

## Approximate Aggregation Methods in Discrete Time Stochastic Population Models

L. Sanz \*, J.A. Alonso

Departamento de Matemáticas, E.T.S.I Industriales, Universidad Politécnica de Madrid

**Abstract.** Approximate aggregation techniques consist of introducing certain approximations that allow one to reduce a complex system involving many coupled variables obtaining a simpler "aggregated system" governed by a few variables. Moreover, they give results that allow one to extract information about the complex original system in terms of the behavior of the reduced one. Often, the feature that allows one to carry out such a reduction is the presence of different time scales in the system under consideration. In this work we deal with aggregation techniques in stochastic discrete time models and their application to the study of multiregional models, i.e., of models for an age structured population distributed amongst different spatial patches and in which migration between the patches is usually fast with respect to the demography (reproduction-survival) in each patch. Stochasticity in population models can be of two kinds: environmental and demographic. We review the formulation and the main properties of the dynamics of the different models for populations evolving in discrete time and subjected to the effects of environmental and demographic stochasticity. Then we present different stochastic multiregional models with two time scales in which migration is fast with respect to demography and we review the main relationships between the dynamics of the original complex system and the aggregated simpler one. Finally, and within the context of models with environmental stochasticity in which the environmental variation is Markovian, we make use these techniques to analyze qualitatively the behavior of two multiregional models in which the original complex system is intractable. In particular we study conditions under which the population goes extinct or grows exponentially.

**Key words:** approximate aggregation, environmental stochasticity, demographic stochasticity, time scales, multiregional models, lognormal distribution.

**AMS subject classification:** 60G99, 92D40

---

\*Corresponding author. E-mail: lsanz@etsii.upm.es

## 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 on one hand and realism on the other. Models which include many biological details in their 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, etc.

Aggregation techniques deal with systems that are complex in the sense of having a large number of variables, and make use of the existence in the model of different time scales i.e., of biological processes which take place with characteristic times very different from each other. By approximating the original system by an auxiliary system in which the fast process has reached equilibrium, one can substitute the global system with a reduced system that retains at least some of the properties of the dynamics of the original system.

We can think of a system with different time scales as an hierarchically structured system with division into subsystems that are weakly coupled and simultaneously exhibit a strong internal dynamics. The idea of aggregation is to choose a few (usually one) global variable for each subsystem and to build a reduced system for it. In many cases the dimension of the aggregated system is considerably lower than that of the original system.

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 reproduction or aging [7]. 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.

Aggregation techniques were introduced originally in economics and have been widely studied in the context of ecology for time continuous systems with different time scales modelled by systems of ODEs. The techniques were introduced heuristically in [3], found rigorous justification in [4, 5] and have been applied to a number of different models in ecology (see [8] for an overview and a list of references). The deterministic discrete time case has been dealt with in linear autonomous [15, 50, 52, 11, 55], linear non-autonomous [51, 54, 12] and density dependent contexts [16, 44, 58].

Currently, stochasticity is one of the main items in the field of ecological modelling. In general, ecosystems are influenced by a large number of factors and this makes it impractical to include all of them in a mathematical model. On the other hand, their exclusion generates unexplained

variation whose consideration implies the use of stochasticity.

Concentrating in the discrete time case, 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 that consider a deterministic model in which a random error is added [49]. These models present some inconveniences to study population dynamics [60] and have not been treated with much profusion in the field.

The second kind of models consider "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 probability distribution. These models are conceptually adequate and lead to projections that agree with the variability observed in large populations.

The third kind of models deal with "demographic stochasticity", which models the deviations in the behavior of each individual with respect to the global vital rates. 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 not very useful in moderate to large populations, for they are not capable to explain their observed variability, but they are crucial in the study of extinction of small populations [32, 67, 17, 39].

The purpose of this work is threefold. On one hand we review the basic models used in ecology to model populations evolving in discrete time and affected by demographic or environmental stochasticity: we construct the models and review some of the most important features of their dynamics from the point of view of population dynamics. In the second place we present the different aggregation techniques the literature offers in this field: for each kind of stochasticity we construct a quite general model with two time scales, show how it can be reduced and give conditions under which one can guarantee that different features of the behavior of the original system can be described in terms of the behavior of the reduced system. Although the techniques are more general, emphasis is given to multiregional models in which migration is fast with respect to demography. Last, and in the context of models with environmental stochasticity, we show (Sections 3.3. and 3.4.) how the reduction procedure allows one to analyze otherwise intractable models and to draw conclusions about how the different parameters of the model, those defining demography, those defining migration and those defining the random temporal variations, affect the fate of the population.

The paper is structured as follows: in Section 2. we present the reduction procedure of discrete time deterministic models and some relationships between the original and the reduced systems. This section is important to understand the more elaborate stochastic settings. The general technique is applied to the reduction of a multiregional model in which migration is fast with respect to demography.

The reduction of models with environmental stochasticity is presented in Section 3.. We present a general model with two time scales and reduce it, and then we proceed to review the relationships between the behaviour of both systems. In order to do so we review the basic features of models of this kind and their application to population dynamics when the pattern of environmental vari-

ation is markovian. Finally, we present two applications in which the reduction technique allows one to perform a qualitative analysis of an otherwise untractable model. First we consider an age structured population distributed in two patches, and assuming that only migration is affected by stochasticity we arrive to a deterministic reduced system that can be easily analyzed. In the second application we deal with a population which is unstructured by age but live in a multipatch environment; without introducing any assumptions regarding how stochasticity affects demography and migration, we obtain a scalar stochastic system that we analyze and draw several conclusion regarding its asymptotic behavior.

In Section 4. we present the reductions of models that incorporate demographic stochasticity. We present the reduction technique for the case of a population structured in two age classes and living in a two patch environment, review some features of the behaviour of models with demographic stochasticity and present results relating the behavior of the original and reduced systems. Finally, section 5. recapitulates over the reduction techniques and its applications to population dynamics and points to future lines of work.

## 2. Aggregation of deterministic linear models

### 2.1. The model and its reduction

We suppose a stage-structured population in which the population is classified into stages or groups in terms of any characteristic of the life cycle. Moreover, 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. The model is in principle general in the sense that we do not state in detail the nature of the population or the subpopulations.

We consider the population being subdivided into  $d$  populations (or groups). Each group is subdivided into subpopulations (subgroups) in such a way that for each  $i = 1, 2, \dots, d$ , group  $i$  has  $N^i$  subgroups. Therefore, the total number of subgroups is  $N = N^1 + N^2 + \dots + N^d$ .

We will denote  $x^{ij}(n)$  the density of subpopulation  $j$  of population  $i$  at time  $n$ , with  $i = 1, 2, \dots, d$  and  $j = j(i) = 1, 2, \dots, N^i$ . In order to describe the population of group  $i$  we will use vector  $\mathbf{x}^i(n) = (x^{i1}(n), x^{i2}(n), \dots, x^{iN^i}(n)) \in \mathbb{R}^{N^i}$ ,  $i = 1, 2, \dots, d$ . The composition of the total population is then given by vector  $\mathbf{x}(n) = (\mathbf{x}^1(n), \mathbf{x}^2(n), \dots, \mathbf{x}^d(n))^T \in \mathbb{R}^N$  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 in our model both time scales we will model these two processes, to which we will refer as the fast and the slow dynamics, by two different matrices  $\mathbf{P}$  and  $\mathbf{M}$  respectively. The characteristic time scale of the fast process gives the projection interval associated to matrix  $\mathbf{P}$ , i.e., the state of the population, due to the fast process, after one fast time unit is  $\mathbf{P}\mathbf{x}(n)$ . Analogously, the effect of the slow process after one slow time unit is calculated multiplying by matrix  $\mathbf{M}$ .

In order to write a single discrete model combining both processes, and therefore their different time scales, we have to choose its time unit. Two possible and reasonable choices are the time units associated to each one of the two processes. If we choose the time unit of the fast process, than we

need to approximate the effect of the slow process over a time step which is shorter than its own. In order to do that we can model its effect through matrix

$$\epsilon \mathbf{M} + (1 - \epsilon) \mathbf{I}_N, \quad 0 < \epsilon < 1$$

where  $\epsilon$  is the quotient between the time steps of the fast and slow process. This approach, which is followed in [15], can be justified in deterministic linear settings but is difficult to extend to non-linear or stochastic settings. This is the reason why the vast majority of works in the field follow the other approach, i.e., the projection interval of the global model is chosen to be that corresponding to the slow dynamics. In this case we need to approximate the effect of the fast dynamics over a time interval much longer than its own. In order to do so we will suppose that during each projection interval, matrix  $\mathbf{P}$  has operated a number  $k$  of times, where  $k$  is a big enough integer that can be interpreted as the ratio between the projection intervals corresponding to the slow and fast dynamics. Therefore, the proposed model will consist of the following system of  $N$  linear difference equations that we will denote as *original system*:

$$\mathbf{x}_k(n+1) = \mathbf{M}\mathbf{P}^k \mathbf{x}_k(n) \quad (2.1)$$

where we use the subscript  $k$  to denote the dependence of the variables on this parameter.

In principle, we will make no special assumptions regarding the characteristics of the slow dynamics. Thus,  $\mathbf{M}$  will be a general nonnegative projection matrix  $\mathbf{M} \in \mathbb{R}^{N \times N}$  which we consider divided into blocks  $\mathbf{M}_{ij}$ ,  $1 \leq i, j \leq d$  in such a way that

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} & \cdots & \mathbf{M}_{1d} \\ \mathbf{M}_{21} & \mathbf{M}_{22} & \cdots & \mathbf{M}_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{M}_{d1} & \mathbf{M}_{d2} & \cdots & \mathbf{M}_{dd} \end{bmatrix} \quad (2.2)$$

where each block  $\mathbf{M}_{ij} = [M_{ij}^{ml}]$  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 order to reduce the system we must make some assumptions on the fast dynamics. We suppose that for each group  $i = 1, \dots, d$  the fast dynamics is internal, with an asymptotically stable distribution among the subgroups. These assumptions are met if we represent the fast dynamics for each group  $i$  by a  $N^i \times N^i$  non-negative projection matrix  $\mathbf{P}_i$  which is primitive, and has 1 as a dominant eigenvalue. Note that these hypotheses are met if each  $\mathbf{P}_i$  is primitive and column stochastic, in which case the total population is invariant for the fast process. The matrix  $\mathbf{P}$  that represents the fast dynamics for the whole population is then  $\mathbf{P} := \text{diag}(\mathbf{P}_1, \dots, \mathbf{P}_d)$ .

For each  $i$ , the Perron-Frobenius theorem [33] applies to  $\mathbf{P}_i$  and therefore this matrix has, associated to eigenvalue 1, positive right and left eigenvectors,  $\mathbf{v}_i = (v_i^1, \dots, v_i^{N^i})^T$  and  $\mathbf{u}_i = (u_i^1, \dots, u_i^{N^i})$ , respectively column and row vectors, verifying  $\mathbf{P}_i \mathbf{v}_i = \mathbf{v}_i$ ,  $\mathbf{u}_i \mathbf{P}_i = \mathbf{u}_i$  and normalized so that  $\mathbf{u}_i \cdot \mathbf{v}_i = 1$ . If we denote  $\mathbf{P}_i := \lim_{k \rightarrow \infty} \mathbf{P}_i^k = \mathbf{v}_i \mathbf{u}_i$  and  $\mathbf{P} := \text{diag}(\mathbf{P}_1, \dots, \mathbf{P}_d)$ , we have that:

$$\bar{\mathbf{P}} = \lim_{k \rightarrow \infty} \mathbf{P}^k = \mathbf{V}\mathbf{U}. \quad (2.3)$$

where  $\mathbf{V} := \text{diag}(\mathbf{v}_1, \dots, \mathbf{v}_d)_{N \times d}$  and  $\mathbf{U} := \text{diag}(\mathbf{u}_1, \dots, \mathbf{u}_d)_{d \times N}$ . Note that vector  $\mathbf{v}_i$  can be interpreted as the equilibrium distribution of migration in group  $i$ . Similarly, vector  $\mathbf{u}_i$  contains the reproductive values associated to migration in group  $i$ . These reproductive values play a key role in many biological situations (see [36]) and, as we will see, in this context they represent the weight of each initial variable  $x_{ij}(n)$  in the global variable  $y_i(n)$ .

If we think that the ratio of slow to fast time scale tends to infinity, i.e.  $k \rightarrow \infty$ , or, in other words, that the fast process is instantaneous in relation to the slow process, we can approximate system (2.1) by the following so-called *auxiliary system* with variables  $\hat{\mathbf{x}}(n)$  defined through

$$\hat{\mathbf{x}}(n+1) = \mathbf{M}\bar{\mathbf{P}}\hat{\mathbf{x}}(n), \quad (2.4)$$

which using (2.3) can be written as

$$\hat{\mathbf{x}}(n+1) = \mathbf{M}\mathbf{V}\mathbf{U}\hat{\mathbf{x}}(n).$$

Here we see that the evolution of the system depends on  $\mathbf{U}\hat{\mathbf{x}}(n) \in \mathbb{R}^d$ , what suggests that dynamics of the system could be described in terms of a lesser number of variables, the global variables or macro variables defined by

$$\begin{aligned} \mathbf{y}(n) & : = \mathbf{U}\hat{\mathbf{x}}(n) = (\mathbf{u}_1 \cdot \hat{\mathbf{x}}^1(n), \dots, \mathbf{u}_d \cdot \hat{\mathbf{x}}^d(n)) = \\ & = \left( u_1^1 \hat{x}^{11}(n) + \dots + u_1^{N^1} \hat{x}^{1N^1}(n), \dots, u_d^1 \hat{x}^{d1}(n) + \dots + u_d^{N^d} \hat{x}^{dN^d}(n) \right)^T, \end{aligned} \quad (2.5)$$

i.e., the macro variables are linear combinations of the population vector corresponding to each group, where the coefficients of the combination are given by the left eigenvectors of the  $\mathbf{P}_i$ . Therefore the auxiliary system (2.4) can be easily transformed into a  $d$ -dimensional system pre-multiplying by matrix  $\mathbf{U}$ , giving rise to the so-called *aggregated system* or *macro system*  $\mathbf{y}(n+1) = \mathbf{U}\mathbf{M}\mathbf{V}\mathbf{y}(n)$ , where we denote  $\bar{\mathbf{M}} = \mathbf{U}\mathbf{M}\mathbf{V}$  and obtain

$$\mathbf{y}(n+1) = \bar{\mathbf{M}}\mathbf{y}(n). \quad (2.6)$$

The solutions to the auxiliary system can be obtained from the solutions to the aggregated system. It is straightforward to check that the solution  $\{\mathbf{x}(n)\}_{n \in \mathbb{N}}$  of system (2.4) for the initial condition  $\mathbf{x}_0$  is related to the solution  $\{\mathbf{y}(n)\}_{n \in \mathbb{N}}$  of system (2.6) for the initial condition  $\mathbf{y}_0 = \mathbf{U}\mathbf{x}_0$  through the expression  $\hat{\mathbf{x}}(n) = \mathbf{M}\mathbf{V}\mathbf{y}(n-1)$  for every  $n \geq 1$ . Therefore, the passage from the auxiliary system to the reduced system is an example of perfect aggregation in the sense of [34].

Once the task of building up a reduced system is carried out, the important issue is to see if the dynamics of the general system (2.1) can also be studied by means of the aggregated system (2.6). In [50] it is proved that the asymptotic elements defining the long term behavior of system (2.1) can be approximated by those of the corresponding aggregated system when the matrix associated to the latter is primitive. In order to explore this further let us introduce the following hypothesis:

**Hypothesis (H1):**  $\bar{\mathbf{M}}$  is a primitive matrix.

Assuming hypothesis (H1), let  $\bar{\lambda} > 0$  be the (strictly) dominant eigenvalue of  $\bar{\mathbf{M}}$ , and  $\bar{\mathbf{l}}$  and  $\bar{\mathbf{r}}$  its associated left and right eigenvectors, respectively. We then have that, given any non negative initial condition  $\mathbf{y}_0$ , system (2.6) verifies

$$\lim_{n \rightarrow \infty} \frac{\mathbf{y}(n)}{\bar{\lambda}^n} = \frac{\bar{\mathbf{l}}^T \mathbf{y}_0}{\bar{\mathbf{l}}^T \bar{\mathbf{r}}} \bar{\mathbf{r}}$$

Concerning the asymptotic behavior of the auxiliary system (2.4), it is proved that  $\bar{\lambda} > 0$  is also the strictly dominant eigenvalue of  $\mathbf{M}\bar{\mathbf{P}}$ , and that  $\mathbf{U}^T\bar{\mathbf{1}}$  and  $\mathbf{M}\mathbf{V}\bar{\mathbf{r}}$  are associated left and right eigenvectors respectively.

The asymptotic behavior of the general system (2.1) can be expressed in terms of the asymptotic elements of the auxiliary system (2.4) by considering  $\mathbf{M}\mathbf{P}^k$  as a perturbation of  $\mathbf{M}\bar{\mathbf{P}}$ . To be precise, let us order the eigenvalues of  $\mathbf{P}$  (the union of those of the  $\mathbf{P}_i$ ) according to decreasing modulus in the following way:  $\lambda_1 = \dots = \lambda_d = 1 > |\lambda_{d+1}| \geq \dots \geq |\lambda_N|$ . So, if  $\|\cdot\|$  is any norm in the space of  $N \times N$  matrices, then for every  $\alpha > |\lambda_{d+1}|$  we have  $\|\mathbf{M}\mathbf{P}^k - \mathbf{M}\bar{\mathbf{P}}\| = o(\alpha^k)$  ( $k \rightarrow \infty$ ). This last result implies [59] that matrix  $\mathbf{M}\mathbf{P}^k$  has a strictly dominant eigenvalue of the form  $\bar{\lambda} + o(\alpha^k)$ , an associated left eigenvector  $\mathbf{U}^T\bar{\mathbf{1}} + o(\alpha^k)$  and an associated right eigenvector  $\mathbf{M}\mathbf{V}\bar{\mathbf{r}} + o(\alpha^k)$ . Having in mind that  $\alpha$  can be chosen to be less than 1, we see that the elements defining the asymptotic behavior of the aggregated and the general systems can be related in a precise way as a function of the separation between the two time scales. One can also build bounds for the error we incur in when we approximate the eigenvalue of the original system through  $\bar{\lambda}$ , and the variables  $\mathbf{x}_k(n)$  through  $\hat{\mathbf{x}}(n) = \mathbf{M}\mathbf{V}\mathbf{y}(n-1)$  for finite values of  $n$  and  $k$  [55].

## 2.2. Multiregional model with fast migration and slow demography

As an illustration of the general technique above, we consider the case of multiregional models and will justify that, 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 for the study of human populations [48] and in ecology [18]. In many situations migration can be considered to be fast in comparison with demography [15, 50], and there are cases regarding the dynamics of human populations in which demography can be considered a fast process in relation to migration [42, 43, 47].

In this work, and in order to be consistent with the applications we propose in the sections regarding the reduction of stochastic models, we will concentrate in the case in which migration is fast with respect to demography. We refer the reader to [52] for the case in which the contrary happens, as well as for the reduction procedure when matrices  $\mathbf{P}_i$  are not primitive reflecting the fact that fast process can be asymptotically cyclic instead of tending to an equilibrium.

Now we consider a multiregional model for population structured in juveniles (age class 1) and adults (age class 2) in a two-patch environment under the assumption that migration between the patches is fast with respect to demography, i.e., with respect to the reproduction-growth of the population.

Let  $x^{ij}(n)$  be the density of the subpopulation aged  $i$  (group) on patch  $j$  (subgroup) at time  $n$ . This corresponds to the general setting with  $d = 2$ ,  $N^1 = N^2 = 2$ . On each patch, the population grows according to a Leslie model and individuals belonging to a given age-class also move from patch to patch.

Let  $s_j$  be the survival rate of juveniles on patch  $j$  and  $f_{ij}$  the fertility rate of age class  $i$  on patch  $j$ . Then the population vector is  $\mathbf{x}(n) = (x^{11}(n), x^{12}(n), x^{21}(n), x^{22}(n))^T$ . Assuming that the time

step of the model corresponds to that of demography and that no individuals of age 2 survive after a time step of the model, the matrix describing the demography of the population in both patches is

$$\mathbf{M} = \begin{pmatrix} f_{11} & 0 & f_{21} & 0 \\ 0 & f_{12} & 0 & f_{22} \\ s_1 & 0 & 0 & 0 \\ 0 & s_2 & 0 & 0 \end{pmatrix}$$

Migration of individuals of age  $i$  is described by the following migration rates:  $p_i$  (resp.  $t_i$ ) is the migration rate from patch 1 to patch 2 (resp. from patch 2 to patch 1). So the matrix describing migration for the population is

$$\mathbf{P} = \text{diag}(\mathbf{P}_1, \mathbf{P}_2) = \begin{pmatrix} 1 - p_1 & t_1 & 0 & 0 \\ p_1 & 1 - t_1 & 0 & 0 \\ 0 & 0 & 1 - p_2 & t_2 \\ 0 & 0 & p_2 & 1 - t_2 \end{pmatrix}$$

which, if  $0 < p_i, t_i < 1$ ,  $i = 1, 2$ , meets the hypotheses above. Indeed, each  $\mathbf{P}_i$  is primitive and moreover is column stochastic, so it has a dominant eigenvalue equal to one and a stationary equilibrium distribution. Then, the model reads

$$\mathbf{x}_k(n+1) = \mathbf{M}\mathbf{P}^k \mathbf{x}_k(n) \quad (2.7)$$

where  $k$  is the ratio between the characteristic times of demography and migration.

In this case the vectors that represent the equilibrium distribution of migration are

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

and the left eigenvectors of matrices  $\mathbf{P}_i$  are

$$\mathbf{u}_1 = (1, 1) ; \quad \mathbf{u}_2 = (1, 1)$$

Moreover, matrices  $\bar{\mathbf{P}}$ ,  $\mathbf{V}$  and  $\mathbf{U}$  read

$$\bar{\mathbf{P}} = \begin{pmatrix} \frac{t_1}{p_1+t_1} & \frac{t_1}{p_1+t_1} & 0 & 0 \\ \frac{p_1}{p_1+t_1} & \frac{p_1}{p_1+t_1} & 0 & 0 \\ 0 & 0 & \frac{t_2}{p_2+t_2} & \frac{t_2}{p_2+t_2} \\ 0 & 0 & \frac{p_2}{p_2+t_2} & \frac{p_2}{p_2+t_2} \end{pmatrix} ; \quad \mathbf{V} = \begin{pmatrix} \frac{t_1}{p_1+t_1} & 0 \\ \frac{p_1}{p_1+t_1} & 0 \\ 0 & \frac{t_2}{p_2+t_2} \\ 0 & \frac{p_2}{p_2+t_2} \end{pmatrix} ; \quad \mathbf{U} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Since the  $\mathbf{u}_i$  have all components equal to one we have that the global variables are the total population in each age class

$$\mathbf{y}(n) = (y^1(n), y^2(n))^T = (x^{11}(n) + x^{12}(n), x^{21}(n) + x^{22}(n))^T$$

and the aggregated system has the form

$$\mathbf{y}(n+1) = \bar{\mathbf{M}}\mathbf{y}(n)$$

with

$$\bar{\mathbf{M}} = \begin{bmatrix} \frac{t_1 f_{11} + p_1 f_{12}}{p_1 + t_1} & \frac{t_2 f_{21} + p_2 f_{22}}{p_2 + t_2} \\ \frac{t_1 s_1 + p_1 s_2}{p_1 + t_1} & 0 \end{bmatrix}$$

### 3. Models with environmental stochasticity and their reduction

#### 3.1. The model and its reduction

Let us now consider the study and the reduction of time discrete linear models incorporating the so called 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.

Let us first examine the general form of linear models that incorporate environmental stochasticity. Let us consider a structured population, represented by vector  $\mathbf{z}(n) \in \mathbb{R}^{m \times 1}$ , living in an habitat in which that there are  $l$  different environmental conditions that we consider numbered in the set  $\mathcal{I} = \{1, \dots, l\}$ . In all what follows,  $\|\cdot\|$  will denote the 1-norm in the corresponding space. Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be the probability space and let  $\tau_n$  be the random variable taking values in  $\mathcal{I}$  that defines the environment under which the population lives in the time step  $[n-1, n)$ . Now let matrix  $\mathbf{A}(\xi) \in \mathbb{R}^{m \times m}$  be the matrix of vital rates for the population in environment  $\xi \in \mathcal{I}$ . Then the model for the population reads

$$\mathbf{z}(n+1) = \mathbf{A}(\tau_{n+1})\mathbf{z}(n) \quad (3.1)$$

In order to formulate a model with two time scales that incorporates the effect of environmental stochasticity, we can consider (2.1) as a starting point and allow the temporal variation to be random. Specifically, let us assume that there are  $l$  different environmental conditions  $\mathcal{I} = \{1, \dots, l\}$  and let  $\tau_n$  be the random variable that defines the environment under which the population lives in the time step  $[n-1, n)$ . Now let  $\mathbf{M}(\xi)$  and  $\mathbf{P}(\xi)$  denote the matrices of vital rates for the slow and the fast process respectively, in environment  $\xi \in \mathcal{I}$ . Therefore, the model with two time scales reads

$$\mathbf{x}_k(n+1) = \mathbf{M}(\tau_{n+1})\mathbf{P}^k(\tau_{n+1})\mathbf{x}_k(n). \quad (3.2)$$

The reduction of these kind of models is dealt with in [53] under quite general conditions that basically imply that each one of the matrices  $\mathbf{P}_i(\xi)$  is primitive and with dominant eigenvalue equal to one. In order to simplify this exposition we concentrate in the particular although very important case in applications in which the  $\mathbf{P}_i(\xi)$  are primitive stochastic matrices. For each  $i = 1, \dots, l$ , we define vectors  $\mathbf{v}_i(\xi) > 0$  through the relationships  $\mathbf{P}_i(\xi)\mathbf{v}_i(\xi) = \mathbf{v}_i(\xi)$  and  $\|\mathbf{v}_i(\xi)\| = 1$  and then consider matrices  $\bar{\mathbf{P}}_i(\xi) := \lim_{k \rightarrow \infty} \mathbf{P}_i^k(\xi) = \mathbf{v}_i(\xi)\mathbf{1}$ , where  $\mathbf{1}$  denotes a row vector with all its components equal to 1, and  $\bar{\mathbf{P}}(\xi) := \text{diag}(\bar{\mathbf{P}}_1(\xi), \dots, \bar{\mathbf{P}}_d(\xi))$ .

In order to reduce the model, we proceed in an analogous way to the deterministic case. Letting  $k \rightarrow \infty$  in (3.2) we obtain the auxiliary system

$$\hat{\mathbf{x}}(n+1) = \mathbf{M}(\tau_{n+1})\bar{\mathbf{P}}(\tau_{n+1})\hat{\mathbf{x}}(n). \quad (3.3)$$

that we can express in the form

$$\hat{\mathbf{x}}(n+1) = \mathbf{M}(\tau_{n+1})\mathbf{V}(\tau_{n+1})\mathbf{U}\hat{\mathbf{x}}(n). \quad (3.4)$$

where  $\mathbf{V}(\xi) := \text{diag}(\mathbf{v}_1(\xi), \dots, \mathbf{v}_d(\xi))_{N \times d}$  and  $\mathbf{U} := \text{diag}(\mathbf{1}, \dots, \mathbf{1})_{d \times N}$ . Now premultiplying both sides of (3.4) by  $\mathbf{U}$  and defining the vector of  $d$  global variables through

$$\begin{aligned} \mathbf{y}(n) &: = (y^1(n), \dots, y^d(n)) = \mathbf{U}\hat{\mathbf{x}}(n) = \\ &= \left( \hat{x}^{11}(n) + \dots + \hat{x}^{1N^1}(n), \dots, \hat{x}^{d1}(n) + \dots + \hat{x}^{dN^d}(n) \right) \end{aligned}$$

so that  $y^i(n)$  corresponds to the total population in group  $i$ , we obtain the reduced system

$$\mathbf{y}(n+1) = \bar{\mathbf{M}}(\tau_n)\mathbf{y}(n) \quad (3.5)$$

where

$$\bar{\mathbf{M}}(\xi) = \mathbf{U}\mathbf{M}(\xi)\mathbf{V}(\xi)$$

### 3.2. Behavior of the original and the reduced models

Let us examine the relationships between the original system (3.2) and the reduced one (3.5). As it was the case in the non-autonomous setting, for any finite value of  $n$  the population vector of the original system can be approximated through the knowledge of the population vector in the aggregated system

$$\mathbf{x}_k(n) = \mathbf{M}(\tau_n)\mathbf{V}(\tau_n)\mathbf{y}(n-1) + \mathbf{o}_n(\delta^k), \quad k \rightarrow \infty \quad (3.6)$$

where  $\alpha$  verifies  $0 < \alpha < 1$  and  $\mathbf{o}_n(\alpha^k)$  depends on  $n$ . In [57] bounds are obtained for the "error" we incur in when we approximate  $\mathbf{x}_k(n)$  by  $\mathbf{M}(\tau_n)\mathbf{V}(\tau_n)\mathbf{y}(n-1)$  for finite values of  $n$  and  $k$ .

So far, the pattern of environmental variation defined by  $\tau_n$  did not have to meet any special requirements. Subsequently we will explore the relationships between the original system and the aggregated system in the case that the pattern of temporal variation for the former is an homogeneous Markov chain. Indeed, a great part of the models of the kind (3.1) found in the literature represent environmental change through Markov chains [24, 18], for Markov chains can represent sequential dependence between environments and yet are simple enough to be analyzed with great detail. Besides, the systems under Markovian temporal variation usually have some desirable properties regarding their asymptotic behavior. Indeed, an appropriate Markovian temporal variation combined with certain conditions on the set of environmental matrices guarantee on one hand the presence of strong stochastic ergodicity, i.e., the probability distribution of the population structure converges to a stationary probability distribution independent of initial conditions [24] and, on the other hand, that almost all realizations of the process ultimately grow or decay exponentially with

the same growth rate. This approach contemplates, as a particular case, the possibility of the  $\tau_n$  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.

Thus, let  $\tau_n$  be a homogeneous Markov chain, defined in terms of a initial probability vector and a matrix  $\mathbf{Q} \in \mathbb{R}^{l \times l}$  of transition probabilities, where  $q_{\alpha\beta} := \mathbb{P}\{\tau_{n+1} = \alpha \mid \tau_n = \beta\}$ ,  $1 \leq \alpha, \beta \leq l$ . The i.i.d. setting corresponds to the case in which  $q_{\alpha\beta}$  is independent of  $\beta$ .

To start with, let us review some of the most important features of model (3.1) when  $\tau_n$  is a homogeneous Markov chain. The mathematical details can be found in [65, 27, 31, 23]. We refer the application-oriented reader to [26] for a practical application to this kind of models to a fish population and to [18, 66] for an overview and a list of references.

In the first place, let us examine the asymptotic behavior of the expected value of the population size  $\mathbb{E} \|\mathbf{z}(n)\|$ . One can show that the limit

$$\mu := \lim_{n \rightarrow \infty} \frac{\log \mathbb{E} \|\mathbf{z}(n)\|}{n} \quad (3.7)$$

exists so that  $\mathbb{E} \|\mathbf{z}(n)\|$  grows asymptotically in an exponential fashion. Moreover [53],  $\mu$  can be computed through  $\mu = \log \gamma$  where  $\gamma$  is the spectral radius of matrix

$$\mathbf{B} := \begin{pmatrix} q_{11}\mathbf{A}(1) & q_{12}\mathbf{A}(1) & \cdots & q_{1l}\mathbf{A}(1) \\ q_{21}\mathbf{A}(2) & q_{22}\mathbf{A}(2) & \cdots & q_{2l}\mathbf{A}(2) \\ \vdots & \vdots & \ddots & \vdots \\ q_{l1}\mathbf{A}(l) & q_{l2}\mathbf{A}(l) & \cdots & q_{ll}\mathbf{A}(l) \end{pmatrix}$$

Let us turn our attention to the behavior of the realizations of  $\|\mathbf{z}(n)\|$ . One has that under some hypotheses each realization of the system grows (or decays) exponentially with the same rate. More specifically, let us consider system (3.1) and let us assume the following hypothesis:

**Hypothesis (H2):**  $\tau_n$  is an homogenous Markov chain such that it is irreducible and ergodic (i.e., its matrix of transition probabilities is primitive). Moreover, the set  $\{\mathbf{A}(1), \dots, \mathbf{A}(l)\}$  is "ergodic", i.e., there exists a positive integer  $r$  such that the product of any  $r$  matrices drawn from this set is a positive matrix.

Given hypothesis H2 (see [63]) there exists the following limit with probability one

$$a := \lim_{n \rightarrow \infty} \frac{\log \|\mathbf{z}(n)\|}{n} \quad (3.8)$$

where  $a$ , the so called stochastic growth rate (s.g.r.) of the system, is a fixed (non random) number independent of the initial probabilities of the chain and of the initial (non-zero) population vector  $\mathbf{z}_0 \geq \mathbf{0}$ . The s.g.r. is the most important parameter in order to characterize the behavior of models of the kind (3.1), and plays a role which is similar to that of the dominant eigenvalue in deterministic linear models. If  $a < 0$  every realization of the population goes extinct, and if  $a > 0$  every realization of the process grows exponentially with rate  $e^a$ . Note that in general  $a \neq \mu$ . In fact because of the Jensen inequality  $a \leq \mu$  and so the expected value of the population grows/decays at a different rate than that of the realizations. For example one can have a situation

in which  $\mu > 0$  (so the expected value of the population grows to infinity) and however  $a < 0$  (the population goes extinct with probability one). This stresses the importance of dealing directly with the stochastic model instead of just taking expected values in order to study the population.

Under hypothesis H2 one can define the "scaled logarithmic variance" (s.l.v.) of the system as

$$\sigma^2 := \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{V} [\log \|\mathbf{z}(n)\|]$$

where  $\mathbb{V}$  denotes variance and where  $\sigma^2$  is independent of the initial probabilities of the chain and of the initial (non-zero) population vector  $\mathbf{z}_0 \geq \mathbf{0}$  [63]. Therefore  $\sigma^2$  is a measure of the rate  $n\sigma^2$  at which the variance of  $\log \|\mathbf{z}(n)\|$  grows for large values of  $n$ .

Moreover, if  $\sigma^2 > 0$  the population size is asymptotically lognormal in the sense that

$$\frac{\log \|\mathbf{z}(n)\| - na}{n^{1/2}\sigma} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$$

where  $\mathcal{N}(0, 1)$  denotes a normal distribution of zero mean and unit variance and  $\mathcal{L}$  denotes convergence in distribution. Therefore for large values of  $n$  the distribution of  $\log \|\mathbf{z}(n)\|$  is approximately normal with a mean value of  $na$  and variance  $n\sigma^2$ .

Moreover, H2 guarantees that system (3.1) is stochastically strongly ergodic, i.e., the vector of population structure  $\mathbf{z}(n)/\|\mathbf{z}(n)\|$  converges to a distribution  $\mathcal{Y}$  independent of the initial conditions [23, 24, 25].

Its important to keep in mind that in most practical situations it is not possible to compute the s.g.r. and the s.l.v. of a model analytically and one must resort to computer simulations [65, 66], the reason being that although there are explicit expressions for  $a$  and  $\sigma^2$  in terms of expectations, they involve the use of the stationary distribution of age structure  $\mathcal{Y}$  which can not be computed except in a few particular cases [64, 65].

Coming back to the study of the relationships between the systems (3.2) and (3.5), it is important to study the conditions for the existence of a s.g.r., s.l.v. and the relationship between them in both systems. Moreover, since these parameters can not be computed explicitly in most practical cases, it would be interesting to study whether there are situations in which the above reduction procedure allows one to render systems for which the s.g.r. and the s.l.v. can be computed. This would make the aggregation procedure especially useful.

Let us assume that  $\tau_n$  is a homogenous Markov chain with a primitive matrix of transition probabilities, and let us assume that the set of matrices  $\{\bar{\mathbf{M}}(1), \dots, \bar{\mathbf{M}}(l)\}$  for the reduced system (3.5) is an ergodic set. Therefore the reduced system (3.5) verifies the sufficient conditions (H2) for the existence of a s.g.r.  $\bar{a} := \lim_{n \rightarrow \infty} \log \|\mathbf{y}(n)\|/n$ , where the limit is with probability one, and a s.l.v.  $\bar{\sigma}^2 := \lim_{n \rightarrow \infty} \mathbb{V} [\log \|\mathbf{y}(n)\|]/n$ . In [57] and [1] it is shown that if matrices  $\mathbf{M}(1), \dots, \mathbf{M}(l)$  have no zero rows, then for large enough  $k$  the set  $\{\mathbf{M}(1)\mathbf{P}^k(1), \dots, \mathbf{M}(l)\mathbf{P}^k(l)\}$  is also ergodic and therefore the original system (3.2) verifies that there exists a s.g.r.  $a_k := \lim_{n \rightarrow \infty} \log \|\mathbf{x}_k(n)\|/n$  and a s.l.v.  $\sigma_k^2 := \lim_{n \rightarrow \infty} \mathbb{V} [\log \|\mathbf{x}_k(n)\|]/n$ . Besides, it is proved that  $\lim_{k \rightarrow \infty} a_k = \bar{a}$  and  $\lim_{k \rightarrow \infty} \sigma_k^2 = \bar{\sigma}^2$  and therefore, if  $k$  is large enough, we can approximate  $a_k$  and  $\sigma_k^2$  by  $\bar{a}$  and  $\bar{\sigma}^2$  respectively. Moreover, [57] derives upper bounds for  $|a_k - \bar{a}|$  for finite values of  $k$ , so we can estimate the error we incur in when we approximate the study of (3.2) through that of (3.5). Finally, [53] shows that for any positive integer  $m$ , the asymptotic behavior when  $n \rightarrow \infty$  of the

moments of order  $m$  of system (3.2) can be approximated through the behavior of the corresponding moments for the reduced system (3.5). In particular, if  $\bar{\mu} := \lim_{n \rightarrow \infty} \log \mathbb{E} \|\mathbf{y}(n)\| / n$  and  $\mu_k := \lim_{n \rightarrow \infty} \log \mathbb{E} \|\mathbf{x}_k(n)\| / n$  are the growth rates of the expected value of the population for the reduced and the original system respectively, then we have  $\lim_{n \rightarrow \infty} \mu_k = \bar{\mu}$ .

### 3.3. A structured population distributed between two patches

As an application of the technique above, we present a model, introduced in [53] for an age structured population distributed between two spatial patches in a random varying environment under the assumption that migration is fast with respect to demography. We assume that there are  $l$  different environments  $1, 2, \dots, l$  and we set up the model by reasoning analogously to the deterministic case but allowing the fecundity, survival and migration rates to be dependent on the environment. In this way the matrices of demography and migration in each environment  $\xi$  are

$$\mathbf{M}(\xi) = \begin{bmatrix} f_{11}(\xi) & 0 & f_{21}(\xi) & 0 \\ 0 & f_{12}(\xi) & 0 & f_{22}(\xi) \\ s_1(\xi) & 0 & 0 & 0 \\ 0 & s_2(\xi) & 0 & 0 \end{bmatrix}$$

$$\mathbf{P}(\xi) = \begin{bmatrix} 1 - p_1(\xi) & t_1(\xi) & 0 & 0 \\ p_1(\xi) & 1 - t_1(\xi) & 0 & 0 \\ 0 & 0 & 1 - p_2(\xi) & t_2(\xi) \\ 0 & 0 & p_2(\xi) & 1 - t_2(\xi) \end{bmatrix},$$

where  $s_j(\xi)$  stands for the survival rate of juveniles in patch  $j$ ,  $p_i(\xi)$  denotes the migration rates in environment  $\xi$  of individuals of age  $i$  from patch 1 to patch 2, and  $t_i(\xi)$  has an analogous meaning for migration from patch 2 to patch 1. Therefore the complete model reads

$$\mathbf{x}_k(n+1) = \mathbf{M}(\tau_{n+1})\mathbf{P}^k(\tau_{n+1})\mathbf{x}_k(n)$$

where  $k$  is the ratio between the characteristic times of demography and migration.

If we assume that  $p_1(\xi)$ ,  $p_2(\xi)$ ,  $t_1(\xi)$  and  $t_2(\xi)$  are different from 0 or 1 for all  $\xi \in \{1, 2, \dots, l\}$ , then the matrices  $\mathbf{P}_1(\xi)$  and  $\mathbf{P}_2(\xi)$  governing migration for each age class are stochastic and primitive for all  $\xi$ , and so the hypotheses to reduce the system are met. We have

$$\mathbf{y}(n+1) = \bar{\mathbf{M}}(\tau_{n+1})\mathbf{y}(n) \quad (3.9)$$

with

$$\mathbf{y}(n) = (y^1(n), y^2(n))^T = (\hat{x}^{11}(n) + \hat{x}^{12}(n), \hat{x}^{21}(n) + \hat{x}^{22}(n))^T$$

and

$$\bar{\mathbf{M}}(\xi) = \begin{bmatrix} \bar{f}_1(\xi) & \bar{f}_2(\xi) \\ \bar{s}(\xi) & 0 \end{bmatrix}; \quad \bar{s}(\xi) = \frac{t_1(\xi)s_1(\xi) + p_1(\xi)s_2(\xi)}{p_1(\xi) + t_1(\xi)}$$

$$\bar{f}_i(\xi) = \frac{t_i(\xi)f_{i1}(\xi) + p_i(\xi)f_{i2}(\xi)}{p_i(\xi) + t_i(\xi)}, \quad i = 1, 2; \quad \xi = 1, 2, \dots, l$$

Let us now consider a particular case of environmental variation for which we will be able to give exact results for system (3.9). Let us assume that the parameters that govern demography are independent of the environment, i.e.,  $f_{ij}(\xi) = f_{ij}$ ,  $s_j(\xi) = s_j$ , where  $f_{ij}, s_j > 0$ ,  $s_j < 1$ . Moreover, let

$$p_i(\tau_n) = \eta_n p_i ; t_i(\tau_n) = \eta_n t_i$$

where  $p_i, t_i \in (0, 1)$  are non random and  $\eta_n$  is a homogeneous Markov chain with values in the set  $\{\delta_1, \dots, \delta_l\}$ ,  $0 < \delta_i < 1$ . This model corresponds to the situation in which reproduction and survival are not affected by the environment and migration rates depend on the environment through a common multiplicative random variable. A good/bad environment for migration increases/decreases all migration rates by the same factor.

The original system is a stochastic model for which exact calculations for the s.g.r. and the s.l.v. are not feasible. However, when aggregating we have

$$\bar{\mathbf{M}}(\xi) = \bar{\mathbf{M}} = \begin{bmatrix} \bar{f}_1 & \bar{f}_2 \\ \bar{s} & 0 \end{bmatrix}, \quad \xi = 1, 2, \dots, l$$

where

$$\bar{s} := \frac{t_1 s_1 + p_1 s_2}{p_1 + t_1} ; \bar{f}_i := \frac{t_i f_{i1} + p_i f_{i2}}{p_i + t_i}, \quad i = 1, 2,$$

i.e., the reduced system is the deterministic system  $\mathbf{y}(n+1) = \bar{\mathbf{M}}\mathbf{y}(n)$ . Since the set of environmental matrices for the reduced system consists only on matrix  $\bar{\mathbf{M}}$  and this matrix is primitive, the set is ergodic and hypothesis  $H1$  is met. Moreover, if we assume all the fertility and survival coefficients  $f_{ij}(\xi), s_i(\xi)$  to be non-zero for all  $\xi$ , matrices  $\mathbf{M}(1), \dots, \mathbf{M}(l)$  have no zero rows and so we have that both the reduced and the original system have a s.g.r. and a s.l.v. and those of the latter can be approximated by those of the former. Since the reduced system is deterministic we have that its s.g.r.  $\bar{a}$  is given by the spectral radius of matrix  $\bar{\mathbf{M}}$  and the s.l.v.  $\bar{\sigma}^2$  is zero. Therefore for large enough  $k$  the original system behaves almost deterministically, with a s.g.r. and a s.l.v. that can be approximated through

$$\bar{a} = \log \frac{\bar{f}_1 + \sqrt{(\bar{f}_1)^2 + 4\bar{f}_2\bar{s}}}{2} ; \bar{\sigma}^2 = 0$$

As a result, the population goes extinct (resp. grows exponentially) with probability one if  $\bar{f}_1 + \sqrt{(\bar{f}_1)^2 + 4\bar{f}_2\bar{s}} < 2$  (resp.  $> 2$ ).

### 3.4. Non-structured population living in a multipatch environment. Exact results and qualitative study

Let us now consider a multiregional model with fast migration in which the population is not structured by age ( $d = 1$ ), is distributed amongst  $N$  spatial patches and there are  $l$  possible environmental conditions that vary according to a Markov chain  $\tau_n$  with values in the set  $\mathcal{I} = \{1, 2, \dots, l\}$ .

The population vector is  $\mathbf{x}(n) = (x^1(n), \dots, x^N(n)) \in \mathbb{R}^N$  where  $x^i(n)$  is the population in patch  $i$ . Demography in each environment  $\xi \in \mathcal{I}$  is modeled by matrix

$$\mathbf{M}(\xi) = \text{diag}(m_1(\xi), m_2(\xi), \dots, m_N(\xi))$$

where  $m_i(\xi)$ ,  $i = 1, \dots, N$ ;  $\xi = 1, \dots, l$  is the growth rate in patch  $i$  and environment  $\xi$ . We assume that migration amongst the different patches in each environment  $\xi \in \mathcal{I}$  is modeled by a column stochastic primitive matrix  $\mathbf{P}(\xi) \in \mathbb{R}^{N \times N}$ .

The Markov chain  $\tau_n$  is defined by a certain vector  $\mathbf{q} \in \mathbb{R}^l$  of initial probabilities and a column stochastic matrix  $\mathbf{Q} \in \mathbb{R}^{l \times l}$  of transition probabilities. We assume that  $\tau_n$  is irreducible and ergodic, i.e., that  $\mathbf{Q}$  is a primitive matrix and so  $\tau_n$  has a stationary probability distribution, that we denote  $\pi = (\pi_1, \pi_2, \dots, \pi_l)^T$ . The model then reads

$$\mathbf{x}_k(n+1) = \mathbf{M}(\tau_{n+1})\mathbf{P}^k(\tau_{n+1})\mathbf{x}_k(n). \quad (3.10)$$

Let  $\mathbf{v}(\xi) > 0$  be the equilibrium distribution for migration in environment  $\xi$  normalized so that  $\|\mathbf{v}(\xi)\| = 1$ . Then the reduced system takes the form of a scalar stochastic system

$$y(n+1) = \bar{m}(\tau_{n+1})y(n) \quad (3.11)$$

where  $y(n) = x^1(n) + \dots + x^N(n)$  is the total population and

$$\bar{m}(\xi) := \mathbf{1}^T \mathbf{M}(\xi) \mathbf{v}(\xi) = \sum_{i=1}^N v_i(\xi) m_i(\xi) > 0$$

Clearly the set  $\{\bar{m}_1, \dots, \bar{m}_l\}$  of vital rates for the reduced system is ergodic and matrices  $\mathbf{M}(\xi)$  have no zero rows, so both the reduced system (3.11) and the original system (for large enough  $k$ ) meet the sufficient hypotheses for the existence of a s.g.r. and a s.l.v.. Moreover, we can approximate these parameters for the original system through those of the reduced scalar system. Now, the computation of the s.g.r. and the s.l.v. for a system like (3.10) is not feasible, but it is easily done for a scalar system like (3.11). Let us assume that the transition probability matrix  $\mathbf{Q}$  is diagonalizable so that  $\mathbf{Q} = \mathbf{W}^{-1} \mathbf{D} \mathbf{W}$ , with  $\mathbf{D} = \text{diag}(1, \lambda_2, \dots, \lambda_l)$  and  $\mathbf{W} = (\mathbf{1}, \mathbf{w}_2, \dots, \mathbf{w}_l)^T$  where  $1, \lambda_2, \dots, \lambda_l$  are the eigenvalues of  $\mathbf{Q}$  and  $\mathbf{1}, \mathbf{w}_2, \dots, \mathbf{w}_l$  are corresponding left eigenvectors  $\mathbf{w}_j^T \mathbf{Q} = \lambda_j \mathbf{w}_j^T$ ,  $2 \leq j \leq l$  and  $\mathbf{1}^T \mathbf{Q} = \mathbf{1}^T$ . Adapting to system (3.11) the results of [1] we have that the s.g.r. is given by

$$\bar{a} = \mathbb{E}_\pi(\log \bar{m}_{\tau_n}) = \sum_{\xi=1}^l \pi_\xi \log \bar{m}_\xi = \sum_{\xi=1}^l \pi_\xi \log \left( \sum_{i=1}^N v_i(\xi) m_i(\xi) \right) \quad (3.12)$$

(note that  $a$  depends only on the equilibrium distribution of  $\tau_n$  and not on any other information defining  $\tau_n$ ) and the s.l.v. is

$$\bar{\sigma}^2 = \mathbb{E}_\pi(\log \bar{m}_{\tau_n} - \bar{a})^2 + 2\mathbf{z}^T \mathbf{W}^{-1} \left( 0, \frac{\lambda_2}{1 - \lambda_2} \mathbf{z}_\pi^T \mathbf{w}_2, \frac{\lambda_3}{1 - \lambda_3} \mathbf{z}_\pi^T \mathbf{w}_3, \dots, \frac{\lambda_l}{1 - \lambda_l} \mathbf{z}_\pi^T \mathbf{w}_l \right)^T \quad (3.13)$$

where  $\mathbb{E}_\pi$  denotes expectation with respect to the distribution  $\pi$ ,

$$\mathbf{z} = (z_1, z_2, \dots, z_l)^T \quad ; \quad \mathbf{z}_\pi = (z_1\pi_1, z_2\pi_2, \dots, z_l\pi_l)^T$$

and  $z_\xi := \log \bar{m}_\xi - \bar{a}$ ,  $1 \leq \xi \leq l$ .

As we have seen before, the rate of growth of the expected value of the population for the reduced system  $\bar{\mu} := \lim_{n \rightarrow \infty} \log \mathbb{E} \|\mathbf{y}(n)\| / n$  can be computed through  $\bar{\mu} = \log \gamma$  where  $\gamma$  is the spectral radius of matrix

$$\mathbf{B} := \begin{pmatrix} q_{11}\bar{m}_1 & \cdots & q_{1l}\bar{m}_l \\ \vdots & \ddots & \vdots \\ q_{l1}\bar{m}_l & \cdots & q_{ll}\bar{m}_l \end{pmatrix}$$

and then we have that  $\lim_{n \rightarrow \infty} \log \mathbb{E} y(n) = \infty$  if  $\bar{\mu} > 0$  and  $\lim_{n \rightarrow \infty} \log \mathbb{E} y(n) = 0$  if  $\bar{\mu} < 0$ .

Therefore we have explicit expressions that approximate the s.g.r. and the s.l.v. of the multi-regional model (3.10) and a matrix whose spectral radius approximates the rate of growth of the expected value of its population size. With them it is possible to carry out a qualitative study of the model, studying how migration and growth rates in each environment, as well as the characteristics of  $\tau_n$ , affect the fate (extinction or exponential growth) of the population. In order to see how it can be done, let us simplify our model and assume that there are only two environmental states ( $l = 2$ ) and so  $\tau_n$  takes values in the set  $\{1, 2\}$ . Let us express matrix  $\mathbf{Q}$  in the convenient form

$$\mathbf{Q} = \begin{pmatrix} \pi_1 + \rho(1 - \pi_1) & \pi_1(1 - \rho) \\ (1 - \rho)(1 - \pi_1) & 1 - \pi_1(1 - \rho) \end{pmatrix}$$

where  $\pi_1 \in (0, 1)$  and for each value of  $\pi_1$  parameter  $\rho$  varies between  $\rho_{inf}$  and 1 where

$$\rho_{inf} = -\min\{\pi_1/(1 - \pi_1), (1 - \pi_1)/\pi_1\}$$

Then, the equilibrium distribution of  $\tau_n$  is given by  $\pi = (\pi_1, 1 - \pi_1)$  and  $\rho$  is a measure of the serial correlation between environments.  $\rho = 0$  corresponds to the i.i.d. case,  $\rho$  close to one implies that a given environment is very likely to be followed by the same environment, and  $\rho$  close to  $\rho_{inf}$  implies the contrary happens. Matrix  $\mathbf{Q}$  can be diagonalized in the following way

$$\mathbf{Q} = \mathbf{W}^{-1} \mathbf{D} \mathbf{W} = \frac{1}{1 - \rho} \begin{pmatrix} \pi_1(1 - \rho) & 1 \\ (1 - \rho)(1 - \pi_1) & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \rho \end{pmatrix} \begin{pmatrix} 1 & 1 \\ (1 - \rho)(1 - \pi_1) & -\pi_1(1 - \rho) \end{pmatrix}$$

and then, after some computations we obtain from (3.12) and (3.13)

$$\begin{aligned} \bar{a} &= \pi_1 \log \bar{m}_1 + (1 - \pi_1) \log \bar{m}_2 \\ \bar{\sigma}^2 &= \frac{1 + \rho}{1 - \rho} \pi_1(1 - \pi_1) (\log \bar{m}_1 - \log \bar{m}_2)^2 \end{aligned}$$

Clearly,  $\bar{a}$  is independent of  $\rho$  and is an increasing function of  $\pi_1$  if  $\bar{m}_1 \geq \bar{m}_2$  and a decreasing function of  $\pi_1$  if  $\bar{m}_2 \geq \bar{m}_1$ .

Assuming in what follows that  $\bar{m}_1 \geq \bar{m}_2$  the condition for the extinction/exponential growth of the population is

$$\bar{a} < 0 \text{ (resp. } \bar{a} > 0) \Leftrightarrow \pi_1 < \alpha := -\frac{\log \bar{m}_2}{\log \bar{m}_1 - \log \bar{m}_2} \text{ (resp. } \pi_1 > \alpha)$$

Note that if  $\bar{m}_2 > 1$  then necessarily  $\bar{a} > 0$  and if  $\bar{m}_1 < 1$  then necessarily  $\bar{a} < 0$ . Regarding  $\bar{\sigma}^2$ , we have that it tends to  $\infty$  when  $\rho \uparrow 1$ . Moreover, for fixed  $\rho$  its maximum is reached when  $\pi_1 = 1/2$ .

We can easily compute  $\bar{\mu}$  obtaining

$$\bar{\mu} = \log \left( \frac{1}{2} \left[ \bar{b} + \sqrt{\bar{b}^2 - 4\bar{m}_1\bar{m}_2\rho} \right] \right)$$

where

$$\bar{b} := \bar{m}_1 + \bar{m}_2 - (1 - \rho) [\bar{m}_1(1 - \pi_1) + \bar{m}_2\pi_1]$$

For all values of the parameters involved we have already seen that  $\bar{a} \leq \bar{\mu}$ . In fact we can have a situation in which  $\bar{\mu} > 0$ , and so the expected value of the population grows to infinity, and however  $\bar{a} < 0$  and so the population goes extinct with probability one. Indeed let us assume that

$$\bar{m}_2 < 1 \quad ; \quad \bar{m}_1 > 1 \tag{3.14}$$

and let us place ourselves in the i.i.d. case, i.e.,  $\rho = 0$ , for which

$$\bar{\mu}_{iid} = \log(\pi_1\bar{m}_1 + (1 - \pi_1)\bar{m}_2)$$

Therefore,  $\bar{\mu}_{iid} > 0$  if and only if

$$\pi_1 > \beta := (1 - \bar{m}_2)/(\bar{m}_1 - \bar{m}_2).$$

Now using (3.14), we have  $0 < \alpha < \beta < 1$  so that whenever  $\beta < \pi_1 < \alpha$  we have  $\bar{a} < 0$  and  $\bar{\mu} > 0$ . This situation, that shows how even in the relatively simple i.i.d. case the stochastic model can have dynamics very different from the deterministic model built by taking expectations, can be explained using the properties of the lognormal distribution of population size. Indeed this distribution is very skewed and its long tale makes the expected value be very high although most of the probability is concentrated around zero [63].

## 4. Models incorporating demographic stochasticity and their reduction

### 4.1. A multiregional model for an age structured population

As we pointed out in the introduction, demographic stochasticity, of great importance in the study of the extinction of small populations, refers to the variability in a population arising from random differences among individuals in survival and reproduction.

In this section we contemplate the multiregional model with two time scales of section 2.2. but now we assume that both migration and demography are affected by the effects of demographic stochasticity. Therefore these two processes will be modeled by a so called multitype Bienaymé-Galton-Watson branching processes (MBGWBP). We start by presenting some basics about the formulation of the models. The reader can refer to [45, 30] for a full mathematical analysis and [18, 37, 29] for biological applications and extensions.

Let  $\mathbb{Z}_+^m$  denote the set of  $m$ -dimensional vectors whose components are non-negative integers. Let us assume a population with  $m$  types of individuals, and let the population vector at time  $n$  be

$$\mathbf{z}(n) = (z^1(n), \dots, z^m(n)) \in \mathbb{Z}_+^m, \quad n = 0, 1, 2, \dots \quad (4.1)$$

MBGWBP's arise from the following modelling assumptions: a) Each particle behaves independently of the rest (no density dependence) and independently of time (autonomous model) and b) at the end of each time step of the model each individual of type  $i$  "disappears" and leaves an "offspring" according to certain probabilities that, given certain initial conditions, completely characterize the dynamics of the population.

Note that due to the independence assumption, the behavior of a MBGWBP for any (deterministic) initial condition can be known if we know its behavior for the initial conditions  $\mathbf{e}^j$  ( $j$ -th canonical vector in  $\mathbb{R}^m$ ) for  $j = 1, \dots, m$ , which usually simplifies the mathematical treatment of the model.

We denote

$$p_{\alpha_1, \dots, \alpha_m}^i := \text{probability that an individual of type } i \text{ has offspring } (\alpha_1, \dots, \alpha_m) \in \mathbb{Z}_+^m$$

In order to deal with MBGWBP it is useful to work with the so called probability generating function (in what follows p.g.f.) of the process, defined as the function

$$\mathbf{G}(\mathbf{s}) := (G^1(\mathbf{s}), \dots, G^m(\mathbf{s})) \quad (4.2)$$

where

$$G^i(s_1, \dots, s_m) := \sum_{\alpha \in \mathbb{Z}_+^m} p_{\alpha_1, \dots, \alpha_m}^i s_1^{\alpha_1} s_2^{\alpha_2} \dots s_m^{\alpha_m}, \quad i = 1, \dots, m \quad (4.3)$$

from which one can derive all the properties of the model.

The matrix  $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{m \times m}$  of expected values of the process is defined through

$$a_{ij} := \mathbb{E}(z^i(n+1) \mid \mathbf{z}(n) = \mathbf{e}^j), \quad i, j = 1, \dots, m$$

where  $\mathbf{e}^i$  corresponds to a population composed of one individual of type  $i$  and where we are assuming that the expectations are finite.

Now we are ready to present the analog of the multiregional model considered in section 2.2. but incorporating demographic stochasticity. As we did in that section we will restrict ourselves to the case in which there are only 2 age classes and 2 spatial patches, although the technique is valid in the general setting of [56], which, under some assumptions can be applied to populations

governed by two stochastic processes (not necessarily corresponding to demography and migration) with different time scales. In particular this covers the case of a multiregional model with any number of age classes and spatial patches.

Therefore we consider a population evolving in discrete time and structured in two age classes, young and adults, that we will denote 1 and 2 respectively. Moreover the population is distributed among two spatial patches, T1 and T2. Let  $x^{ij}(n)$ ,  $i, j = 1, 2$  be the number of individuals of age  $i$  living in patch  $j$  at time  $n$ . So the population vector at time  $n$  is

$$\mathbf{x}(n) = (x^{11}(n), x^{12}(n), x^{21}(n), x^{22}(n))^T$$

The evolution of the population is governed by demography in each patch and migration among the patches and we assume that both processes are influenced by the effects of demographic stochasticity, and that both processes are density independent (which is not very unrealistic for low populations sizes, which is the domain in which this kind of models is particularly useful). This results in demography being modelled by a discrete time birth-death process, whilst migration is modelled by a Markov chain. We also assume that migration is a fast process with respect to demography.

Given a vector  $\alpha = (\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}) \in \mathbb{Z}_+^4$ , let  $p_S^{ij}(\alpha)$  denote the probability of an individual of age  $i$  living at patch  $j$  to produce an offspring  $\alpha$  (i.e.,  $\alpha_{11}$  young at T1,  $\alpha_{12}$  at T2,  $\alpha_{21}$  adults at T1 and  $\alpha_{22}$  at T2) after an iteration of the birth-death process. Then, the p.g.f. for the slow process is the function  $\mathbf{G}_S : \mathbb{R}^4 \rightarrow \mathbb{R}^4$

$$\mathbf{G}_S(\mathbf{s}) = (G_S^{11}(\mathbf{s}), G_S^{12}(\mathbf{s}), G_S^{21}(\mathbf{s}), G_S^{22}(\mathbf{s}))$$

given by

$$G_S^{ij}(\mathbf{s}) = \sum_{\alpha \in \mathbb{Z}_+^4} p_S^{ij}(\alpha) \mathbf{s}^\alpha \quad ; \quad i, j = 1, 2 \quad (4.4)$$

where  $\mathbf{s} = (s_{11}, s_{12}, s_{21}, s_{22})$  and we are denoting  $\mathbf{s}^\alpha := \prod_{ij} s_{ij}^{\alpha_{ij}} = s_{11}^{\alpha_{11}} s_{12}^{\alpha_{12}} s_{21}^{\alpha_{21}} s_{22}^{\alpha_{22}}$ . We will assume that for each  $i$  and  $j$  there is only a finite number of  $\alpha$  for which  $p_S^{ij}(\alpha)$  is non zero, and so the sum in (4.4) is in fact finite.

In order to write a more specific example, and for the sake of simplicity in the exposition, sometimes we will consider the particular case, to which we will refer as (S), in which each individual can produce at most two offspring in a time step and besides adult individuals do not survive after the time step of the model. Assuming this situation and taking into account that the newborn offspring from a father inhabits the patch of the father we have that  $p_S^{i1}(\alpha) = 0$  if  $\alpha_{12} > 0$  or  $\alpha_{22} > 0$  and that  $p_S^{i2}(\alpha) = 0$  if  $\alpha_{11} > 0$  or  $\alpha_{21} > 0$ . Besides,  $p_S^{i2}(\alpha) = 0$  if  $\alpha_{11} > 0$  or  $\alpha_{12} > 2$ . On the other hand, a young individual may produce an adult by surviving to the next generation but it will never produce more than one, so  $p_S^{1j}(\alpha) = 0$  if  $\alpha_{2j} > 1$ . Finally, an adult can just produce young individuals, so  $p_S^{2j}(\alpha) = 0$  if  $\alpha_{2j} > 0$ . Therefore the p.g.f. associated to the birth-death process in our example has the form

$$\begin{aligned} G_S^{11}(\mathbf{s}) &= p_{0000}^{11} + p_{1000}^{11} s_{11} + p_{2000}^{11} s_{11}^2 + p_{0010}^{11} s_{21} + p_{1010}^{11} s_{11} s_{21} + p_{2010}^{11} s_{11}^2 s_{21} \\ G_S^{12}(\mathbf{s}) &= p_{0000}^{12} + p_{0100}^{12} s_{12} + p_{0200}^{12} s_{12}^2 + p_{0001}^{12} s_{22} + p_{0101}^{12} s_{12} s_{22} + p_{0201}^{12} s_{12}^2 s_{22} \\ G_S^{21}(\mathbf{s}) &= p_{0000}^{21} + p_{1000}^{21} s_{11} + p_{2000}^{21} s_{11}^2 \\ G_S^{22}(\mathbf{s}) &= p_{0000}^{22} + p_{0100}^{22} s_{12} + p_{0200}^{22} s_{12}^2 \end{aligned}$$

where, in order to keep the notation compact, we are denoting  $p_{1010}^{11} := p_S^{11}(1, 0, 1, 0)$  and so on.

We denote by  $\mathbf{M} \in \mathbb{R}^{4 \times 4}$  the matrix of expected values for demography. We can consider  $\mathbf{M}$  divided into four blocks  $\mathbf{M}_{ij} = [M_{ij}^{tt}] \in \mathbb{R}^{2 \times 2}$ ,  $i, j = 1, 2$  in such a way that  $M_{ij}^{tt}$  is the expected value at time  $n + 1$  of the individuals of age  $i$  living in patch  $t$  given that the population at time  $n$  is composed of a single individual of age  $j$  and living in patch  $l$ . Therefore,

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{bmatrix}, \quad (4.5)$$

where each block  $\mathbf{M}_{ij}$  contains the expected values for the transference of individuals of age  $j$  to age  $i$ . In the case (S) matrix  $\mathbf{M}$  takes the form

$$\mathbf{M} = \begin{pmatrix} m_{11} & 0 & m_{13} & 0 \\ 0 & m_{22} & 0 & m_{24} \\ m_{31} & 0 & 0 & 0 \\ 0 & m_{42} & 0 & 0 \end{pmatrix}$$

where

$$\begin{aligned} m_{11} & : = M_{11}^{11} = p_{1000}^{11} + 2p_{2000}^{11} + p_{1010}^{11} + 2p_{2010}^{11} \\ m_{31} & : = M_{21}^{11} = p_{0010}^{11} + p_{1010}^{11} + p_{2010}^{11} \\ m_{22} & : = M_{11}^{22} = p_{0100}^{12} + 2p_{0200}^{12} + p_{0101}^{12} + 2p_{0201}^{12} \\ m_{42} & : = M_{21}^{22} = p_{0001}^{12} + p_{0101}^{12} + p_{0201}^{12} \\ m_{13} & : = M_{12}^{11} = p_{1000}^{21} + 2p_{2000}^{21} \\ m_{24} & : = M_{12}^{22} = p_{0100}^{22} + 2p_{0200}^{22}, \end{aligned}$$

Regarding migration, since this process is conservative of the total number of individuals of each age, it is modelled by a Markov chain with two states corresponding to the two patches.

For each vector  $\alpha = (\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}) \in \mathbb{Z}_+^4$ , let  $p_F^{ij}(\alpha)$  denote the probability of an  $i$ -aged individual living in patch  $j$  to produce an "offspring"  $\alpha$  after an iteration of the migration process. Then, the p.g.f. associated to migration is the function  $\mathbf{G}_F : \mathbb{R}^4 \rightarrow \mathbb{R}^4$

$$\mathbf{G}_F(\mathbf{s}) = (G_F^{11}(\mathbf{s}), G_F^{12}(\mathbf{s}), G_F^{21}(\mathbf{s}), G_F^{22}(\mathbf{s}))$$

given by

$$G_F^{ij}(\mathbf{s}) = \sum_{\alpha \in \mathbb{Z}_+^4} p_F^{ij}(\alpha) \mathbf{s}^\alpha, \quad i, j = 1, 2 \quad (4.6)$$

Let  $\mathbf{e}^{ij}$  be the vector that describes a population composed of a single individual with age  $i$  and living in patch  $j$ . Since  $p_F^{ij}(\alpha) = 0$  if  $\alpha \notin \{\mathbf{e}^{i1}, \mathbf{e}^{i2}\}$  we have

$$\begin{aligned} G_F^{11}(\mathbf{s}) & = p_F^{11}(\mathbf{e}^{11}) s_{11} + p_F^{11}(\mathbf{e}^{12}) s_{12}; \quad G_F^{12}(\mathbf{s}) = p_F^{12}(\mathbf{e}^{11}) s_{11} + p_F^{12}(\mathbf{e}^{12}) s_{12} \\ G_F^{21}(\mathbf{s}) & = p_F^{21}(\mathbf{e}^{21}) s_{21} + p_F^{21}(\mathbf{e}^{22}) s_{22}; \quad G_F^{22}(\mathbf{s}) = p_F^{22}(\mathbf{e}^{21}) s_{21} + p_F^{22}(\mathbf{e}^{22}) s_{22}, \end{aligned}$$

The matrices of expected values for migration in each age class are

$$\mathbf{P}_i = \begin{pmatrix} p_F^{i1}(\mathbf{e}^{i1}) & p_F^{i2}(\mathbf{e}^{i1}) \\ p_F^{i1}(\mathbf{e}^{i2}) & p_F^{i2}(\mathbf{e}^{i2}) \end{pmatrix}; \quad i = 1, 2$$

and for the whole population we have matrix  $\mathbf{P} = \text{diag}\{\mathbf{P}_1, \mathbf{P}_2\}$ . In this case with only two patches, we can make notation simpler. If we let  $p_i$  be the probability of an  $i$ -aged individual living in patch 1 migrating to patch 2 after an iteration of the migration process, and  $t_i$  has the same meaning but in this case for the migration from patch 2 to 1 we have

$$\begin{aligned} G_F^{11}(\mathbf{s}) &= (1 - p_1)s_{11} + p_1s_{12} & ; & \quad G_F^{12}(\mathbf{s}) = t_1s_{11} + (1 - t_1)s_{12} \\ G_F^{21}(\mathbf{s}) &= (1 - p_2)s_{21} + p_2s_{22} & ; & \quad G_F^{22}(\mathbf{s}) = t_2s_{21} + (1 - t_2)s_{22} \end{aligned}$$

and

$$\mathbf{P}_i = \begin{pmatrix} 1 - p_i & t_i \\ p_i & 1 - t_i \end{pmatrix}; \quad i = 1, 2$$

We formulate a global model that takes into account the joint effect of both demography and migration in the following way: we set the time step of the model as the projection interval of the birth-death process, and we assume that, in each one of these time steps, migration acts  $k$  times before the birth-death process does. Here  $k$  can be interpreted as the ratio of the characteristic times of demography and migration, and in applications it will be a large integer. Then [13, 56] the p.g.f. of the global system is

$$\mathbf{G}_k(\mathbf{s}) = \mathbf{G}_F \circ \overset{(k)}{\dots} \circ \mathbf{G}_F \circ \mathbf{G}_S, \quad (4.7)$$

and the matrix of expectations for this global model is  $\mathbf{M}\mathbf{P}^k$ .

## 4.2. Reduction of the model

In order to reduce the system we assume that migration of individuals of each age class reaches an equilibrium among the patches, i.e., that the limit

$$\mathbf{G}_{\bar{F}} := \lim_{k \rightarrow \infty} \mathbf{G}_F \circ \overset{(k)}{\dots} \circ \mathbf{G}_F$$

exists. This holds if  $\mathbf{P}_1$  and  $\mathbf{P}_2$  are primitive matrices, which is the case if  $0 < p_i, t_i < 1$ ,  $i = 1, 2$ . Then migration of individuals of each age class reaches an equilibrium given by the stationary distribution

$$\mathbf{v}_i = (v_i^1, v_i^2)^T = \left( \frac{p_F^{i2}(\mathbf{e}^{i1})}{p_F^{i1}(\mathbf{e}^{i2}) + p_F^{i2}(\mathbf{e}^{i1})}, \frac{p_F^{i1}(\mathbf{e}^{i2})}{p_F^{i1}(\mathbf{e}^{i2}) + p_F^{i2}(\mathbf{e}^{i1})} \right)^T = \left( \frac{t_i}{p_i + t_i}, \frac{p_i}{p_i + t_i} \right)^T, \quad i = 1, 2$$

in such a way that  $\mathbf{G}_{\bar{F}} = (G_{\bar{F}}^{11}, G_{\bar{F}}^{12}, G_{\bar{F}}^{21}, G_{\bar{F}}^{22})$  is given by

$$G_{\bar{F}}^{ij} = \frac{t_i}{p_i + t_i} s_{i1} + \frac{p_i}{p_i + t_i} s_{i2}, \quad i, j = 1, 2$$

Note that  $G_{\bar{F}}^{ij}$  does not depend of  $j$ .

Now, we consider the auxiliary system, with variables

$$\hat{\mathbf{x}}(n) = (\hat{x}^{11}(n), \hat{x}^{12}(n), \hat{x}^{21}(n), \hat{x}^{22}(n))^T$$

that can be interpreted as the original system (4.7) when  $k \rightarrow \infty$ , i.e., when we let migration reach equilibrium. Therefore its p.g.f. is defined by

$$\hat{\mathbf{G}}(\mathbf{s}) = \left( \hat{G}^{11}(\mathbf{s}), \hat{G}^{12}(\mathbf{s}), \hat{G}^{21}(\mathbf{s}), \hat{G}^{22}(\mathbf{s}) \right) := \lim_{k \rightarrow \infty} \mathbf{G}_k(\mathbf{s}) = \mathbf{G}_{\bar{F}} \circ \mathbf{G}_S(\mathbf{s})$$

so that

$$\begin{aligned} \hat{G}^{1j}(\mathbf{s}) &= v_1^1 G_S^{11}(\mathbf{s}) + v_1^2 G_S^{12}(\mathbf{s}) = \frac{t_1}{p_1 + t_1} G_S^{11}(\mathbf{s}) + \frac{p_1}{p_1 + t_1} G_S^{12}(\mathbf{s}) \\ \hat{G}^{2j}(\mathbf{s}) &= v_2^1 G_S^{21}(\mathbf{s}) + v_2^2 G_S^{22}(\mathbf{s}) = \frac{t_2}{p_2 + t_2} G_S^{21}(\mathbf{s}) + \frac{p_2}{p_2 + t_2} G_S^{22}(\mathbf{s}), \quad j = 1, 2 \end{aligned}$$

and in particular  $\hat{G}^{ij}$  does not depend on  $j$ . The matrix of expected values for the auxiliary system is  $\mathbf{M}\bar{\mathbf{P}}$  where

$$\bar{\mathbf{P}} = \begin{pmatrix} v_1^1 & v_1^1 & 0 & 0 \\ v_1^2 & v_1^2 & 0 & 0 \\ 0 & 0 & v_2^1 & v_2^1 \\ 0 & 0 & v_2^2 & v_2^2 \end{pmatrix} = \begin{pmatrix} \frac{t_1}{p_1 + t_1} & \frac{t_1}{p_1 + t_1} & 0 & 0 \\ \frac{p_1}{p_1 + t_1} & \frac{p_1}{p_1 + t_1} & 0 & 0 \\ 0 & 0 & \frac{t_2}{p_2 + t_2} & \frac{t_2}{p_2 + t_2} \\ 0 & 0 & \frac{p_2}{p_2 + t_2} & \frac{p_2}{p_2 + t_2} \end{pmatrix}$$

Now we define the aggregated or reduced system as the system with variables  $\mathbf{y}_n = (y_n^1, y_n^2)^T$  corresponding to the total population in each class for the auxiliary system, i.e.,

$$y_n^i := \hat{x}_n^{i1} + \hat{x}_n^{i2}; \quad i = 1, 2$$

In [56] it is shown that the p.g.f.

$$\bar{G}(s_1, s_2) = (\bar{G}^1(s_1, s_2), \bar{G}^2(s_1, s_2)) \quad (4.8)$$

of the aggregated system is given by

$$\begin{aligned} \bar{G}^1(s_1, s_2) &= G'^{1j}(s_1, s_1, s_2, s_2) = v_1^1 G_S^{11}(s_1, s_1, s_2, s_2) + v_1^2 G_S^{12}(s_1, s_1, s_2, s_2) \\ \bar{G}^2(s_1, s_2) &= G'^{2j}(s_1, s_1, s_2, s_2) = v_2^1 G_S^{21}(s_1, s_1, s_2, s_2) + v_2^2 G_S^{22}(s_1, s_1, s_2, s_2) \end{aligned}$$

In the particular case (S) we have

$$\begin{aligned} \bar{G}^1(s_1, s_2) &= \bar{p}_{00}^1 + \bar{p}_{10}^1 s_1 + \bar{p}_{20}^1 s_1^2 + \bar{p}_{01}^1 s_2 + \bar{p}_{11}^1 s_1 s_2 + \bar{p}_{21}^1 s_1^2 s_2 \\ \bar{G}^2(s_1, s_2) &= \bar{p}_{00}^2 + \bar{p}_{10}^2 s_1 + \bar{p}_{20}^2 s_1^2 \end{aligned}$$

where

$$\begin{aligned}
\bar{p}_{00}^{-1} &= \frac{t_1}{p_1 + t_1} p_{0000}^{11} + \frac{p_1}{p_1 + t_1} p_{0000}^{12}; \bar{p}_{10}^{-1} = \frac{t_1}{p_1 + t_1} p_{1000}^{11} + \frac{p_1}{p_1 + t_1} p_{0100}^{12} \\
\bar{p}_{20}^{-1} &= \frac{t_1}{p_1 + t_1} p_{2000}^{11} + \frac{p_1}{p_1 + t_1} p_{0200}^{12}; \bar{p}_{01}^{-1} = \frac{t_1}{p_1 + t_1} p_{0010}^{11} + \frac{p_1}{p_1 + t_1} p_{0001}^{12} \\
\bar{p}_{11}^{-1} &= \frac{t_1}{p_1 + t_1} p_{1010}^{11} + \frac{p_1}{p_1 + t_1} p_{0101}^{12}; \bar{p}_{21}^{-1} = \frac{t_1}{p_1 + t_1} p_{2010}^{11} + \frac{p_1}{p_1 + t_1} p_{0201}^{12} \\
\bar{p}_{00}^{-2} &= \frac{t_2}{p_2 + t_2} p_{0000}^{21} + \frac{p_2}{p_2 + t_2} p_{0000}^{22}; \bar{p}_{10}^{-2} = \frac{t_2}{p_2 + t_2} p_{1000}^{21} + \frac{p_2}{p_2 + t_2} p_{0100}^{22} \\
\bar{p}_{20}^{-2} &= \frac{t_2}{p_2 + t_2} p_{2000}^{21} + \frac{p_2}{p_2 + t_2} p_{0200}^{22}
\end{aligned}$$

Note that the probabilities of "offspring production" for the aggregated system are linear combinations of the probabilities corresponding to the birth-death process modelling demography in which the coefficients of the combination are the equilibrium probabilities corresponding to migration.

The matrix of expected values for the aggregated system is shown to be

$$\bar{\mathbf{M}} = \mathbf{U}\mathbf{M}\mathbf{V}$$

where

$$\mathbf{U} := \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}; \mathbf{V} = \begin{pmatrix} v_1^1 & 0 \\ v_1^2 & 0 \\ 0 & v_2^1 \\ 0 & v_2^2 \end{pmatrix}$$

In case (S), matrix  $\bar{\mathbf{M}}$  takes the form

$$\bar{\mathbf{M}} = \begin{pmatrix} \frac{t_1}{p_1+t_1} m_{11} + \frac{p_1}{p_1+t_1} m_{22} & \frac{t_2}{p_2+t_2} m_{13} + \frac{p_2}{p_2+t_2} m_{24} \\ \frac{t_1}{p_1+t_1} m_{31} + \frac{p_1}{p_1+t_1} m_{42} & 0 \end{pmatrix}$$

### 4.3. Some generalities about the asymptotic behavior of MBGWBP

We are now interested in establishing results that guarantee that the behavior of the original system (4.7) can be approximated in terms of the behavior of the reduced system (4.8) when  $k$  is large enough. In contrast with the situation in section 2. (see equation (3.6)), in this setting it is not feasible to give "straightforward" results to relate the random variables  $\mathbf{x}_k(n)$  and  $\mathbf{y}(n)$  for fixed  $n$  pathwise, i.e., for each  $\omega \in \Omega$ . However one can show that for any finite value of  $n$ , we have

$$\|\mathbf{x}_k(n)\| \xrightarrow[k \rightarrow \infty]{\mathcal{L}} \|\mathbf{y}(n)\|$$

where  $\mathcal{L}$  denotes convergence in distribution. So the distribution of the total population of the original system can be approximated, for large enough  $k$ , by the distribution of the total population in the reduced system.

Let us now turn our attention to the relationships between the asymptotic ( $n \rightarrow \infty$ ) behavior of the models (4.7) and (4.8). In order to do that, let us review some of the most important features of the asymptotic behavior of a MBGWBP. Let us consider the system with variables (4.1), defined by the p.g.f. (4.2). If we fix an initial population  $\mathbf{z}_0 \geq 0$ ,  $\mathbf{z} \neq \mathbf{0}$  the expected value of the population vector verifies

$$\mathbb{E}\mathbf{z}(n) = \mathbf{A}^n \mathbf{z}_0$$

i.e. the expected value of the population behaves as in a classical linear matrix model: it grows (or decays) exponentially with rate given by  $\lambda :=$ dominant eigenvalue of  $\mathbf{A}$ , and moreover

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}\mathbf{z}(n)}{\lambda^n} = (\mathbf{1}^T \mathbf{z}_0) \mathbf{r} \quad (4.9)$$

where  $\mathbf{r}$  and  $\mathbf{l}$  are, respectively, right and left eigenvectors of matrix  $\mathbf{A}$  associated to eigenvalue  $\lambda$  and normalized so that  $\|\mathbf{r}\| = 1$ ,  $\mathbf{l}^T \mathbf{r} = 1$ . This allows one to classify the processes regarding the asymptotic behavior of the mean population as supercritical/critical/subcritical when  $\lambda$  is greater/equal/less than one.

As we have already mentioned, the effects of demographic stochasticity are particularly useful in studying the extinction of the population. Let us then define  $q_n(\mathbf{z}_0) = \mathbb{P}(\mathbf{z}(n) = \mathbf{0} \mid \mathbf{z}(0) = \mathbf{z}_0)$  as the probability of extinction up to time  $n$  for an initial population  $\mathbf{z}(0) = \mathbf{z}_0$ . Then one can show that  $q_n(\mathbf{z}_0)$  tends as  $n \rightarrow \infty$  to a limit  $q(\mathbf{z}_0) = \mathbb{P}(\mathbf{z}(n) = \mathbf{0} \text{ for some } n \mid \mathbf{z}(0) = \mathbf{z}_0)$  which is the probability of ultimate extinction given  $\mathbf{z}(0) = \mathbf{z}_0$ . By independence, we can compute  $\mathbf{q}(\mathbf{z}_0)$  if we know vector  $\mathbf{q} = (q^1, \dots, q^m)$  where  $q^i := \lim_{n \rightarrow \infty} \mathbb{P}(\mathbf{z}(n) = \mathbf{0} \mid \mathbf{z}(0) = \mathbf{e}^i)$ ,  $i = 1, \dots, m$ .

In all what follows we will assume the "generic case", i.e., that the MBGWBP is positive regular (i.e., matrix  $\mathbf{A}$  is primitive) and non-singular (i.e., its p.g.f. is not a linear function). Under those conditions, classical results [35, 45, 30] guarantee that:

a) In the subcritical and critical cases  $\mathbf{q} = \mathbf{1}$  so that  $q(\mathbf{z}_0) = 1$  for all  $\mathbf{z}_0$ , and so the population eventually goes extinct with probability one for any initial condition. In the supercritical case  $\mathbf{0} \leq \mathbf{q} < \mathbf{1}$  (i.e.,  $0 \leq q^i < 1$ ,  $i = 1, \dots, m$ ) and therefore  $0 \leq q(\mathbf{z}_0) < 1$  for all  $\mathbf{z}_0$ , so even although the mean population size grows to infinity, there is always a positive probability of extinction. Moreover,  $\mathbf{q}$  is the only solution of the fixed point equation  $\mathbf{G}(\mathbf{q}) = \mathbf{q}$  that belongs to the set  $\mathbf{0} \leq \mathbf{q} < \mathbf{1}$

b) Let us fix an initial condition  $\mathbf{z}_0$  and let us examine the asymptotic behavior of the second order moments  $\mathbb{E}(z^i(n)z^j(n))$  of the population vector, that we group in the matrix  $\mathbf{C}(n) = \mathbb{E}(\mathbf{z}(n)\mathbf{z}(n)^T)$ . In the subcritical (resp. supercritical) case,  $\mathbf{C}(n)$  decays (resp. grows) geometrically with an asymptotic growth rate  $\lambda$  (resp.  $\lambda^2$ ) and  $\mathbf{C}(n)/\lambda^n$  (resp.  $\mathbf{C}(n)/\lambda^{2n}$ ) converges when  $n \rightarrow \infty$  to a matrix that can be calculated explicitly in terms of  $\mathbf{A}$ ,  $\mathbf{r}$ ,  $\mathbf{l}$  and the so called covariance matrices of offspring production  $\mathbf{D}^j = [D_{rl}^j] \in \mathbb{R}^{m \times m}$  defined through  $D_{rl}^i = \text{Cov}(z^r(n+1), z^l(n+1) \mid \mathbf{z}(n) = \mathbf{e}^i)$ ;  $i, r, l = 1, \dots, m$ , that can be directly computed from the p.g.f. of the process.

c) If the process is supercritical then, with probability one

$$\frac{\mathbf{z}(n)}{\lambda^n} \xrightarrow{n \rightarrow \infty} W \mathbf{r} \quad (4.10)$$

where  $W$  is a non-negative scalar random variable. Therefore almost all realizations verify that the total population normalized by the factor  $\lambda^n$  converges to a scalar random variable  $W$ , i.e., for each  $\omega \in \Omega$  the total population behaves like  $W(\omega)\lambda^n + o_\omega(\lambda^n)$  for large  $n$ . Moreover, for every realization for which there is no ultimate extinction the structure  $\mathbf{z}(n)/\|\mathbf{z}(n)\|$  of the population vector converges to  $\mathbf{r}$ , a fixed (non random) vector.

d) If the process is subcritical, there exists an asymptotic stationary distribution for the population conditional on non-extinction, which is independent of the initial population. In other words, for each  $\alpha \in \mathbb{Z}_+^m$  and each  $\mathbf{z}(0) \neq \mathbf{0}$ ,  $\lim_{n \rightarrow \infty} \mathbb{P}(\mathbf{z}(n) = \alpha \mid \mathbf{z}(n) \neq \mathbf{0}, \mathbf{z}(0) = \mathbf{z}_0)$  exists and is independent of  $\mathbf{z}_0$ . Moreover under very general conditions that are trivially satisfied if the p.g.f. has only a finite number of summands, we have

$$\lim_{n \rightarrow \infty} \mathbb{E}(\mathbf{z}(n) \mid \mathbf{z}(n) \neq \mathbf{0}, \mathbf{z}(0) = \mathbf{z}_0) = \frac{1}{\xi} \mathbf{r} \quad (4.11)$$

for any  $\mathbf{z}_0 \neq \mathbf{0}$ , where  $\xi > 0$ . Therefore, the expected value of the population conditional on non-extinction converges to a vector proportional to  $\mathbf{r}$ .

#### 4.4. Relationships between the original and the reduced system

Keeping in mind the previous general results for MBGWBP, let us go back to investigate the asymptotic relationships between systems (4.7) and (4.8). As starting hypotheses we will assume:

i) The reduced system (4.8) is positively regular, i.e., matrix  $\bar{\mathbf{M}}$  is primitive, and non-singular. Let  $\bar{\lambda}$  be the dominant eigenvalue of  $\bar{\mathbf{M}}$  and let  $\bar{\mathbf{r}}$  and  $\bar{\mathbf{l}}$  be the right and left associated positive eigenvectors normalized so that  $\|\bar{\mathbf{r}}\| = 1$ ,  $\bar{\mathbf{l}}^T \bar{\mathbf{r}} = 1$ .

ii) The matrix  $\mathbf{M}$  (4.5) of expected values for demography has no zero rows.

In these conditions, one can show [56] that if  $k$  is large enough, the original system is non-singular and the matrix  $\mathbf{M}\mathbf{P}^k$  of expected values for the original system is primitive, so this system is also positively regular. Moreover the dominant eigenvalue  $\lambda_k$  of  $\mathbf{M}\mathbf{P}^k$  can be expressed in the form

$$\lambda_k = \bar{\lambda} + o(\gamma^k), \quad (4.12)$$

where  $0 < \gamma < 1$ , and associated to  $\lambda_k$  there are right and left eigenvectors  $\mathbf{r}_k$  and  $\mathbf{l}_k$  that can be written in the form

$$\mathbf{r}_k = \mathbf{M}\mathbf{V}\bar{\mathbf{r}} + o(\gamma^k); \quad \mathbf{l}_k = \mathbf{U}^T\bar{\mathbf{l}} + o(\gamma^k). \quad (4.13)$$

and so in particular, using (4.9) we have that the expected value of the normalized population vector for the original system can be expressed in terms of  $\bar{\mathbf{r}}$  and  $\bar{\mathbf{l}}$  in the following way

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\mathbb{E}\mathbf{x}_k(n)}{(\lambda_k)^n} &= \frac{(\bar{\mathbf{l}}^T \mathbf{U} + o(\gamma^k)) \mathbf{x}_k(0) (\mathbf{M}\mathbf{V}\bar{\mathbf{r}} + o(\gamma^k))}{(\bar{\mathbf{l}}^T \mathbf{U} + o(\gamma^k)) (\mathbf{M}\mathbf{V}\bar{\mathbf{r}} + o(\gamma^k))} = \frac{(\bar{\mathbf{l}}^T \mathbf{y}_0) \mathbf{M}\mathbf{V}\bar{\mathbf{r}} + o(\gamma^k)}{\bar{\mathbf{l}}^T \bar{\mathbf{M}}\bar{\mathbf{r}} + o(\gamma^k)} \\ &= \frac{\bar{\mathbf{l}}^T \mathbf{y}_0}{\bar{\lambda}} \mathbf{M}\mathbf{V}\bar{\mathbf{r}} + o(\gamma^k) \end{aligned}$$

From (4.12) we have that if  $\bar{\lambda} > 1$  (resp.  $\bar{\lambda} < 1$ ) and if  $k$  is large enough then  $\lambda_k > 1$  (resp.  $\lambda_k < 1$ ) and so if the reduced system is supercritical (resp. subcritical) then so is the original

system. Therefore, in the case in which the reduced system is subcritical we have that for any initial condition the original system will go extinct with probability one. Now, let us assume that the reduced system is supercritical and let  $\bar{q}^i := \mathbb{P}(\mathbf{y}(n) = \mathbf{0} \text{ for some } n \mid \mathbf{y}(0) = \mathbf{e}^i)$ ,  $i = 1, 2$  and  $q_k^{ij} = \mathbb{P}(\mathbf{x}_k(n) = \mathbf{0} \text{ for some } n \mid \mathbf{x}_k(0) = \mathbf{e}^{ij})$ ,  $i, j = 1, 2$ , be the ultimate extinction probabilities associated to the reduced and original system respectively. Then it can be proved that  $\lim_{k \rightarrow \infty} q_k^{ij} = \bar{q}^i$  and so we can approximate the extinction probabilities of the original system through those corresponding to the reduced system. Recall that in order to calculate the latter, we must solve the fixed point equation  $\bar{\mathbf{G}}(\bar{\mathbf{q}}) = \bar{\mathbf{q}}$  in the set  $\mathbf{0} \leq \bar{\mathbf{q}} < \mathbf{1}$  of  $\mathbb{R}^2$ , which is much easier than solving an equation in  $\mathbb{R}^4$ .

Regarding the asymptotic behavior of the second order moments  $\mathbb{E}(\mathbf{x}_k(n)\mathbf{x}_k(n)^T)$  of (4.7) we have that in the supercritical (resp. subcritical) case they grow (resp. decay) exponentially at a rate  $\bar{\lambda}^2 + o(\gamma^k)$  (resp.  $\bar{\lambda} + o(\gamma^k)$ ) and  $\mathbb{E}(\mathbf{x}_k(n)\mathbf{x}_k(n)^T) / (\bar{\lambda}^2 + o(\gamma^k))$  (resp.  $\mathbb{E}(\mathbf{x}_k(n)\mathbf{x}_k(n)^T) / (\bar{\lambda} + o(\gamma^k))$ ) converges to a matrix that can be calculated explicitly in terms of  $\bar{\mathbf{r}}$ ,  $\bar{\mathbf{I}}$  and the covariance matrices of offspring production corresponding to demography.

From (4.10) we know that in the supercritical case  $\bar{\lambda} > 1$  we have that, with probability one,

$$\frac{\|\mathbf{y}(n)\|}{\bar{\lambda}^n} \xrightarrow[n \rightarrow \infty]{} \bar{W}$$

and

$$\frac{\|\mathbf{x}_k(n)\|}{(\lambda_k)^n} \xrightarrow[n \rightarrow \infty]{} W_k$$

where  $\bar{W}$  and  $W_k$  are non-negative scalar random variables. One can show [46] that  $W_k$  converges when  $k \rightarrow \infty$  to  $\bar{W}$  in distribution. Therefore we have, taking into account (4.13), that the asymptotic behavior of the scaled population vector for the original system is characterized by

$$\frac{\mathbf{x}_k(n)}{(\lambda_k)^n} \xrightarrow[n \rightarrow \infty]{} W_k \left( \frac{\mathbf{M}\mathbf{V}\bar{\mathbf{r}}}{\|\mathbf{M}\mathbf{V}\bar{\mathbf{r}}\|} + o(\gamma^k) \right) \text{ with probability 1}$$

where the distribution of  $W_k$  can be approximated by that of  $\bar{W}$ . Specifically, if  $k$  is large enough the asymptotic structure of the population of the original system can be approximated by vector  $\mathbf{M}\mathbf{V}\bar{\mathbf{r}} / \|\mathbf{M}\mathbf{V}\bar{\mathbf{r}}\|$  and for all  $t \geq 0$  the probability  $\mathbb{P}(\|\mathbf{x}_k(n)\| \leq (\lambda_k)^n t)$  can be approximated for large  $n$  by  $\mathbb{P}(\bar{W} \leq t)$ .

Let us now examine the relation between the original and the reduced system when the latter is subcritical ( $\bar{\lambda} < 1$ ). We know, using (4.11) and (4.13) that for all  $\mathbf{x}_0 \in \mathbb{R}^N$ ,  $\mathbf{x}_0 \geq \mathbf{0}$

$$\lim_{n \rightarrow \infty} \mathbb{E}(\mathbf{y}(n) \mid \mathbf{y}(n) \neq \mathbf{0}, \mathbf{y}(0) = \mathbf{y}_0) = \frac{1}{\bar{\xi}} \bar{\mathbf{r}}$$

where  $\mathbf{y}_0 = \mathbf{U}\mathbf{x}_0$  and

$$\lim_{n \rightarrow \infty} \mathbb{E}(\mathbf{x}_k(n) \mid \mathbf{x}_k(n) \neq \mathbf{0}, \mathbf{x}_k(0) = \mathbf{x}_0) = \frac{1}{\xi_k} \mathbf{M}\mathbf{V}\bar{\mathbf{r}} + o(\gamma^k)$$

where  $\bar{\xi}$  and  $\xi_k$  are scalar constants. It is possible to show [56] that  $\lim_{k \rightarrow \infty} \xi_k = \bar{\xi}$  and therefore we have that the expected value of the asymptotic population vector conditional on non extinction for the original system can be approximated through the knowledge of  $\bar{\mathbf{r}}$  and  $\bar{\xi}$ , that pertain to the reduced system.

## 5. Conclusion and perspectives

The last years in the literature of population dynamics have seen a dramatic increase in the number of works that incorporate stochasticity. Whenever the model is a discrete time one and the processes involved in it have different time scales, as in the case of multiregional models, the techniques presented in this work can be useful in order to reduce their dimension and therefore simplify their study. In some cases, like in the models in sections 3.3. and 3.4., the reduction procedure makes feasible the analytical study of the otherwise intractable original complex model. In particular, the reduced model makes it possible to study the influence of the parameters of the two processes in the global dynamics.

The multiregional models of section 2. have been applied to study a trout fish population in a multipatch setting corresponding to an arborescent river network when migration between the patches is fast with respect to demography [19, 20, 21, 22]), in deterministic linear and non-linear contexts. Similarly they have been studied, in a deterministic density-dependent context, to model the effect of migrations processes on host-parasitoid systems [40, 41, 61, 62]. Applications to real populations in stochastic settings remain to be done.

Regarding future lines of work, from the theoretical point of view and within the discrete time context, it would be interesting to work in the reduction of models that incorporate both demographic and environmental stochasticity, in the so called multitype branching processes in random environments (MBPRE) [10].

Most biological models are density dependent, and therefore it would be interesting to extend aggregation techniques to non linear stochastic settings. This would allow one to study stochastic models of great ecological interest in which the fast or the slow process can correspond to non linear dispersal, competition, epidemic dynamics, game dynamics, etc (see for example [6]). The difficulty here lies in the great analytical complexity of non linear stochastic models. In this direction, and restricting to the discrete time case, we will consider the aggregation of density-dependent branching models [14, 38]

The aggregation of systems in the context of stochastic time continuous models has been studied only in the case of stochastic differential equations (SDEs). [28] explores the (very restrictive) conditions under which it is possible to carry out the perfect aggregation of SDEs and [9] carries out the study of the approximate reduction of SDEs in the case that the random noise is small enough. It would be interesting to extend those results to deal with the case in which the noise intensity is not restricted. Moreover, the reduction of systems governed by time continuous branching processes [2] deserves to be explored, as this would allow one to deal with birth-death processes for a population living in a multi-patch environment.

## Acknowledgements

This work has been partially supported by Proyecto de Investigación MTM2008-06462-C02-01 (MCI).

## References

- [1] J.A. Alonso, L. Sanz. *Aproximating the Distribution of Population Size in Stochastic Multi-regional Matrix Models with Fast Migration*. Phil. Trans. R. Soc. A, 367 (2009), 4801-4827.
- [2] K. B. Athreya, P. E. Ney. *Branching processes*. Springer-Verlag, Berlin, 1972.
- [3] P. Auger. *Dynamics and Thermodynamics in Hierarchically Organized Systems, Applications in Physics, Biology and Economics*. Pergamon Press, Oxford, 1989.
- [4] P. Auger, R. Roussarie. *Complex ecological models with simple dynamics: from individuals to populations*. Acta Biotheoretica, 42 (1994), 111-136.
- [5] P. Auger, J.C. Poggiale. *Aggregation and Emergence in Systems of Ordinary Differential Equations*. Mathematical Computer Modelling, 27 (1998), 1-22.
- [6] P. Auger, M Rachid, C. Tanmay, S. Gauthier, T. Maurice, C. Joydev. *Effects of a disease affecting a predator on the dynamics of a predator-prey system*. Journal of theoretical biology, 258 (2009), 344-351.
- [7] P. Auger, C. Lett. *Integrative Biology: Linking Levels of Organization*. Comptes Rendus de l'Académie des Sciences Biologies, 326 (2003):517-522.
- [8] P. Auger, R. Bravo de la Parra, J.C. Poggiale, E. Sánchez, L. Sanz. *Aggregation methods in dynamical systems variables and applications in population and community dynamics*. Physics of Life Reviews, 5 (2008), 79-105.
- [9] N. Berglund, B. Gentz. *Geometric singular perturbation theory for stochastic differential equations*. J. Diff. Equat., 191 (2003), 1-54.
- [10] J.D. Biggins, H. Cohn, O. Nerman. *Multi-type branching in varying environment*. Stoc. Proc. Appl., 83 (1999), 357-400.
- [11] A. Blasco, L. Sanz, P. Auger, R. Bravo de la Parra. *Linear Discrete Population Models with Two Time Scales in Fast Changing Environments I: Autonomous Case*. Acta Biotheoretica, 49 (2001), 261-276.
- [12] A. Blasco, L. Sanz, P. Auger, R. Bravo de la Parra. *Linear Discrete Population Models with Two Time Scales in Fast Changing Environments II: Non Autonomous Case*. Acta Biotheoretica, 1 (2002), 15-38.
- [13] A. Blasco, L. Sanz, R. Bravo de la Parra. *Approximate reduction of multiregional birth-death models with fast migration*. Mathematical and Computer Modelling, 36 (2002), 47-65.
- [14] G. L. Block, L. J. S. Allen. *Population extinction and quasi-stationary behavior in stochastic density-dependent structured models*. Bull. Math. Bio., 62 (2000), 199-228.

- [15] R. Bravo, P. Auger, E. Sánchez. *Aggregation methods in discrete Models*, J. Biol. Sys., 3 (1995), 603-612.
- [16] R. Bravo de la Parra, E. Sánchez, O. Arino, P. Auger. *A Discrete Model with Density Dependent Fast Migration*. Mathematical Biosciences, 157 (1999), 91-110.
- [17] H. Caswell, M. Fujiwara, S. Brault. *Declining survival probability threatens the North Atlantic right whale*. Proc. Natl. Acad. Sci. USA, 96 (1999), 3308-3313.
- [18] H. Caswell. *Matrix population models* (2<sup>nd</sup> ed.). Sinauer Associates, Sunderland, Massachusetts, 2001.
- [19] 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 Biotheoretica, 46 (1998), 223-234.
- [20] S. Charles, R. Bravo de la Parra , J.P. Mallet, H. Persat, P. Auger. *A density dependent model describing Salmo trutta population dynamics in an arborescent river network: effects of dams and channelling*. C. R. Acad. Sci. Paris, Sciences de la vie, 321 (1998), 979-990.
- [21] S. Charles, R. Bravo de la Parra , J.P. Mallet, H. Persat, P. Auger. *Annual spawning migrations in modeling brown trout population dynamics inside an arborescent river network*. Ecological Modelling, 133 (2000), 15-31.
- [22] A. Chaumot, S. Charles, P. Flammarion, P. Auger. *Do Migratory or Demographic Disruptions Rule the Population Impact of Pollution in Spatial Networks?*. Theoretical Population Biology, 64 (2003), 473-480.
- [23] J. E. Cohen. *Ergodicity of Age Structure in Populations with Markovian Vital Rates, II, General States*. Advances in Appl. Probability, 9 (1977), 18-37.
- [24] J. E. Cohen. *Ergodics Theorems of Demography*. Bulletin of the American Mathematical Society N.S., 1 (1979), 275-295.
- [25] J. E. Cohen. *Multiregional age structured populations with changing vital rates: weak and strong stochastic ergodic theorems*. In Land, K.C., A. Rogers editors. Multiregional mathematical demography. Academic Press, New York, 477-503, 1982.
- [26] J.E. Cohen, S.W. Christensen, C.P. Goodyear. *A stochastic age-structured model of Striped Bass (Morone saxatilis) in the Potomac River*. Can. J. Fish. Aquat. Sci., 40 (1983), 2170-2183.
- [27] H. Furstenberg, H. Kesten. *Products of Random Matrices*. Ann. Math. Statist. 31 (1960), 457-469.
- [28] T.C. Gard. *Aggregation in stochastic ecosystem models*. Ecol. Modelling, 44 (1988), 153-164.

- [29] P. Haccou, P. Jagers, V. Vatutin. *Branching processes: Variation, growth, and extinction of populations*, Cambridge University Press, 2005.
- [30] T. Harris. *The theory of branching processes*, Springer-Verlag, Berlín, 1963.
- [31] C.C. Heyde, J. E. Cohen. Confidence Intervals for Demographic Projections Based on Products of Random Matrices. *Theoretical Population Biology* 27 (1985), 120-153.
- [32] K. E. Holsinger. *Demography and extinction in small populations*, in: *Genetics, demography and the viability of fragmented populations*, eds. A. Young, G. Clarke, Cambridge University Press, 2000.
- [33] R. Horn, C. Johnson. *Matrix Analysis*. Cambridge Univ. Press, 1985.
- [34] Y. Iwasa, V. Andreasen, S. Levin. *Aggregation in model ecosystems I: Perfect Aggregation*. *Ecological Modeling*, 37 (1987), 287-302.
- [35] A. Joffe, F. Spitzer. *On multitype branching processes with  $\rho \leq 1$* . *Journal of Mathematical Analysis and Its Applications*, 19 (1967), 409-430.
- [36] M. Khaladi M., V. Grosbois, J. D. Lebreton. *An explicit approach to evolutionarily stable dispersal strategies with a cost of dispersal*. *Nonlinear Anal.: Real World Appl.*, 1 (2000), 137-144.
- [37] M. Kimmel, D.E. Axelrod. *Branching Processes in Biology*, Springer, New York, 2002.
- [38] F. Klebaner, Population Size Dependent Processes. In: *Branching Processes: Variation, Growth and Extinction of Populations*, P. Haccou, P. Jagers and V.A. Vatutin, 133-135, Cambridge University Press, 2005.
- [39] S. Legendre, J. Clobert, A. P. Moller, G. Sorci. *Demographic stochasticity and social mating system in the process of extinction of small populations: the case of passerines introduced to New Zealand*. *The American Naturalist*, 153 (1999), 449-463.
- [40] C. Lett, P. Auger, R. Bravo de la Parra. *Migration Frequency and the Persistence of Host-Parasitoid Interactions*. *Journal of Theoretical Biology*, 221 (2003), 639-654.
- [41] C. Lett, P. Auger and F. Fleury. *Effects of asymmetric dispersal and environmental gradients on the stability of host-parasitoid systems*. *Oikos*, 109 (2005), 603-613.
- [42] K.L. Liaw. *Multistate dynamics: the convergence of an age-by-region population system*. *Environment and Planning A*, 12 (1980), 589-613.
- [43] L. Liaw. *Spatial Popuylation Dynamics*. In *Migration and Settlement: A multiregional comparative study*; A. Rogers, Willekens eds, 419-455. Dordrecht, D. Reidel, 1986.
- [44] M. Marva, E. Sanchez, R. Bravo de la Parra, L. Sanz. *Reduction of slow-fast discrete models coupling migration and demography*. *Journal of Theoretical Biology*, 258 (2009), 371-379.

- [45] C.J. Mode. *Multitype Branching Processes. Theory and Applications*. American Elsevier Publishing Co., Inc., New York, 1971.
- [46] M. A. Rincón, J.A. Alonso, L. Sanz. *Supercritical multiregional stochastic models with fast migration*. *Acta Biotheoretica*, 57 (2009), 479-500.
- [47] A. Rogers. *Shrinking large-scale population projection models by aggregation and decomposition*. *Environment and Planning A*, 8 (1976), 515-541.
- [48] A. Rogers. *Multiregional Demography*, Chichester, New York, 1995.
- [49] J. M. Saboia. *Arima models for birth forecasting*. *Journal of the American Statistical Association*, 72 (1977), 264-270.
- [50] E. Sánchez, R. Bravo de la Parra, P. Auger. *Linear discrete models with different time scales*. *Acta Bio.*, 43 (1995), 465-479.
- [51] L. Sanz, R. Bravo de la Parra. *Variables Aggregation in Time Varying Discrete Systems*. *Acta Biotheoretica*, 46 (1998), 273-297.
- [52] L. Sanz, R. Bravo de la Parra. *Variables aggregation in a time discrete linear model*. *Math. Biosc.* 157 (1999), 111-146.
- [53] L. Sanz, R. Bravo de la Parra. *Time scales in stochastic multiregional models*. *Nonlinear Analysis: Real World Applications*, 1 (2000), 89-122.
- [54] L. Sanz, R. Bravo de la Parra. *Time scales in a non autonomous linear discrete model*. *Mathematical Models and Methods in Applied Sciences*, 11 (2001), 1-33.
- [55] L. Sanz, R. Bravo de la Parra. *Approximate Reduction Techniques in Population Models with Two Time Scales: Study of the Approximation*, *Acta Biotheoretica*, 50 (2002), 297-322.
- [56] L. Sanz, A. Blasco and R. Bravo de la Parra, Approximate reduction of Galton-Watson processes with two time scales, *Mathematical Models and Methods in Applied Sciences* 13:491-525, 2003.
- [57] L. Sanz, R. Bravo de la Parra. *Approximate reduction of multiregional models with environmental stochasticity*. *Mathematical Biosciences*, 206 (2007), 134-154.
- [58] L. Sanz, R. Bravo de la Parra, E. Sánchez. *Approximate Reduction of Non-Linear Discrete Models with Two Time Scales*. *Journal of Difference Equations and Applications*, 14 (2008), 607-627.
- [59] G.W. Stewart, J.I. Guang Sun. *Matrix Perturbation Theory*, Academic Press, Boston, 1990.
- [60] Z. M. Sykes. *Some stochastic versions of the matrix model for population dynamics*. *J. Amer. Statist. Assoc.*, 64 (1969), 111-130.

- [61] T. Nguyen-Huu, C. Lett, P. Auger P. J.C. Poggiale. *Spatial synchrony in host-parasitoid models using aggregation of variables*. Mathematical Biosciences, 203 (2006), 204-221.
- [62] T. Nguyen-Huu, P. Auger, C. Lett, M. Marva. *Emergence of global behaviour in a host-parasitoid model with density-dependent dispersal in a chain of patches*. Ecological Complexity, 5 (2008), 9-21.
- [63] S. Tuljapurkar, S. Orzack. *Population dynamics in variable environments. I. Long-run growth rates and extinction*. Theor. Popul. Biol., 18 (1980) 314–342.
- [64] S. Tuljapurkar. *Demography in stochastic environments. I. Exact distributions of age structure*. J. Math. Biol., 19 (1984), 335-350.
- [65] S. Tuljapurkar. Population Dynamics in Variable Environments, Springer-Verlag, Berlin, 1990.
- [66] S. Tuljapurkar, H. Caswell (eds). Structured-Population Models in Marine, Terrestrial, and Freshwater Systems, Chapman and Hall, New York, 1997.
- [67] G. Wang, W. D. Edge, J. O. Wolff. *Demographic uncertainty in ecological risk assessments*. Ecological Modelling, 136 (2001), 95-102.