.i.d. cases

In this section we further explore the relationships between the microsystem and the aggregated system in the case the pattern of temporal variation for the former is a Markov chain. Indeed, a great part of the models of the kind (2.1) found in the literature represent environmental change through Markov chains (see among others [13–15]), for Markov chains can represent sequential dependence between environments and yet are simple enough to be analyzed in detail. 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.

Section 4.1 showed that, given hypothesis (D) holds, the number of environmental conditions and the kind of environmental change for the original and aggregated systems coincide. However, in the general case, neither the number of environmental states nor the pattern of temporal variation for the microsystem coincide with those corresponding to the aggregated system. In this section we study the relationships between those patterns of environmental stochasticity in the case in which the τ_n are a homogeneous Markov chain.

Assume the chain τ_n which characterizes the environmental change for the microsystem is defined by the following transition and initial probabilities:

$$p(\tau_{n+1} = i \mid \tau_n = j) = q_{ij}, \quad i, j = 1, \ldots, s, \ n = 1, 2, \ldots,$$

$$p(\tau_1 = i) = q^i, \quad i = 1, \ldots, s. \tag{6.4}$$

We define the stochastic matrix $\mathbf{Q} = [q_{ij}] \in \mathbb{R}^{s \times s}$ and the probability normed vector $\mathbf{q} = (q^1, \ldots, q^s)^{\mathrm{T}}$.

The following proposition, whose proof is straightforward, shows that under the above conditions, the random variables which characterize environmental variation for the aggregated system also constitute a Markov chain.

Proposition 6.3. *Assuming τ_n constitute a Markov chain defined by conditions* (6.4), *then $\bar{\tau}_n = (\tau_{n+1}, \tau_n)$ are also a Markov chain with transition and initial probabilities given by*

$$p(\bar{\tau}_{n+1} = (i,j) \mid \bar{\tau}_n = (\alpha, \beta)) = \delta_{j\alpha} q_{ij},$$

$$p(\bar{\tau}_1 = (i,j)) = q_{ij} q^j.$$

For the environmental states of the aggregated system, we consider the ordering consistent with that of (4.13), i.e.,

$$(1,1), (2,1), \ldots, (s,1), (1,2), (2,2), \ldots, (s,2), \ldots, (1,s), (2,s), \ldots, (s,s), \tag{6.5}$$

and to each pair (i,j) we associate the number l_{ij} corresponding to its place in the above ordering, i.e., $l_{ij} = (j-1)s + i$.

Then, matrix $\mathbf{R} \in \mathbb{R}^{s^2 \times s^2}$ and vector $\mathbf{t} \in \mathbb{R}^{s^2}$ which characterize, respectively, the transition probabilities and the initial probabilities for the chain corresponding to the

aggregated system, are given by

$$r(l_{ij}, l_{\alpha\beta}) = p(\bar{\tau}_{n+1} = (i,j) \mid \bar{\tau}_n = (\alpha, \beta)) = \delta_{j\alpha} q_{ij}, \quad i, j, \alpha, \beta = 1, \ldots, s, \tag{6.6}$$

$$\mathbf{t} = (q_{11}q^1, q_{21}q^1, \ldots, q_{s1}q^1, q_{12}q^2, q_{22}q^2, \ldots, q_{s2}q^2, \ldots, q_{1s}q^s, q_{2s}q^s, \ldots, q_{ss}q^s),$$

where $r(l_{ij}, l_{\alpha\beta})$ denotes the element of row l_{ij} and column $l_{\alpha\beta}$ of $\mathbf{R}$.

Regarding the environmental states for the aggregated system, the following two comments deserve attention:

(1) if i and j are such that $q_{ij} = 0$, then the state (i, j) is a forbidden state in the evolution of the environment.

(2) for all $j \neq \alpha$ it follows $p(\bar{\tau}_{n+1} = (i,j) \mid \bar{\tau}_n = (\alpha, \beta)) = 0$. Therefore, at most s^3 out of the s^4 elements of $\mathbf{R}$ will be non-zero.

As a particular case of the above situation, let us consider the case in which the pattern of environmental variation in the original system is i.i.d., in such a way that

$$p(\tau_n = \sigma) = q_\sigma, \quad \sigma = 1, \ldots, s, \ n = 1, 2, \ldots . \tag{6.7}$$

Then, the environmental variation for the aggregated system is not i.i.d, for we have from Proposition 6.3:

$$p(\bar{\tau}_{n+1} = (i,j) \mid \bar{\tau}_n = (\alpha, \beta)) = \delta_{j\alpha} q_i.$$

This is a direct consequence of the fact that the environmental state at time n in the aggregated system depends on the state of the environment in the original system at times n and $n+1$, and so the $\bar{\tau}_n$ are correlated even though the τ_n are not.

6.2.1. Moments of the population vector for original and aggregated systems

Now we are interested in relating the statistical moments of the population vector for microsystem and aggregated system. In particular, we want to obtain the dynamics of the moments corresponding to the microsystem in terms of information pertaining the aggregated system. Note that even though we know the relationships between microvariables and global variables through Proposition 6.2, we cannot derive the relationships between their moments directly from that proposition. Indeed, even in the case the τ_n are i.i.d., variables $\mathbf{M}_{\tau_n} \mathbf{V}_{\tau_n}$ and $\mathbf{Y}_{n-1}$ are correlated in (6.3), and so we can not obtain the expectation of the microvariables in terms of the expectation of the global variables.

In this section we will relate the behavior of the statistical moments of the population vector corresponding to the original and aggregated systems under the hypothesis that the environmental change in the original system is characterized by a Markov chain. In order to do so we will use the fact that the evolution of the moments of the population vector in systems of the kind (2.1) with the τ_n being an homogeneous Markov chain and the initial population vector being fixed, can be known explicitly in terms of the powers of certain matrices.

A useful tool for this task is the Kronecker matrix product, which for two matrices $\mathbf{A} = [a_{ij}] \in \mathbb{C}^{m \times n}$ and $\mathbf{B} = [b_{ij}] \in \mathbb{C}^{r \times s}$ is defined as the matrix of size $mr \times ns$

given by

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{21}\mathbf{B} & \cdots & a_{2n}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{pmatrix}, \tag{6.8}$$

i.e., $\mathbf{A} \otimes \mathbf{B}$ is a matrix with mn blocks in which the block in position i, j has the form $a_{ij}\mathbf{B}$. A notable property of the Kronecker product that we will frequently use in the sequel is (see [25]):

$$(\mathbf{A}_1 \otimes \mathbf{A}_2 \otimes \cdots \otimes \mathbf{A}_m)(\mathbf{B}_1 \otimes \mathbf{B}_2 \otimes \cdots \otimes \mathbf{B}_m) = (\mathbf{A}_1\mathbf{B}_1) \otimes (\mathbf{A}_2\mathbf{B}_2) \otimes \cdots \otimes (\mathbf{A}_m\mathbf{B}_m). \tag{6.9}$$

In order to shorten some expressions we will use the notation $\mathbf{A}^{\otimes m} = \mathbf{A} \otimes \mathbf{A} \otimes \overset{(m)}{\cdots} \otimes \mathbf{A}$ when necessary.

We will denote the expectation of a random variable z as $E(z)$. If $\mathbf{Z} = (z^1, \ldots, z^N)^{\mathrm{T}}$ is a vector random variable, the (non-central) moments of any order of the components of $\mathbf{Z}$, i.e., $E((z^i)^s(z^j)^t)$ for $s, t = 0, 1, \ldots$ can be expressed in a compact form through the use of the Kronecker product. Indeed, the moments of order m of $\mathbf{Z}$ will be given by $E(\mathbf{Z} \otimes \mathbf{Z} \otimes \overset{(m)}{\cdots} \otimes \mathbf{Z}) \in \mathbb{R}^{N^m \times 1}$. In particular, the first order moments of $\mathbf{Z}$ provide the expected value of its components, while the second order moments allow us to compute the variances and covariances of such components. For example, if $\mathbf{Z} = (z^1, z^2)^{\mathrm{T}}$, the moments of first order of $\mathbf{Z}$ will be $E(\mathbf{Z}) = (E(z^1), E(z^2))^{\mathrm{T}}$, meanwhile the moments of second order are given by $E(\mathbf{Z} \otimes \mathbf{Z}) = (E(z^1z^1), E(z^1z^2), E(z^2z^1), E(z^2z^2))^{\mathrm{T}}$.

Suppose that the chain τ_n is defined by conditions (6.4). Then (see [37]), if we assume that the initial population vector is some fixed $\mathbf{z}_0$ we have that, for all natural m and n

$$E(\mathbf{Z}_n \otimes \overset{(m)}{\cdots} \otimes \mathbf{Z}_n) = \mathbf{J}_m\mathbf{D}_m^{n-1}\mathbf{F}_m\mathbf{S}_m(\mathbf{z}_0 \otimes \overset{(m)}{\cdots} \otimes \mathbf{z}_0) \in \mathbb{R}^{N^m \times N^m}, \tag{6.10}$$

where

$$\mathbf{D}_m = \mathbf{F}_m(\mathbf{Q} \otimes \mathbf{I}_{N^m}) \in \mathbb{R}^{sN^m \times sN^m},$$

$$\mathbf{F}_m = diag(\mathbf{A}_1 \otimes \overset{(m)}{\cdots} \otimes \mathbf{A}_1, \ldots, \mathbf{A}_s \otimes \overset{(m)}{\cdots} \otimes \mathbf{A}_s) \in \mathbb{R}^{sN^m \times sN^m},$$

$$\mathbf{J}_m = (\mathbf{I}_{N^m}|\mathbf{I}_{N^m}|\overset{(s)}{\cdots}|\mathbf{I}_{N^m}) \in \mathbb{R}^{N^m \times sN^m},$$

$$\mathbf{S}_m = (q^1\mathbf{I}_{N^m}|q^2\mathbf{I}_{N^m}|\overset{(s)}{\cdots}|q^s\mathbf{I}_{N^m})^{\mathrm{T}} \in \mathbb{R}^{sN^m \times N^m}.$$

Therefore, for each m the evolution of the moments of order m of $\mathbf{Z}_n$ obeys a linear dynamics characterized by the powers of the non-negative matrix $\mathbf{D}_m$. This matrix is composed of $s \times s$ blocks in such a way that the block element α, β is $(\mathbf{D}_m)_{\alpha\beta} = q_{\alpha\beta}\mathbf{A}_\alpha \otimes \overset{(m)}{\cdots} \otimes \mathbf{A}_\alpha$. So, if $\mathbf{D}_m$ has a simple and strictly dominant eigenvalue λ associated to right and left eigenvectors $\mathbf{r}_m$ and $\mathbf{l}_m$ (both of size $sN^m \times 1$), we have that the asymptotic behavior of the moments of order m of the system is given by

$$\lim_{n\to\infty} \frac{E(\mathbf{Z}_n \otimes \ldots^{(m)} \otimes \mathbf{Z}_n)}{\lambda^n} = \frac{1}{\lambda}\frac{\langle \mathbf{l}_m, \mathbf{F}_m\mathbf{S}_m(\mathbf{z}_0 \otimes \ldots^{(m)} \otimes \mathbf{z}_0)\rangle}{\langle \mathbf{l}_m, \mathbf{r}_m\rangle}\mathbf{J}_m\mathbf{r}_m. \tag{6.11}$$

Therefore we have: (a) The vector of moments grows asymptotically in an exponential fashion with growth rate λ. (b) If we consider $\mathbf{r}_m$ divided in s column blocks $\mathbf{r}_m^i$ each of them of dimensions $N^m \times 1$, the asymptotic structure of the vector of moments is determined by vector $\mathbf{J}_m\mathbf{r}_m = \mathbf{r}_m^1 + \cdots + \mathbf{r}_m^s$. In particular, the expected population size grows asymptotically at a rate given by the dominant eigenvalue of $\bar{\mathbf{D}}_1$.

In the remainder of this section we will relate the evolution and the asymptotic behavior of the moments of the original, aggregated and auxiliary systems. As an initial condition for the original and auxiliary systems we will impose that the initial population vector is some fixed $\mathbf{a} \in \mathbb{R}^{N \times 1}$, i.e., $\mathbf{X}_0(\omega) = \mathbf{a}$ for all $\omega \in \Omega$. Note that for the aggregated system we will have $\mathbf{Y}_0 = \mathbf{U}_{\tau_1}\mathbf{a}$ and therefore the initial population for the original system is a (non-constant) random variable.

Let m be fixed. Following (6.10), the evolution of the moments of order m of the original and auxiliary systems is given by

$$\text{Original system}: E(\mathbf{X}_n \otimes \overset{(m)}{\cdots} \otimes \mathbf{X}_n) = \mathbf{J}_m\mathbf{D}_m^{n-1}(k)\mathbf{F}_m(k)\mathbf{S}_m(\mathbf{a} \otimes \overset{(m)}{\cdots} \otimes \mathbf{a}),$$
$$\text{Auxiliary system}: E(\mathbf{X}_n \otimes \overset{(m)}{\cdots} \otimes \mathbf{X}_n) = \mathbf{J}_m\mathbf{D}_m'^{n-1}\mathbf{F}_m'\mathbf{S}_m(\mathbf{a} \otimes \overset{(m)}{\cdots} \otimes \mathbf{a}), \tag{6.12}$$

where the different matrices are defined by

$$\mathbf{F}_m(k) = diag(\mathbf{M}_1\mathbf{P}_1^k \otimes \overset{(m)}{\cdots} \otimes \mathbf{M}_1\mathbf{P}_1^k, \ldots, \mathbf{M}_s\mathbf{P}_s^k \otimes \overset{(m)}{\cdots} \otimes \mathbf{M}_s\mathbf{P}_s^k) \in \mathbb{R}^{sN^m \times sN^m},$$
$$\mathbf{F}_m' = diag(\mathbf{M}_1\bar{\mathbf{P}}_1 \otimes \overset{(m)}{\cdots} \otimes \mathbf{M}_1\bar{\mathbf{P}}_1, \ldots, \mathbf{M}_s\bar{\mathbf{P}}_s \otimes \overset{(m)}{\cdots} \otimes \mathbf{M}_s\bar{\mathbf{P}}_s) \in \mathbb{R}^{sN^m \times sN^m},$$
$$\mathbf{D}_m(k) = \mathbf{F}_m(k)(\mathbf{Q} \otimes \mathbf{I}_{N^m}) \in \mathbb{R}^{sN^m \times sN^m},$$
$$\mathbf{D}_m' = \mathbf{F}_m'(\mathbf{Q} \otimes \mathbf{I}_{N^m}) \in \mathbb{R}^{sN^m \times sN^m},$$
$$\mathbf{J}_m = (\mathbf{I}_{N^m}|\mathbf{I}_{N^m}|\overset{(s)}{\cdots}|\mathbf{I}_{N^m}) \in \mathbb{R}^{N^m \times sN^m},$$
$$\mathbf{S}_m = (q^1\mathbf{I}_{N^m}|q^2\mathbf{I}_{N^m}|\overset{(s)}{\cdots}|q^s\mathbf{I}_{N^m})^{\mathrm{T}} \in \mathbb{R}^{sN^m \times N^m}. \tag{6.13}$$

In this way, the powers of matrices $\mathbf{D}_m(k)$ and $\mathbf{D}_m'$ characterize the evolution of the moments of order m for the original and auxiliary systems, respectively.

In order to characterize the vector of moments for the aggregated system (4.10) we could proceed in the same way and build appropriate matrices according to (6.10). However, and due to the great number of environmental states in the aggregated system (s^2 states), this would imply working with matrices of great dimensions. For example, the matrix whose powers would define the behavior of the moments of the system, would have size $s^2q^m \times s^2q^m$. We will follow an alternative approach and will deduce a way of calculating the moments of the population vector for the aggregated system through the powers of matrices of size $sq^m \times sq^m$. Indeed we have

$$E(\mathbf{Y}_n \otimes \overset{(m)}{\cdots} \otimes \mathbf{Y}_n) =$$
$$\sum_{\omega} prob(\omega)\bar{\mathbf{M}}_{\bar{\tau}_n(\omega)}^{\otimes m} \ldots \bar{\mathbf{M}}_{\bar{\tau}_2(\omega)}^{\otimes m}\bar{\mathbf{M}}_{\bar{\tau}_1(\omega)}^{\otimes m}(\mathbf{Y}_0(\omega) \otimes \overset{(m)}{\cdots} \otimes \mathbf{Y}_0(\omega))$$

$$= \sum_{i_1,\dots,i_{n+1}=1}^{s} q_{i_{n+1}i_n} q_{i_n i_{n-1}} \dots q_{i_2 i_1} q^{i_1} \bar{\mathbf{M}}_{i_{n+1}i_n}^{\otimes m} \bar{\mathbf{M}}_{i_n i_{n-1}}^{\otimes m} \dots \bar{\mathbf{M}}_{i_2 i_1}^{\otimes m} \mathbf{U}_{i_1}^{\otimes m} (\mathbf{a} \otimes \overset{(m)}{\dots} \otimes \mathbf{a})$$

$$= \sum_{i_1,\dots,i_{n+1}=1}^{s} (q_{i_{n+1}i_n} \bar{\mathbf{M}}_{i_{n+1}i_n}^{\otimes m})(q_{i_n i_{n-1}} \bar{\mathbf{M}}_{i_n i_{n-1}}^{\otimes m}) \dots (q_{i_2 i_1} \bar{\mathbf{M}}_{i_2 i_1}^{\otimes m})(q^{i_1} \mathbf{U}_{i_1}^{\otimes m})(\mathbf{a} \otimes \overset{(m)}{\dots} \otimes \mathbf{a}).$$

$$(6.14)$$

We define matrix $\bar{\mathbf{C}}_m \in \mathbb{R}^{sq^m \times sq^m}$ as a matrix of $s \times s$ blocks in such a way that the block element α, β is $(\bar{\mathbf{C}}_m)_{\alpha\beta} = q_{\alpha\beta} \bar{\mathbf{M}}_{\alpha\beta}^{\otimes m} = q_{\alpha\beta} \bar{\mathbf{M}}_{\alpha\beta} \otimes \overset{(m)}{\dots} \otimes \bar{\mathbf{M}}_{\alpha\beta}$ and matrices

$$\bar{\mathbf{J}}_m = (\mathbf{I}_{q^m} | \mathbf{I}_{q^m} | \overset{(s)}{\dots} | \mathbf{I}_{q^m}) \in \mathbb{R}^{q^m \times sq^m},$$

$$\mathbf{G}_m = diag(\mathbf{U}_1 \otimes \overset{(m)}{\dots} \otimes \mathbf{U}_1, \dots, \mathbf{U}_s \otimes \overset{(m)}{\dots} \otimes \mathbf{U}_s) \in \mathbb{R}^{sq^m \times sN^m}.$$

Then, it is easy to check that we can write (6.14) in the form

$$E(\mathbf{Y}_n \otimes \overset{(m)}{\dots} \otimes \mathbf{Y}_n) = \bar{\mathbf{J}}_m \bar{\mathbf{C}}_m^n \mathbf{G}_m \mathbf{S}_m (\mathbf{a} \otimes \overset{(m)}{\dots} \otimes \mathbf{a}) \tag{6.15}$$

and therefore the moments of order m of the aggregated system can be calculated in terms of the powers of matrix $\bar{\mathbf{C}}_m$.

The following paragraphs will deal with some relationships (in particular spectral relationships) between matrices $\mathbf{D}_m(k)$, $\mathbf{D}'_m$ and $\bar{\mathbf{C}}_m$ which will allow us to relate the moments of the population vectors for the three systems.

We introduce two matrices that will be helpful in the subsequent developments:

$$\mathbf{T}_m = \mathbf{G}_m(\mathbf{Q} \otimes \mathbf{I}_{N^m}) \in \mathbb{R}^{sq^m \times sN^m},$$

$$\mathbf{H}_m = diag(\mathbf{M}_1 \mathbf{V}_1 \otimes \overset{(m)}{\dots} \otimes \mathbf{M}_1 \mathbf{V}_1, \dots, \mathbf{M}_s \mathbf{V}_s \otimes \overset{(m)}{\dots} \otimes \mathbf{M}_s \mathbf{V}_s) \in \mathbb{R}^{sN^m \times sq^m}.$$

Next proposition relates some properties of matrices $\bar{\mathbf{C}}_m$ and $\mathbf{D}'_m$ including relationships between their eigenelements (i.e., eigenvalues and eigenvectors).

Proposition 6.4. *Matrices $\bar{\mathbf{C}}_m$ and $\mathbf{D}'_m$ verify*:
(a) $\bar{\mathbf{C}}_m = \mathbf{T}_m \mathbf{H}_m$; $\mathbf{D}'_m = \mathbf{H}_m \mathbf{T}_m$,
(b) $\bar{\mathbf{C}}_m^n = \mathbf{T}_m \mathbf{D}_m'^{n-1} \mathbf{H}_m$ *and* $\mathbf{D}_m'^n = \mathbf{H}_m \bar{\mathbf{C}}_m^{n-1} \mathbf{T}_m$,
(c) $\det(\lambda \mathbf{I}_{sN^m} - \mathbf{D}'_m) = \lambda^{s(N^m-q^m)} \det(\lambda \mathbf{I}_{sq^m} - \bar{\mathbf{C}}_m)$; *in particular, the dominant eigenvalues of both matrices, including multiplicities, coincide.*
(d) *If* $\mathbf{r}$ *and* $\mathbf{l}$ *are, respectively, right and left eigenvalues of* $\bar{\mathbf{C}}_m$ *associated to* $\lambda \neq 0$ *then* $\mathbf{H}_m \mathbf{r}$ *and* $\mathbf{T}_m^{\mathrm{T}} \mathbf{l}$ *are, respectively, right and left eigenvectors of* $\mathbf{D}'_m$ *associated to* λ.

Proof. (a) Straightforward from the expressions of $\mathbf{T}_m$ and $\mathbf{H}_m$, property (6.9) and the fact that $\mathbf{V}_\sigma \mathbf{U}_\sigma = \bar{\mathbf{P}}_\sigma$ for all σ. (b) Trivial from (a). (c) Direct consequence of the fact that the non-zero eigenvalues, including multiplicities, of $\mathbf{T}_m \mathbf{H}_m$ coincide with those of $\mathbf{H}_m \mathbf{T}_m$ (see [19]). (d) We know $\bar{\mathbf{C}}_m \mathbf{r} = \lambda \mathbf{r} \neq \mathbf{0}$, i.e., $\mathbf{T}_m \mathbf{H}_m \mathbf{r} = \lambda \mathbf{r} \neq \mathbf{0}$ (*) so it must be $\mathbf{H}_m \mathbf{r} \neq \mathbf{0}$. Multiplying on the left by $\mathbf{H}_m$ and using (a) we have

$$\mathbf{H}_m \mathbf{T}_m \mathbf{H}_m \mathbf{r} = \mathbf{D}'_m \mathbf{H}_m \mathbf{r} = \lambda \mathbf{H}_m \mathbf{r}.$$

Analogously, we know $\mathbf{l}^T\bar{\mathbf{C}}_m = \lambda\mathbf{l}^T \neq \mathbf{0}$, i.e., $\mathbf{l}^T\mathbf{T}_m\mathbf{H}_m = \lambda\mathbf{l}^T \neq \mathbf{0}$, so $\mathbf{l}^T\mathbf{U}_{gm} \neq \mathbf{0}$. Multiplying on the right by $\mathbf{T}_m$ and using (a) we have

$$\mathbf{l}^T\mathbf{T}_m\mathbf{H}_m\mathbf{T}_m = \mathbf{l}^T\mathbf{T}_m\mathbf{H}_m = \lambda\mathbf{l}^T\mathbf{T}_m$$

as we wanted to show. $\quad\square$

Next, we relate the spectral properties of matrices $\mathbf{D}_m(k)$ and $\mathbf{D}'_m$ considering the former as a perturbation of the latter. In order to see this intuitively, we write

$$\mathbf{D}_m(k) - \mathbf{D}'_m = (\mathbf{F}_m(k) - \mathbf{F}'_m)(\mathbf{Q} \otimes \mathbf{I}_{N^2}).$$

Matrix $\mathbf{F}_m(k) - \mathbf{F}'_m$ is block diagonal. Bearing in mind property (6.9), its σth diagonal block has the form

$$(\mathbf{M}_\sigma \otimes \overset{(m)}{\cdots} \otimes \mathbf{M}_\sigma)(\mathbf{P}_\sigma^k \otimes \overset{(m)}{\cdots} \otimes \mathbf{P}_\sigma^k - \bar{\mathbf{P}}_\sigma \otimes \overset{(m)}{\cdots} \otimes \bar{\mathbf{P}}_\sigma).$$

The basic idea is that the second factor can be made arbitrarily small for big k using the fact that $\mathbf{P}_\sigma^k$ can be made arbitrarily close to $\bar{\mathbf{P}}_\sigma$, and therefore $\mathbf{D}_m(k)$ is indeed a perturbation of $\mathbf{D}'_m$.

Proposition 6.5. *For all* $m = 1, 2, \ldots$

$$\mathbf{D}_m(k) - \mathbf{D}'_m = o(\alpha^k), \quad k \to \infty,$$

$$\mathbf{F}_m(k) - \mathbf{F}'_m = o(\alpha^k), \quad k \to \infty,$$

where α is any number verifying (6.2).

Proof. See the appendix. $\quad\square$

The following proposition shows that the moments of the population vector corresponding to the original and auxiliary systems can be evaluated in terms of the powers of matrix $\bar{\mathbf{C}}_m$ that characterizes the evolution of those moments for the aggregated system.

Proposition 6.6. *For all* $m, n = 1, 2, \ldots$ *we have*
(1) *Aggregated system*: $E(\mathbf{Y}_n \otimes \overset{(m)}{\cdots} \otimes \mathbf{Y}_n) = \bar{\mathbf{J}}_m\bar{\mathbf{C}}_m^n\mathbf{G}_m\mathbf{S}_m(\mathbf{a} \otimes \overset{(m)}{\cdots} \otimes \mathbf{a})$.
(2) *Auxiliary system*: $E(\mathbf{X}_n \otimes \overset{(m)}{\cdots} \otimes \mathbf{X}_n) = \mathbf{J}_m\mathbf{H}_m\bar{\mathbf{C}}_m^{n-1}\mathbf{G}_m\mathbf{S}_m(\mathbf{a} \otimes \overset{(m)}{\cdots} \otimes \mathbf{a})$.
(3) *Original system*: $E(\mathbf{X}_n \otimes \overset{(m)}{\cdots} \otimes \mathbf{X}_n) = \mathbf{J}_m\mathbf{H}_m\bar{\mathbf{C}}_m^{n-1}\mathbf{G}_m\mathbf{S}_m(\mathbf{a} \otimes \overset{(m)}{\cdots} \otimes \mathbf{a}) + o(\alpha^k)$ *where α is any number verifying (6.2).*

Proof. We already have (1). Condition (2) follows trivially using Proposition 6.4. In order to obtain (3) we write, for the original system and n and m fixed:

$$E(\mathbf{X}_n \otimes \overset{(m)}{\cdots} \otimes \mathbf{X}_n) = \mathbf{J}_m\mathbf{D}_m^{n-1}(k)\mathbf{F}_m(k)\mathbf{S}_m(\mathbf{a} \otimes \overset{(m)}{\cdots} \otimes \mathbf{a})$$

$$= \mathbf{J}_m(\mathbf{D}'_m + o(\alpha^k))^{n-1}(\mathbf{F}'_m + o(\alpha^k))\mathbf{S}_m(\mathbf{a} \otimes \overset{(m)}{\cdots} \otimes \mathbf{a})$$

$$= \mathbf{J}_m \mathbf{D}_m'^{n-1} \mathbf{F}_m' \mathbf{S}_m (\mathbf{a} \otimes \overset{(m)}{\ldots} \otimes \mathbf{a}) + o(\alpha^k)$$

$$= \mathbf{J}_m \mathbf{H}_m \bar{\mathbf{C}}_m^{n-1} \mathbf{G}_m \mathbf{S}_m (\mathbf{a} \otimes \overset{(m)}{\ldots} \otimes \mathbf{a}) + o(\alpha^k),$$

where we have used Proposition 6.5. □

The asymptotic behavior of the moments of order m of the original, auxiliary and aggregated systems is determined by the dominant eigenelements of matrices $\mathbf{D}_m(k)$, $\mathbf{D}_m'$ and $\bar{\mathbf{C}}_m$, respectively. In the following, we will obtain results that allow us to approximate the asymptotic behavior of the moments of any order for the original system in terms of the dominant eigenelements corresponding to matrices $\bar{\mathbf{C}}_m$.

We introduce an assumption that guarantees that the moments of order m of the aggregated system tend to an asymptotic structure:

H1. *Matrix* $\bar{\mathbf{C}}_m$ *has an eigenvalue* λ *which is simple and strictly dominant, with associated right and left non-negative eigenvectors* $\mathbf{r}$ *and* $\mathbf{l}$, *respectively.*

Assume the original system has a non-negative initial condition $\mathbf{a}$. Using (6.15) we have that hypothesis H1 guarantees that the aggregated system verifies

$$\lim_{n \to \infty} \frac{E(\mathbf{Y}_n \otimes \ldots^{(m)} \otimes \mathbf{Y}_n)}{\lambda^n} = \lim_{n \to \infty} \bar{\mathbf{J}}_m \left(\frac{\bar{\mathbf{C}}_m}{\lambda} \right)^n \mathbf{G}_m \mathbf{S}_m \mathbf{a}$$

$$= \frac{\langle \mathbf{l}, \mathbf{G}_m \mathbf{S}_m (\mathbf{a} \otimes \ldots^{(m)} \otimes \mathbf{a}) \rangle}{\langle \mathbf{l}, \mathbf{r} \rangle} \bar{\mathbf{J}}_m \mathbf{r}_m.$$

Now, the asymptotic behavior of the moments of order m of for the auxiliary system is given by the following result:

Proposition 6.7. *Let m be such that H1 holds. Then, for any initial condition $\mathbf{a} \geq 0$ for the auxiliary system, the asymptotic behavior of the moments of order m of this system is characterized by*

$$\lim_{n \to \infty} \frac{E(\mathbf{X}_n \otimes \ldots^{(m)} \otimes \mathbf{X}_n)}{\lambda^n} = \frac{1}{\lambda} \frac{\langle \mathbf{l}, \mathbf{G}_m \mathbf{S}_m (\mathbf{a} \otimes \ldots^{(m)} \otimes \mathbf{a}) \rangle}{\langle \mathbf{l}, \mathbf{r} \rangle} \mathbf{J}_m \mathbf{H}_m \mathbf{r}_m.$$

Proof. Absolutely analogous to the proof of Proposition 6.9 below. □

The following proposition characterizes the dominant eigenelements of matrix $\mathbf{C}_m(k)$.

Proposition 6.8. *Let m be such that $\bar{\mathbf{C}}_m$ verifies H1. Then, for k big enough, matrix $\mathbf{D}_m(k)$ has a simple and strictly dominant eigenvalue $\lambda(k)$ that can be expressed in the form*

$$\lambda(k) = \lambda + o(\alpha^k), \quad \alpha \to \infty$$

with α given by (6.2). Besides, associated to $\lambda(k)$ there are right and left eigenvectors that can be written, respectively, in the form

$$\mathbf{H}_m \mathbf{r} + o(\alpha^k), \quad \mathbf{T}_m^{\mathrm{T}} \mathbf{l} + o(\alpha^k).$$

Proof. See the appendix. □

The proposition above guarantees that the asymptotic growth rate and the structure of the moments of any order of the population vector for the microsystem can be approximated if we know its analogues for the aggregated system. The larger the separation of time scales is, the sharper the approximation will be. Indeed we have:

Proposition 6.9. *Let m be such that* H1 *holds. Then, for k big enough and for any initial condition* $\mathbf{a} \geq 0$ *for the original system, the asymptotic behavior of the moments of order m of this system is characterized by*

$$\lim_{n \to \infty} \frac{E(\mathbf{X}_n \otimes \ldots^{(m)} \otimes \mathbf{X}_n)}{(\lambda + o(\alpha^k))^n} = \frac{1}{\lambda} \frac{\langle \mathbf{l}, \mathbf{G}_m \mathbf{S}_m(\mathbf{a} \otimes \ldots^{(m)} \otimes \mathbf{a}) \rangle}{\langle \mathbf{l}, \mathbf{r} \rangle} \mathbf{J}_m \mathbf{H}_m \mathbf{r}_m + o(\alpha^k).$$

Proof. From (6.11) and Propositions 6.5 and 6.8 we have that, for big enough k,

$$\lim_{n \to \infty} \frac{E(\mathbf{X}_n \otimes \ldots^{(m)} \otimes \mathbf{X}_n)}{(\lambda(k))^n}$$

$$= \frac{1}{\lambda + o(\alpha^k)} \mathbf{J}_m \frac{(\mathbf{H}_m \mathbf{r} + o(\alpha^k))(\mathbf{l}^\mathsf{T} \mathbf{T}_m + o(\alpha^k))}{(\mathbf{l}^\mathsf{T} \mathbf{T}_m + o(\alpha^k))(\mathbf{H}_m \mathbf{r} + o(\alpha^k))} (\mathbf{F}'_m + o(\alpha^k)) \mathbf{S}_m(\mathbf{a} \otimes \overset{(m)}{\ldots} \otimes \mathbf{a})$$

$$= \left(\frac{1}{\lambda} + o(\alpha^k) \right) \mathbf{J}_m \frac{\mathbf{H}_m \mathbf{r} \mathbf{l}^\mathsf{T} \mathbf{T}_m + o(\alpha^k)}{\mathbf{l}^\mathsf{T} \mathbf{T}_m \mathbf{H}_m \mathbf{r} + o(\alpha^k)} (\mathbf{F}'_m \mathbf{S}_m(\mathbf{a} \otimes \overset{(m)}{\ldots} \otimes \mathbf{a}) + o(\alpha^k))$$

$$= \left(\frac{1}{\lambda} + o(\alpha^k) \right) \mathbf{J}_m \left(\frac{\mathbf{H}_m \mathbf{r} \mathbf{l}^\mathsf{T} \mathbf{T}_m}{\mathbf{l}^\mathsf{T} \bar{\mathbf{C}}_m \mathbf{r}} + o(\alpha^k) \right) (\mathbf{H}_m \mathbf{G}_m \mathbf{S}_m(\mathbf{a} \otimes \overset{(m)}{\ldots} \otimes \mathbf{a}) + o(\alpha^k))$$

$$= \frac{1}{\lambda} \mathbf{J}_m \frac{\mathbf{H}_m \mathbf{r} \mathbf{l}^\mathsf{T} \bar{\mathbf{C}}_m}{\lambda \mathbf{l}^\mathsf{T} \mathbf{r}} \mathbf{G}_m \mathbf{S}_m(\mathbf{a} \otimes \overset{(m)}{\ldots} \otimes \mathbf{a}) + o(\alpha^k)$$

$$= \frac{1}{\lambda} \mathbf{J}_m \frac{\lambda \mathbf{H}_m \mathbf{r} \mathbf{l}^\mathsf{T}}{\lambda \mathbf{l}^\mathsf{T} \mathbf{r}} \mathbf{G}_m \mathbf{S}_m(\mathbf{a} \otimes \overset{(m)}{\ldots} \otimes \mathbf{a}) + o(\alpha^k)$$

$$= \frac{1}{\lambda} \frac{\langle \mathbf{l}, \mathbf{G}_m \mathbf{S}_m(\mathbf{a} \otimes \ldots^{(m)} \otimes \mathbf{a}) \rangle}{\langle \mathbf{l}, \mathbf{r} \rangle} \mathbf{J}_m \mathbf{H}_m \mathbf{r}_m + o(\alpha^k),$$

where we have used $\bar{\mathbf{C}}_m = \mathbf{T}_m \mathbf{H}_m$ and, as it is trivial to show, $\mathbf{F}'_m = \mathbf{H}_m \mathbf{G}_m$. □

To illustrate how one can study the influence of the fast and slow processes in the system through the dynamics of the aggregated system, let us consider the multiregional model with fast demography (5.4) presented in Section 5.2. Migration is a conservative process for the total number of individuals in the population, so if we consider the stochastic multiregional model in the case migration is the only process acting, the system will keep its total population constant. However, if we consider also

demography, the picture considerably changes. For example, given a Markovian pattern of environmental variation, the matrix $\bar{\mathbf{C}}_1$ corresponding to the aggregated system (5.5) will not in general have 1 as its dominant eigenvalue. Consequently, the multiregional model will have an expected population size that, depending on the value of the migratory and demographic rates in the different environments, either decays to zero or grows exponentially.

7. Discussion

This work develops a method which allows one to treat a stochastic complex system in which there are two time scales involved to obtain a reduced model. The variables which govern the aggregated system are certain linear combinations of the state variables of the original system, being the coefficients dependent of the characteristics of the fast process. The parameters of the aggregated system can be easily expressed as functions of the slow dynamics and of the equilibrium proportions of individuals corresponding to the fast process. The variables corresponding to the original and reduced systems can be related in an easy way, therefore allowing one to study the dynamics of the former by that of the latter. Moreover, in the Markovian case it is possible to relate the moments of the population vector for both systems. In this way, it is possible to study how changes in both the fast and slow processes affect the dynamics of the aggregated system.

There are some features related to the dynamics of systems of kind (2.1) that would be interesting to address in future contributions to check whether we can be related for the original and reduced models. The first is the existence of stochastic ergodicity, roughly the capacity of the distribution of the structure of the population to forget initial conditions [13,14]. Besides, we plan to study parameters as the asymptotic expectation and variance for the logarithm of total population size [36], which approximately characterize the probability of extinction of the population [21]. The former, also known as the stochastic growth rate, characterizes the expected growth rate over the different realizations of the process and plays a major role in the study of these kinds of systems [37,38].

Appendix

Proof of Proposition 6.1. Let σ be fixed and let us consider a Jordan canonical decomposition of $\mathbf{P}_\sigma$. Eigenvalue 1 is simple and strictly dominant for each of the $\mathbf{P}_i(\sigma)$ and is associated to right and left eigenvectors $\mathbf{v}_i(\sigma)$ and $\mathbf{u}_i(\sigma)$. Therefore, we have that for matrix $\mathbf{P}_\sigma$ eigenvalue 1 is strictly dominant, semisimple and has multiplicity q. Besides the columns of $\mathbf{V}_\sigma$ and the rows of $\mathbf{U}_\sigma$ are bases of its associated right and left eigenspaces, respectively. Since $\mathbf{U}_\sigma \mathbf{V}_\sigma = \mathbf{I}_q$ we have that a Jordan decomposition of $\mathbf{P}_\sigma$ with eigenvalues ordered by decreasing modulus will have the form

$$\mathbf{P}_\sigma = (\mathbf{V}_\sigma \mid \mathbf{V}'_\sigma)\, diag(\mathbf{I}_q, \mathbf{H}) \left(\frac{\mathbf{U}_\sigma}{\mathbf{U}'_\sigma} \right)$$

where $\mathbf{V}'_\sigma$ and $\mathbf{U}'_\sigma$ are appropriate matrices and $\mathbf{H}$ corresponds to Jordan blocks associated to eigenvalues $\lambda_{q+1}(\sigma),\ldots,\lambda_N(\sigma)$ (of modulus strictly less than α). Therefore, taking into account that $\bar{\mathbf{P}}_\sigma = \mathbf{V}_\sigma \mathbf{U}_\sigma$ we have

$$\mathbf{P}^k_\sigma = \bar{\mathbf{P}}_\sigma + (\mathbf{V}_\sigma \mid \mathbf{V}'_\sigma)\,diag(\mathbf{0}, \mathbf{H}^k)\left(\frac{\mathbf{U}_\sigma}{\mathbf{U}'_\sigma}\right)$$

and so $\mathbf{P}^k_\sigma - \bar{\mathbf{P}}_\sigma = o(\alpha^k)$, $k \to \infty$. Since the state space of τ_n is $\{1,\ldots,s\}$, it follows that for all n and all $\omega \in \Omega$, $\mathbf{P}^k_{\tau_n(\omega)} = \bar{\mathbf{P}}_{\tau_n(\omega)} + o(\alpha^k)$; $k \to \infty$ as we wanted to show. $\square$

Proof of Proposition 6.5. Let $\sigma \in \{1,\ldots,s\}$ be fixed. The σ-th diagonal block of $\mathbf{F}_m(k) - \mathbf{F}'_m$ is $(\mathbf{M}_\sigma \otimes \overset{(m)}{\cdots} \otimes \mathbf{M}_\sigma)(\mathbf{P}^k_\sigma \otimes \overset{(m)}{\cdots} \otimes \mathbf{P}^k_\sigma - \bar{\mathbf{P}}_\sigma \otimes \overset{(m)}{\cdots} \otimes \bar{\mathbf{P}}_\sigma)$. Using Lemma 3.1 and proceeding by induction, it is trivial to show that the second factor can be expressed in the form $(\mathbf{P}_\sigma \otimes \overset{(m)}{\cdots} \otimes \mathbf{P}_\sigma - \bar{\mathbf{P}}_\sigma \otimes \overset{(m)}{\cdots} \otimes \bar{\mathbf{P}}_\sigma)^k$. We denote $\Delta_\sigma = (\mathbf{P}_\sigma \otimes \overset{(m)}{\cdots} \otimes \mathbf{P}_\sigma - \bar{\mathbf{P}}_\sigma \otimes \overset{(m)}{\cdots} \otimes \bar{\mathbf{P}}_\sigma)$. Let us consider a Jordan canonical decomposition of $\mathbf{P}_\sigma$ with eigenvalues ordered by decreasing modulus. Bearing (6.1) in mind, it follows that this decomposition will have the form $\mathbf{P}_\sigma = \mathbf{S}\,diag(\mathbf{I}_q, \mathbf{H}_\sigma)\mathbf{S}^{-1}$ where $\mathbf{S}$ is an appropriate matrix and $\mathbf{H}_\sigma$ corresponds to upper triangular Jordan blocks associated to eigenvalues $\lambda_{q+1}(\sigma),\ldots,\lambda_N(\sigma)$ (of modulus strictly less than α). Since $\bar{\mathbf{P}}_\sigma = \lim_{k\to\infty}\mathbf{P}^k_\sigma$ we have $\bar{\mathbf{P}}_\sigma = \mathbf{S}\,diag(\mathbf{I}_q, \mathbf{0})\mathbf{S}^{-1}$ and consequently it follows

$$\Delta_\sigma = (\mathbf{S} \otimes \overset{(m)}{\cdots} \otimes \mathbf{S})(\mathbf{A}_\sigma - \mathbf{B}_\sigma)(\mathbf{S} \otimes \overset{(m)}{\cdots} \otimes \mathbf{S})^{-1}, \tag{7.1}$$

where

$$\mathbf{A}_\sigma = diag(\mathbf{I}_q, \mathbf{H}_\sigma) \otimes \overset{(m)}{\cdots} \otimes diag(\mathbf{I}_q, \mathbf{H}_\sigma); \quad \mathbf{B}_\sigma = diag(\mathbf{I}_q, \mathbf{0}) \otimes \overset{(m)}{\cdots} \otimes diag(\mathbf{I}_q, \mathbf{0})$$

and where we have used that the inverse of the Kronecker product is the Kronecker product of the inverses [16, p. 138]. The direct product of upper triangular matrices is upper triangular and its diagonal elements are the product "all with all" and ordered in a certain sense of the diagonal elements of the factors of the product [16, p. 138]. Therefore, since $\mathbf{H}_\sigma$ has diagonal elements $\lambda_{q+1}(\sigma), \lambda_{q+2}(\sigma),\ldots,\lambda_N(\sigma)$, matrix $\mathbf{A}_\sigma$ will be upper triangular and its diagonal entries will consist (in a certain order which is nor relevant for our discussion) in q^m ones and $N^m - q^m$ elements of the kind $\lambda_{i_1}(\sigma)\lambda_{i_2}(\sigma)\ldots\lambda_{i_m}(\sigma)$ where the $i_j \in \{1,\ldots,N\}$; $j = 1,\ldots,m$ and at most $m-1$ of the i_j are equal to one. In the same way, matrix $\mathbf{B}_\sigma$ will be upper triangular and its diagonal entries will be q^m ones, which will be placed in the same positions that those corresponding to $\mathbf{A}_\sigma$, and $N^m - q^m$ zeroes. Therefore, $\mathbf{A}_\sigma - \mathbf{B}_\sigma$ is upper triangular and its diagonal elements, i.e., its eigenvalues are q^m zeroes and $N^m - q^m$ elements of the kind $\lambda_{i_1}(\sigma)\lambda_{i_2}(\sigma)\ldots\lambda_{i_m}(\sigma)$ where the $i_j \in \{1,\ldots,N\}$; $j = 1,\ldots,m$ and at most $m-1$ of the i_j are equal to one. Since all $|\lambda_i(\sigma)| < 1$ and $|\lambda_i(\sigma)| < \alpha$ for $i \in \{q+1,\ldots,N\}$ we have that the spectral radius of $\mathbf{A}_\sigma - \mathbf{B}_\sigma$ is less than α. From (7.1) we have Δ_σ is similar to $\mathbf{A}_\sigma - \mathbf{B}_\sigma$, so $\rho(\Delta_\sigma) < \alpha$ and therefore $\lim_{k\to\infty}(\Delta_\sigma/\alpha)^k = 0$ for all σ. We have then

$$\mathbf{D}_m(k) - \mathbf{D}'_m = diag(\mathbf{M}_1 \otimes \cdots \otimes \mathbf{M}_1 \Delta^k_1,\ldots, \mathbf{M}_s \otimes \cdots \otimes \mathbf{M}_s \Delta^k_s)(\mathbf{Q} \otimes \mathbf{I}_{N^2})$$

and dividing by α^k and taking limits $k \to \infty$, the desired result follows. $\square$

Proof of Proposition 6.8. We know (Proposition 6.5) that for any consistent matrix norm $\| * \|$ we have $\|\mathbf{D}_m(k) - \mathbf{D}'_m\| = \mathbf{E}$ with $\mathbf{E} = o(\alpha^k)$. Therefore, using Theorem 7.1 (appendix) and Proposition 6.4 we have that, for k big enough, $\mathbf{D}_m(k)$ has an eigenvalue $\lambda(k)$ which is simple and strictly dominant in the form

$$\lambda(k) = \lambda + \frac{\langle \mathbf{T}_m^{\mathrm{T}}\mathbf{l}, (\mathbf{D}_m(k) - \mathbf{D}'_m)\mathbf{H}_m\mathbf{r} \rangle}{\langle \mathbf{T}_m^{\mathrm{T}}\mathbf{l}, \mathbf{H}_m\mathbf{r} \rangle} + O(o(\alpha^k)^2).$$

Associated to λ there are right and left eigenvectors $\mathbf{H}_m\mathbf{r} + O(o(\alpha^k))$ and $\mathbf{T}_m^{\mathrm{T}}\mathbf{l} + O(o(\alpha^k))$ respectively. Since $O(o(\alpha^k)) = o(\alpha^k)$ all we have to do is to show that

$$\frac{\langle \mathbf{T}_m^{\mathrm{T}}\mathbf{l}, (\mathbf{D}_m(k) - \mathbf{D}'_m)\mathbf{H}_m\mathbf{r} \rangle}{\langle \mathbf{T}_m^{\mathrm{T}}\mathbf{l}, \mathbf{H}_m\mathbf{r} \rangle} = o(\alpha^k),$$

which is trivial for the scalar product is a continuous mapping and $\mathbf{D}'_m(k) - \mathbf{D}' = o(\alpha^k)$. $\square$

Below we state a result about matrix perturbation theory that is used in this work and which can be found in [34].

Theorem 7.1. *Let λ be a simple eigenvalue of a matrix $\mathbf{A}$ of dimensions $N \times N$ with associated right and left eigenvectors $\mathbf{x}_r$ and $\mathbf{x}_l$, respectively. Let $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{E}$ be a perturbation of matrix $\mathbf{A}$, and $\| * \|$ any consistent matrix norm in the space of matrices $N \times N$. Then*

(a) *there exists a unique eigenvalue $\tilde{\lambda}$ of $\tilde{\mathbf{A}}$ such that*

$$\tilde{\lambda} = \lambda + \frac{\mathbf{x}_l^{\mathrm{T}}\mathbf{E}\mathbf{x}_r}{\mathbf{x}_l^{\mathrm{T}}\mathbf{x}_r} + O(\|\mathbf{E}\|^2).$$

(b) *associated to $\tilde{\lambda}$ there exist right and left eigenvectors $\tilde{\mathbf{x}}_r$ and $\tilde{\mathbf{x}}_l$ respectively such that*

$$\tilde{\mathbf{x}}_r = \mathbf{x}_r + O(\|\mathbf{E}\|); \quad \tilde{\mathbf{x}}_l = \mathbf{x}_l + O(\|\mathbf{E}\|).$$

(c) *for small enough $\|\mathbf{E}\|$, $\tilde{\lambda}$ is the only eigenvalue of $\tilde{\mathbf{A}}$ in a certain neighborhood of λ (therefore, if λ is strictly dominant for $\mathbf{A}$, so will be $\tilde{\lambda}$ for $\tilde{\mathbf{A}}$).*

References

[1] P. Auger, Dynamics and Thermodynamics in Hierarchically Organized Systems, Applications in Physics, Biology and Economics, Pergamon Press, Oxford, 1989.

[2] P. Auger, E. Benoit, A Prey-Predator model in a multipatch environment with different time scales, J. Biol. Systems 1 (1993) 187–197.

[3] P. Auger, R. Roussarie, Complex ecological models with simple dynamics: from individuals to populations, Acta Biotheor. 42 (1994) 111–136.

[4] P. Auger, J.C. Poggiale, Aggregation and emergence in hierarchically organized systems: population dynamics, Acta Biotheor. 44 (1996) 301–316.

[5] P. Auger, J.C. Poggiale, Aggregation and emergence in systems of ordinary differential equations, Math. Comput. Modelling 27 (1998) 1–21.

[6] R. Bravo de la Parra, P. Auger, E. Sánchez, Aggregation methods in discrete models, J. Biol. Systems 3 (2) (1995) 603–612.

[7] R. Bravo de la Parra, E. Sánchez, P. Auger, Time scales in density dependent discrete models, J. Biol. Systems 5 (1997) 111–129.

[8] R. Bravo de la Parra, E. Sánchez, Aggregation methods in population dynamics discrete models, Math. Comput. Modelling 4 (1998) 23–39.

[9] R. Bravo de la Parra, E. Sánchez, O. Arino, P. Auger, Discrete model with density dependent fast migration, Math. Biosc. 157 (1999) 91–109.

[10] H. Caswell, Matrix Population Models, Sinauer Associates, Sunderland, 1989.

[11] 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 Biotheor. (1997), in press.

[12] C. Chatfield, The Analysis of Time Series, Chapman & Hall, London, 1989.

[13] J.E. Cohen, Ergodicity of age structure in populations with Markovian vital rates, II, General states, Adv. in Appl. Probability 9 (1977) 18–37.

[14] J.E. Cohen, Ergodics theorems of demography, Bull. Amer. Math. Soc. N.S. 1 (2) (1979) 275–295.

[15] 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, pp. 477–503.

[16] M. Fiedler, Special Matrices and Their Applications in Numerical Mathematics, Martinus Nijhoff, Dordrecht, 1986.

[17] T.C. Gard, Aggregation in stochastic ecosystem models, Ecol. Modelling 44 (1988) 153–164.

[18] T.E. Harris, The Theory of Branching Processes, Springer, Berlin, 1963.

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

[20] Y. Iwasa, V. Andreasen, S.A. Levin, Agregation in model ecosystems, (I) Perfect Aggregation, Ecol. Modelling 37 (1987) 287–302.

[21] R. Lande, S.H. Orzack, Extinction dynamics of age structured populations in a fluctuating environment, Proc. Natl. Acad. Sci. USA 85 (1988) 7418–7421.

[22] K.L. Liaw, Multistate dynamics: the convergence of an age-by-region population system, Environ. Planning A 12 (1980) 589–613.

[23] K.L. Liaw, Spatial population dynamics, in: A. Rogers, F. Willekens (Eds.), Migration and Settlement: A Multiregional Comparative Study, D. Reidel, Dordrecht, 1986, pp. 419–455.

[24] D.O. Logofet, Matrices and graphs. Stability Problems in Mathematical Ecology, CRC press, Boca Raton, 1993.

[25] J.H. Pollard, A note on multitype Galton-Watson processes with random branching probabilities, Biometrika 55 (1968) 589–590.

[26] J.H. Pollard, The multi-type Galton-Watson processes in a genetical context, Biometrics 24 (1968) 147–158.

[27] J.H. Pollard, Mathematical Models for the Growth of Human Populations, Cambridge University Press, New York, 1973.

[28] A. Rogers, Shrinking large-scale population projection models by aggregation and decomposition, Environ. Planning A 8 (1976) 515–541.

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

[30] J.M. Saboia, Arima models for birth forecasting, J. Amer. Statist. Assoc. 72 (1977) 264–270.

[31] E. Sánchez, R. Bravo de la Parra, P. Auger, Linear discrete models with different time scales, Acta Biotheor. 43 (1995) 465–479.

[32] L. Sanz, R. Bravo de la Parra, Variables aggregation in time varying discrete systems, Acta Biotheor. 46 (1998) 273–297.

[33] L. Sanz, R. Bravo de la Parra, Variables aggregation in a time discrete linear model, Math. Biosc. 157 (1999) 111–146.

[34] G.W. Stewart, J.I. Guang Sun, Matrix Perturbation Theory, Academic Press, New York, 1990.

[35] Z.M. Sykes, Some stochastic versions of the matrix model for population dynamics, J. Amer. Statist. Assoc. 64 (1969) 111–130.

[36] S. Tuljapurkar, S. Orzack, Population dynamics in variable environments, I, Long-run growth rates and extinction, Theor. Popul. Biol. 18 (1980) 314–342.

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

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