



# Approximate Reduction of Multiregional Birth-Death Models with Fast Migration

A. BLASCO\*

Departamento de Matemáticas, Universidad de Alcalá  
28871 Alcalá de Henares, Madrid, Spain  
angel.blasco@uah.es

L. SANZ

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

R. BRAVO DE LA PARRA

José Gutiérrez Abascal, 2, 28006 Madrid, Spain

(Received March 2002; accepted April 2002)

**Abstract**—In this work, we deal with the reduction of a time discrete model for a population distributed among  $N$  spatial patches and whose dynamics is controlled both by reproduction and by migration. These processes take place at different time scales in the sense of the latter being much faster than the former. We incorporate the effect of demographic stochasticity into the population, which results in both dynamics being modelled by multitype Galton-Watson branching processes. We present a multitype global model that incorporates the effect of the two processes and develop a method that takes advantage of the difference of time scales to reduce the model obtaining a unitype “aggregated” process that approximates the evolution of the total size of the population. We show that, given the separation of time scales between the birth-death process and the migration process is sufficiently high, we can obtain both qualitative and quantitative information about the behavior of the multitype global model through the study of this simple aggregated model. © 2002 Elsevier Science Ltd. All rights reserved.

**Keywords**—Approximate aggregation, Multiregional models, Demographic stochasticity.

## 1. INTRODUCTION

As a consequence of the intrinsic complexity of many ecological systems, their modelling implies dealing with systems involving a large number of variables. Indeed, in general, populations are not homogeneous, but they are composed of individuals with different ages, sizes, activities, locations, etc. Models that incorporate this inherent complexity of the populations often lead to systems with a large number of variables. These describe the biological system in detail, but have the severe disadvantage of being difficult to handle analytically.

A common approach to deal with such complex models is the use of computer simulations, but this involves dealing with restrictions, generally unknown, concerning robustness of solutions

\* Author to whom all correspondence should be addressed.

This work has been partially supported by Proyecto de Investigación PB98-0702 (Spanish M.E.C.).

with respect to parameters and initial conditions. Another possibility consists in simplifying the model by ignoring some characteristics of the individuals, but in this way, we lose information about the internal structure of the population which may be crucial in its dynamics.

Frequently, in detailed models, we can distinguish several processes that take place at different time scales. Indeed, changes taking place at the population level, such as the growth of the population, are frequently slow in comparison with those taking place at the individual level, such as migration, changes of activity, etc., see [1]. The so-called aggregation techniques have been developed to reduce the dimension of such models without losing the relevant information corresponding to the various processes. The “aggregated system” is a reduced model that takes into account the effects of both the fast and the slow dynamics and evolves at the time scale corresponding to the slow process. The fast dynamics is reflected in the coefficients of this reduced model allowing us to carry out the study of the interaction between the different processes through the analysis of the aggregated model.

Aggregation techniques were first applied to economical models [2] and later to ecology [3,4] in a general context, without requiring the presence of different time scales. More recently, they have been studied in the context of systems of ordinary differential equations [5–8] and partial differential equations [9]. These techniques have also been extended to deal with time discrete systems in both autonomous [10–13] and nonautonomous cases [14,15].

The literature in population dynamics contemplates two sources of stochasticity: environmental [16] and demographic [17]. The former has to do with the uncertainty induced in the vital rates by random environmental variations. Demographic stochasticity refers to the variability in a population arising from random differences among individuals in survival and reproduction. In this latter case, the number of offspring produced by an individual in each time interval is a discrete random variable (binomial, Poisson, etc.) with parameters depending on the individual’s stage. In this approach, any population (but in some degenerated cases) has a chance to eventually become extinct even if its mean number grows indefinitely. Models that take into account demographic stochasticity are particularly useful to study the problem of extinction of small populations. Indeed, it is in the context of small populations where the effect of demographic stochasticity is really relevant [18].

The reduction of systems for populations subjected to environmental stochasticity has been addressed in [19]. On the contrary, there are no works concerning the reduction of systems with different time scales and subjected to demographic stochasticity. In this work, we undertake the reduction of one such model.

The mathematical tool to study structured populations subjected to demographic stochasticity and evolving in discrete time is the multitype Galton-Watson branching process [20,21] that describes the dynamics of a population in which the offspring produced by each individual in a time step (or generation) is a random variable whose distribution depends on the current stage of the individual. In spite of their simplicity, this kind of model has frequently been used in practical applications [22,23].

This work addresses the approximate reduction of a multiregional model for a spatially distributed population that is governed by two processes, growth and migration, which are supposed to have different characteristic times.

In the first place, we present a multitype Galton-Watson model for a spatially distributed population. The population is supposed to be unstructured by age and the individuals are distributed among a set of  $N$  spatial patches. Its dynamics is controlled by two processes, both of which are subjected to demographic stochasticity: a birth-death process that models the growth of the population in each one of the patches and, on the other hand, a Markov chain that models migration among patches. We assume, as is the case in many practical biological situations [1,7,9], that the growth of the population is slow in comparison with migration, which results in the projection interval of the birth-death process being much larger than that of migration. We

build a global multitype Galton-Watson process with  $N$  types and time step corresponding to the projection interval of the birth-death process, that incorporates the effect of both dynamics.

The Markov chain modelling migration is referred to the characteristic time of this process. In order to approximate the effect of migration over the projection interval of the birth-death process, which is much longer than the characteristic time of migration, we assume that in each time step of the model, migration acts a large number  $k$  of times. Here  $k$  can be interpreted as the ratio of the characteristic times for the growth of the population and for migration. In this way, the model that takes into account the joint effect of growth and migration referred to the characteristic time of the former consists of a Galton-Watson process in which in each time step, there are  $k$  iterations of the migration process followed by one iteration of the birth-death process.

By making use of the existence of different time scales, we can approximate the resulting model, governed by  $N$  variables, by a unitype process for the total size of the population. In order to do so, we first approximate the original system by one in which migration has reached equilibrium in each time step and, next, we reduce the resulting system. Due to the approximation introduced to reduce the model, one cannot expect to be able to describe exactly the dynamics of the original system in terms of that of the reduced one. However, we show that we can obtain both qualitative and quantitative information about the behavior of the original process through the study of this simple aggregated model given the separation of time scales between the birth-death process and the migration process is sufficiently high. In particular, the behavior of the expected value of the population vector can be approximated in terms of that corresponding to the aggregated system. Moreover, the original system is shown to be supercritical (subcritical) if and only if the aggregated system is supercritical (subcritical) and the probability of ultimate extinction for the original system can be approximated by that corresponding to the aggregated system. In this way, the problem of extinction in the original multiregional model can be studied through the study of the simpler reduced unitype model.

## 2. ORGANIZATION OF THE WORK AND MAIN RESULTS

The paper is organized as follows. Section 3 presents a multitype Galton-Watson process (g.w.p.) with  $N$  types for the spatially distributed population according to the comments made in the last section. In the first place, we separately introduce the Galton-Watson processes corresponding to the birth-death process and to migration through their probability generating functions (p.g.f.). Each of these p.g.f.s is referred to the projection interval of the corresponding process.

As described in the introduction, we formulate a multitype g.w.p. that takes into account the effect of both dynamics and is referred to the characteristic time of the birth-death process. This model can be interpreted as the composition of  $k$  iterations of the migration process and one iteration of the birth-death process. Lemma 1 characterizes mathematically the composition of multitype g.w.p.s and allows us to build the p.g.f. of the resulting model. In Section 4, we approximate the original system by an auxiliary system in which the Markov chain modeling migration reaches its equilibrium distribution in each time step of the model. Let  $v$  denote the vector corresponding to that spatial distribution among the patches. By defining a new variable corresponding to the total population, the auxiliary system is reduced to obtain a unitype g.w.p. Proposition 1 allows us to construct the p.g.f. of this reduced system, which is a weighted sum of the p.g.f.s corresponding to the birth-death processes in the different patches, being the weights the components of vector  $v$ . Moreover, this result gives the growth rate  $\lambda$  of the mean population size for the reduced process in terms of vector  $v$  and the growth rates  $m_{ii}$ ,  $i = 1, \dots, N$  of the mean population size for the birth-death process in the different patches.

The behaviors of the original multiregional model and the reduced model are compared in Section 5. Proposition 3 states that the expected growth rate  $\lambda_k$  of the population size for

the global process can be obtained as a perturbation of  $\lambda$ , i.e., as a perturbation of the expected growth rate for the reduced model. Moreover, the dominant eigenvectors of the matrix of expected values for the global system are obtained as a perturbation of certain vectors only depending on  $v$  and the  $m_{ii}$ . These perturbations are characterized asymptotically when  $k$ , that measures the separation of time scales between the growth process and migration, tends to infinity. The size of this perturbation is related to the modulus of the subdominant eigenvalue  $\mu_2$  of the matrix of expected values for migration, in such a way that the faster migration reaches its equilibrium distribution (i.e., the lower  $|\mu_2|$  is), the smaller the perturbation is. As a direct consequence of Proposition 3, we can approximate the asymptotic behavior of the mean population vector in terms of the  $m_{ii}$ ,  $v$ , and  $\mu_2$ .

Proposition 5 shows that the probability of extinction in finite time in the global process can be obtained as a perturbation of the corresponding probability of extinction in the reduced system.

Theorem 1, that constitutes our main result, deals with the comparison of the probabilities of ultimate extinction in the global process and in the reduced process. For a sufficiently high separation of time scales, the global system is subcritical (supercritical) if and only if the aggregated process is subcritical (supercritical). Besides, in the supercritical case, the probability of ultimate extinction for the global process converges to that of the aggregated system. In this way, Theorem 1 guarantees that if the reduced process is subcritical, then so is the global process for big enough  $k$  and the probabilities of ultimate extinction for both process are one. Analogously, if the reduced process is supercritical (and so the probability of ultimate extinction is lower than one), then so is the global process for big enough  $k$ , and the probabilities of ultimate extinction of the latter can be approximated by that of the former.

An estimation of the accuracy of the approximations above, i.e., of the error we incur in when studying the global system in terms of the reduced one, has been carried out through computer simulations. These show the dependence of the error on the separation of time scales  $k$  and on  $|\mu_2|$ . Moreover, they show that this error is small even for moderate values of  $k$  and so our reduction technique can be useful in the study of real biological populations.

### 3. A GALTON-WATSON PROCESS FOR A SPATIALLY DISTRIBUTED POPULATION

We consider an unstructured population distributed among  $N$  spatial patches and evolving in discrete time. The number of individuals living in patch  $i$  at time  $t$  is denoted by  $x_t^i$ ,  $i = 1, \dots, N$ ;  $t = 0, 1, 2, \dots$ . Therefore, the population at a given time  $t$  is described by vector

$$\mathbf{x}_t = (x_t^1, \dots, x_t^N)^\top, \quad (1)$$

where  $\top$  denotes transposition.

We assume that the evolution of the population is governed by two processes: a birth-death process and a migration process. The former determines the number of offspring produced by each individual in each time interval. Here we use the term offspring in the sense of [17], i.e., an individual who survives a time interval is considered as part of its offspring in the next interval. On the other hand, migration determines the transference of individuals among the patches. Note that this process does not produce changes in the population size, but only in the distribution of the individuals among the patches.

Each of these processes is considered to be subjected to demographic stochasticity and, consequently, will be modelled by means of a multitype g.w.p. We assume that the characteristic time scale corresponding to migration is very short compared to that of the birth-death process. We will sometimes refer to these processes as fast and slow dynamics, respectively.

The projection interval of our model, i.e., the time elapsed between times  $t$  and  $t + 1$ , will be that corresponding to the birth-death process. For notational convenience, we will denote the time span  $[t, t + 1)$  as  $\Delta_t$ .

### 3.1. Probability Generating Function of the Birth-Death Process

The birth-death process determines the number of net offspring produced by each individual in each time interval.

We assume the following.

- (H1) An individual can produce offspring only in its current patch.
- (H2) In each patch, the probability of having no descendance is lower than one and there is only a finite number of possible values for the offspring production.

Let us initially assume that the system is exclusively controlled by the birth-death process, i.e., there is no migration, and let  $\mathbf{h}_t = (h_t^1, \dots, h_t^N)^\top$  denote the population vector for such a system at time  $t$ .

Throughout this work, let  $\mathbb{Z}_+^m$  denote the set of  $m$ -dimensional vectors whose components are nonnegative integers and let  $\mathbf{e}^j$  be the  $j^{\text{th}}$  canonical vector of  $\mathbb{R}^m$ . Unless stated otherwise, all vectors in this work are considered as column vectors. In addition, we will use the notation  $\mathbf{s}^\alpha := s_1^{\alpha_1} s_2^{\alpha_2} \dots s_m^{\alpha_m}$  for any pair of vectors  $\mathbf{s} = (s_1, \dots, s_m)^\top \in \mathbb{R}^m$  and  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)^\top \in \mathbb{Z}_+^m$ . We will denote the probability of an individual living in patch  $j$  to produce an offspring  $\alpha$  after one time step by  $p_S^j(\alpha)$ , i.e.,

$$p_S^j(\alpha) = \text{pr}(\mathbf{h}_{t+1} = \alpha \mid \mathbf{h}_t = \mathbf{e}^j), \quad j = 1, \dots, N.$$

The probability generating function (p.g.f.) which characterizes the birth-death process is an  $N$ -dimensional function  $\mathbf{G}_S(\mathbf{s}) = (G_S^1(\mathbf{s}), \dots, G_S^N(\mathbf{s}))^\top$ , where

$$G_S^j(\mathbf{s}) = \sum_{\alpha} p_S^j(\alpha) \mathbf{s}^\alpha, \quad j = 1, \dots, N,$$

and the summation is extended to the (finite) set of possible values of  $\alpha$ .

We will denote by  $\mathbf{M} \in \mathbb{R}^{N \times N}$  the matrix of expected values associated to the birth-death process. Taking into account that each individual can produce offspring only in its current patch, we have that  $\mathbf{M} = \text{diag}(m_{11}, \dots, m_{NN})$ , where

$$m_{jj} = E(h_{t+1}^j \mid \mathbf{h}_t = \mathbf{e}^j) = \frac{\partial G_S^j}{\partial s_j}(\mathbf{1}),$$

being  $\mathbf{1} = (1, \dots, 1)^\top$ . Note that, according to Hypothesis H2,  $m_{jj} > 0$ , ( $j = 1, \dots, N$ ).

### 3.2. Probability Generating Function of Migration

Migration determines the transference of individuals among the patches and is conservative of the total population size. Since migration is supposed to be subjected to demographic stochasticity, it is modeled through a Markov chain which, from the point of view of branching processes, is a singular g.w.p., i.e., a process with a linear p.g.f. For the sake of coherence with previous works in the field of aggregation in a deterministic context [12], we will denote by  $\mathbf{P} = (p_{ij}) \in \mathbb{R}^{N \times N}$  the matrix of transition probabilities associated to the migration process,  $p_{ij}$  being the probability of an individual in patch  $j$  to migrate to patch  $i$ . This notation has already been used in the context of population dynamics [17].

Let us assume the system is controlled by migration exclusively, and let  $\mathbf{w}_t = (w_t^1, \dots, w_t^N)^\top$  denote the population vector at time  $t$ . Then, given an initial population consisting of a single individual in patch  $j$  ( $j = 1, \dots, N$ ), the probability of obtaining a population  $\alpha$  after a migration period is

$$p_F^j(\alpha) = \text{pr}(\mathbf{w}_{t+1} = \alpha \mid \mathbf{w}_t = \mathbf{e}^j) = \begin{cases} p_{ij}, & \text{if } \alpha = \mathbf{e}^i \text{ for some } i = 1, \dots, N, \\ 0, & \text{otherwise.} \end{cases}$$

The corresponding p.g.f. is  $\mathbf{G}_F(\mathbf{s}) = (G_F^1(\mathbf{s}), \dots, G_F^N(\mathbf{s}))^\top$ , where

$$G_F^j(\mathbf{s}) = \sum_{\alpha} p_F^j(\alpha) \mathbf{s}^\alpha = \sum_{i=1}^N p_F^j(\mathbf{e}^i) s^i = \sum_{i=1}^N p_{ij} s^i, \quad j = 1, \dots, N,$$

and therefore,  $\mathbf{G}_F(\mathbf{s}) = \mathbf{P}^\top \mathbf{s}$ . The matrix of expected values corresponding to migration is  $\mathbf{P}$ , since

$$E(w_{t+1}^i \mid \mathbf{w}_t = \mathbf{e}^j) = \frac{\partial G_F^j}{\partial s_i}(\mathbf{1}) = p_{ij}, \quad i, j = 1, \dots, N.$$

Analogously, the probability that an individual initially in patch  $j$  is in patch  $i$  after  $k$  periods of migration is given by the corresponding entry of the matrix  $\mathbf{P}^k$ , i.e.,

$$\text{pr}(\mathbf{w}_{t+k} = \mathbf{e}^i \mid \mathbf{w}_t = \mathbf{e}^j) = (\mathbf{P}^k)_{ij},$$

and, as a consequence, given an initial population consisting of a single individual in patch  $j$ , the probability of obtaining a population  $\alpha$  after  $k$  periods of migration is

$$p_{F,k}^j(\alpha) = \text{pr}(\mathbf{w}_{t+k} = \alpha \mid \mathbf{w}_t = \mathbf{e}^j) = \begin{cases} (\mathbf{P}^k)_{ij}, & \text{if } \alpha = \mathbf{e}^i, \text{ for some } i=1, \dots, N, \\ 0, & \text{otherwise,} \end{cases} \quad j = 1, \dots, N.$$

Hence, the p.g.f. corresponding to  $k$  periods of migration is  $\mathbf{G}_{F,k}(\mathbf{s}) = (G_{F,k}^1(\mathbf{s}), \dots, G_{F,k}^N(\mathbf{s}))^\top$ , where

$$G_{F,k}^j(\mathbf{s}) = \sum_{\alpha} p_{F,k}^j(\alpha) \mathbf{s}^\alpha = \sum_{i=1}^N p_{F,k}^j(\mathbf{e}^i) s^i = \sum_{i=1}^N (\mathbf{P}^k)_{ij} s^i, \quad j = 1, \dots, N,$$

and therefore,

$$\mathbf{G}_{F,k}(\mathbf{s}) = (\mathbf{P}^k)^\top \mathbf{s}. \quad (2)$$

As above, differentiating these functions, we have

$$E(w_{t+k}^i \mid \mathbf{w}_t = \mathbf{e}^j) = \frac{\partial G_{F,k}^j}{\partial s_i}(\mathbf{1}) = (\mathbf{P}^k)_{ij}, \quad i, j = 1, \dots, N,$$

so the matrix of expected values for  $k$  iterations of migration is  $\mathbf{P}^k$ .

### 3.3. Probability Generating Function of the Global Process

As we stated before, we choose the projection interval of the birth-death process as the time step  $\Delta_t$  of the model. Therefore, we need to approximate the effect of migration over a period much longer than its corresponding projection interval. We will assume that during each  $\Delta_t$ , migration operates a number  $k$  of times before the birth-death process acts. In this way,  $k$  can be interpreted as the ratio between the projection intervals corresponding to the growth of the population and to migration. Since these projection intervals are supposed to be very different from each other,  $k$  is a big number which, in addition, we assume to be an integer.

In order to determine the p.g.f. of this ‘‘global’’ model that takes into account growth and migration, we will study the ‘‘composition’’ of Galton-Watson processes. Let us consider a population governed by two different g.w.p.s,  $\mathbf{x}_t$  and  $\mathbf{y}_t$  characterized by p.g.f.s  $\mathbf{G}_X$  and  $\mathbf{G}_Y$ , respectively. We assume that in each time step of the model, the population is first subjected to an iteration of process  $\mathbf{x}_t$  followed by an iteration of process  $\mathbf{y}_t$ . Therefore, we can think of the population as governed by a g.w.p.,  $\mathbf{z}_t$  (with offspring production  $\mathbf{Z}$ ) which can be interpreted as the ‘‘composition’’ of  $\mathbf{x}_t$  and  $\mathbf{y}_t$ . The mathematical description of process  $\mathbf{z}_t$  in the unitype case has been addressed in [21, p. 21]. The following lemma is a straightforward generalization of that result to deal with the multitype case.

LEMMA 1. *Under the assumptions above, we have the following.*

1. *The p.g.f. of process  $\mathbf{z}_t$  verifies  $\mathbf{G}_Z = \mathbf{G}_X \circ \mathbf{G}_Y$ .*
2. *Let the matrices of expected values for the processes  $\mathbf{x}_t$  and  $\mathbf{y}_t$  be  $\mathbf{A}_X$  and  $\mathbf{A}_Y$ , respectively. Then the matrix  $\mathbf{A}_Z$  of expected values for the process  $\mathbf{z}_t$  is  $\mathbf{A}_Z = \mathbf{A}_Y \mathbf{A}_X$ .*

PROOF. The p.g.f.  $\mathbf{G}_Z(\mathbf{s}) = (G_Z^1(\mathbf{s}), \dots, G_Z^N(\mathbf{s}))^\top$  of a process  $\mathbf{z}_t$  has components given by  $G_Z^j(\mathbf{s}) = E[\mathbf{s}^Z \mid \text{parent type } j]$ , where  $\mathbf{Z}$  is a random variable determining the offspring production of a single individual. These expectations can be written as

$$G_Z^j(\mathbf{s}) = \sum_{\alpha} E[\mathbf{s}^Z \mid \mathbf{X} = \alpha] P(\mathbf{X} = \alpha \mid \text{parent type } j),$$

being  $E[\mathbf{s}^Z \mid \mathbf{X} = \alpha] \equiv E[s_1^{Z_1} \dots s_m^{Z_m} \mid X_1 = \alpha_1, \dots, X_m = \alpha_m]$ . Here,  $Z_i$  represents the offspring of type  $i$  produced in a given time step and is defined as the sum of independent random variables  $Z_i = Y_i^{1,1} + \dots + Y_i^{1,\alpha_1} + \dots + Y_i^{m,1} + \dots + Y_i^{m,\alpha_m}$ . Each  $Y_i^{j,l}$  ( $l = 1, \dots, \alpha_l$ ) is a realization of the random variable  $Y_i^j$ , which represents the number of type  $i$  offspring produced by an individual of type  $j$  after an iteration of process  $\mathbf{y}_t$ . Consequently, for each  $i$  and  $j$ , there are  $\alpha_j$  realizations  $\{Y_i^{j,l}, l = 1, \dots, \alpha_j\}$ , corresponding to the  $\alpha_j$  individuals of type  $j$  produced by  $\mathbf{x}_t$ . These realizations are independent and have a common probability distribution given by that of  $Y_i^j$ . Therefore,

$$\begin{aligned} E[\mathbf{s}^Z \mid \mathbf{X} = \alpha] &= E\left[s_1^{Y_1^{1,1} + \dots + Y_1^{1,\alpha_1} + \dots + Y_1^{m,1} + \dots + Y_1^{m,\alpha_m}} \dots s_m^{Y_m^{1,1} + \dots + Y_m^{1,\alpha_1} + \dots + Y_m^{m,1} + \dots + Y_m^{m,\alpha_m}}\right] \\ &= E\left[s_1^{Y_1^1} \dots s_m^{Y_m^1}\right]^{\alpha_1} \dots E\left[s_1^{Y_1^m} \dots s_m^{Y_m^m}\right]^{\alpha_m} = G_Y^1(\mathbf{s})^{\alpha_1} \dots G_Y^m(\mathbf{s})^{\alpha_m}. \end{aligned}$$

As a consequence, we have

$$\begin{aligned} G_Z^j(\mathbf{s}) &= \sum_{\alpha_1, \dots, \alpha_m} (G_Y^1(\mathbf{s}))^{\alpha_1} \dots (G_Y^m(\mathbf{s}))^{\alpha_m} P(\mathbf{X} = \alpha \mid \text{parent type } j) \\ &= E\left[G_Y(\mathbf{s})^{\mathbf{X}} \mid \text{parent type } j\right] = G_{\mathbf{X}}^j(G_Y(\mathbf{s})), \end{aligned}$$

as we wanted to prove. Part 2 follows easily by differentiating in  $\mathbf{G}_Z = \mathbf{G}_X \circ \mathbf{G}_Y$ .  $\blacksquare$

Therefore, the global model proposed above is a multitype branching process  $\mathbf{x}_0, \mathbf{x}_1, \dots$ , that can be considered as the composition of  $k$  iterations of migration followed by one iteration of the birth-death process. Consequently, its p.g.f. is given by

$$\mathbf{G}_k(\mathbf{s}) = \mathbf{G}_{F,k} \circ \mathbf{G}_S(\mathbf{s}). \quad (3)$$

Now, using (2), we have

$$\mathbf{G}_k(\mathbf{s}) = (\mathbf{P}^k)^\top \mathbf{G}_S(\mathbf{s}), \quad (4)$$

i.e.,  $\mathbf{G}_k(\mathbf{s}) = (G_k^1(\mathbf{s}), \dots, G_k^N(\mathbf{s}))^\top$ , where

$$G_k^j(\mathbf{s}) = \sum_{i=1}^N G_S^i(\mathbf{s}) (\mathbf{P}^k)_{ij}. \quad (5)$$

Hence, we have

$$G_k^j(\mathbf{s}) = \sum_{i=1}^N \left( \sum_{\alpha} p_S^i(\alpha) \mathbf{s}^\alpha \right) (\mathbf{P}^k)_{ij} = \sum_{\alpha} \left( \sum_{i=1}^N p_S^i(\alpha) \alpha (\mathbf{P}^k)_{ij} \right) \mathbf{s}^\alpha,$$

i.e.,  $G_k^j(\mathbf{s})$  can also be expressed as  $G_k^j(\mathbf{s}) = \sum_{\alpha} p_k^j(\alpha) \mathbf{s}^\alpha$ , where  $p_k^j(\alpha)$  is the probability of an individual in patch  $j$  to produce an offspring  $\alpha$  after an iteration of the global process and has the form

$$p_k^j(\alpha) = \text{pr}(\mathbf{x}_{t+1} = \alpha \mid \mathbf{x}_t = \mathbf{e}^j) = \sum_{i=1}^N p_S^i(\alpha) (\mathbf{P}^k)_{ij}. \quad (6)$$

From Lemma 1, we have that the matrix of expected values of the global process is  $\mathbf{M}\mathbf{P}^k$ .

## 4. AGGREGATION OF THE GLOBAL MODEL

In this section, we will obtain a simple unitype Galton-Watson branching process which we will refer to as “aggregated process” that reflects, in a certain way, the effects of both the birth-death process and migration. In this way, we extend the work presented in [12] which deals with the reduction of deterministic linear models in which there are two time scales involved.

When the reduction of a system to a simpler one can be done in such a way that the dynamics of the complex system can be known exactly in terms of the dynamics of the simpler system, the reduction process is called perfect aggregation [3,24]. Biological systems can be perfectly aggregated only in some cases and for very particular values of the parameters involved, so perfect aggregation has only a theoretical interest. In particular, a multitype g.w.p. can be perfectly reduced to a unitype g.w.p. only if all the types have the same offspring distribution.

Therefore, in the general case, it is necessary to resort to approximate aggregation [4], in which some approximations need to be introduced to build the reduced system.

The aim of this section is to apply these approximate aggregation techniques to the original model (4). In order to do so, it is necessary that the Markov chain corresponding to migration approaches an equilibrium in each time step of the model, in the sense that the probability distribution of the individuals among the patches tends to fixed values. This requirement is met with the following hypothesis.

(H3)  $\mathbf{P}$  is primitive, i.e., there exists a positive integer  $h$  such that  $\mathbf{P}^h$  is a positive matrix.

Under this assumption, migration will tend to a stationary distribution given by  $\mathbf{v}$ , where  $\mathbf{v}$  is the right probability normed Perron-Frobenius eigenvector of  $\mathbf{P}$ , i.e.,  $\mathbf{v}$  is the unique positive vector verifying  $\mathbf{P}\mathbf{v} = \mathbf{v}$  and  $\mathbf{1}^\top \mathbf{v} = 1$  [25]. Note that this equilibrium distribution is independent of the initial state, so migration is a strongly ergodic process [26].

REMARK 1. Vector  $\mathbf{v}$  may be interpreted in terms of the behavior of the migration process. Consider the hypothetical situation in which the system was governed by the migration process exclusively, and assume that  $\Delta_t$  is long enough with respect to the projection interval corresponding to migration for this to reach its equilibrium distribution during  $\Delta_t$ . Then, for a population consisting of one individual in patch  $j$  at time  $t$ , we have that, at the end of  $\Delta_t$ , the probability of that individual to live in patch  $i$  will be  $v_i$  (note that  $v_i$  is independent of  $j$ ).

In the sequel, we will refer to this situation of equilibrium for migration as “equilibrium migration process”. We can think of it as the result of applying an infinity of times the migration process on the population. Its p.g.f. is given by  $\mathbf{G}_{\bar{F}}(\mathbf{s}) = \lim_{k \rightarrow \infty} \mathbf{G}_{F,k}(\mathbf{s})$ .

On the other hand, since  $\mathbf{P}$  is a primitive column-stochastic matrix [25], we have that  $\lim_{k \rightarrow \infty} \mathbf{P}^k = \mathbf{v}\mathbf{1}^\top \equiv \bar{\mathbf{P}}$ , and therefore, taking limits on (2), we have

$$\mathbf{G}_{\bar{F}}(\mathbf{s}) = \bar{\mathbf{P}}^\top \mathbf{s},$$

i.e.,  $\mathbf{G}_{\bar{F}}(\mathbf{s}) = (G_{\bar{F}}^1(\mathbf{s}), \dots, G_{\bar{F}}^N(\mathbf{s}))^\top$ , where

$$G_{\bar{F}}^j(\mathbf{s}) = \lim_{k \rightarrow \infty} G_{F,k}^j(\mathbf{s}) = \mathbf{v}^\top \mathbf{s},$$

for all values of  $j$ .

Reasoning as in the previous section, we deduce that the matrix of expected values of the equilibrium migration process is  $\bar{\mathbf{P}}$ .

### 4.1. Auxiliary Model

In order to obtain the aggregated model, we will first introduce a so-called auxiliary system, which can be thought of as the original model when the migration process reaches equilibrium in each time step of the model, i.e., it can be seen as the composition of the equilibrium migration process and the birth-death process.



Let  $\mathbf{x}'_t = (x'_t{}^1, \dots, x'_t{}^N)^\top$  be the population vector associated to the auxiliary model. Using Lemma 1, we obtain its probability generating function

$$\mathbf{G}'(\mathbf{s}) = \mathbf{G}_{\bar{F}} \circ \mathbf{G}_S(\mathbf{s}) = \bar{\mathbf{P}}^\top \mathbf{G}_S(\mathbf{s}), \quad (7)$$

i.e.,  $\mathbf{G}'(\mathbf{s}) = (G'^1(\mathbf{s}), \dots, G'^N(\mathbf{s}))^\top$ , where

$$G'^j(\mathbf{s}) = \mathbf{v}^\top \mathbf{G}_S(\mathbf{s}) = \sum_{i=1}^N G_S^i(\mathbf{s}) v_i, \quad j = 1, \dots, N. \quad (8)$$

Note that  $G'^j(\mathbf{s})$  is independent of  $j$ , and therefore, we can write

$$\mathbf{G}'(\mathbf{s}) = (G'(\mathbf{s}), \dots, G'(\mathbf{s}))^\top,$$

where

$$G'(\mathbf{s}) = G'^j(\mathbf{s}), \quad j = 1, \dots, N.$$

We have  $G'(\mathbf{s}) = \sum_{\alpha} p'(\alpha) \mathbf{s}^\alpha$ , where  $p'(\alpha) = \text{pr}(\mathbf{x}'_{t+1} = \alpha \mid \mathbf{x}'_t = \mathbf{e}^j)$  is the probability of an individual (whatever its patch) to produce an offspring  $\alpha$  after an iteration of the auxiliary process. From (8), we obtain

$$p'(\alpha) = \sum_{i=1}^N p_S^i(\alpha) v_i. \quad (9)$$

Again, using Lemma 1, we deduce that the matrix of expected values associated to the auxiliary process is  $\mathbf{M}\bar{\mathbf{P}}$ .

**REMARK 2.** Since all the component functions of  $\mathbf{G}'(\mathbf{s})$  are identical, the information provided by the auxiliary model is redundant. In this way, for any  $t$ , the distribution of  $\mathbf{x}'_t$  does not depend on the spatial distribution of the original population. As we will see, a consequence of this is that the auxiliary model can be perfectly aggregated.

## 4.2. Aggregated Model

The aggregated model is a unitype Galton-Watson branching process for the total size of a population in the auxiliary process or, in other words, for the total size of the population of the global model under the assumption that migration reaches its equilibrium distribution in each time step.

The variables associated to this model have the form

$$y_t = \mathbf{1}^\top \mathbf{x}'_t = x'_t{}^1 + \dots + x'_t{}^N, \quad (10)$$

and its p.g.f. is a scalar function  $\bar{G}(s) = \sum_{l=0}^{\infty} \bar{p}(l) s^l$ , where  $\bar{p}(l) = \text{pr}(y_{t+1} = l \mid y_t = 1)$  and, according to the assumptions of the model, there is only a finite number of nonzero coefficients  $\bar{p}(l)$ .

The following proposition allows us to obtain the p.g.f. and the asymptotic growth rate of the mean population size for the aggregated model.

**PROPOSITION 1.** *The p.g.f. of the aggregated process is given by*

$$\bar{G}(s) = G' \left( s, s, \binom{N}{\cdot}, s \right) = \sum_{i=1}^N v_i G_S^i \left( s, s, \binom{N}{\cdot}, s \right).$$

Moreover, the asymptotic growth rate of the mean population size for the aggregated system is  $\lambda = \mathbf{1}^\top \mathbf{M}\mathbf{v}$ .

PROOF. The p.g.f. of the aggregated process is

$$\begin{aligned}\bar{G}(s) &= \sum_{l=1}^{\infty} \bar{p}(l) s^l = \sum_{l=1}^{\infty} \text{pr}(y_{t+1} = l \mid y_t = 1) s^l \\ &= \sum_{l=1}^{\infty} \text{pr} \left( \sum_{i=1}^N x'_{t+1}{}^i = l \mid \bigcup_{j=1}^N (\mathbf{x}'_t = \mathbf{e}^j) \right) s^l,\end{aligned}$$

where we have used that  $y_t = x_t^{\prime 1} + \dots + x_t^{\prime N}$ . The distribution of  $\mathbf{x}'_{t+1}$  conditioned to  $(\mathbf{x}'_t = \mathbf{e}^j)$  does not depend on  $j$ , so we have

$$\bar{G}(s) = \sum_{l=1}^{\infty} \text{pr} \left( \sum_{i=1}^N x'_{t+1}{}^i = l \mid \mathbf{x}'_t = \mathbf{e}^j \right) s^l, \quad (11)$$

for any value of  $j$  ( $\star^1$ ). On the other hand, the event  $(\sum_{i=1}^N x'_{t+1}{}^i = l)$  can be expressed as the union of the events  $(\mathbf{x}'_{t+1} = \boldsymbol{\alpha})$ , where  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N)$  and  $\alpha_1 + \dots + \alpha_N = l$ , i.e.,

$$\begin{aligned}\text{pr} \left( \sum_{i=1}^N x'_{t+1}{}^i = l \mid \mathbf{x}'_t = \mathbf{e}^j \right) &= \text{pr} \left( \bigcup_{\alpha_1 + \dots + \alpha_N = l} (\mathbf{x}'_{t+1} = (\alpha_1, \dots, \alpha_N)^\top) \mid \mathbf{x}'_t = \mathbf{e}^j \right) \\ &= \sum_{\alpha_1 + \dots + \alpha_N = l} \text{pr} (\mathbf{x}'_{t+1} = (\alpha_1, \dots, \alpha_N)^\top \mid \mathbf{x}'_t = \mathbf{e}^j),\end{aligned}$$

where we have used that the previous events are mutually exclusive. Coming back to (11), we can write

$$\begin{aligned}\bar{G}(s) &= \sum_{l=1}^{\infty} \left( \sum_{\alpha_1 + \dots + \alpha_N = l} \text{pr} (\mathbf{x}'_{t+1} = (\alpha_1, \dots, \alpha_N)^\top \mid \mathbf{x}'_t = \mathbf{e}^j) \right) s^l \\ &= \sum_{l=1}^{\infty} \sum_{\alpha_1 + \dots + \alpha_N = l} \text{pr} (\mathbf{x}'_{t+1} = (\alpha_1, \dots, \alpha_N)^\top \mid \mathbf{x}'_t = \mathbf{e}^j) s^l \\ &= \sum_{\alpha_1, \dots, \alpha_N} \text{pr} (\mathbf{x}'_{t+1} = (\alpha_1, \dots, \alpha_N)^\top \mid \mathbf{x}'_t = \mathbf{e}^j) s^{\alpha_1 + \dots + \alpha_N} \\ &= \sum_{\boldsymbol{\alpha}} \text{pr} (\mathbf{x}'_{t+1} = \boldsymbol{\alpha} \mid \mathbf{x}'_t = \mathbf{e}^j) s^{\alpha_1} \dots s^{\alpha_N} = \sum_{\boldsymbol{\alpha}} p'(\boldsymbol{\alpha}) s^{\alpha_1} \dots s^{\alpha_N} = G' \left( \overset{(N)}{s, \dots, s} \right),\end{aligned}$$

as we wanted to prove.

Let us now show that the asymptotic growth rate  $\lambda$  of the mean population size is  $\mathbf{1}^\top \mathbf{M} \mathbf{v}$ . We have, by definition,  $\lambda = E[y_{t+1} \mid y_t = 1]$ . Reasoning as in ( $\star^1$ ), we obtain  $\lambda = E[x'_{t+1}{}^1 + \dots + x'_{t+1}{}^N \mid \mathbf{x}'_t = \mathbf{e}^j] = \sum_{i=1}^N E[x'_{t+1}{}^i \mid \mathbf{x}'_t = \mathbf{e}^j]$  for any value of  $j$ . For each  $i, j = 1, \dots, N$ , we have  $E[x'_{t+1}{}^i \mid \mathbf{x}'_t = \mathbf{e}^j] = (\mathbf{M} \bar{\mathbf{P}})_{ij} = (\mathbf{M} \mathbf{v})_i$  and then,  $\lambda = \sum_{i=1}^N (\mathbf{M} \mathbf{v})_i = \mathbf{1}^\top \mathbf{M} \mathbf{v}$ .  $\blacksquare$

REMARK 3. Notice that  $\lambda$  is a weighted mean of the growth rates associated to the different patches being the weights the equilibrium probabilities for migration given by vector  $\mathbf{v}$ .

## 5. RELATIONSHIPS BETWEEN THE GLOBAL MODEL AND THE AGGREGATED MODEL

In the previous section, we have constructed a unitype process from a multitype branching process with  $N$  types. Now we explore the relationships between both models and, in particular, the way to obtain information about the original global model through the study of this new aggregated system.

We will proceed as follows. The behavior of the auxiliary process is deduced from that of the aggregated process with exactitude, and then the characteristics of the global process are studied considering it as a perturbation of the auxiliary process.

The following proposition summarizes the spectral characteristics of the matrix  $\mathbf{M}\bar{\mathbf{P}}$  associated to the expected values of the auxiliary process.

**PROPOSITION 2.** *The only nonzero eigenvalue of  $\mathbf{M}\bar{\mathbf{P}}$  is  $\lambda = \mathbf{1}^\top \mathbf{M}\mathbf{v}$ . It is associated to positive right and left eigenvectors  $\mathbf{r} = \mathbf{M}\mathbf{v}$  and  $\mathbf{l} = \mathbf{1}$ , respectively.*

**PROOF.** The rank of this matrix is clearly 1, because  $\mathbf{M}\bar{\mathbf{P}} = \mathbf{M}\mathbf{v}\mathbf{1}^\top$ , where  $\mathbf{M}\mathbf{v} > \mathbf{0}$ , so it has a unique nonzero eigenvalue. Now,  $(\mathbf{M}\bar{\mathbf{P}})\mathbf{r} = (\mathbf{M}\mathbf{v}\mathbf{1}^\top)(\mathbf{M}\mathbf{v}) = (\mathbf{M}\mathbf{v})(\mathbf{1}^\top \mathbf{M}\mathbf{v}) = \mathbf{r}\lambda$  and  $\mathbf{l}^\top(\mathbf{M}\bar{\mathbf{P}}) = \mathbf{1}^\top(\mathbf{M}\mathbf{v}\mathbf{1}^\top) = (\mathbf{1}^\top \mathbf{M}\mathbf{v})\mathbf{1}^\top = \lambda\mathbf{l}^\top$ . ■

Therefore, the asymptotic growth rate of the mean population for the auxiliary model coincides with that of the aggregated model.

Let the eigenvalues of  $\mathbf{P}$ , ordered by decreasing modulus, be

$$1 = \mu_1 > |\mu_2| \geq \dots \geq |\mu_N|$$

(recall that  $\mathbf{P}$  is primitive, and therefore, 1 is a strictly dominant eigenvalue).

**LEMMA 2.** *Let  $\gamma$  be any number such that  $|\mu_2| < \gamma < 1$ . Then we have  $\mathbf{P}^k = \bar{\mathbf{P}} + \mathbf{o}(\gamma^k)$ ;  $k \rightarrow \infty$ .*

**PROOF.** Since  $\mathbf{P}$  is column-stochastic, it has a strictly dominant eigenvalue 1 which is associated to positive right and left eigenvectors,  $\mathbf{v}$  and  $\mathbf{1}$ , respectively. Let us consider a Jordan canonical decomposition of  $\mathbf{P}$ ,

$$\mathbf{P} = (\mathbf{v} \mid \mathbf{V}) \text{diag}(1, \mathbf{H}) \begin{pmatrix} \mathbf{1}^\top \\ \mathbf{U} \end{pmatrix},$$

where  $\mathbf{V}$  and  $\mathbf{U}$  are appropriate matrices and  $\mathbf{H}$  corresponds to Jordan blocks associated to eigenvalues  $\mu_2, \dots, \mu_N$  (of modulus strictly less than  $\gamma$ ). Therefore, taking into account that  $\bar{\mathbf{P}} = \mathbf{v}\mathbf{1}^\top$ , we have

$$\mathbf{P}^k = \bar{\mathbf{P}} + (\mathbf{v} \mid \mathbf{V}) \text{diag}(\mathbf{0}, \mathbf{H}^k) \begin{pmatrix} \mathbf{1}^\top \\ \mathbf{U} \end{pmatrix},$$

so

$$\frac{\mathbf{P}^k - \bar{\mathbf{P}}}{\gamma^k} = (\mathbf{v} \mid \mathbf{V}) \text{diag}\left(\mathbf{0}, \left(\frac{\mathbf{H}}{\gamma}\right)^k\right) \begin{pmatrix} \mathbf{1}^\top \\ \mathbf{U} \end{pmatrix},$$

and taking limits  $k \rightarrow \infty$ , the desired result follows. ■

In order to obtain the asymptotic behavior of the mean population for the global model, we will make use of perturbation theory to relate the dominant spectral elements of matrices  $\mathbf{M}\bar{\mathbf{P}}$  and  $\mathbf{M}\mathbf{P}^k$ .

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

**PROPOSITION 3.** *For big enough  $k$ , matrix  $\mathbf{M}\mathbf{P}^k$  has a simple and strictly dominant eigenvalue  $\lambda_k$  that can be expressed in the form*

$$\lambda_k = \mathbf{1}^\top \mathbf{M}\mathbf{v} + \mathbf{o}(\gamma^k).$$

*Besides, associated to  $\lambda_k$ , there are right and left eigenvectors  $\mathbf{r}_k$  and  $\mathbf{l}_k$ , respectively, that can be written in the form*

$$\mathbf{r}_k = \mathbf{M}\mathbf{v} + \mathbf{o}(\gamma^k), \quad \mathbf{l}_k = \mathbf{1} + \mathbf{o}(\gamma^k).$$

**PROOF.** Let  $\lambda$  be a simple and strictly dominant eigenvalue of a matrix  $\mathbf{A}$  with associated right and left eigenvectors  $\mathbf{x}_r$  and  $\mathbf{x}_l$ , respectively. Let  $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{E}$  be a perturbation of matrix  $\mathbf{A}$ . In Chapter V of [27], it is shown that  $\tilde{\mathbf{A}}$  has a simple and strictly dominant eigenvalue in the

form  $\tilde{\lambda} = \lambda + O(\|\mathbf{E}\|)$ . Moreover, there exist right and left eigenvectors  $\tilde{\mathbf{x}}_r$  and  $\tilde{\mathbf{x}}_l$ , respectively, associated to  $\tilde{\lambda}$  such that

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

Let  $\mathbf{A} = \mathbf{M}\tilde{\mathbf{P}}$ . Then, from Proposition 2, its dominant eigenvalue is  $\lambda = \mathbf{1}^\top \mathbf{M}\mathbf{v}$ , which has associated eigenvectors  $\mathbf{r} = \mathbf{M}\mathbf{v}$  and  $\mathbf{l} = \mathbf{1}$ . Matrix  $\mathbf{M}\mathbf{P}^k$  can be considered as a perturbation of  $\mathbf{M}\tilde{\mathbf{P}}$ . Then  $\mathbf{E} = \mathbf{M}(\mathbf{P}^k - \tilde{\mathbf{P}})$  and, from Lemma 2, we have  $\|\mathbf{E}\| = o(\gamma^k)$  and the result follows. ■

As a consequence of this result, the asymptotic behavior of the mean population vector for the global process given an initial population  $\mathbf{x}_0$  is characterized by

$$\lim_{t \rightarrow \infty} \frac{E(\mathbf{x}_t)}{(\lambda_k)^t} = \frac{\mathbf{l}_k^\top \mathbf{x}_0}{\mathbf{l}_k^\top \mathbf{r}_k} \mathbf{r}_k = \frac{\mathbf{1}^\top \mathbf{x}_0}{\mathbf{1}^\top \mathbf{M}\mathbf{v}} \mathbf{M}\mathbf{v} + o(\gamma^k).$$

Let us now focus on the relationships between the probability of extinction in finite time for the different systems. Given an initial population consisting of a single individual in patch  $j$  ( $j = 1, \dots, N$ ), the probability to be extinct at time  $t$  in the global process is defined as  $q_k^j(t) := \text{pr}(\mathbf{x}_t = \mathbf{0} \mid \mathbf{x}_0 = \mathbf{e}^j)$ . Analogously, the corresponding probabilities for the auxiliary and the aggregated models are defined as  $q'(t) := \text{pr}(\mathbf{x}'_t = \mathbf{0} \mid \mathbf{x}'_0 = \mathbf{e}^j)$  (independent of  $j$ ) and  $\bar{q}(t) := \text{pr}(y_t = 0 \mid y_0 = 1)$ , respectively. From (10), it is clear that  $y_t = 0$  if and only if  $\mathbf{x}'_t = \mathbf{0}$ , and so

$$q'(t) = \bar{q}(t), \quad (12)$$

for all  $t$ .

In order to relate  $q_k^j(t)$  and  $q'(t)$ , we will make use of the following result, which will be also used to relate the probabilities of ultimate extinction for the three systems.

**PROPOSITION 4.** *Given an initial population consisting of a single individual in patch  $j$ , we have the following for any vector  $\boldsymbol{\alpha} \in Z_+^N$ .*

1.  $p_k^j(\boldsymbol{\alpha}) = p'(\boldsymbol{\alpha}) + o(\gamma^k)$ .
2.  $\mathbf{G}_k(s) = G'(s) + o(\gamma^k)$  uniformly in  $\mathbf{0} \leq \mathbf{s} \leq \mathbf{1}$ .
3. For a fixed time  $t$ ,  $\text{pr}(\mathbf{x}_t = \boldsymbol{\alpha} \mid \mathbf{x}_0 = \mathbf{e}^j) = \text{pr}(\mathbf{x}'_t = \boldsymbol{\alpha} \mid \mathbf{x}'_0 = \mathbf{e}^j) + o(\gamma^k)$ .

**PROOF.**

- (1) A direct consequence of (6), (9), and Lemma 2, taking into account that  $(\tilde{\mathbf{P}})_{ij} = v_i$ .
- (2) Let  $\|\cdot\|$  denote any matrix norm in  $\mathbb{R}^{N \times N}$  and also its associated vector norm in  $\mathbb{R}^N$ .  $\mathbf{G}_S(\mathbf{s})$  is a continuous function in  $\mathbb{R}^N$  so its norm is bounded in the compact set  $\mathbf{0} \leq \mathbf{s} \leq \mathbf{1}$  by, say,  $K$ . Then from (4) and (7), we have

$$\begin{aligned} \frac{\|\mathbf{G}_k(\mathbf{s}) - \mathbf{G}'(\mathbf{s})\|}{\gamma^k} &= \frac{\|(\mathbf{P}^k - \tilde{\mathbf{P}})^\top \mathbf{G}_S(\mathbf{s})\|}{\gamma^k} = \frac{\|\mathbf{G}_S(\mathbf{s})^\top (\mathbf{P}^k - \tilde{\mathbf{P}})\|}{\gamma^k} \\ &\leq \frac{\|\mathbf{G}_S(\mathbf{s})\| \|\mathbf{P}^k - \tilde{\mathbf{P}}\|}{\gamma^k} \leq K \frac{\|\mathbf{P}^k - \tilde{\mathbf{P}}\|}{\gamma^k}, \end{aligned}$$

and the result follows by taking limits and using Lemma 2.

- (3) We denote by  $\mathbf{G}_{t,k}(\mathbf{s})$ , the p.g.f. of the process at time  $t$ . Using (2), we have

$$\mathbf{G}_{t,k}(\mathbf{s}) = \mathbf{G}_k \circ \dots \circ \mathbf{G}_k^{(t)}(\mathbf{s}) = \mathbf{G}_k \circ \dots \circ \mathbf{G}_k^{(t-1)} \left[ \mathbf{G}'(\mathbf{s}) + o(\gamma^k) \right].$$

Since  $\mathbf{G}_k$  is differentiable in any point, we have, applying Taylor's theorem around  $\mathbf{G}'(\mathbf{s})$ , that

$$\mathbf{G}_k [\mathbf{G}'(\mathbf{s}) + \mathbf{o}(\gamma^k)] = \mathbf{G}_k [\mathbf{G}'(\mathbf{s})] + \mathbf{J}_k(\mathbf{c}) \mathbf{o}(\gamma^k),$$

$\mathbf{c}$  being a point in the segment joining  $\mathbf{G}'(\mathbf{s})$  and  $\mathbf{G}'(\mathbf{s}) + \mathbf{o}(\gamma^k)$  and  $\mathbf{J}_k(\mathbf{c})$  the Jacobian matrix of  $\mathbf{G}_k$  evaluated in  $\mathbf{c}$ . Since  $\mathbf{G}_k$  is a  $C^\infty$  function in  $\mathbb{R}^N$ , its derivatives are bounded in any compact set, so  $\mathbf{o}(\gamma^k)\mathbf{J}_k(\mathbf{c}) = \mathbf{o}(\gamma^k)$ . Now

$$\mathbf{G}_{t,k}(\mathbf{s}) = \mathbf{G}_k \circ \overset{(t-1)}{\dots} \circ \mathbf{G}_k \circ [\mathbf{G}'(\mathbf{s}) + \mathbf{o}(\gamma^k)] = \mathbf{G}_k \circ \overset{(t-2)}{\dots} \circ \mathbf{G}_k \circ [\mathbf{G}_k [\mathbf{G}'(\mathbf{s})] + \mathbf{o}(\gamma^k)].$$

But, from (2), we have  $\mathbf{G}_k[\mathbf{G}'(\mathbf{s})] + \mathbf{o}(\gamma^k) = \mathbf{G}'[\mathbf{G}'(\mathbf{s})] + \mathbf{o}(\gamma^k) + \mathbf{o}(\gamma^k) = \mathbf{G}' \circ \mathbf{G}'(\mathbf{s}) + \mathbf{o}(\gamma^k)$ , so

$$\mathbf{G}_{t,k}(\mathbf{s}) = \mathbf{G}_k \circ \overset{(t-2)}{\dots} \circ \mathbf{G}_k \circ [\mathbf{G}' \circ \mathbf{G}'(\mathbf{s}) + \mathbf{o}(\gamma^k)],$$

and iterating this reasoning, we obtain

$$\mathbf{G}_{t,k}(\mathbf{s}) = \mathbf{G}' \circ \overset{(t)}{\dots} \circ \mathbf{G}'(\mathbf{s}) + \mathbf{o}(\gamma^k) = \mathbf{G}'_t(\mathbf{s}) + \mathbf{o}(\gamma^k),$$

for all  $\mathbf{0} \leq \mathbf{s} \leq \mathbf{1}$ .

Finally, we have that  $\text{pr}(\mathbf{x}_t = \boldsymbol{\alpha} \mid \mathbf{x}_0 = \mathbf{e}^j)$  and  $\text{pr}(\mathbf{x}'_t = \boldsymbol{\alpha} \mid \mathbf{x}'_0 = \mathbf{e}^j)$  are the coefficients corresponding to  $\mathbf{s}^\alpha$  in  $\mathbf{G}_{t,k}^j(\mathbf{s})$  and  $\mathbf{G}'_t(\mathbf{s})$ , respectively. The result follows taking into account that two polynomials which coincide in an infinite set must have the same coefficients. ■

The next proposition relates the probabilities of extinction in the global and aggregated models.

**PROPOSITION 5.** *Given an initial population consisting of a single individual in patch  $j$  ( $j = 1, \dots, N$ ), and any  $t$  fixed, we have  $q_k^j(t) = \bar{q}(t) + \mathbf{o}(\gamma^k)$ ;  $k \rightarrow \infty$ .*

**PROOF.** Straightforward from Proposition 4 (Part 3) in the case  $\boldsymbol{\alpha} = \mathbf{0}$  and the fact that  $q'(t) = \bar{q}(t)$ . ■

Given an initial population that consists of a single individual in patch  $j$ , the probabilities of ultimate extinction for the three processes are, respectively,

$$q_k^j = \lim_{t \rightarrow \infty} q_k^j(t), \quad q' = \lim_{t \rightarrow \infty} q'(t), \quad \bar{q} = \lim_{t \rightarrow \infty} \bar{q}(t).$$

Taking limits in (12), it follows that  $q' = \bar{q}$ .

Let us recall that a multitype process is said to be positively regular if its matrix of expected values is primitive. Besides, it is subcritical, critical, or supercritical depending on the asymptotic growth rate of the mean population being, respectively, less, equal, or greater than one. In critical and subcritical positively regular processes, the probability of ultimate extinction for an initial population of any type is always one. In a supercritical process, this probability is strictly lower than one.

The next theorem, which constitutes the main result of this work, relates the probabilities of ultimate extinction for the original and aggregated systems.

**THEOREM 1.**

- (1) *The global process is positively regular for  $k$  big enough.*
- (2) *The original system is subcritical (supercritical) for  $k$  big enough if and only if the aggregated process is subcritical (supercritical).*
- (3) *Moreover, in the supercritical case,*

$$\lim_{k \rightarrow \infty} q_k^j = \bar{q}.$$

PROOF.

- (1) Since  $\mathbf{P}$  is primitive, it follows  $\mathbf{P}^k > \mathbf{0}$  for  $k$  big enough. Since  $m_{ii} > 0$  for all  $i$ , we have that  $\mathbf{M}$  is row-allowable, i.e.,  $\mathbf{M}$  has at least a nonzero element in each one of its rows [25]. It is easy to check that if  $\mathbf{A}$  is row-allowable and  $\mathbf{z}$  is a positive vector, then  $\mathbf{Az}$  is also positive as long as the product is defined. Consequently,  $\mathbf{MP}^k$  is positive and, therefore, primitive.
- (2) It is an immediate consequence of Propositions 1 and 3.
- (3) Since  $\bar{q} = q'$ , we only need to show  $\lim_{k \rightarrow \infty} q_k^j = q'$ . Let us define the vectors of probabilities of extinction for the auxiliary and original systems  $\mathbf{q}' = (q', q', \dots, q')^\top$  and  $\mathbf{q}_k = (q_k^1, q_k^2, \dots, q_k^N)^\top$  and we will show  $\lim_{k \rightarrow \infty} \mathbf{q}_k = \mathbf{q}'$ .

Let  $T$  be the subset of  $\mathbb{R}^N$  given by  $T = [0, 1) \times [0, 1) \times \dots \times [0, 1)$  and let  $\bar{T}$  denote the adherence of  $T$ . Since  $\bar{T}$  is a compact set and  $\mathbf{q}_k \in \bar{T}$  for any value of  $k$ , then the sequence  $\mathbf{q}_k$  must have, at least, one accumulation point  $\mathbf{a} \in \bar{T}$ . Let  $\mathbf{q}_{\phi(k)} = \mathbf{q}_m$  be a subsequence of the  $\mathbf{q}_k$  that converges to  $\mathbf{a}$ , i.e.,  $\lim_{m \rightarrow \infty} \mathbf{q}_m = \mathbf{a}$ .

Now we will use a well-known theorem [28, p. 41] which states that, for a supercritical process, the vector  $\mathbf{q}$  of probabilities of ultimate extinction for the different types, verifies  $\mathbf{0} \leq \mathbf{q} < \mathbf{1}$  and  $\mathbf{G}(\mathbf{q}) = \mathbf{q}$ . Moreover, the unique solutions of this equation in the unit cube  $\mathbf{0} \leq \mathbf{q} \leq \mathbf{1}$  are  $\mathbf{q}$  and  $\mathbf{1}$  (\*<sup>2</sup>).

We know that both the original and the aggregated processes are supercritical for big enough  $m$ . Then we have that  $\mathbf{q}_m < \mathbf{1}$  and  $\mathbf{q}_m = \mathbf{G}_m(\mathbf{q}_m)$ .

Applying Proposition 4, it follows  $\mathbf{q}_m = \mathbf{G}'(\mathbf{q}_m) + \mathbf{o}(\gamma^m)$  which is verified uniformly in  $\mathbf{0} \leq \mathbf{q}_m \leq \mathbf{1}$ . Hence,  $\mathbf{a} = \lim_{m \rightarrow \infty} \mathbf{q}_m = \lim_{m \rightarrow \infty} \mathbf{G}'(\mathbf{q}_m)$  and, since  $\mathbf{G}'$  is a continuous function,  $\mathbf{a} = \mathbf{G}'(\mathbf{a})$ , so  $\mathbf{a}$  is a fixed point for the transformation defined by  $\mathbf{G}'$ . As stated before (\*<sup>2</sup>) the only fixed points of  $\mathbf{G}'$  in  $\bar{T}$  are  $\mathbf{q}'$  and  $\mathbf{1}$ ; we will show that  $\mathbf{a} \neq \mathbf{1}$ .

In the following, let  $\|*\|$  denote the  $\infty$ -norm in  $\mathbb{R}^N$  and assume  $\mathbf{a} = \mathbf{1}$ . We know that  $\mathbf{q}_m = \lim_{t \rightarrow \infty} \mathbf{q}_m(t)$ , so  $\mathbf{a} = \lim_{m \rightarrow \infty} \lim_{t \rightarrow \infty} \mathbf{q}_m(t) = \mathbf{1}$ , i.e.,  $\forall \varepsilon > 0, \exists m_0 \in \mathbb{N}$  such that for all  $m \geq m_0, \exists t_0(m) \in \mathbb{N}$ , such that, for  $t \geq t_0$ ,  $\|\mathbf{1} - \mathbf{q}_m(t)\| < \varepsilon$ . On the other hand,  $\mathbf{q}_m(t) = \mathbf{G}_{t,m}(\mathbf{0}) = \mathbf{G}_m \circ \dots \circ \mathbf{G}_m(\mathbf{0})$  so, given any  $\varepsilon > 0$ , we can find numbers  $m \geq m_0(\varepsilon)$  and  $t \geq t_0(m)$  such that  $\|\mathbf{1} - \mathbf{G}_{t,m}(\mathbf{0})\| < \varepsilon$  (\*<sup>3</sup>).

We know that the Jacobian matrix of  $\mathbf{G}_{t,m}$  evaluated at  $\mathbf{s} = \mathbf{1}$  is  $\mathbf{A}_m^t = (\mathbf{MP}^m)^t$ . Hence, for any  $\mathbf{s} \geq \mathbf{0}, \mathbf{s} \neq \mathbf{0}$ , we have the Taylor expansion  $\mathbf{G}_{t,m}(\mathbf{1} - \mathbf{s}) = \mathbf{1} - \mathbf{A}_m^t \mathbf{s} + \mathbf{o}_{t,m}(\|\mathbf{s}\|)$ , where we have used that, since  $\mathbf{G}_{t,m}$  is a p.g.f., then  $\mathbf{G}_{t,m}(\mathbf{1}) = \mathbf{1}$ . Taking norms in this expression, we have

$$\begin{aligned} \|\mathbf{1} - \mathbf{G}_{t,m}(\mathbf{1} - \mathbf{s})\| &= \|\mathbf{A}_m^t \mathbf{s} - \mathbf{o}_{t,m}(\|\mathbf{s}\|)\| \\ &\geq \|\|\mathbf{A}_m^t \mathbf{s}\| - \|\mathbf{o}_{t,m}(\|\mathbf{s}\|)\|\|. \end{aligned} \quad (13)$$

Let us now study the terms  $\|\mathbf{A}_m^t \mathbf{s}\|$  and  $\|\mathbf{o}_{t,m}(\|\mathbf{s}\|)\|$ .

- (a) Matrix  $\mathbf{A}_m$  can be decomposed as follows:  $\mathbf{A}_m = \mathbf{MP}^m = \mathbf{M}\bar{\mathbf{P}} + \mathbf{M}(\mathbf{P}^m - \bar{\mathbf{P}})$ . From Lemma 2, we know that the term  $\mathbf{M}(\mathbf{P}^m - \bar{\mathbf{P}})$  tends to zero when  $m \rightarrow \infty$ .

Besides, since  $\mathbf{M}$  is row-allowable and  $\bar{\mathbf{P}}$  is positive, we have that  $\mathbf{M}\bar{\mathbf{P}}$  is a positive matrix (and, therefore, primitive). Since the aggregated process is supercritical, so is the auxiliary system (Proposition 2) and so the dominant eigenvalue of  $\mathbf{M}\bar{\mathbf{P}}$  is  $\lambda > 1$ . Now, from Lemma 3 (see the appendix), there exists a number  $t_1 \in \mathbb{N}$  such that, if  $t \geq t_1$ , then  $\|(\mathbf{M}\bar{\mathbf{P}})^t \mathbf{s}\|_\infty > 3\|\mathbf{s}\|_\infty$  for any vector  $\mathbf{s} \geq \mathbf{0}, \mathbf{s} \neq \mathbf{0}$ .

In the remainder of the proof, let  $\tau \geq t_1$  be fixed. Matrix  $\mathbf{A}_m^\tau$  can be expressed as  $\mathbf{A}_m^\tau = (\mathbf{M}\bar{\mathbf{P}})^\tau + \mathbf{B}_{\tau,m}$  where  $\mathbf{B}_{\tau,m}$  is a sum of terms that have the common factor  $\mathbf{M}(\mathbf{P}^m - \bar{\mathbf{P}})$ . Clearly, for a fixed  $\tau$ , we have  $\mathbf{B}_{\tau,m} \xrightarrow{m \rightarrow \infty} \mathbf{0}$ , so there exists a number  $m_1(\tau) \in \mathbb{N}$  such that, if  $m \geq m_1$ ,  $\|\mathbf{B}_{\tau,m}\|_\infty < 1$ . As a consequence,

$$\begin{aligned}\|\mathbf{A}_m^\tau \mathbf{s}\|_\infty &= \|(\mathbf{M}\bar{\mathbf{P}})^\tau \mathbf{s} + \mathbf{B}_{\tau,m} \mathbf{s}\|_\infty \geq \|(\mathbf{M}\bar{\mathbf{P}})^\tau \mathbf{s}\|_\infty - \|\mathbf{B}_{\tau,m} \mathbf{s}\|_\infty \\ &> 3 \|\mathbf{s}\|_\infty - \|\mathbf{s}\|_\infty = 2 \|\mathbf{s}\|_\infty,\end{aligned}$$

for any  $m \geq m_1$  and any  $\mathbf{s} \geq \mathbf{0}$ ,  $\mathbf{s} \neq \mathbf{0}$ .

- (b) We focus on the term  $\mathbf{o}_{\tau,m}(\|\mathbf{s}\|)$ . From Lemma 4 (see the appendix), we have that there exist  $m_2(\tau) \in \mathbb{N}$  and  $\delta(\tau) > 0$  such that, if  $m \geq m_2$  and  $\|\mathbf{s}\| < \delta$ ,  $\mathbf{s} \neq \mathbf{0}$ , then  $(\|\mathbf{o}_{\tau,m}(\|\mathbf{s}\|)\|)/(\|\mathbf{s}\|) < 1$ , i.e.,  $\|\mathbf{o}_{\tau,m}(\|\mathbf{s}\|)\| < \|\mathbf{s}\|$ .

Now let  $m \geq \max\{m_1, m_2\}$  and  $\|\mathbf{s}\| < \delta$ . Then, as a consequence of (a) and (b), we have that  $\|\mathbf{1} - G_{\tau,m}(\mathbf{1} - \mathbf{s})\| > \|\mathbf{s}\|$  (\*4). Let us define  $\mathbf{s} = \mathbf{1} - G_{t,m}(\mathbf{0})$ . From (\*3), if  $m \geq m_0(\delta)$  and  $t \geq t_0(m)$ , we have  $\|\mathbf{s}\| < \delta$  so, choosing  $m = \max\{m_0, m_1, m_2\}$  and replacing  $\mathbf{s}$  in (\*4), we get  $\|\mathbf{1} - G_{\tau,m}(G_{t,m}(\mathbf{0}))\| > \|\mathbf{1} - G_{t,m}(\mathbf{0})\|$ . Recall that  $G_{\tau,m}(G_{t,m}(\mathbf{0})) = G_{\tau+t,m}(\mathbf{0})$  and that  $G_{t,m}(\mathbf{0}) = \mathbf{q}_m(t)$ , so it follows  $\|\mathbf{1} - \mathbf{q}_m(\tau + t)\| > \|\mathbf{1} - \mathbf{q}_m(t)\|$  and, hence, there exists at least one value for  $i = 1, \dots, N$  such that  $\mathbf{q}_m^i(\tau + t) < \mathbf{q}_m^i(t)$ . At this point, we have reached a contradiction because  $\mathbf{q}_m^i(t)$  must be a monotone nondecreasing function of  $t$ . Therefore, we deduce that  $\mathbf{a} \neq \mathbf{1}$  as we wanted to prove and, hence, the sequence  $\mathbf{q}_k$  has a unique accumulation point given by  $\mathbf{a} = \mathbf{q}'$ , i.e.,  $\lim_{k \rightarrow \infty} \mathbf{q}_k = \mathbf{q}'$ . ■

REMARK 4. When the aggregated system is critical, the previous results do not allow one to decide the character (super/sub/critical) of the original process.

In order to illustrate the utility of our aggregation procedure to study the multiregional model, let us make the simplifying assumption that each individual can produce zero, one, or two offspring in each time step. We denote by  $p_S^j(n)$  the probability of an individual living in patch  $j = 1, \dots, N$  to produce  $n = 0, 1, 2$  offspring (necessarily in patch  $j$ ) after one iteration of the birth-death process and define  $\mathbf{p}_S(n) := (p_S^1(n), \dots, p_S^N(n))^\top$ . The p.g.f. of this process is then given by  $G_S^j(s) = p_S^j(0) + p_S^j(1)s_j + p_S^j(2)s_j^2$  ( $j = 1, \dots, N$ ) and the original system (5) has the p.g.f.

$$G_k^j(\mathbf{s}) = \sum_{i=1}^N (p_S^i(0) + p_S^i(1)s_i + p_S^i(2)s_i^2) (\mathbf{P}^k)_{ij}, \quad j = 1, \dots, N.$$

Our reduction procedure renders the aggregated system with p.g.f. and growth rate of the mean population given by

$$\begin{aligned}\bar{G}(\mathbf{s}) &= c_0 + c_1 \mathbf{s} + c_2 \mathbf{s}^2, \\ \lambda &= \sum_{i=1}^N v_i (p_S^i(1) + 2p_S^i(2)) = 1 + (c_2 - c_0),\end{aligned}$$

where

$$c_n = \sum_{i=1}^N v_i p_S^i(n) = \mathbf{v}^\top \mathbf{p}_S(n).$$

Consequently, we have that the aggregated process (and, for  $k$  big enough, also the global process) is subcritical if  $c_0 > c_2$  and supercritical if  $c_0 < c_2$ . In the latter case, the probability of ultimate extinction  $\bar{q}$  for the aggregated system is the unique solution of  $\bar{G}(\bar{q}) = \bar{q}$  in the interval  $[0, 1)$  which is easily seen to be

$$\bar{q} = \frac{c_0}{c_2} = \frac{\mathbf{v}^\top \mathbf{p}_S(0)}{\mathbf{v}^\top \mathbf{p}_S(2)}.$$

Now, we know that the probabilities of ultimate extinction for the original system can be approximated by  $\bar{q}$ . Note that calculating this probability directly from the expression of the original system would imply dealing with a system of  $N$  second-order polynomial equations.

On the other hand, note that the preceding expressions allow one to study in a very simple way the interactions between the birth-death process and migration in relevant parameters of the model such as  $\lambda$  or  $\bar{q}$ . In this way, we can study how changes in these processes affect the behavior of the population.

We have performed some simulations to experimentally estimate the error generated when approximating the growth rate of the original system by means of the aggregated system. Lemma 2 and Proposition 3 suggest the existence of a relationship between the accuracy of the approximations provided by the aggregated system and the quantities  $|\mu_2|$  (the modulus of the subdominant eigenvalue of  $P$ ) and  $k$  (i.e., the number that characterizes the separation between the time scales of the birth-death process and migration).

For  $k = 5$ , we have assigned to  $|\mu_2|$  different values ranging from 0 to 1. In addition, for each selection of  $|\mu_2|$  in that range, we have randomly generated 50 examples of original systems, all of them having one single age class and two patches. For each of these instances, we have built the corresponding aggregated system and we have compared the growth rate of both systems by computing the relative error

$$\frac{|\lambda_k - \lambda|}{\lambda_k} 100\%.$$

Figure 1a represents the maximum relative error obtained in each sample, i.e., for each value of  $|\mu_2|$ . This plot suggests that the resulting relative error in the approximation of  $\lambda_k$  increases when the value of  $|\mu_2|$  grows. As one can see, for  $|\mu_2| = 0.9$ , the maximum relative error is about 45%; however, in 88% of the instances that error was less than 5% and the mean error was about 3.8%.

We have repeated the same computations for  $k = 15$  and  $k = 25$ . The results are shown in Figures 1b and 1c, respectively. Note that for  $k = 25$ , the maximum relative error in the worst case ( $|\mu_2| = 0.9$ ) has been drastically reduced to roughly 5%.

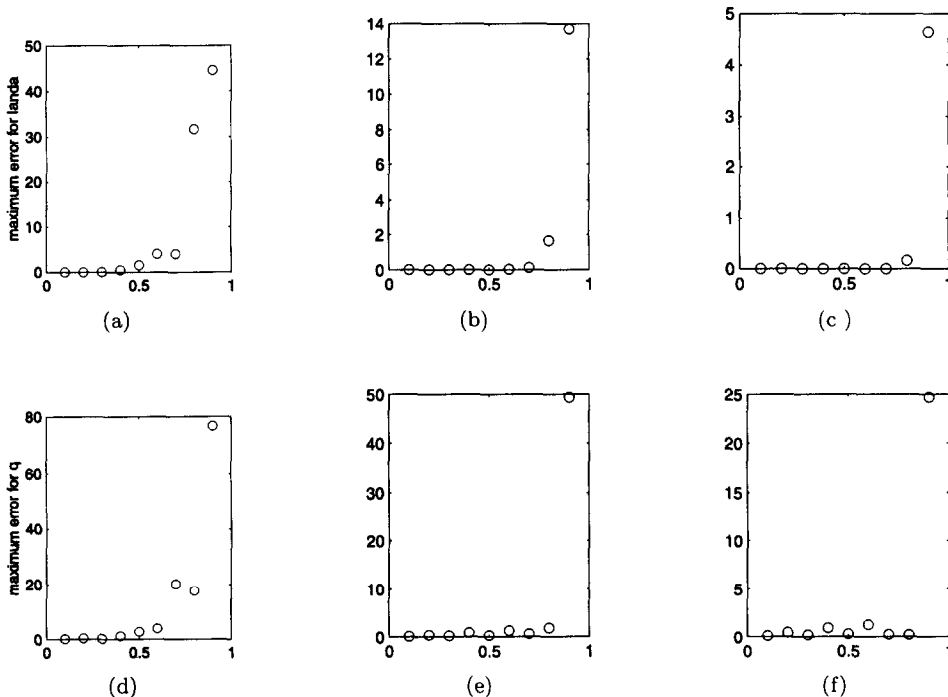


Figure 1. Maximum relative errors obtained in the approximation of the characteristics of the original system by means of the aggregated system. For  $|\mu_2|$  taking values: 0.1, 0.2, ..., 0.9, the maximum was computed over samples of 50 randomly generated systems. Figures 1a–1c compare the growth rate of the aggregated system with that of the original system for different values of  $k$ . Figures 1d–1f do the same regarding the probability of extinction.



Analogous computations on the probability of ultimate extinction yield similar results (see Figures 1d–1f), but now the quality of the approximations decreases moderately. Figure 1f shows that, with  $k = 25$ , the maximum relative error obtained was about 25%. However, again in the 90% of the instances, this error was less than 5%.

These results show numerically that

- (a) the accuracy of the approximations depends directly on the value of  $k$  and inversely on the modulus of the subdominant value of matrix  $P$ ;
- (b) except for very high values of  $|\mu_2|$ , the magnitude of the error we incur does not exceed 10% and so our reduction technique may be useful to study real biological populations.

## 6. CONCLUSION

The aggregation method that we present allows us to simplify the treatment of a significant type of complex stochastic models. We start from a complex setting corresponding to a population distributed among  $N$  spatial patches and subjected to demographic stochasticity. The evolution of the population is controlled by two processes, a birth-death process and a migration process, which take place at different time scales. This situation is modelled by means of a global process resulting from the composition, in a certain sense, of two multitype Galton-Watson branching processes. From this global model, we derive a simpler aggregated model corresponding to a unitype branching process for the total population that approximates the behavior of the global one.

It is shown that we can obtain essential information about the asymptotic behavior of the global model through the study of the aggregated one. In fact, we show that the greater the difference between the time scales corresponding to the two processes involved, the closer some characteristics such as asymptotic growth rate of the mean population size and probability of extinction are for both models.

In further works, we aim to generalize this technique to the study of more complex models such as age or size-structured multiregional models. Moreover, it would be interesting to relate characteristics of the global and reduced models such as the second-order moments or, in the subcritical case, the distribution of population size conditioned to nonextinction.

In addition, our attention is now focused on the development of reduction techniques for some kinds of branching processes that yield more realistic models. In particular, we think of density-dependent models [29] and multitype branching processes in random environments (MBPRE) [30]. This latter kind takes into account the effects of both environmental and demographic stochasticity.

## APPENDIX

LEMMA 3. *Let  $\mathbf{A}$  be a primitive matrix with dominant eigenvalue  $\lambda > 1$ . For any  $M > 0$ , there exists  $t_0(M) \in \mathbb{N}$  such that, if  $t \geq t_0$ , then  $\|\mathbf{A}^t \mathbf{s}\|_\infty > M \|\mathbf{s}\|_\infty$  for any vector  $\mathbf{s} \geq \mathbf{0}$ ,  $\mathbf{s} \neq \mathbf{0}$ .*

PROOF. Let  $m(t) = \min_{i,j} (\mathbf{A}^t)_{ij}$ . Since  $\mathbf{A}$  is primitive and  $\lambda > 1$ , then  $m(t) \xrightarrow{t \rightarrow \infty} \infty$ . In particular, there exists  $t_0 \in \mathbb{N}$  such that, if  $t \geq t_0$ , then  $m(t) > M$ . Now, since  $\mathbf{s} \geq \mathbf{0}$ ,  $\|\mathbf{A}^t \mathbf{s}\|_\infty \geq \|m(t) \mathbf{1}^\top \mathbf{1} \mathbf{s}\|_\infty = m(t) \|\mathbf{s}\|_\infty > M \|\mathbf{s}\|_\infty$ . ■

LEMMA 4. *Let us define  $\mathbf{o}_{t,m}(\|\mathbf{s}\|)$  as in (13). Given a fixed  $t \in \mathbb{N}$ , there exists  $m_0(t) \in \mathbb{N}$  such that*

$$\sup_{m \geq m_0} \frac{\|\mathbf{o}_{t,m}(\|\mathbf{s}\|)\|}{\|\mathbf{s}\|} \xrightarrow{\mathbf{s} \rightarrow \mathbf{0}} 0.$$

PROOF. Let  $A = \{\mathbf{s} : \mathbf{0} \leq \mathbf{s} \leq \mathbf{1}\}$  and let  $t \in \mathbb{N}$  be fixed. From Taylor's theorem,  $\mathbf{o}_{t,m}(\|\mathbf{s}\|) = (o_{t,m}^1(\|\mathbf{s}\|), \dots, o_{t,m}^N(\|\mathbf{s}\|))^\top$ , where  $o_{t,m}^j(\|\mathbf{s}\|) = \mathbf{s}^\top \mathbf{H}_{m,t}^j \mathbf{s}$ ,  $\mathbf{H}_{m,t}^j$  being the Hessian matrix of  $G_{m,t}^i$  evaluated in a certain point of the segment joining  $\mathbf{1}$  and  $\mathbf{1} - \mathbf{s}$  (therefore, belonging to  $A$ ). Hence,

$|\sigma_{t,m}^j(\|\mathbf{s}\|)| = |\mathbf{s}^\top \mathbf{H}_{m,t}^j \mathbf{s}| \leq \|\mathbf{H}_{m,t}^j\| \|\mathbf{s}\|^2 (*5)$ , for each  $j = 1, \dots, N$ . The entries of  $\mathbf{H}_{m,t}^j$  are the second derivatives of  $G_{m,t}^j$ . We will show that, for any  $l, r = 1, \dots, N$ ,

$$\frac{\partial^2 G_{m,t}^j}{\partial s_r \partial s_l}(\mathbf{s}) \xrightarrow{m \rightarrow \infty} \frac{\partial^2 G_t^{l,j}}{\partial s_r \partial s_l}(\mathbf{s}),$$

uniformly in  $\mathbf{s} \in A$ . Indeed, we have, for  $r \neq l$ ,

$$\begin{aligned} \left| \frac{\partial^2 G_{m,t}^j}{\partial s_r \partial s_l}(\mathbf{s}) - \frac{\partial^2 G_t^{l,j}}{\partial s_r \partial s_l}(\mathbf{s}) \right| &\leq \sum_{\alpha=(\alpha_1, \dots, \alpha_N)} \left| p_{m,t}^j(\alpha) - p_t^{l,j}(\alpha) \right| \alpha_l \alpha_r s_1^{\alpha_1} \dots s_l^{\alpha_l - 1} \dots s_r^{\alpha_r - 1} \dots s_N^{\alpha_N} \\ &\leq \sum_{\alpha=(\alpha_1, \dots, \alpha_N)} \left| p_{m,t}^j(\alpha) - p_t^{l,j}(\alpha) \right| \alpha_l \alpha_r, \end{aligned}$$

and Proposition 4 guarantees that this sum tends to zero when  $m \rightarrow \infty$ , so the convergence of the second derivatives is uniform in  $\mathbf{s} \in A$ . Notice that the result also stands for  $l = r$ .

On the other hand,  $G_t^{l,j}$  is  $C^\infty$  in  $\mathbb{R}^N$  and  $A$  is a compact set, so its second derivatives are bounded in  $A$ . Using this and the uniform convergence above, it follows that, for  $m$  big enough, the second-order derivatives of  $G_{m,t}^j$  are bounded in  $A$  uniformly in  $m$ , i.e., there exist values  $m_0(t) \in \mathbb{N}$  and  $C_t > 0$  such that, if  $m \geq m_0$ , then

$$\left| \frac{\partial^2 G_{m,t}^j}{\partial s_r \partial s_l}(\mathbf{s}) \right| \leq C_t,$$

for all  $r, l, j = 1, \dots, N$  and  $\mathbf{s} \in A$ .

Finally, coming back to (\*5), if  $m \geq m_0(t)$ , we have that  $|\sigma_{t,m}^j(\|\mathbf{s}\|)| \leq \|\mathbf{H}_{m,t}^j\| \|\mathbf{s}\|^2 \leq C_t \|\mathbf{s}\|^2$  and, hence,

$$\sup_{m \geq m_0} \frac{|\sigma_{t,m}^j(\|\mathbf{s}\|)|}{\|\mathbf{s}\|} \leq C_t \|\mathbf{s}\| \xrightarrow{\mathbf{s} \rightarrow 0} 0,$$

for every  $j$ , from where the result follows. ■

## REFERENCES

1. P. Auger, *Dynamics and Thermodynamics in Hierarchically Organized Systems, Applications in Physics, Biology and Economics*, Pergamon, Oxford, (1989).
2. H.A. Simon and A. Ando, Aggregation of variables in dynamic systems, *Econometrica* **29**, 111–135 (1961).
3. Y. Iwasa, V. Andreassen and S.A. Levin, Aggregation in model ecosystems I. Perfect aggregation, *Ecol. Modelling* **37**, 287–302 (1987).
4. Y. Iwasa, S.A. Levin and V. Andreassen, Aggregation in model ecosystems II. Approximate aggregation, IMA, *Journal of Mathematics Applied in Medicine and Biology* **6**, 1–23 (1989).
5. P. Auger and R. Roussarie, Complex ecological models with simple dynamics: From individuals to populations, *Acta Biotheoretica* **42**, 111–136 (1994).
6. P. Auger and J.C. Poggiale, Emergence of population growth models: Fast migration and slow growth, *Journal of Theoretical Biology* **182**, 99–108 (1996).
7. P. Auger and J.C. Poggiale, Aggregation and emergence in systems of ordinary differential equations, *Mathl. Comput. Modelling* **27** (4), 1–21 (1998).
8. P. Auger and R. Bravo de la Parra, Methods of aggregation of variables in population dynamics, *C.R. Acad. Sci. Paris, Life Sciences* **323**, 665–674 (2000).
9. R. Bravo de la Parra, O. Arino, E. Sánchez and P. Auger, A model for an age-structured population with two time scales, *Mathl. Comput. Modelling* **31** (4/5), 17–26 (2000).
10. R. Bravo de la Parra and E. Sánchez, Aggregation methods in population dynamics discrete models, *Mathl. Comput. Modelling* **27** (4), 23–39 (1998).
11. R. Bravo de la Parra, E. Sánchez, O. Arino and P. Auger, A discrete model with density dependent fast migration, *Math. Biosc.* **157**, 91–109 (1999).
12. L. Sanz and R. Bravo de la Parra, Variables aggregation in a time discrete linear model, *Math. Biosc.* **157**, 111–146 (1999).

13. A. Blasco, L. Sanz, P. Auger and R. Bravo de la Parra, Linear discrete population models with two time scales in fast changing environments I: Autonomous case, *Acta Biotheoretica* **49** (4), 261–276 (2001).
14. L. Sanz and R. Bravo de la Parra, Variables aggregation in time varying discrete systems, *Acta Biotheoretica* **46** (3), 273–297 (1998).
15. L. Sanz and R. Bravo de la Parra, Time scales in a nonautonomous linear discrete model, *Mathematical Models and Methods in Applied Sciences* **11** (7), 1–33 (2001).
16. S. Tuljapurkar, Population dynamics in variable environments, In *Lecture Notes in Biomathematics*, (Edited by S. Levin), Springer-Verlag, Berlin, (1990).
17. H. Caswell, *Matrix Population Models*, Second Edition, Sinauer Associates, Sunderland, (2001).
18. K.E. Holsinger, Demography and extinction in small populations, In *Genetics, Demography and the Viability of Fragmented Populations*, (Edited by A. Young and G. Clarke), pp. 55–74, Cambridge University Press, Cambridge, (2000).
19. L. Sanz and R. Bravo de la Parra, Time scales in stochastic multiregional models, *Nonlinear Analysis: Real World Applications* **1**, 89–122 (2000).
20. C.J. Mode, Multitype branching processes, In *Modern Analytic and Computational Methods in Science and Mathematics*, (Edited by R. Bellman), American Elsevier, New York, (1971).
21. P. Jagers, *Branching Processes with Biological Applications*, John Wiley & Sons, London, (1975).
22. H. Caswell, M. Fujiwara and S. Brault, Declining survival probability threatens the North Atlantic right whale, *Proc. Natl. Acad. Sci. USA* **96**, 3308–3313 (1999).
23. S. Legendre, J. Clobert, A.P. Moller and 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**, 449–463 (1999).
24. T.C. Gard, Aggregation in stochastic ecosystem models, *Ecol. Modelling* **44**, 153–164 (1988).
25. E. Seneta, *Nonnegative Matrices and Markov Chains*, Springer Verlag, New York, (1981).
26. J.E. Cohen, Ergodic theorems in demography, *Bulletin of the American Mathematical Society N.S.* **1**, 275–295 (1979).
27. G.W. Stewart and J.I-G. Sun, *Matrix Perturbation Theory*, Academic Press, (1990).
28. T.E. Harris, *The Theory of Branching Processes*, Dover, New York, (1963).
29. G.L. Block and L.J.S. Allen, Population extinction and quasi-stationary behavior in stochastic density-dependent structured models, *Bull. Math. Bio.* **62**, 199–228 (2000).
30. J.D. Biggins, H. Cohn and O. Nerman, Multi-type branching in varying environment, *Stoc. Proc. Appl.* **83**, 357–400 (1999).