THE RELIABILITY OF APPROXIMATE REDUCTION TECHNIQUES IN POPULATION MODELS WITH TWO TIME SCALES

Luis Sanz[1] and Rafael Bravo de la Parra[2]

[1]Departamento de Matemáticas, E.T.S.I. Industriales, Universidad Politécnica de Madrid, c/ José Gutiérrez Abascal, 2, 28006 Madrid, Spain. Email: lsanz@etsii.upm.es (corresponding author).
[2]Departamento de Matemáticas, Universidad de Alcalá, 28871 Alcalá de Henares (Madrid), Spain. Email: rafael.bravo@uah.es

ABSTRACT

As a result of the complexity inherent in some natural systems, mathematical models employed in ecology are often governed by a large number of variables. For instance, in the study of population dynamics we often find multiregional models for structured populations in which individuals are classified regarding their age and their spatial location. Dealing with such structured populations leads to high dimensional models. Moreover, in many instances the dynamics of the system is controlled by processes whose time scales are very different from each other. For example, in multiregional models migration is often a fast process in comparison to the growth of the population.

Approximate reduction techniques take advantage of the presence of different time scales in a system to introduce approximations that allow one to transform the original system into a simpler low dimensional system. In this way, the dynamics of the original system can be approximated in terms of that of the reduced system. This work deals with the study of that approximation. In particular, we work with a non-autonomous discrete time model previously presented in the literature and obtain different bounds for the error we incur when we describe the dynamics of the original system in terms of the reduced one.

The results are illustrated by some numerical simulations corresponding to the reduction of a Leslie type model for a population structured in two age classes and living in a two patch system.

Keywords: Approximate aggregation, population dynamics, time scales, non-autonomous models.

1. INTRODUCTION

Nature offers many examples of systems with an inherent complexity. For example, communities are systems 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, etc.

The study of these complex systems leads to mathematical models with a large number of state variables whose analytical study is, in most cases, not feasible. In order to be able to extract important information about the behavior of some of these

Acta Biotheoretica **50**: 297–322, 2002.

complex models, one can resort to system reduction techniques. These techniques, called "aggregation methods" transform the system under consideration into a reduced system with a lesser number of variables, called "global variables" which can be more easily studied.

Among aggregation techniques we can distinguish "perfect" and "approximate" aggregation. The former transforms the complex system into a reduced system without introducing any approximations, i.e., a set of global variables is defined as a function of the state variables in such a way that the original system depends solely on these global variables (Iwasa *et al*., 1987). Biological systems can be perfectly aggregated only in some cases and for very particular values of the parameters involved, so perfect aggregation has only a theoretical interest.

In general, in order to reduce the complex original system one has to resort to approximate aggregation (see Auger and Roussarie, 1994 for a seminal paper), in which some approximations are introduced in order to build the reduced system. In this way, the behavior of the original system can be approximated, but not known with exactitude, in terms of the knowledge of the behavior of the reduced system.

A property of complex systems that allows one to use approximate aggregation techniques is the existence of two time scales, i.e., of processes whose characteristic times are very different from each other. In addition, the fast process is required to tend, at least in a generalized sense that depends on the context, to an equilibrium. The main idea of these techniques is to approximate the original system by an "auxiliary" system in which the fast dynamics has reached equilibrium and then, taking advantage of the redundancies present in this auxiliary system, carry out its exact reduction, obtaining an aggregated system whose dynamics allows one to approximate that of the original system.

The presence of different time scales is common to many of the complex systems found in nature. For example, it is usually the case that processes such as migration or changes of activity are fast with respect to reproduction or aging. 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 on 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.

Approximate aggregation techniques have been widely studied in the context of time continuous systems with different time scales for both linear and density dependent models (see among others Auger and Roussarie, 1994; Poggiale and Auger, 1995; Auger and Poggiale, 1996a, 1996b; Auger and Bravo de la Parra, 2000). The discrete time case has also been thoroughly explored in linear, non-linear, non-autonomous and stochastic contexts (Bravo de la Parra and Sánchez, 1998; Sanz and Bravo de la Parra, 1998, 1999, 2000, 2001; Blasco *et al.*, 2001, 2002; Bravo de la Parra *et al.*, 1999).

Concentrating on the discrete time case, the above-mentioned works carry out the reduction of systems and proves that the original and the aggregated models can be related when the separation of time scales between the two processes tends to infinity. For example, in the linear autonomous case it has been shown (Sanz and Bravo de la Parra, 1999) that the asymptotic growth rate of the original system tends to that of the

aggregated system as the separation of time scales grows. Similarly, in a non-autonomous context (Sanz and Bravo de la Parra, 2001), the variables of the original system can be approximated in terms of those of the reduced system, and the approximation is exact when the difference of time scales tends to infinity. However, these works do not deal with the problem of estimating the precision of the approximation for a given finite separation of time scales between the slow and the fast process.

The purpose of this paper, working in the context of non-autonomous linear models, is to derive bounds that allow one to estimate the error we make when we study the original system in terms of the reduced one. In particular, we will work with a very general model introduced in Sanz and Bravo de la Parra (2001). Section 2 presents this model, a non-autonomous discrete time model with two time scales, and describes the aggregation procedure. The general method is illustrated by reducing a Leslie type model for an age structured population living in a two patch environment.

In Section 3 we present some of the results previously obtained in the field that are relevant to our study and define some quantities that characterize the magnitude of the error. In Section 4 we obtain different bounds for the error. These bounds differ both in the kind of approach to the study of the error and in the degree of complexity involved in obtaining them.

The tightness of the different bounds for the Leslie type model introduced in Section 2 is illustrated by some numerical simulations in Section 5. The proofs of the results can be found in the Appendix, where the mathematically oriented reader can find some insight into the kind of mathematical approach used to obtain each bound.

2. APPROXIMATE REDUCTION OF DISCRETE TIME NON-AUTONOMOUS SYSTEMS WITH TWO TIME SCALES

In this work we deal with a discrete time model for non-autonomous systems with two time scales that has been proposed in Sanz and Bravo de la Parra (2001). In this section we introduce the basics of the model and the reduction procedure and refer to that work for further details on the properties of the model and for illustrations of their applicability to practical situations.

Our model contemplates a stage-structured population in which the population is classified into stages or "groups" in terms of a certain characteristic of the life cycle, for example age or size. Furthermore, each of these groups is divided into several subgroups that can correspond to different spatial patches, different individual activities or any other characteristic that could change the life cycle parameters. In this way, we assume that the population is subdivided into q populations (or groups). Each group is subdivided into subpopulations (subgroups) in such a way that for each $i=1,2,...,q$, group i has N_i subgroups. Thus, the total number of subgroups is $N = N_1 + N_2 + \cdots + N_q$.

We denote by x_n^{ij} the density of subpopulation j of population i at time $n=0,1,2,\ldots$, with $i=1,2,...,q$ and $j=1,2,...,N_i$. In order to describe the population of group i we will use the vector $\mathbf{x}_n^i = (x_n^{i1}, x_n^{i2}, \ldots, x_n^{iN_i}) \in \mathbf{R}^{N_i}$, $i=1,2,...,q$. The composition of the total population is then given by the vector $\mathbf{X}_n = (\mathbf{x}_n^1, \mathbf{x}_n^2, \ldots, \mathbf{x}_n^q)^T \in \mathbf{R}^N$ where T denotes transposition.

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

The projection interval of our model is 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, and suppose that the coefficients of the model are constant during each of the intervals Δ_n. This implies that the characteristics of the fast process are not allowed to vary with its projection interval but only with each time step corresponding to the model, i.e., with each projection interval associated with the slow process. This hypothesis has been relaxed in Blasco *et al.* (2002) to allow the fast process to vary with its own projection interval. We will not explore this case for it introduces several mathematical complications.

In principle, we introduce no special assumptions regarding the characteristics of the slow process and so, for a certain fixed projection interval, the slow dynamics will be represented at time n by a non-negative projection matrix $\mathbf{M}_n \in \mathbf{R}^{N\times N}$. We consider $\mathbf{M}_n$ as divided into blocks $\mathbf{M}_{ij}(n)$, $1 \le i, j \le q$, and so we have

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

where each block $\mathbf{M}_{ij}(n) = [M_{ij}^{rl}(n)]$ $(1 \le r \le N_i,\ 1 \le l \le N_j)$ and characterizes the rates of transference of individuals from the subgroups of group j to the subgroups of group *i* at time n.

Regarding the fast process, the following assumptions are made:

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* is represented, during interval Δ_n, by a non-negative matrix $\mathbf{P}_i(n) = [P_i^{rl}(n)]$ of dimensions $N_i \times N_i$ for each n and $i = 1, 2, ..., q$.

b) For each *i* and n, matrix $\mathbf{P}_i(n)$ has spectral radius equal to one, i.e., the modulus of the dominant eigenvalue of $\mathbf{P}_i(n)$ is one.

c) For each *i* and n, matrix $\mathbf{P}_i(n)$ is irreducible and primitive, i.e., the power $\mathbf{P}_i^s(n)$ is a positive matrix for some positive integer s.

The matrix which represents the fast dynamics for the whole population during interval $[n, n+1)$ is

$$\mathbf{P}_n = [P_{\alpha\beta}^n] = diag(\mathbf{P}_1(n), \mathbf{P}_2(n), \ldots, \mathbf{P}_q(n)). \tag{2}$$

As stated above, the projection interval of the model is that corresponding to matrices $\mathbf{M}_n$. We approximate the effect of the fast dynamics over the time step of the

model, which is much longer than its own, supposing that during each interval Δ_n, matrix $\mathbf{P}_n$ operates a k number of times, where k can be interpreted as the ratio between the projection intervals corresponding to the slow and fast dynamics. Since these projection intervals are very different from each other we can assume that k is a large integer. In this way, the fast dynamics during interval Δ_n is modelled by $\mathbf{P}_n^k$ and the resulting model consists of the following system of N linear difference equations that we denote "original system"

$$\mathbf{X}_{n+1} = \mathbf{M}_n \mathbf{P}_n^k \mathbf{X}_n . \tag{3}$$

The primitivity of the $\mathbf{P}_i(n)$, together with assumption (b), guarantees that for a sufficiently high separation of time scales, the fast process during each Δ_n approaches an equilibrium distribution.

Let $i = 1, 2, ..., q$ and n be fixed and consider the positive vectors $\mathbf{v}_i(n)$ and $\mathbf{u}_i(n)$ defined by the conditions

$$\mathbf{P}_i(n)\mathbf{v}_i(n) = \mathbf{v}_i(n);\ \mathbf{u}_i^T(n)\mathbf{P}_i(n) = \mathbf{u}_i^T(n)$$

$$\mathbf{1}^T \mathbf{v}_i(n) = 1;\ \mathbf{u}_i^T(n)\mathbf{v}_i(n) = 1;\ \mathbf{v}_i(n) > \mathbf{0};\ \mathbf{u}_i(n) > \mathbf{0}, \tag{4}$$

where $\mathbf{1} = (1, 1, \ldots, 1)^T$. We use the notation $\mathbf{A} > \mathbf{0}$ (resp. $\mathbf{A} \geq \mathbf{0}$) to denote that $\mathbf{A}$ is a matrix or vector of positive (resp. non-negative) elements. Therefore, $\mathbf{v}_i(n)$ and $\mathbf{u}_i(n)$ are, respectively, the positive right and left eigenvectors of $\mathbf{P}_i(n)$ associated with eigenvalue one and verifying some normalization conditions. Note that, since $\mathbf{P}_i(n)$ is primitive, eigenvalue one is simple and therefore $\mathbf{v}_i(n)$ and $\mathbf{u}_i(n)$ are defined in a unique way by conditions (4). Vectors $\mathbf{v}_i(n)$ and $\mathbf{u}_i(n)$ can be interpreted in the following way in terms of the fast dynamics of group *i* in interval Δ_n. Let us consider a hypothetical situation in which the system would be governed by the fast process exclusively. Suppose, moreover, 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 "initial condition" of the system at time n, the structure of the population of group *i* at the end of Δ_n would be defined by $\mathbf{v}_i(n)$, meanwhile the reproductive value (Caswell, 2001) of the individuals of that group would be characterized by vector $\mathbf{u}_i(n)$.

Therefore, the equilibrium of the fast process in Δ_n for each group *i* and for the whole population is characterized, respectively, by matrices $\overline{\mathbf{P}}_i(n) = [\overline{P}_i^{\alpha\beta}(n)]$ and $\overline{\mathbf{P}}_n = [\overline{P}_{\alpha\beta}^{n}]$ given by

$$\overline{\mathbf{P}}_i(n) = \lim_{k\to\infty} \mathbf{P}_i^k(n) = \mathbf{v}_i(n)\mathbf{u}_i^T(n) > \mathbf{0}$$

$$\overline{\mathbf{P}}_n = diag(\overline{\mathbf{P}}_1(n), \overline{\mathbf{P}}_2(n), \ldots, \overline{\mathbf{P}}_q(n)). \tag{5}$$

Let us define matrices

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

$$\mathbf{U}_n = diag(\mathbf{u}_1^T(n), \mathbf{u}_2^T(n), \ldots, \mathbf{u}_q^T(n)) \tag{6}$$

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

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

Lemma 1. *Matrices* $\mathbf{P}_n$, $\overline{\mathbf{P}}_n$, $\mathbf{V}_n$ *and* $\mathbf{U}_n$ *verify, for all* n*:*

a) $\overline{\mathbf{P}}_n\mathbf{P}_n = \mathbf{P}_n\overline{\mathbf{P}}_n = \overline{\mathbf{P}}_n\overline{\mathbf{P}}_n = \overline{\mathbf{P}}_n$.

b) $\mathbf{P}_n\mathbf{V}_n = \overline{\mathbf{P}}_n\mathbf{V}_n = \mathbf{V}_n$.

c) $\mathbf{U}_n\overline{\mathbf{P}}_n = \mathbf{U}_n$; $\mathbf{U}_n\mathbf{V}_n = \mathbf{I}_q$; $\overline{\mathbf{P}}_n = \mathbf{V}_n\mathbf{U}_n$.

System (3), consisting of N variables (microvariables) associated with the different subgroups can be approximated by a reduced system ("aggregated system") of q variables (global variables), each of them associated with one group. In order to do so, we introduce a so called "auxiliary system" which approximates the dynamics of the general system and can be reduced without introducing any error. This auxiliary system is defined by

$$\mathbf{X}'_{n+1} = \mathbf{M}_n\overline{\mathbf{P}}_n\mathbf{X}'_n \tag{7}$$

where the vector of variables for this new system is denoted by $\mathbf{X}'_n$. The auxiliary system can be interpreted as the microsystem when we substitute the fast process corresponding to each interval Δ_n by the equilibrium characteristics of the fast process in Δ_n. In other words, we are letting $k \to \infty$ in the expression of the microsystem, which is equivalent to supposing that Δ_n is long enough with respect to the projection interval of the fast process for the fast dynamics to reach equilibrium. Multiplying both sides of (7) by matrix $\mathbf{U}_{n+1}$ we have

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

where we have used $\overline{\mathbf{P}}_n = \mathbf{V}_n\mathbf{U}_n$. We define the vector of global variables by

$$\mathbf{Y}_n = \mathbf{U}_n\mathbf{X}'_n \tag{8}$$

and see that (7) can be expressed as a function of the global variables exclusively. In this way, the reduced system is defined by

$$\mathbf{Y}_{n+1} = \overline{\mathbf{M}}_n\mathbf{Y}_n \tag{9}$$

where matrix $\overline{\mathbf{M}}_n \in \mathbf{R}^{q\times q}$ is given by

$$\overline{\mathbf{M}}_n = \mathbf{U}_{n+1}\mathbf{M}_n\mathbf{V}_n.$$

The global variables $\mathbf{Y}_n = (y_n^1, \ldots, y_n^q)^T$, defined by (8), have the following expression in terms of the variables $\mathbf{X}'_n$ of the auxiliary system:

$$y_n^i = \mathbf{u}_i^T(n)\mathbf{x}'^{iT}_n = u_i^1(n)x'^{i1}_n + u_i^2(n)x'^{i2}_n + \ldots + u_i^{N_i}(n)x'^{iN_i}_n;\ i = 1,2,...,q, \tag{10}$$

i.e., each y_n^i is a linear combination of the variables of the auxiliary system corresponding to group *i*, being the coefficients of the combination the components of vector $\mathbf{u}_i(n)$. Moreover, these global variables are easily seen to be conservative for the fast process.

Recall that an approximation is made in order to build the auxiliary system in terms of the original system, while the aggregated system is built exactly in terms of the auxiliary one. As we show in Section 3, this is reflected in the fact that the global variables and the variables of the auxiliary system can be known without error in terms of each other, while the microvariables can only be known approximately in terms of the auxiliary variables.

Reduction of a Leslie type model with fast migration

As an illustration of the above general setting, we will carry out the reduction of a Leslie type model for a space distributed population in which migration is a fast process in relation to demography. This particular setting will also be used in Section 5 to test the accuracy of the different bounds that we derive for the error we incur when aggregating.

We consider a population structured by age in two classes (groups), say juveniles and adults, corresponding to groups. We assume that the population is distributed in two patches (subgroups) among which they can migrate. In this way, the population is structured in four stages, each of them corresponding to an age class and a spatial location. The demographic and migratory processes are responsible for the transference of individuals between the different stages and we suppose that migration is a fast process in comparison with demography. We choose as time step for the model, $\Delta_n = [n, n+1)$, the duration of the age class of the juveniles.

Generalizations of this situation to the case where there is an arbitrary number of age classes and spatial patches, as well as to the case where demography is fast with respect to migration can be found in Sanz and Bravo de la Parra (2001). Charles *et al.* (1998) shows a practical application of the reduction of a Leslie type model to actual populations.

We denote by x_n^{ij} the number of individuals of age i in the j-th spatial patch at time n, $i, j = 1,2$. Vectors $\mathbf{x}_n^i = (x_n^{i1}, x_n^{i2})$; $i = 1,2$ describe the spatial allocation of individuals in age class i thus for the whole population we have vector $\mathbf{X}_n = (\mathbf{x}_n^1, \mathbf{x}_n^2)^T = (x_n^{11}, x_n^{12}, x_n^{21}, x_n^{22})^T$.

The transference of individuals between the different age classes is characterized by the following matrices

$$\mathbf{M}_{11}(n) = \begin{bmatrix} F_1^1(n) & 0 \\ 0 & F_1^2(n) \end{bmatrix}, \ \mathbf{M}_{12}(n) = \begin{bmatrix} F_2^1(n) & 0 \\ 0 & F_2^2(n) \end{bmatrix},$$

$$\mathbf{M}_{21}(n) = \begin{bmatrix} S_1^1(n) & 0 \\ 0 & S_1^2(n) \end{bmatrix}, \ \mathbf{M}_{22}(n) = \begin{bmatrix} S_2^1(n) & 0 \\ 0 & S_2^2(n) \end{bmatrix},$$

where $F_i^j(n)$ and $S_i^j(n)$ are, respectively, the fertility and survival rate for individuals of age i in patch j during interval Δ_n; $i = 1,2$. Note that we are using the fact that the demographic process does not make individuals change patch and so matrices $\mathbf{M}_{ij}(n)$ are diagonal. In addition, we allow the survival coefficients of the second age class to be non-zero to account for the adults that live longer than one time step.

The matrix $\mathbf{M}_n$ modelling demography at time n is then given by

$$\mathbf{M}_n = \begin{bmatrix} F_1^1(n) & 0 & F_2^1(n) & 0 \\ 0 & F_1^2(n) & 0 & F_2^2(n) \\ S_1^1(n) & 0 & S_2^1(n) & 0 \\ 0 & S_1^2(n) & 0 & S_2^2(n) \end{bmatrix}.$$

Let $\mathbf{P}_i(n)$ be the matrix characterizing migration between the two patches for the individuals of age class i; $i=1,2$ at time n. Since migration is a conservative process for the total number of individuals, matrices $\mathbf{P}_i(n)$ are stochastic and therefore their dominant eigenvalue is one. We then have

$$\mathbf{P}_1(n) = \begin{bmatrix} 1-p_1(n) & q_1(n) \\ p_1(n) & 1-q_1(n) \end{bmatrix}, \quad \mathbf{P}_2(n) = \begin{bmatrix} 1-p_2(n) & q_2(n) \\ p_2(n) & 1-q_2(n) \end{bmatrix}$$

where $p_1(n)$ and $p_2(n)$ denote migration rates at time n for individuals of age class 1 and 2, respectively, from patch 1 to patch 2, and $q_1(n)$ and $q_2(n)$ have the analogous meaning for the migration from patch 2 to patch 1.

If we assume that all the migration rates are different from 0 and 1, then matrices $\mathbf{P}_1(n)$ and $\mathbf{P}_2(n)$ are positive and, consequently, primitive. In this way, migration meets the hypotheses required in Section 2 for the fast process.

If the duration of each age class is sufficiently long with respect to the projection interval of migration, this latter process tends, for each age class i and each Δ_n, to an equilibrium distribution among the patches given by the positive eigenvector $\mathbf{v}_i(n)$ of $\mathbf{P}_i(n)$ associated with eigenvalue 1. Vectors $\mathbf{v}_i(n)$ are given by

$$\mathbf{v}_1(n) = \left(\frac{q_1(n)}{p_1(n)+q_1(n)}, \frac{p_1(n)}{p_1(n)+q_1(n)} \right)^T, \quad \mathbf{v}_2(n) = \left(\frac{q_2(n)}{p_2(n)+q_2(n)}, \frac{p_2(n)}{p_2(n)+q_2(n)} \right)^T$$

and, since matrices $\mathbf{P}_i(n)$ are stochastic, the left eigenvectors $\mathbf{u}_i(n)$ associated with eigenvalue 1 verify $\mathbf{u}_i(n) = (1,1)^T = \mathbf{1}^T$. In this way, the matrices that define the migration equilibria are

$$\bar{\mathbf{P}}_i(n) = \mathbf{v}_i(n)\mathbf{1}^T = \begin{bmatrix} \dfrac{q_i(n)}{p_i(n)+q_i(n)} & \dfrac{q_i(n)}{p_i(n)+q_i(n)} \\ \dfrac{p_i(n)}{p_i(n)+q_i(n)} & \dfrac{p_i(n)}{p_i(n)+q_i(n)} \end{bmatrix}, \quad i=1,2.$$

The matrix that characterizes migration for the whole population is $\mathbf{P}_n = diag(\mathbf{P}_1(n), \mathbf{P}_2(n))$ while its equilibrium is determined by $\bar{\mathbf{P}}_n = diag(\bar{\mathbf{P}}_1(n), \bar{\mathbf{P}}_2(n))$.

The original system has the expression

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

where k can be interpreted as the ratio between the projection intervals of demography and migration.

If we let migration reach equilibrium in the original system we obtain the auxiliary system that reads

$$\begin{bmatrix} x'^{11}_{n+1} \\ x'^{12}_{n+1} \\ x'^{21}_{n+1} \\ x'^{22}_{n+1} \end{bmatrix} = \mathbf{M}_n \bar{\mathbf{P}}_n \begin{bmatrix} x'^{11}_{n} \\ x'^{12}_{n} \\ x'^{21}_{n} \\ x'^{22}_{n} \end{bmatrix}$$

where

$$\mathbf{M}_n \bar{\mathbf{P}}_n = \begin{bmatrix} \frac{F_1^1(n)q_1(n)}{p_1(n)+q_1(n)} & \frac{F_1^1(n)q_1(n)}{p_1(n)+q_1(n)} & \frac{F_2^1(n)q_2(n)}{p_2(n)+q_2(n)} & \frac{F_2^1(n)q_2(n)}{p_2(n)+q_2(n)} \\ \frac{F_1^2(n)p_1(n)}{p_1(n)+q_1(n)} & \frac{F_1^2(n)p_1(n)}{p_1(n)+q_1(n)} & \frac{F_2^2p_2(n)}{p_2(n)+q_2(n)} & \frac{F_2^2p_2(n)}{p_2(n)+q_2(n)} \\ \frac{S_1^1(n)q_1(n)}{p_1(n)+q_1(n)} & \frac{S_1^1(n)q_1(n)}{p_1(n)+q_1(n)} & \frac{S_2^1(n)q_2(n)}{p_2(n)+q_2(n)} & \frac{S_2^1(n)q_2(n)}{p_2(n)+q_2(n)} \\ \frac{S_1^2(n)p_1(n)}{p_1(n)+q_1(n)} & \frac{S_1^2(n)p_1(n)}{p_1(n)+q_1(n)} & \frac{S_2^2p_2(n)}{p_2(n)+q_2(n)} & \frac{S_2^2p_2(n)}{p_2(n)+q_2(n)} \end{bmatrix}.$$

In order to reduce the auxiliary system we build matrices $\mathbf{V}_n$ and $\mathbf{U}_n$ that take the form

$$\mathbf{V}_n = diag(\mathbf{v}_1(n), \mathbf{v}_2(n)) = \begin{bmatrix} \frac{q_1(n)}{p_1(n)+q_1(n)} & 0 \\ \frac{p_1(n)}{p_1(n)+q_1(n)} & 0 \\ 0 & \frac{q_2(n)}{p_2(n)+q_2(n)} \\ 0 & \frac{p_2(n)}{p_2(n)+q_2(n)} \end{bmatrix}$$

$$\mathbf{U}_n = diag(\mathbf{u}_1^T, \mathbf{u}_2^T) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

The global variables $y_n^i;\ i=1,2$ are defined by $\mathbf{y}_n = \mathbf{U}_n \mathbf{X}'_n$, i.e.,

$$y_n^1 = \mathbf{u}_1^T (x'^{11}_n, x'^{12}_n)^T = x'^{11}_n + x'^{12}_n$$

$$y_n^2 = \mathbf{u}_2^T (x'^{21}_n, x'^{22}_n)^T = x'^{21}_n + x'^{22}_n$$

so in this case they correspond to the total population in each age class.

The aggregated model is given by

$$\begin{bmatrix} y^1_{n+1} \\ y^2_{n+1} \end{bmatrix} = \overline{\mathbf{M}}_n \begin{bmatrix} y^1_n \\ y^2_n \end{bmatrix}$$

where

$$\overline{\mathbf{M}}_n = \mathbf{U}_{n+1}\mathbf{M}_n\mathbf{V}_n = \begin{bmatrix} \dfrac{q_1(n)F_1^1(n) + p_1(n)F_1^2(n)}{p_1(n) + q_1(n)} & \dfrac{q_2(n)F_2^1(n) + p_2(n)F_2^2(n)}{p_2(n) + q_2(n)} \\ \dfrac{q_1(n)S_1^1(n) + p_1(n)S_1^2(n)}{p_1(n) + q_1(n)} & \dfrac{q_2(n)S_2^1(n) + p_2(n)S_2^2(n)}{p_2(n) + q_2(n)} \end{bmatrix},$$

i.e., the aggregated model is a conventional Leslie type model whose parameters are linear combinations of the parameters defining demography, being the coefficients dependent on the equilibrium spatial distribution for migration.

3. ASYMPTOTIC RELATIONSHIPS BETWEEN THE SYSTEMS

Returning to the general setting of Section 2, we define the following matrix products, for each n and k, which will be useful in the sequel:

$$\Pi_{n,k} = \mathbf{M}_{n-1}\mathbf{P}_{n-1}^k \ldots \mathbf{M}_1\mathbf{P}_1^k\mathbf{M}_0\mathbf{P}_0^k;\ \Pi'_n = \mathbf{M}_{n-1}\overline{\mathbf{P}}_{n-1} \ldots \mathbf{M}_1\overline{\mathbf{P}}_1\mathbf{M}_0\overline{\mathbf{P}}_0;$$

$$\overline{\Pi}_n = \overline{\mathbf{M}}_{n-1} \ldots \overline{\mathbf{M}}_1\overline{\mathbf{M}}_0$$

and we have the following expression for the original, auxiliary and aggregated systems in terms of the vector of initial conditions (note that both the auxiliary and the original system start from the same initial condition)

$$\mathbf{X}_n = \Pi_{n,k}\mathbf{X}_0;\ \ \mathbf{X}'_n = \Pi'_n\mathbf{X}_0;\ \ \mathbf{Y}_n = \overline{\Pi}_n\mathbf{Y}_0$$

where $\mathbf{Y}_0 = \mathbf{U}_0\mathbf{X}_0$. In addition, let

$$\mathbf{E}_{n,k} = \Pi_{n,k} - \Pi'_n.$$

For each n, let the eigenvalues of $\mathbf{P}_n$ (i.e., the union of the eigenvalues of the $\mathbf{P}_i(n)$) ordered by decreasing modulus be

$$1 = \lambda_1(n) = \cdots = \lambda_q(n) > \left|\lambda_{q+1}(n)\right| \geq \cdots \geq \left|\lambda_N(n)\right|$$

and let δ verify the condition

$$\delta = 1 \text{ if } \sup_n\left\{\left|\lambda_{q+1}(n)\right|\right\} = 1, \tag{11}$$

$$1 > \delta > \sup_n\left\{\left|\lambda_{q+1}(n)\right|\right\} \text{ if } \sup_n\left\{\left|\lambda_{q+1}(n)\right|\right\} < 1.$$

From this definition we see that δ is greater than the modulus of the subdominant eigenvalue of matrices $\mathbf{P}_n$.

The following proposition is taken from previous works of the authors (Sanz and Bravo de la Parra, 1999, 2001) where it is shown that the variables (and, in the autonomous case, the dominant eigenvalues) corresponding to the three systems under consideration can be related in a certain way.

Proposition 2. *a) For each n we have*

$$\Pi'_{n+1} = \mathbf{M}_n\mathbf{V}_n\overline{\Pi}_n\mathbf{U}_0;\ \ \overline{\Pi}_n = \mathbf{U}_n\Pi'_n\mathbf{V}_0$$

and therefore the relationship between the variables corresponding to the aggregated system and to the auxiliary system is given by

$$\mathbf{X}'_{n+1} = \mathbf{M}_n \mathbf{V}_n \mathbf{Y}_n; \quad \mathbf{Y}_n = \mathbf{U}_n \mathbf{X}'_n.$$

b) The sequence $\mathbf{E}_{n,k} = \Pi_{n,k} - \Pi'_n$ *verifies, for fixed* n,

$$\mathbf{E}_{n,k} = \mathbf{o}\left(\delta^k\right); \quad k \to \infty$$

where δ *is given by (11) and so* $\mathbf{X}'_n - \mathbf{X}_n = \mathbf{o}\left(\delta^k\right)$. *Therefore, the variables* $\mathbf{X}_n$ *of the original system can be approximated by those of the aggregated system* $\mathbf{Y}_n$ *and reciprocally, in the following way*

$$\mathbf{X}_n = \mathbf{M}_{n-1}\mathbf{V}_{n-1}\mathbf{Y}_{n-1} + \mathbf{E}_{n,k}\mathbf{X}_0 = \mathbf{M}_{n-1}\mathbf{V}_{n-1}\mathbf{Y}_{n-1} + \mathbf{o}\left(\delta^k\right); \quad k \to \infty \tag{12}$$

$$\mathbf{Y}_n = \mathbf{U}_n\mathbf{X}_n + \mathbf{U}_n\mathbf{E}_{n,k}\mathbf{X}_0 = \mathbf{U}_n\mathbf{X}_n + \mathbf{o}\left(\delta^k\right); \; k \to \infty.$$

c) In the autonomous case, $\rho\left(\overline{\mathbf{M}}\right) = \rho\left(\mathbf{M}\overline{\mathbf{P}}\right)$ *and* $\rho(\mathbf{M}\mathbf{P}^k) = \rho\left(\mathbf{M}\overline{\mathbf{P}}\right) + o\left(\delta^k\right)$, $k \to \infty$, *where* ρ *denotes the spectral radius.*

This result shows that we can obtain the population vector of the auxiliary system exactly in terms of that of the aggregated system and reciprocally. In addition, the variables of the original system can be approximated knowing those of the aggregated system, and, for fixed n, the error we make decays geometrically with k. However, none of the works carried out so far in the field of the approximate reduction of systems with different time scales undertake the study of the approximation, i.e., the magnitude of this error, for finite values of k.

We will refer to the discrepancy between the original system and the auxiliary system as the "error" we make when carrying out the study of the original system in terms of that of the reduced system. This error can be measured in the form

$$\left\|\mathbf{X}_n - \mathbf{M}_{n-1}\mathbf{V}_{n-1}\mathbf{Y}_{n-1}\right\| = \left\|\mathbf{X}_n - \mathbf{X}'_n\right\| = \left\|\mathbf{E}_{n,k}\mathbf{X}_0\right\| \le \left\|\mathbf{E}_{n,k}\right\|\left\|\mathbf{X}_0\right\|$$

where $\|*\|$ denotes both a vector norm in $\mathbf{R}^N$ and its associated matrix norm, i.e., $\|\mathbf{A}\| = \max_{\|\mathbf{x}\|=1}\|\mathbf{A}\mathbf{x}\|$ for $\mathbf{A} \in \mathbf{R}^{N\times N}$. The error is thus characterized by $\|\mathbf{E}_{n,k}\|$ and the rest of the paper is basically devoted to obtaining upper bounds for this norm as functions of n and k.

In addition to the study of $\|\mathbf{E}_{n,k}\|$, in order to study the usefulness of the aggregated system to describe the dynamics of the original system, we will pay attention to the relative error which, considering the original system as a perturbation of the auxiliary system, is characterized by the quotient

$$R_{n,k} = \frac{\left\|\mathbf{E}_{n,k}\right\|}{\left\|\Pi'_n\right\|}.$$

Since we are interested in the error we make when estimating the total population, the natural norm to calculate the error is the 1-norm in $\mathbf{R}^N$, i.e., $\|\mathbf{z}\| = |z_1| + |z_2| + \cdots + |z_N|$ and its associated matrix norm in $\mathbf{R}^{N\times N}$, i.e., $\|\mathbf{A}\|_1$ is the

maximum of the 1-norm of the columns of $\mathbf{A}$. However, our work is more general and when we use the symbol $\|*\|$ we refer to any of the 1, 2 and ∞ norms in $\mathbf{R}^{N\times N}$.

4. ERROR BOUNDS

The presentation of the results of this Section is conceived to be useful to biologists and ecologists, and so the mathematical discussion has been kept to a minimum. The Appendix contains, together with the proofs of the results, some of the mathematical ideas behind the derivation of the bounds.

Let us consider the autonomous case, i.e., $\mathbf{M}_n = \mathbf{M}$, $\mathbf{P}_n = \mathbf{P}$ for all n, in order to illustrate the qualitative behavior of $\|\mathbf{E}_{n,k}\|$. Proposition (2) guarantees that if the matrix associated with the aggregated system has spectral radius lower (resp. higher) than one then so does that corresponding to the auxiliary system and, for large enough k, the original system. Therefore, if $\rho(\overline{\mathbf{M}}) < 1$, the population of both the auxiliary and the original system tends to zero when $n \to \infty$ and, consequently, the error $\|\mathbf{E}_{n,k}\|$ will also tend to zero. Analogously, if $\rho(\overline{\mathbf{M}}) > 1$ then, for large enough k the population of both models tends to infinity when $n \to \infty$. Since we are dealing with a multiplicative process, we expect the error we make in each time step to be amplified in time and to tend to infinity. This qualitative behavior is confirmed by numerical simulations. The qualitative behavior of the relative error is not so obvious and can be seen, in some particular cases, in Section 5.

The next result provides a first error bound:

Proposition 3. *Let us define*

$$\beta := \sup_n \|\mathbf{M}_n \overline{\mathbf{P}}_n\|; \quad \phi_k := \sup_n \|\mathbf{M}_n(\mathbf{P}_n^k - \overline{\mathbf{P}}_n)\|. \tag{13}$$

Then we have, for all n and k,

$$\|\mathbf{E}_{n,k}\| \le C_{n,k}^1 := \beta^n \left[\left(1 + \frac{\phi_k}{\beta} \right)^n - 1 \right] \tag{14}$$

and so the relative error $R_{n,k}$ verifies

$$R_{n,k} \le B_{n,k}^1 := \frac{C_{n,k}^1}{\|\Pi_n'\|}.$$

Proof. See Appendix.

Note that in order to obtain the previous bound, the suprema (13) must be finite. This situation presents itself again in the subsequent results. In Lemma 8 we provide sufficient conditions for all those suprema to be finite.

Let us illustrate an important qualitative shorthand of bound $C_{n,k}^1$. For the sake of simplicity let us consider the autonomous case: we can have a situation in which $\rho(\overline{\mathbf{MP}}) < 1$ (and therefore $\rho(\mathbf{MP}^k) < 1$ for large enough k) and however $\|\overline{\mathbf{MP}}\| > 1$. Therefore, for large enough k, $\Pi_n(k)$ and Π_n' tend to zero when n tends to infinity

and, consequently, so does the error $\left\|\mathbf{E}_{n,k}\right\|$. However the quantity $C^1_{n,k}$ tends to infinity, therefore giving useless bounds for practical purposes. For example, it is easy to check that this situation arises in the Leslie type model of Section 2 if we work with the 1-norm, take $\dfrac{F_2^1 q_2 + F_2^2 p_2}{p_2 + q_2} > 1$ and F_1^1, F_1^2 and the survival rates are close enough to zero.

In order to obtain $C^1_{n,k}$ we need to calculate the powers $\mathbf{P}_n^k$ for all n. The following result gives coarser bounds $C^{1a}_{n,k}$ and $C^{1b}_{n,k}$ that can be obtained without calculating those powers.

Proposition 4. *a) For each n and k we have*

$$\left\|\mathbf{M}_n(\mathbf{P}_n^k - \bar{\mathbf{P}}_n)\right\| \le \left\|\mathbf{M}_n\right\|\left\|\mathbf{P}_n^k - \bar{\mathbf{P}}_n\right\| \le \left\|\mathbf{M}_n\right\|\left\|\mathbf{P}_n - \bar{\mathbf{P}}_n\right\|^k \tag{15}$$

and therefore, if we define

$$m = \sup_n \left\|\mathbf{M}_n\right\|;\ \gamma = \sup_n \left\|\mathbf{P}_n - \bar{\mathbf{P}}_n\right\|, \tag{16}$$

it follows

$$\left\|\mathbf{E}_{n,k}\right\| \le C^{1a}_{n,k} := \beta^n \left[\left(1 + \frac{m}{\beta}\gamma^k\right)^n - 1\right] \tag{17}$$

$$R_{n,k} \le B^{1a}_{n,k} := \frac{C^{1a}_{n,k}}{\left\|\Pi'_n\right\|}.$$

b) Let us assume that $\mathbf{P}_i(n)$ is diagonalizable, for each n and each $i = 1,2,...,q$, in the form $\mathbf{P}_i(n) = \mathbf{Q}_i(n)\Sigma_i(n)\mathbf{Q}_i(n)^{-1}$ where $\Sigma_i(n)$ is a diagonal matrix. Then

$$\left\|\mathbf{M}_n(\mathbf{P}_n^k - \bar{\mathbf{P}}_n)\right\| \le \tau_n \left\|\mathbf{M}_n\right\|\left|\lambda_{q+1}(n)\right|^k \tag{18}$$

where $\tau_n := \max_{i=1,...,q}\left\{\left\|\mathbf{Q}_i(n)\right\|\left\|\mathbf{Q}_i(n)^{-1}\right\|\right\} > 1$. Therefore, given (11) and defining

$$\tau := \sup_n \tau_n \tag{19}$$

we have, for all n and k

$$\left\|\mathbf{E}_{n,k}\right\| \le C^{1b}_{n,k} := \beta^n \left[\left(1 + \frac{\tau m}{\beta}\delta^k\right)^n - 1\right] \tag{20}$$

$$R_{n,k} \le B^{1b}_{n,k} := \frac{C^{1b}_{n,k}}{\left\|\Pi'_n\right\|}.$$

c) $C^1_{n,k} \le \min\left\{C^{1a}_{n,k}, C^{1b}_{n,k}\right\}$ for all n and k.

d) $\left|\lambda_{q+1}(n)\right| \le \left\|\mathbf{P}_n - \bar{\mathbf{P}}_n\right\|$ for all n. As a consequence, if k is large enough so that $\tau\delta^k \le \gamma^k$, then $C^{1b}_{n,k} \le C^{1a}_{n,k}$ for all n.

Proof. See Appendix.

Note that the error bound $C^{1,b}_{n,k}$, although better than $C^{1,a}_{n,k}$ for large enough k, involves diagonalizing the diagonal blocks of $\mathbf{P}_n$ and is therefore more complicated to obtain than $C^{1,a}_{n,k}$. Bound $C^{1,b}_{n,k}$ is also useful qualitatively to study the decaying of the error in terms of the subdominant eigenvalues of the matrices that characterize the fast dynamics. Since any diagonalizable matrix can be diagonalized by an infinity of matrices, bound $C^{1,b}_{n,k}$ will depend on the choice of the $\mathbf{Q}_j(n)$.

Note that the bounds $C^{1}_{n,k}$, $C^{1,a}_{n,k}$ and $C^{1,b}_{n,k}$ for the absolute error can be calculated directly in terms of the data of the problem. However, in order to obtain the bounds $B^{1}_{n,k}$, $B^{1,a}_{n,k}$ and $B^{1,b}_{n,k}$ for the relative error, it is necessary to know $\left\|\Pi'_n\right\| = \left\|\mathbf{M}_{n-1}\mathbf{V}_{n-1}\overline{\Pi}_{n-1}\mathbf{U}_0\right\|$, i.e., it is necessary to know beforehand the behavior of the matrix products of the aggregated system.

We will now construct new error bounds that, as shown in Section 5, provide a better estimation of the error than $C^{1}_{n,k}$ in many practical situations. In addition, and unlike for the previous bounds $B^{1}_{n,k}$, $B^{1,a}_{n,k}$ and $B^{1,b}_{n,k}$, we will be able to obtain bounds for the relative error that do not require the knowledge of the dynamics of the aggregated system.

For notational convenience, let us define matrices

$$\mathbf{C}_n = \left[c^n_{\alpha\beta}\right] := \mathbf{M}_n\overline{\mathbf{P}}_n \geq \mathbf{0}; \quad \mathbf{D}_{n,k} = \left[d^{n,k}_{\alpha\beta}\right] := \mathbf{M}_n(\mathbf{P}^k_n - \overline{\mathbf{P}}_n)$$

and now let matrix $\mathbf{W}_{n,k} = \left[w^{n,k}_{\alpha\beta}\right]$ be given by

$$w^{n,k}_{\alpha\beta} := \begin{cases} \dfrac{d^{n,k}_{\alpha\beta}}{c^n_{\alpha\beta}} & \text{if } c^n_{\alpha\beta} \neq 0 \\ 0 & \text{if } c^n_{\alpha\beta} = 0 \end{cases}. \tag{21}$$

In the first place, let us contemplate a particular case of the general situation defined in Section 2 for which we will obtain results that will be both sharper and easier to obtain.

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

Note that condition (H) holds in the Leslie type model of Section 2. Moreover, all the applications that have been developed in the previous works on the field meet this requirement (Sánchez *et al.*, 1995; Sanz and Bravo de la Parra, 1998, 1999).

Now, given condition (H), we can simplify the calculation of matrices $\mathbf{W}_{n,k}$. Indeed we have:

Lemma 5. *Let* n *and* k *be fixed and let us assume hypothesis (H).*

a) Matrix $\mathbf{W}_{n,k}$ *can be obtained as*

$$w^{n,k}_{\alpha\beta} := \begin{cases} \dfrac{(\mathbf{P}^k_n - \overline{\mathbf{P}}_n)_{\alpha\beta}}{(\overline{\mathbf{P}}_n)_{\alpha\beta}} & \text{if } c^n_{\alpha\beta} \neq 0 \\ 0 & \text{if } c^n_{\alpha\beta} = 0 \end{cases}. \tag{22}$$

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

b) Moreover, if $\mathbf{P}_n^k \neq \overline{\mathbf{P}}_n$ *then* $\mathbf{W}_{n,k}$ *does not have a "definite sign", i.e., neither* $\mathbf{W}_{n,k}$ *nor* $-\mathbf{W}_{n,k}$ *are non-negative matrices.*

Proof. See Appendix.

Now, let numbers $\sigma_M(k)$, $\sigma_m(k)$ be defined as follows

$$\sigma_M(k) := \sup_n \left(\max \mathbf{W}_{n,k}\right);\ \sigma_m(k) := -\inf_n \left(\min \mathbf{W}_{n,k}\right) \tag{23}$$

where, for a matrix $\mathbf{A}$, $\max \mathbf{A}$, $\min \mathbf{A}$ denote, respectively, maximum and minimum of the entries of $\mathbf{A}$. Note that, given hypothesis (H) and excluding the trivial case $\mathbf{P}_n^k = \overline{\mathbf{P}}_n$ for all n, Lemma 5 guarantees that both $\sigma_M(k)$ and $\sigma_m(k)$ are positive numbers.

The next proposition provides a second error bound and, in the autonomous case, a relationship between the dominant eigenvalues of the original and reduced systems.

Proposition 6. *a) For all* n *and* k *we have*

$$\left\|\mathbf{E}_{n,k}\right\| \leq C_{n,k}^2 := B_{n,k}^2 \left\|\Pi'_n\right\| \tag{24}$$

where

$$B_{n,k}^2 := \max\left\{\left(1+\sigma_M(k)\right)^n - 1\ ,\ 1-\left(1-\sigma_m(k)\right)^n\right\}. \tag{25}$$

b) In the autonomous case, if we denote by λ_k *and* λ *the spectral radii of matrices* $\mathbf{MP}^k$ *and* $\overline{\mathbf{M}}$ *respectively, we have*

$$\left|\lambda_k - \lambda\right| \leq \max\left\{\sigma_M(k), \sigma_m(k)\right\}\lambda \tag{26}$$

for all k.

Proof. See Appendix.

Note that (24) shows that $B_{n,k}^2$ is a bound for the relative error. Moreover, given hypothesis (H), $B_{n,k}^2$ is independent of the value of the non-zero entries of the matrices defining the slow dynamics.

As in the case of bound $C_{n,k}^1$, in order to obtain $C_{n,k}^2$ it is necessary to calculate the powers $\mathbf{P}_n^k$ for all n. The following result gives a coarser bound $C_{n,k}^{2a}$ that can be more easily obtained.

Proposition 7. *a) Let* $\upsilon(k)$ *be defined by*

$$\upsilon(k) := \sup_n \left\{ \frac{\max \mathbf{M}_n}{\min^+\left(\mathbf{M}_n \overline{\mathbf{P}}_n\right)} \left\|\mathbf{P}_n - \overline{\mathbf{P}}_n\right\|_1^k \right\} \tag{27}$$

where $\min^+$ *denotes the minimum of the positive entries of the corresponding matrix.*

Then we have

$$\max\left\{\sigma_M(k), \sigma_m(k)\right\} \leq \upsilon(k) \tag{28}$$

and consequently

$$R_{n,k} \le B^{2}_{n,k} \le B^{2a}_{n,k} := \max\left\{\left(1+\upsilon(k)\right)^{n}-1\ ,\ 1-\left(1-\upsilon(k)\right)^{n}\right\}. \tag{29}$$

b) Moreover, if hypothesis (H) holds, bound $B^{2a}_{n,k}$ *can be improved by defining*

$$\upsilon_{H}(k) := \sup_{n}\left\{\frac{\left\|\mathbf{P}_{n}-\overline{\mathbf{P}}_{n}\right\|_{1}^{k}}{\min^{+}\overline{\mathbf{P}}_{n}}\right\} \tag{30}$$

which verifies $\upsilon_{H}(k) \le \upsilon(k)$*, and replacing* $\upsilon(k)$ *by* $\upsilon_{H}(k)$ *in (29).*

Proof. See Appendix.

Finally, the following lemma gives sufficient conditions for parameters β, ϕ_k, m, γ, $\sigma_{M}(k)$, $\sigma_{m}(k)$, $\upsilon(k)$ and $\upsilon_{H}(k)$ to be finite.

Lemma 8. *Assume that the non-zero entries of matrices* $\mathbf{M}_n$ *and* $\mathbf{P}_n$ *are bounded away from zero and infinity, i.e., there exist positive constants* ε, K, ε' *and* K' *such that for all* n *we have*

$$\min^{+}\left(\mathbf{M}_{n}\right) \ge \varepsilon;\ \ \min^{+}\left(\mathbf{P}_{n}\right) \ge \varepsilon' \tag{31}$$

$$\max\left(\mathbf{M}_{n}\right) \le K;\ \ \max\left(\mathbf{P}_{n}\right) \le K'.$$

Then, for any value of k*, the suprema of (16), (13), (27), (30) and the supremum and infimum of (23) are finite.*

Proof. See Appendix.

5. NUMERICAL SIMULATIONS

In order to illustrate the use of the bounds we have obtained, we show several numerical simulations corresponding to the Leslie type model of Section 2. In all of the simulations we deal with the autonomous case and we employ the 1-norm as a measure of the approximation.

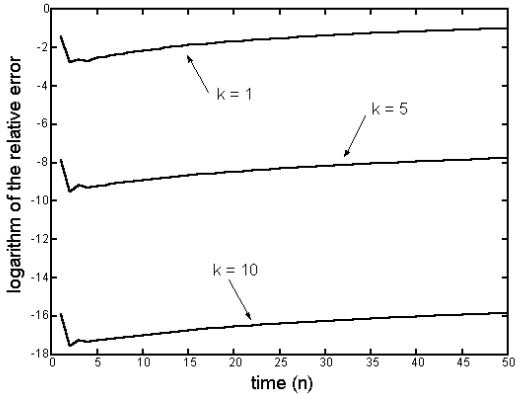

Figure 1. Logarithm of the error $R_{n,k}$ as a function of time for different values of k. The parameter values are those listed under (32).

In the first place, and according to the discussion of Section 4, numerical simulations corroborate that the absolute error we make tends to zero or infinity

depending on whether $\rho\left(\overline{\mathbf{M}}\right) < 1$ or $\rho\left(\overline{\mathbf{M}}\right) > 1$. On the other hand, intensive numerical simulation has shown that, irrespective of whether $\rho\left(\overline{\mathbf{M}}\right) < 1$ or $\rho\left(\overline{\mathbf{M}}\right) > 1$, the relative error always grows with time for large enough values of n.

In order to test the accuracy of the different bounds we have obtained, let us consider case (A) corresponding to the following set of values for the demographic and migratory parameters

$$p_1 = 0.4;\quad p_2 = 0.5;\quad q_1 = 0.7;\quad q_2 = 0.3\ ;\quad F_1^1 = 0.3\ ;\quad F_1^2 = 0.2 \qquad (32)$$

$$F_2^1 = 2\ ;\quad F_2^2 = 3\ ;\quad S_1^1 = 0.4\ ;\quad S_1^2 = 0.6\ ;\quad S_2^1 = 0.2\ ;\quad S_2^2 = 0.3$$

for which $\rho\left(\overline{\mathbf{M}}\right) = 1.3$. Figure 1 shows the evolution of the relative error $R_{n,k}$ with time for different values of k. Since our system is multiplicative we have chosen a logarithmic scale for the error. Note that the relative error grows with time for large enough n.

Figure 2 shows the relative error and the different bounds for it introduced in the paper as a function of time when $k = 5$.

We can see how $B_{n,k}^2$ has an asymptotic growth rate similar to that of the real relative error $R_{n,k}$. On the contrary, the discrepancy between $B_{n,k}^1$ and $R_{n,k}$ increases significantly with time. Note also that $B_{n,k}^2$ works better than $B_{n,k}^{2a}$.

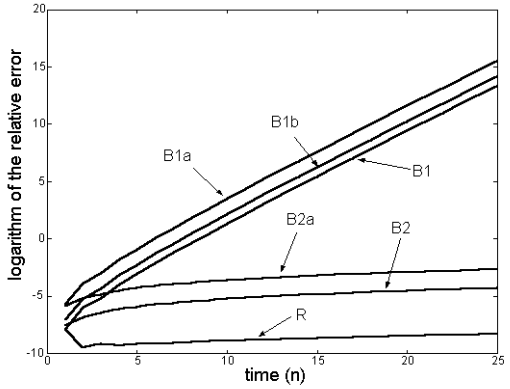

Figure 2. Logarithm of the error $R_{n,k}$ and of the different error bounds for the parameter values defined by (32) and $k = 5$. R, B1, B1a, B1b, B2, B2a stand for $R_{n,k}$, $B_{n,k}^1$, $B_{n,k}^{1a}$, $B_{n,k}^{1b}$, $B_{n,k}^2$ and $B_{n,k}^{2a}$.

We now consider case (B), corresponding to the following set of parameters,

$$k = 5;\ p_1 = 0.45;\quad p_2 = 0.05;\quad q_1 = 0.3;\quad q_2 = 0.45;\quad F_1^1 = 0.8;\quad F_1^2 = 0.4 \qquad (33)$$

$$F_2^1 = 0.5;\quad F_2^2 = 2;\quad S_1^1 = 0.5;\quad S_1^2 = 0.6;\quad S_2^1 = 0.2;\quad S_2^2 = 0.3$$

for which $\rho\left(\overline{\mathbf{M}}\right) = 1.01$. The relative error, together with the different bounds, is shown in Figure 3.

Unlike in situation (A), in case (B) $B_{n,k}^1$ is a more exact estimate of the error than $B_{n,k}^2$.

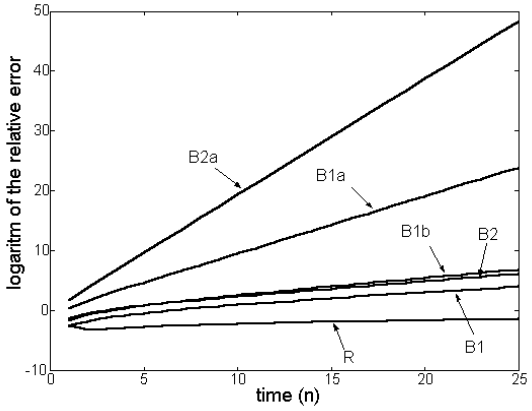

Figure 3. Logarithm of the error $R_{n,k}$ and of the different error bounds for the parameter values defined by (33). R, B1, B1a, B1b, B2, B2a are defined in Figure 2.

In the previous cases, bounds $B^1_{n,k}$ and $B^2_{n,k}$ largely over-estimate the actual error $R_{n,k}$. However, there are situations in which the bounds we have built provide quite sharp estimations of $R_{n,k}$. Indeed, some examples are those corresponding to the set of parameters (34) and (35) for which we have, respectively, Figures 4 and 5:

$$k=5;\ p_1=0.45;\ p_2=0.41;\ q_1=0.89;\ q_2=0.19;\ F_1^1=0.27;\ F_1^2=0.01 \quad (34)$$

$$F_2^1=0.85;\ F_2^2=0.9;\ S_1^1=0.28;\ S_1^2=0.8;\ S_2^1=0.55;\ S_2^2=0.8;$$

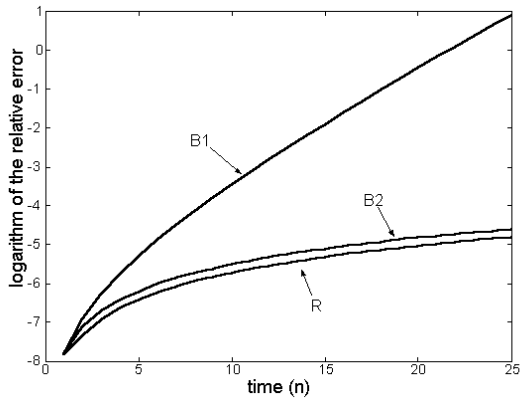

Figure 4. Logarithm of the error $R_{n,k}$, $B^1_{n,k}$ and $B^2_{n,k}$ for the parameter values defined by (34).

$$k=5;\ p_1=0.39;\ p_2=0.54;\ q_1=0.40;\ q_2=0.30;\ F_1^1=1;\ F_1^2=0.02 \quad (35)$$

$$F_2^1=0.32;\ F_2^2=0.32;\ S_1^1=0.01;\ S_1^2=0.01;\ S_2^1=0.1;\ S_2^2=0.1.$$

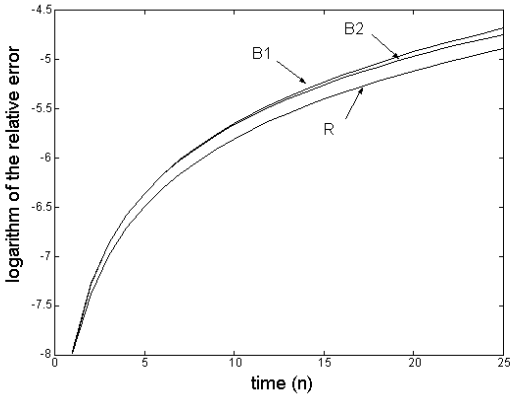

Figure 5. Logarithm of the error $R_{n,k}$, $B^1_{n,k}$ and $B^2_{n,k}$ for the parameter values defined by (35).

6. CONCLUSION

Our work allows one to estimate both the absolute and the relative error we make when carrying out the aggregation of linear non-autonomous models with two time scales. In this way the modeler can gain some insight into the magnitude of the approximation involved in the reduction of each model.

Each of the two main bounds obtained proves to work better than the other in some situations. Therefore, the context will dictate which one to use in each particular case. Moreover, both bounds can be relaxed in order to facilitate their calculation. In this way we can simplify the computation of the bounds at the price of losing some tightness.

As future contributions, we plan to estimate the magnitude of the discrepancy between the original and aggregated system in two different contexts. The first is that corresponding to non-linear models following the approach of Bravo de la Parra *et al.* (1999). The second is that of the systems subjected to environmental stochasticity (Sanz and Bravo de la Parra, 2000) where it would be interesting to obtain a relationship between the stochastic growth rates (Tuljapurkar, 1990) of the original and the aggregated systems.

APPENDIX

Proof of Proposition 3. Throughout the subsequent proofs we will use repeatedly the properties of "subadditivity" and "submultiplicativity" of induced matrix norms of which the norms 1, 2 and ∞ are particular cases. In this way, the norm of a sum (product) of matrices is lower or equal than the sum (product) of the norms of the matrices under consideration.

Let n and k be fixed and let $\mathbf{A}_n := \mathbf{M}_n \overline{\mathbf{P}}_n$ and $\mathbf{B}_n := \mathbf{M}_n \mathbf{S}_{n,k}$. Then

$$\begin{aligned}\Pi_{n,k} &= \left(\mathbf{A}_{n-1} + \mathbf{B}_{n-1}\right)\left(\mathbf{A}_{n-2} + \mathbf{B}_{n-2}\right)\cdots\left(\mathbf{A}_0 + \mathbf{B}_0\right) \\ &= \mathbf{A}_{n-1}\mathbf{A}_{n-2}\cdots\mathbf{A}_0 + \cdots = \Pi'_n + \mathbf{E}_{n,k}\end{aligned}$$

where $\mathbf{E}_{n,k}$ is a sum of 2^n-1 summands of the form $\mathbf{H}_i\mathbf{H}_{i-1}\cdots\mathbf{H}_1$ where, for each *i*, $\mathbf{H}_i \in \{\mathbf{A}_i, \mathbf{B}_i\}$ excluding the case $\mathbf{H}_i = \mathbf{A}_i$ for all *i*. We derive an upper bound for the norm of $\mathbf{E}_{n,k}$ by bounding the norm of the sum by the sum of the norm of the summands and, in each of the 2^n-1 summands, the norm of the product $\mathbf{H}_i\mathbf{H}_{i-1}\cdots\mathbf{H}_1$ by the product of the norms of the factors. Now we use the fact that $\|\mathbf{A}_i\| \le \beta$ and $\|\mathbf{B}_i\| \le \phi_k$ and we finally obtain

$$\begin{aligned}\|\mathbf{E}_{n,k}\| &\le \binom{n}{1}\beta^{n-1}\phi_k^1 + \binom{n}{2}\beta^{n-2}\phi_k^2 + \cdots + \binom{n}{n}\phi_k^n \\ &= (\beta+\phi_k)^n - \beta^n = \beta^n\left[\left(1+\frac{\phi_k}{\beta}\right)^n - 1\right] = C_{n,k}^1\end{aligned}$$

as desired.

Proof of Proposition 4. a) It is straightforward to see

$$\mathbf{P}_n^k - \bar{\mathbf{P}}_n = \left(\mathbf{P}_n - \bar{\mathbf{P}}_n\right)^k. \tag{36}$$

Indeed, proceeding by induction in k, the result is obvious for $k=1$ and, assuming $\mathbf{P}_n^k - \bar{\mathbf{P}}_n = \left(\mathbf{P}_n - \bar{\mathbf{P}}_n\right)^k$ then $\mathbf{P}_n^{k+1} - \bar{\mathbf{P}}_n = \left(\mathbf{P}_n - \bar{\mathbf{P}}_n\right)\left(\mathbf{P}_n^k - \bar{\mathbf{P}}_n\right) = \left(\mathbf{P}_n - \bar{\mathbf{P}}_n\right)\left(\mathbf{P}_n - \bar{\mathbf{P}}_n\right)^k = \left(\mathbf{P}_n - \bar{\mathbf{P}}_n\right)^{k+1}$ where we have used $\bar{\mathbf{P}}_n\bar{\mathbf{P}}_n = \bar{\mathbf{P}}_n$. Now (15) follows from the submultiplicativity of matrix norms and, as a consequence, bound (17) follows directly from (14).

(b) Let n and i be fixed and let us suppose, without loss of generality, that the eigenvalues of $\Sigma_i(n)$ are ordered by decreasing modulus. Then $\Sigma_i(n) = diag(1, \alpha_{i2}(n), \ldots, \alpha_{iN_i}(n))$ where $|\alpha_{ij}(n)| < 1$ for all $j = 2, \ldots, N_i$. Moreover, since the right and left eigenvectors of $\mathbf{P}_i(n)$ associated with 1 are $\mathbf{v}_i(n)$ and $\mathbf{u}_i(n)$ and $\mathbf{u}_i^T(n)\mathbf{v}_i(n) = 1$ then it follows $\bar{\mathbf{P}}_i(n) = \mathbf{v}_i(n)\mathbf{u}_i^T(n) = \mathbf{Q}_i(n)\,diag(1,0,\ldots,0)\mathbf{Q}_i(n)^{-1}$ and consequently, denoting the k-th power of $\mathbf{P}_i(n)$ and $\alpha_{ij}(n)$ by $\mathbf{P}_i^k(n)$ and $\alpha_{ij}^k(n)$ respectively,

$$\begin{aligned}\mathbf{P}_i^k(n) - \bar{\mathbf{P}}_i(n) &= \mathbf{Q}_i(n)\,diag(1, \alpha_{i2}^k(n), \ldots, \alpha_{iN_i}^k(n))\mathbf{Q}_i(n)^{-1} \\ &\quad - \mathbf{Q}_i(n)\,diag(1,0,\ldots,0)\mathbf{Q}_i(n)^{-1} \\ &= \mathbf{Q}_i(n)\,diag(0, \alpha_{i2}^k(n), \ldots, \alpha_{iN_i}^k(n))\mathbf{Q}_i(n)^{-1}\end{aligned}$$

so that

$$\begin{aligned}\left\|\mathbf{P}_i^k(n) - \bar{\mathbf{P}}_i(n)\right\| &\le \|\mathbf{Q}_i(n)\|\|\mathbf{Q}_i(n)^{-1}\|\left\|diag(0, \alpha_{i2}^k(n), \ldots, \alpha_{iN_i}^k(n))\right\| \\ &= \|\mathbf{Q}_i(n)\|\|\mathbf{Q}_i(n)^{-1}\|\left|\alpha_{i2}(n)\right|^k\end{aligned}$$

where we have used the fact that if $\|*\|$ is any of the 1, 2 and ∞ matrix norms, then the norm of a diagonal matrix is the maximum of the modulus of the diagonal elements.

Analogously, the norm of a block diagonal matrix is the maximum of the norms of the diagonal blocks and so

$$\left\|\mathbf{P}_n^k - \overline{\mathbf{P}}_n\right\| \le \left\|diag(\mathbf{P}_1^k(n) - \overline{\mathbf{P}}_1(n), \ldots, \mathbf{P}_q^k(n) - \overline{\mathbf{P}}_q(n))\right\|$$

$$= \max_{i=1,\ldots,q}\left\{\left\|\mathbf{Q}_i(n)\right\|\left\|\mathbf{Q}_i(n)^{-1}\right\|\left|\alpha_{i2}(n)\right|^k\right\}$$

$$\le \max_{i=1,\ldots,q}\left\{\left\|\mathbf{Q}_i(n)\right\|\left\|\mathbf{Q}_i(n)^{-1}\right\|\right\}\max_{i=1,\ldots,q}\left\{\left|\alpha_{i2}(n)\right|^k\right\} = \tau_n\left|\lambda_{q+1}(n)\right|^k,$$

and so (18) follows. Thus, bound (20) is a straightforward consequence of (14). Finally, for each i, $\left\|\mathbf{Q}_i(n)\right\|\left\|\mathbf{Q}_i(n)^{-1}\right\|$ is the condition number of $\mathbf{Q}_i(n)$, and since the condition number of a matrix is always greater or equal to one (Horn and Johnson, 1985, p. 336) we have $\tau_n \ge 1$.

(c) Obvious from the proofs of (a) and (b).

(d) The spectral radius of $\mathbf{P}_n - \overline{\mathbf{P}}_n$ is $\rho\left(\mathbf{P}_n - \overline{\mathbf{P}}_n\right) = \left|\lambda_{q+1}(n)\right|$ and since $\rho(\mathbf{A}) \le \left\|\mathbf{A}\right\|$ for any square matrix $\mathbf{A}$ and any induced matrix norm it follows $\left|\lambda_{q+1}(n)\right| \le \left\|\mathbf{P}_n - \overline{\mathbf{P}}_n\right\|$. The result then follows directly.

Proof of Lemma 5. a) The result easily follows taking into account the block structure of the matrices under consideration. Let us denote $\mathbf{S}_j(n,k) = \left[S_j^{rl}(n,k)\right] := \mathbf{P}_j^k(n) - \overline{\mathbf{P}}_j(n)$ for each $j = 1,\ldots,q$. Due to the block structure of matrices $\mathbf{M}_n$, $\mathbf{P}_n^k$ and $\overline{\mathbf{P}}_n$, matrix $\mathbf{D}_{n,k} = \mathbf{M}_n(\mathbf{P}_n^k - \overline{\mathbf{P}}_n) = \left[d_{\alpha\beta}^{n,k}\right] \in \mathbf{R}^{N\times N}$ (resp. $\mathbf{C}_n = \mathbf{M}_n\overline{\mathbf{P}}_n = \left[c_{\alpha\beta}^n\right] \in \mathbf{R}^{N\times N}$) can be thought of as being composed of q^2 blocks $\mathbf{D}_{ij}(n,k) = \left[d_{ij}^{rl}(n,k)\right] := \mathbf{M}_{ij}(n)(\mathbf{P}_j^k(n) - \overline{\mathbf{P}}_j(n)) = \mathbf{M}_{ij}(n)\mathbf{S}_j(n,k)$ (resp. $\mathbf{C}_{ij}(n) = \left[c_{ij}^{rl}(n)\right] := \mathbf{M}_{ij}(n)\overline{\mathbf{P}}_j(n)$); $i,j = 1,\ldots,q$. Analogously, matrix $\mathbf{W}_{n,k} = \left[w_{\alpha\beta}^{n,k}\right]$ can be thought of as composed of q^2 blocks $\mathbf{W}_{ij} = \left[w_{ij}^{rl}(n,k)\right]$, $i,j = 1,\ldots,q$. Let $i,j \in \{1,\ldots,q\}$ and $r,l \in \{1,\ldots,N_i\}$ be fixed. Since $\mathbf{M}_{ij}(n) = \left[M_{ij}^{rl}(n)\right] \in \mathbf{R}^{N_i\times N_i}$ is diagonal, then $c_{ij}^{rl}(n) = M_{ij}^{rr}(n)\overline{P}_j^{rl}(n)$ and $d_{ij}^{rl}(n,k) = M_{ij}^{rr}(n)S_j^{rl}(n,k)$. Note that the fact that the $\overline{\mathbf{P}}_j(n)$ are positive matrices implies $c_{ij}^{rl}(n) = 0$ if and only if $M_{ij}^{rr}(n)$. Then, by definition (21) we have that $c_{ij}^{rl}(n) = 0$ implies $w_{ij}^{rl}(n,k) = 0$. Moreover, if $c_{ij}^{rl}(n) \ne 0$ then $M_{ij}^{rr}(n) \ne 0$ and $w_{ij}^{rl}(n,k) = \dfrac{M_{ij}^{rr}(n)S_j^{rl}(n,k)}{M_{ij}^{rr}(n)\overline{P}_j^{rl}(n)} = \dfrac{S_j^{rl}(n,k)}{\overline{P}_j^{rl}(n)}$. Result (22) is then proved.

b) According to the definition of $\mathbf{W}_{n,k}$, and taking into account that matrix $\overline{\mathbf{P}}_n$ is non-negative, it suffices to prove that neither $\mathbf{P}_n^k - \overline{\mathbf{P}}_n$ nor $\overline{\mathbf{P}}_n - \mathbf{P}_n^k$ are non-negative matrices. Since $\mathbf{P}_n^k - \overline{\mathbf{P}}_n = diag\left(\mathbf{P}_1^k(n) - \overline{\mathbf{P}}_1(n), \ldots, \mathbf{P}_q^k(n) - \overline{\mathbf{P}}_q(n)\right)$ we will deal with the diagonal blocks. Let $i = 1,\ldots,q$ be fixed. We assume that $\mathbf{P}_i^k(n) \ge \overline{\mathbf{P}}_i(n)$ and will reach

a contradiction. Matrices $\mathbf{P}_i^k(n)$ and $\overline{\mathbf{P}}_i(n)$ are primitive (therefore irreducible) and, for both, the spectral radius is equal to one. Now we will use the following result (Horn and Johnson, 1985, p. 509): Let $\mathbf{A}$, $\mathbf{B}$ be non-negative irreducible square matrices of the same size such that $\mathbf{A} \geq \mathbf{B}$ and $\rho(\mathbf{A}) = \rho(\mathbf{B})$. Then, if $\lambda = \mathrm{e}^{\mathrm{i}\phi}\rho(\mathbf{B})$ is an eigenvalue of $\mathbf{B}$ then there exist $\theta_1, \ldots, \theta_N \in \mathbf{R}$ such that $\mathbf{B} = \mathrm{e}^{i\phi}\mathbf{D}\mathbf{A}\mathbf{D}^{-1}$ where $\mathbf{D} = diag(\mathrm{e}^{\mathrm{i}\theta_1}, \ldots, \mathrm{e}^{\mathrm{i}\theta_N})$. Now we apply this theorem with $\mathbf{A} = \mathbf{P}_i^k(n)$, $\mathbf{B} = \overline{\mathbf{P}}_i(n)$ and $\phi = 0$ and we have as a consequence $\overline{\mathbf{P}}_i(n) = \mathbf{D}^{-1}\mathbf{P}_i^k(n)\mathbf{D}$ for a certain regular matrix $\mathbf{D}$, i.e., $\mathbf{P}_i^k(n)$ and $\overline{\mathbf{P}}_i(n)$ are similar matrices. Therefore both matrices have the same eigenvalues and so those of $\mathbf{P}_i^k(n)$ must be 1,0,0,...,0. But then, since $\mathbf{v}_i(n)$ and $\mathbf{u}_i(n)$ are, respectively, right and left eigenvectors of $\mathbf{P}_n^k$ associated with eigenvalue 1 that verify the normalization condition $\mathbf{u}_i^T(n)\mathbf{v}_i(n) = 1$ then $\mathbf{P}_i^k(n) = \mathbf{v}_i(n)\mathbf{u}_i^T(n)$, i.e., $\mathbf{P}_i^k(n) = \overline{\mathbf{P}}_i(n)$ which is impossible by hypothesis. Therefore $\mathbf{P}_i^k(n)$ is not greater than or equal to $\overline{\mathbf{P}}_i(n)$. Applying the same theorem in the case $\mathbf{B} = \mathbf{P}_i^k(n)$, $\mathbf{A} = \overline{\mathbf{P}}_i(n)$ and reasoning analogously we obtain that $\overline{\mathbf{P}}_i(n)$ is not greater than or equal to $\mathbf{P}_i^k(n)$ and the result is proved.

Proof of Proposition 6. a) We define, for all n and k, $\mathbf{T}_{n,k} = \left[t_{\alpha\beta}^{n,k}\right] := \mathbf{M}_n\mathbf{P}_n^k$ and then we have directly from (21) that if $c_{\alpha\beta}^n \neq 0$ then $t_{\alpha\beta}^{n,k} = c_{\alpha\beta}^n\left(1 + w_{\alpha\beta}^{n,k}\right)$(*). Let us show that $c_{\alpha\beta}^n = 0 \Rightarrow t_{\alpha\beta}^{n,k} = 0$ and so the expression (*) will be valid for all values of α and β. Indeed, matrix $\mathbf{C}_n = \mathbf{M}_n\overline{\mathbf{P}}_n$ (resp. $\mathbf{T}_{n,k} = \mathbf{M}_n\mathbf{P}_n^k$) is composed of q^2 blocks $\mathbf{M}_{ij}(n)\overline{\mathbf{P}}_j(n)$ (resp. $\mathbf{M}_{ij}(n)\mathbf{P}_j^k(n)$); $i, j = 1, \ldots, q$. Since matrices $\overline{\mathbf{P}}_j(n)$ are positive for all j, that implies that if the element (r, l) of $\mathbf{M}_{ij}(n)\overline{\mathbf{P}}_j(n)$ is zero then the r-th row of $\mathbf{M}_{ij}(n)$ must be zero and therefore the element (r, l) of $\mathbf{M}_{ij}(n)\mathbf{P}_j^k(n)$ is also zero. Thus $c_{\alpha\beta}^n = 0 \Rightarrow t_{\alpha\beta}^{n,k} = 0$ as required.

Let n and k be fixed and let $\Pi_{n,k} = \left[\Pi_{\alpha\beta}^{n,k}\right]$ and $\Pi'_n = \left[\Pi'^n_{\alpha\beta}\right]$. Let us fix $\alpha, \beta \in \{1, 2, \ldots, N\}$. We can write

$$\begin{aligned}
\Pi_{\alpha\beta}^{n,k} &= \sum_{I_{\alpha,\beta}} t_{\alpha h_{n-1}}^{n-1} \cdots t_{h_2 h_1}^{1} t_{h_1\beta}^{0} \\
&= \sum_{I_{\alpha,\beta}} c_{\alpha h_{n-1}}^{n-1} \cdots c_{h_2 h_1}^{1} c_{h_1\beta}^{0}\left(1 + w_{\alpha h_{n-1}}^{n-1,k}\right)\cdots\left(1 + w_{h_2 h_1}^{1,k}\right)\left(1 + w_{h_1\beta}^{0,k}\right) \\
\Pi'^n_{\alpha\beta} &= \sum_{I_{\alpha,\beta}} c_{\alpha h_{n-1}}^{n-1} c_{h_{n-1}h_{n-2}}^{n-2} \cdots c_{h_2 h_1}^{1} c_{h_1\beta}^{0}
\end{aligned}$$

where $I_{\alpha,\beta}$ is the following set of indexes:

$$I_{\alpha,\beta} := \left\{(h_{n-1}, \ldots, h_2, h_1) : h_i = 1, \ldots, N;\ i = 1, \ldots, n-1\right\}.$$

Therefore

$$\Pi_{\alpha\beta}^{n,k} \leq \sum_{I_{\alpha,\beta}} c_{\alpha h_{n-1}}^{n-1} \cdots c_{h_2 h_1}^{1} c_{h_1\beta}^{0}\left(1 + \sigma_M(k)\right)^n = \left(1 + \sigma_M(k)\right)^n \Pi'^n_{\alpha\beta}$$

$$\Pi_{\alpha\beta}^{n,k} \geq \sum_{l_{\alpha,\beta}} c_{\alpha h_{n-1}}^{n-1} \cdots c_{h_2 h_1}^{1} c_{h_1 \beta}^{0} (1-\sigma_m(k))^n = (1-\sigma_m(k))^n \Pi_{\alpha\beta}'^n$$

i.e.,

$$\Pi_{\alpha\beta}^{n,k} - \Pi_{\alpha\beta}'^n \leq ((1+\sigma_M(k))^n - 1)\Pi_{\alpha\beta}'^n$$

$$\Pi_{\alpha\beta}'^n - \Pi_{\alpha\beta}^{n,k} \leq (1-(1-\sigma_m(k))^n)\Pi_{\alpha\beta}'^n$$

from where $\left|\Pi_{\alpha\beta}^{n,k} - \Pi_{\alpha\beta}'^n\right| \leq \max\left\{(1+\sigma_M(k))^n - 1, 1-(1-\sigma_m(k))^n\right\}\Pi_{\alpha\beta}'^n$. Now (24) follows by taking norms in the last expression and using that, for any matrix norm induced by a monotone vector norm (of which the 1, 2 and ∞ norms are particular cases) (Horn and Johnson, 1985, p. 310), $|\mathbf{A}| \leq \mathbf{B} \Rightarrow \|\mathbf{A}\| \leq \|\mathbf{B}\|$ ($|\mathbf{A}|$ denotes the matrix whose elements are the absolute value of the elements of $\mathbf{A}$).

b) In the first place, the spectral radii of matrices $\mathbf{M}\overline{\mathbf{P}}$ and $\overline{\mathbf{M}}$ coincide (Proposition 2). From (*) we have, in the autonomous case, $t_{\alpha\beta}^n = c_{\alpha\beta}^n(1+w_{\alpha\beta}^n)$ from where it follows $t_{\alpha\beta}^n \leq c_{\alpha\beta}^n(1+\sigma_M(k))$ for all (α,β), i.e., $\mathbf{M}\mathbf{P}^k \leq (1+\sigma_M(k))\mathbf{M}\overline{\mathbf{P}}$ and since $|\mathbf{A}| \leq \mathbf{B} \Rightarrow \|\mathbf{A}\| \leq \|\mathbf{B}\|$ (Horn and Johnson, 1985, p. 491) we have $\lambda_k \leq (1+\sigma_M(k))\lambda$, i.e., $\lambda_k - \lambda \leq \sigma_M(k)\lambda$ (**). Analogously we have $t_{\alpha\beta}^n \geq c_{\alpha\beta}^n(1-\sigma_m(k))$ from where it follows $\mathbf{M}\mathbf{P}^k \geq (1-\sigma_m(k))\mathbf{M}\overline{\mathbf{P}}$ and then $\lambda_k \geq (1-\sigma_m(k))\lambda$, i.e., $\lambda - \lambda_k - \leq \sigma_m(k)\lambda$ (***). The desired result follows from (**) and (***).

Proof of Proposition 7. a) In the first place let us show that if $\mathbf{A}$ and $\mathbf{B}$ are square matrices of the same order then $|\mathbf{AB}| \leq \max|\mathbf{A}|\|\mathbf{B}\|_1$ (I). Indeed,

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

and so $\max_{ij}\left|(\mathbf{AB})_{ij}\right| \leq \max|\mathbf{A}|\max_j \sum_l |B_{lj}| = \max|\mathbf{A}|\|\mathbf{B}\|_1$ as required.

Let n and k be fixed. By definition (21) we have that if α and β are such that $c_{\alpha\beta}^n = 0$ then $|w_{\alpha\beta}^{n,k}| = 0$. In the case $c_{\alpha\beta}^n \neq 0$ we can write

$$|w_{\alpha\beta}^{n,k}| = \left|\frac{d_{\alpha\beta}^{n,k}}{c_{\alpha\beta}^n}\right| \leq \frac{\max\left|\mathbf{M}_n\left(\mathbf{P}_n^k - \overline{\mathbf{P}}_n\right)\right|}{\min^+\left(\mathbf{M}_n\overline{\mathbf{P}}_n\right)} \leq \frac{\max \mathbf{M}_n}{\min^+\left(\mathbf{M}_n\overline{\mathbf{P}}_n\right)}\left\|\mathbf{P}_n^k - \overline{\mathbf{P}}_n\right\|_1$$

$$\leq \frac{\max \mathbf{M}_n}{\min^+\left(\mathbf{M}_n\overline{\mathbf{P}}_n\right)}\left\|\mathbf{P}_n - \overline{\mathbf{P}}_n\right\|_1^k \leq \upsilon(k)$$

where in the second inequality we have used (I) and, in the third one, (36) and the submultiplicativity of the 1-matrix norm. From the last expression we have $\sup_n \max|\mathbf{W}_{n,k}| \leq \upsilon(k)$ and, since

$$\begin{aligned}\sup_n \max\left|\mathbf{W}_{n,k}\right| &= \sup_n \max\left\{\max \mathbf{W}_{n,k}, -\min \mathbf{W}_{n,k}\right\} \\ &= \max\left\{\sup_n\left(\max \mathbf{W}_{n,k}\right), -\inf_n\left(\min \mathbf{W}_{n,k}\right)\right\} = \max\left\{\sigma_M(k), \sigma_m(k)\right\}\end{aligned}$$

then $\max\left\{\sigma_M(k), \sigma_m(k)\right\} \le \upsilon(k)$. Therefore, $\left(1+\sigma_M(k)\right)^n - 1 \le \left(1+\upsilon(k)\right)^n - 1$ and $1-\left(1-\sigma_m(k)\right)^n \le 1-\left(1-\upsilon(k)\right)^n$ and so the bound (29) now follows directly from (24) and (25).

b) Given hypothesis (H) we use (22) and, if α and β are such that $c_{\alpha\beta}^n \neq 0$, it follows

$$\left|w_{\alpha\beta}^{n,k}\right| = \left|\frac{\left(\mathbf{P}_n^k - \overline{\mathbf{P}}_n\right)_{\alpha\beta}}{\left(\overline{\mathbf{P}}_n\right)_{\alpha\beta}}\right| \le \frac{\max\left|\mathbf{P}_n^k - \overline{\mathbf{P}}_n\right|}{\min^+ \overline{\mathbf{P}}_n} \le \frac{\left\|\mathbf{P}_n^k - \overline{\mathbf{P}}_n\right\|_1}{\min^+ \overline{\mathbf{P}}_n} \le \frac{\left\|\mathbf{P}_n - \overline{\mathbf{P}}_n\right\|_1^k}{\min^+ \overline{\mathbf{P}}_n}$$

where in the second inequality we have used $\max \mathbf{A} \le \left\|\mathbf{A}\right\|_1$ and, in the third one, (36) and the submultiplicativity of the 1-matrix norm. Now $\sup_n \max\left|\mathbf{W}_{n,k}\right| \le \upsilon_H(k)$ and, reasoning similarly to the proof of (a), $\max\left\{\sigma_M(k), \sigma_m(k)\right\} \le \upsilon_H(k)$ from where (29) with $\upsilon(k)$ replaced by $\upsilon_H(k)$ follows immediately.

Lastly, let us prove $\upsilon_H(k) \le \upsilon(k)$ for all k. Let n be fixed. Reasoning as in the proof of Lemma 5 on the block structure of matrix $\mathbf{C}_n$ we have

$$c_{ij}^{rl}(n) = M_{ij}^{rr}(n)\overline{P}_j^{rl}(n) \le \max \mathbf{M}_{ij}(n)\overline{P}_j^{rl}(n)$$

for all $i, j \in \left\{1, \ldots, q\right\}$ and $r, l \in \left\{1, \ldots, \mathcal{N}_i\right\}$. From here we have

$$\min^+ \mathbf{C}_{ij}(n) \le \max \mathbf{M}_{ij}(n) \min \overline{\mathbf{P}}_j(n) \le \max \mathbf{M}_n \min \overline{\mathbf{P}}_j(n).$$

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

$$\frac{\max \mathbf{M}_n}{\min^+\left(\mathbf{M}_n \overline{\mathbf{P}}_n\right)}\left\|\mathbf{P}_n - \overline{\mathbf{P}}_n\right\|_1^k \ge \frac{\left\|\mathbf{P}_n - \overline{\mathbf{P}}_n\right\|_1^k}{\min^+ \overline{\mathbf{P}}_n}$$

and the result follows by taking the supremum in both sides.

Proof of Lemma 8. In the first place, we make use of Proposition 3 in Sanz and Bravo de la Parra (2001). In the proof of that result the authors show that, given conditions (31), there exist numbers ε'' and K'' such that $\min^+\left(\overline{\mathbf{P}}_n\right) \ge \varepsilon''$ and $\max\left(\overline{\mathbf{P}}_n\right) \le K''$ (*).

Since for any square matrix of size $\mathcal{N}$ we have $\left\|\mathbf{A}\right\|_1 \le \mathcal{N} \max \mathbf{A}$, and all norms in $\mathbf{R}^{\mathcal{N}\times\mathcal{N}}$ are equivalent, it follows directly from the hypotheses and (*) that the sequences $\left\|\mathbf{M}_n\right\|$, $\left\|\mathbf{P}_n\right\|$ and $\left\|\overline{\mathbf{P}}_n\right\|$ are bounded for any matrix norm. Now, from the

submultiplicativity and subadditivity of induced matrix norms, $\mathbf{M}_n \bar{\mathbf{P}}_n \le \|\mathbf{M}_n\| \|\bar{\mathbf{P}}_n\|$ and $\|\mathbf{P}_n^k - \bar{\mathbf{P}}_n\| \le \|\mathbf{P}_n\|^k + \|\bar{\mathbf{P}}_n\|$ and so the suprema of (16) and (13) are finite.

It is straightforward to check the validity of the following result: if $\mathbf{A} \in \mathbf{R}^{r\times r}$ and $\mathbf{B} \in \mathbf{R}^{r\times l}$ are non-negative and $\mathbf{AB} \neq \mathbf{0}$, then $\min^+(\mathbf{AB}) \ge \min^+(\mathbf{A}) \min^+(\mathbf{B})$. Then, using conditions (31) and (*) we have

$$\frac{\max \mathbf{M}_n}{\min^+\left(\mathbf{M}_n \bar{\mathbf{P}}_n\right)} \left\|\mathbf{P}_n - \bar{\mathbf{P}}_n\right\|_1^k \le \frac{K}{\varepsilon\varepsilon''}\left(\|\mathbf{P}_n\|^k + \|\bar{\mathbf{P}}_n\|\right)$$

and so the supremum (27) is finite. Finally, the supremum and infimum of (23) and the supremum of (30) are finite since $\upsilon_H(k) \le \upsilon(k)$ and $\max\{\sigma_M(k), \sigma_m(k)\} \le \upsilon(k)$.

REFERENCES

Auger, P. and R. Bravo de la Parra (2000). Methods of aggregation of variables in population dynamics. Comptes Rendus de l'Academie des Sciences de Paris, Sciences de la vie 323: 665-674.

Auger, P. and J.C. Poggiale (1996a). Emergence of population growth models: fast migration and slow growth. Journal of Theoretical Biology 182: 99-108.

Auger, P. and J.C. Poggiale (1996b). Aggregation and emergence inhierarchically organized systems: population dynamics. Acta Biotheoretica 44: 301-316.

Auger, P. and R. Roussarie (1994). Complex ecological models with simple dynamics: From individuals to populations. Acta Biotheoretica 42: 111-136.

Blasco, A., L. Sanz, P. Auger and R. Bravo de la Parra (2001). Linear Discrete Population Models with Two Time Scales in Fast Changing Environments I: Autonomous Case. Acta Biotheoretica 49: 261-276.

Blasco, A., L. Sanz, P. Auger and R. Bravo de la Parra (2002). Linear Discrete Population Models with Two Time Scales in Fast Changing Environments II: Non-Autonomous Case. Acta Biotheoretica 50: 15-38.

Bravo de la Parra, R. and E. Sánchez (1998). Aggregation methods in population dynamics discrete models. Mathematical and Computer Modelling 27: 23-39.

Bravo de la Parra, R., E. Sánchez, O. Arino and P. Auger (1999). Discrete model with density dependent fast migration. Mathematical Biosciences 157: 91-109.

Caswell, H. (2001). Matrix population models. Sinauer Associates Inc., Sunderland, Massachusetts.

Charles, S, R. Bravo de la Parra, J. P. Mallet, H. Persat and P. Auger, (1998). Population dynamics modelling in an hierarchical arborescent river network: an attempt with Salmo trutta. Acta Biotheoretica 46: 223-234.

Horn, R. and C. Johnson. (1985). Matrix Analysis. Cambridge University Press, Cambridge.

Iwasa, Y., V. Andreasen and S.A. Levin (1987). Aggregation in model ecosystems I: Perfect aggregation. Ecological Modelling 37: 287-302.

Poggiale, J.C., and P. Auger, (1995). Fast oscillating migrations in a prey-predator model. Mathematical Models and Methods in Applied Sciences 6: 217-226.

Sánchez, E., R. Bravo de la Parra and P. Auger (1995). Linear discrete models with different time scales. Acta Biotheoretica 43: 465-479.

Sanz, L. and R. Bravo de la Parra (1998). Variables aggregation in time varying discrete systems. Acta Biotheoretica 46: 273-297.

Sanz, L. and R. Bravo de la Parra (1999). Variables aggregation in a time discrete linear model. Mathematical Biosciences 157: 111-146.

Sanz, L. and R. Bravo de la Parra (2000). Time scales in stochastic multiregional models. Nonlinear Analysis: Real World Applications 1: 89-122.

Sanz, L., and R. Bravo de la Parra (2001). Time scales in a nonautonomous linear discrete model. Mathematical Models and Methods in Applied Sciences 11: 1203-1235.

Tuljapurkar, S. (1990). Population Dynamics in Variable Environments. Springer-Verlag, Berlin.