Individual and patch behaviour in structured metapopulation models

Barbour, A D ; Luczak, M J

Individual and patch behaviour in structured metapopulation models

A. D. Barbour[*] and Malwina Luczak[†]

Universität Zürich and Queen Mary University, London

Abstract

Density dependent Markov population processes with countably many types can often be well approximated over finite time intervals by the solution of the differential equations that describe their average drift, provided that the total population size is large. They also exhibit diffusive stochastic fluctuations on a smaller scale about this deterministic path. Here, it is shown that the individuals in such processes experience an almost deterministic environment. Small groups of individuals behave almost independently of one another, evolving as Markov jump processes, whose transition rates are prescribed functions of time. In the context of metapopulation models, we show that 'individuals' can represent either patches or the individuals that migrate among the patches; in host–parasite systems, they can represent both hosts and parasites.

Keywords: Markov population processes, propagation of chaos,
metapopulation, host parasite systems
AMS subject classification: 92D30, 60J27, 60B12
Running head: Structured metapopulation models

1 Introduction

In a series of papers motivated by models of structured metapopulations (Levins 1969, Hanski & Gilpin 1991) and parasitic disease transmission (Kretzschmar 1993), the authors have extended Kurtz's (1970, 1971) theory to provide laws of large numbers and central limit theorems for Markov population processes with countably many types of individual, together with estimates of the approximation errors: see Barbour & Luczak [BL] (2008, 2012a,b). These theorems provide a good description of the overall behaviour of such processes, when the population size is large. However, as observed by Léonard (1990), many ecological models, when seen from the perspective of the individuals themselves, can be interpreted as interacting particle systems. It is then of interest to be able to describe the behaviour of (small groups of) individuals within the large system. Under very stringent assumptions on the transition rates, in particular requiring that they be uniformly bounded, he proves a 'propagation of chaos' theorem, showing that individuals evolve almost independently of one another, as Markov processes whose transition rates are determined by the bulk behaviour of the system.

In this paper, we establish an analogous result for systems with countably many types, under much less restrictive conditions. We formulate a model that is general enough to encompass many host parasite systems and structured metapopulation models. The main tool used in showing the asymptotic independence of individuals in such processes is to couple the process describing the evolution of individuals in the original system with one in which they evolve independently. The coupling is constructed by matching the transition rates in the two processes, and the argument is described in Section 2.

In order to show that the coupling is close, we rely on the quantitative law of large numbers proved in [BL] (2012a). The conditions needed for the law of large numbers have already been shown to be satisfied for a number of examples from the literature, including the models of Arrigoni (2003), Barbour & Kafetzaki (1993), Kretzschmar (1993) and Luchsinger (2001a,b). However, some work is required to find explicit conditions based on the parameters of our general model under which the law holds; this is accomplished in Section 3. The paper concludes with examples taken from Metz & Gyllenberg (2001) and from Kretzschmar (1993).

2 Main results

We begin by formulating our models in a way which explicitly reflects their origins in metapopulation and parasitic disease modelling. The basic description is in terms of the numbers of patches of each of a countable number of types. The type of a patch is determined by the numbers of animals of each of d different varieties present in the patch, indexed by $\mathbf{i} = (i_1, \ldots, i_d) \in \mathbb{Z}_+^d$. For instance, a patch may represent a host, and its type the numbers of parasites of various different species that it harbours. However, an animal's variety may also indicate its developmental stage, or its infection status, so that its variety may change over its lifetime. We also define d further types, to account for animals of the different varieties that are in transit between patches. Thus the possible patch types are indexed by $\mathcal{Z} := \mathcal{Z}_1 \cup \mathcal{Z}_2$, where $\mathcal{Z}_1 = \mathbb{Z}_+^d$ and $\mathcal{Z}_2 = \{1, \ldots, d\}$. In these terms, the state space is expressed as $\mathcal{X} := \{X \in \mathbb{Z}_+^{\mathcal{Z}}, \sum_{z \in \mathcal{Z}} X_z < \infty\}$. The interpretation is that $X_{\mathbf{i}}$ records the number of patches of type $\mathbf{i}$, $\mathbf{i} \in \mathcal{Z}_1$, whereas X_l, $1 \leq l \leq d$, denotes the number of migrating animals of variety l. The restriction $\sum_{z \in \mathcal{Z}} X_z < \infty$ in the definition of $\mathcal{X}$ constrains total numbers of patches and animals to be finite. Our model for the evolution of the metapopulation consists of a family $X^N := (X^N(t),\ t \geq 0)$ of pure jump Markov processes on $\mathcal{X}$, indexed by $N \in \mathbb{N}$, with N to be thought of as a typical number of patches in the process X^N. Writing $e(z)$ for the z-coordinate vector in $\mathbb{R}_+^{\mathcal{Z}}$, $z \in \mathcal{Z}$, and e_l for the l-th coordinate vector in $\mathbb{Z}^d$, the transition rates for X^N are assumed to be given by

$$
\begin{array}{rlll}
\text{I}: & X \;\to\; X + e(\mathbf{j}) - e(\mathbf{i}) & & \\
& \qquad\qquad \text{at rate} \quad X_{\mathbf{i}}\{\bar{\lambda}_{\mathbf{ij}} + \lambda_{\mathbf{ij}}(x)\}, & \mathbf{i}, \mathbf{j} \in \mathcal{Z}_1; \\[4pt]
\text{II}: & X \;\to\; X + e(\mathbf{i}) \quad \text{at rate} \quad N\beta_{\mathbf{i}}(x), & \mathbf{i} \in \mathcal{Z}_1; \\[4pt]
\text{III}: & X \;\to\; X - e(\mathbf{i}) \quad \text{at rate} \quad X_{\mathbf{i}}\{\bar{\delta}_{\mathbf{i}} + \delta_{\mathbf{i}}(x)\}, & \mathbf{i} \in \mathcal{Z}_1; \\[4pt]
\text{IV}: & X \;\to\; X + e(l) + e(\mathbf{i} - e_l) - e(\mathbf{i}) & & \\
& \qquad\qquad \text{at rate} \quad X_{\mathbf{i}}\{\bar{\gamma}_{\mathbf{i}l} + \gamma_{\mathbf{i}l}(x)\}, & \mathbf{i} \in \mathcal{Z}_1,\ 1 \leq l \leq d; \\[4pt]
\text{IV}': & X \;\to\; X + e(l) \quad \text{at rate} \quad \sum_{\mathbf{j} \in \mathcal{Z}_1} X_{\mathbf{j}}\{\bar{\gamma}'_{\mathbf{j}l} + \gamma'_{\mathbf{i}l}(x)\}, & 1 \leq l \leq d; \\[4pt]
\text{V}: & X \;\to\; X + e(\mathbf{i} + e_l) - e(\mathbf{i}) - e(l) & & \\
& \qquad\qquad \text{at rate} \quad X_l x_{\mathbf{i}} \sigma_{l\mathbf{i}}(x), & \mathbf{i} \in \mathcal{Z}_1,\ 1 \leq l \leq d, \\[4pt]
\text{VI}: & X \;\to\; X - e(l) \quad \text{at rate} \quad X_l\{\bar{\zeta}_l + \zeta_l(x)\}, & 1 \leq l \leq d,
\end{array}
$$

where $x := N^{-1}X \in \{x' \in R_+^{\mathcal{Z}},\ \|x'\|_1 < \infty\} =: \mathcal{X}'$, and $\|x\|_1 := \sum_{z \in \mathcal{Z}} x_z$.

The transitions I correspond to changes in the type of a patch, because of births, deaths and changes of status involving animals within the patch, or as a result of infection or catastrophe, or of immigration from outside the metapopulation, and we set $\bar{\lambda}_{\mathbf{ii}} = \lambda_{\mathbf{ii}}(\cdot) = 0$, $\mathbf{i} \in \mathcal{Z}_1$. Then II and III correspond to the creation and destruction of patches, IV and V concern the migration of animals of the different varieties between patches, and VI the deaths of animals during migration. The transitions IV' allow for the possibility of an individual being born as a migrant, as is allowed in our first example, in Section 4. More complicated transitions of this kind could have been incorporated, but the biological motivation for doing so does not seem compelling. The parameters $\bar{\lambda}_{\mathbf{ij}}$, $\bar{\delta}_{\mathbf{i}}$, $\bar{\gamma}_{\mathbf{i}l}$, $\bar{\gamma}'_{\mathbf{i}l}$ and $\bar{\zeta}_l$ represent fixed rates of transition per patch. To ensure that the overall rate of jumps is finite at any $x \in N^{-1}\mathcal{X}$, it is necessary to have $\sum_{\mathbf{j} \in \mathcal{Z}_1} \bar{\lambda}_{\mathbf{ij}} < \infty$ for all $\mathbf{i} \in \mathcal{Z}_1$. The corresponding quantities without the bars, together with $\sigma_{l\mathbf{i}}(\cdot)$ and $\beta_{\mathbf{i}}(\cdot)$, represent state dependent components of the transition rates. For each $x \in \mathcal{X}'$, it is then also necessary to have

$$\sum_{\mathbf{j} \in \mathcal{Z}_1} \lambda_{\mathbf{ij}}(x) < \infty, \quad \sum_{\mathbf{j} \in \mathcal{Z}_1} \beta_{\mathbf{j}}(x) < \infty \quad \text{and} \quad \sum_{\mathbf{j} \in \mathcal{Z}_1} x_{\mathbf{i}} \sigma_{l\mathbf{i}}(x) < \infty; \qquad (2.1)$$

further assumptions are added in Section 3. In transition IV, we require $\bar{\gamma}_{\mathbf{i}l} = \gamma_{\mathbf{i}l}(x) = 0$ whenever $i_l = 0$, to avoid ever having $i_l < 0$, which would be biologically meaningless.

Let $T > 0$ be a constant; we study the evolution of the metapopulation over the interval $[0, T]$. Under further assumptions on the transition rates I–VI and on the initial condition $x^N(0)$, it can be shown that, with high probability, $x^N(t)$ is uniformly close to the solution x of a deterministic integral equation, which is the analogue of the usual deterministic drift differential equations found in finite dimensional problems. In Section 3, we illustrate how to use the results of [BL] (2012a) to justify this. For the rest of this section, we assume that

$$\mathbb{P}\left[\sup_{0 \leq t \leq T} \|x^N(t) - x(t)\|_\mu > \varepsilon_N \right] \leq P_T(N, \varepsilon_N), \qquad (2.2)$$

for some (small) ε_N and $P_T(N, \varepsilon_N)$, and for some norm $\| \cdot \|_\mu$, and show how (2.2) can be used to establish the joint behaviour of groups of individuals in the process X^N.

We begin by investigating the behaviour over time of the type of a single patch $\mathcal{P}$. The transitions I, IV and V each contain elements corresponding

4

to the rate of change of type of a patch that is currently of type $\mathbf{i}$, with the rates depending on the current state of the whole system, and the death rate of such a patch is given in III. Thus we can single out the transition rates for the patch $\mathcal{P}$, with its evolution only being Markovian if the current state x of the whole system is adjoined. For any $\mathbf{i}, \mathbf{j} \in \mathcal{Z}_1$ and $1 \le l \le d$, these take the form

$$
\begin{array}{llll}
\mathbf{i} & \to & \mathbf{j} & \text{at rate} \quad \bar{\lambda}_{\mathbf{ij}} + \lambda_{\mathbf{ij}}(x), & \|\mathbf{j} - \mathbf{i}\|_1 \ge 2; \\
\mathbf{i} & \to & \mathbf{j} & \text{at rate} \quad \bar{\lambda}_{\mathbf{ij}} + \lambda_{\mathbf{ij}}(x) + \bar{\gamma}_{\mathbf{i}l} + \gamma_{\mathbf{i}l}(x); & \mathbf{j} = \mathbf{i} - e_l \\
\mathbf{i} & \to & \mathbf{j} & \text{at rate} \quad \bar{\lambda}_{\mathbf{ij}} + \lambda_{\mathbf{ij}}(x) + x_l \sigma_{l\mathbf{i}}(x); & \mathbf{j} = \mathbf{i} + e_l \\
\mathbf{i} & \to & \Delta & \text{at rate} \quad \bar{\delta}_{\mathbf{i}} + \delta_{\mathbf{i}}(x),
\end{array}
\tag{2.3}
$$

with Δ a state to represent that the patch has been destroyed. We let Y_N denote the process describing the time evolution of the type assigned to $\mathcal{P}$, with $Y_N(t)$ taking values in $\mathcal{Z}_1 \cup \Delta$; the N-dependence reflects that its transition rates are as described in (2.3), but with $x^N(t)$ in place of x for the rates at time t.

Analogously, we could define a process representing the life history of an animal $\mathcal{A}$ in the metapopulation. The migration transitions IV, V and VI are easy to interpret, and the destruction of a patch in III implies the death of any animals in that patch. The transitions I are more complicated. Considering an animal of variety l, its death is typically recorded in a transition in which $j_l \le i_l - 1$ (several animals of the same variety may die as a result of the same event), but a change of developmental stage, for instance, may also result in $j_l = i_l - 1$. Then, for unicellular animals, division is recorded most simply as $j_l = i_l + 1$, though it may be useful to interpret the same event as the death of the original animal at the same time as the birth of two offspring. Furthermore, transitions in which i_l does not change may represent births of animals that are directly associated with the particular animal of variety l being considered, as when an adult gives birth to juveniles that are represented as a distinct variety; such events are naturally to be recorded in a life history. This suggests defining a life history process $Z_N := \{(Z_{N0}(t), \ldots, Z_{Nd}(t)), t \ge 0\}$ for an animal $\mathcal{A}$, whose statespace is

$$
((\mathcal{Z}_1 \times \{1, 2, \ldots, d\}) \cup \{1, 2, \ldots, d\} \cup \Delta) \times \mathbb{Z}_+^d.
$$

A value $Z_{N0}(t) \in \mathcal{Z}_1 \times \{1, 2, \ldots, d\}$ denotes the the type of patch in which $\mathcal{A}$ is living and its current variety. Then $Z_{N0}(t) = l$ if $\mathcal{A}$ is of variety l and in migration, and, if $Z_{N0}(t) = \Delta$, the animal $\mathcal{A}$ has died before time t. The

5

values $Z_{Nl}(t)$, $1 \leq l \leq d$, record the numbers of children of the different varieties to which $\mathcal{A}$ has given birth up to time t. For $\mathbf{i} \in \mathcal{Z}_1$, $l, l' \in \{1, 2, \ldots, d\}$ and $m, s \in \mathbb{Z}_+^d$, the transition rates can be represented in the form

$$
\begin{array}{llll}
((\mathbf{i}, l), m) & \rightarrow & ((\mathbf{i} + s, l), m + s) & \text{at rate} & \bar{\lambda}_{\mathbf{i}ls}^{(1)} + \lambda_{\mathbf{i}ls}^{(1)}(x); \\
((\mathbf{i}, l), m) & \rightarrow & ((\mathbf{j}, l), m) & \text{at rate} & \bar{\lambda}_{\mathbf{ij}}^{(2)} + \lambda_{\mathbf{ij}}^{(2)}(x); \\
((\mathbf{i}, l), m) & \rightarrow & ((\mathbf{i} - e_l + e_{l'}, l'), m) & \text{at rate} & \bar{\lambda}_{\mathbf{i}ll'}^{(3)} + \lambda_{\mathbf{i}ll'}^{(3)}(x); \\
((\mathbf{i}, l), m) & \rightarrow & ((\mathbf{i}, l), m + e_{l'}) & \text{at rate} & \bar{\lambda}_{\mathbf{i}ll'}^{(4)} + \lambda_{\mathbf{i}ll'}^{(4)}(x); \\
((\mathbf{i}, l), m) & \rightarrow & (\Delta, m) & \text{at rate} & \bar{\delta}_{\mathbf{i}l}' + \delta_{\mathbf{i}l}'(x); \\
((\mathbf{i}, l), m) & \rightarrow & (l, m) & \text{at rate} & i_l^{-1}\{\bar{\gamma}_{\mathbf{i}l} + \gamma_{\mathbf{i}l}(x)\}; \\
(l, m) & \rightarrow & ((\mathbf{i} + e_l, l), m) & \text{at rate} & x_{\mathbf{i}}\sigma_{l\mathbf{i}}(x); \\
(l, m) & \rightarrow & (\Delta, m) & \text{at rate} & \bar{\zeta}_l + \zeta_l(x).
\end{array}
\tag{2.4}
$$

Here, the quantities $\bar{\lambda}_{\mathbf{i}ls}^{(1)}$ and $\lambda_{\mathbf{i}ls}^{(1)}(x)$ represent the rates at which, in a type $\mathbf{i}$ patch, an animal of variety l produces offspring in the composition s, and they would form a part of the rates $\bar{\lambda}_{\mathbf{i},\mathbf{i}+s}$ and $\lambda_{\mathbf{i},\mathbf{i}+s}(x)$; they are assumed not to depend on m. Similar considerations apply to the quantities $\bar{\lambda}_{\mathbf{ij}}^{(2)}$ and $\lambda_{\mathbf{ij}}^{(2)}(x)$, which relate to events changing the composition of the patch containing $\mathcal{A}$ that do not result in offspring for $\mathcal{A}$ or a change in its variety, including migration of other animals from the patch or the arrival of migrants. Thus, for instance, one might have $\bar{\lambda}_{\mathbf{i},\mathbf{i}+e_l} = \varphi_{1l}i_l$, $\bar{\lambda}_{\mathbf{i},\mathbf{i}-e_l} = \varphi_{2l}i_l$, $\bar{\gamma}_{\mathbf{i}l} = i_l\varphi_{3l}$ and $\sigma_{l\mathbf{i}}(x) = \sigma_{l\mathbf{i}}$, $1 \leq l \leq d$, corresponding to constant *per capita* birth, death, migration and immigration rates φ_{1l}, φ_{2l}, φ_{3l} and $\sigma_{l\mathbf{i}}$ of individuals of variety l. These would imply $\bar{\lambda}_{\mathbf{i}le_l}^{(1)} = \varphi_{1l}$, $\bar{\lambda}_{\mathbf{i},\mathbf{i}+e_l}^{(2)} = (i_l-1)\varphi_{1l}$, $\lambda_{\mathbf{i},\mathbf{i}+e_l}^{(2)}(x) = x_l\sigma_{l\mathbf{i}}$, and $\bar{\lambda}_{\mathbf{i},\mathbf{i}-e_l}^{(2)} = (i_l - 1)(\varphi_{2l} + \varphi_{3l})$ for transitions only involving l-animals, and, for $l' \neq l$, $\bar{\lambda}_{\mathbf{i},\mathbf{i}+e_{l'}}^{(2)} = i_{l'}\varphi_{1l'}$, $\lambda_{\mathbf{i},\mathbf{i}+e_{l'}}^{(2)}(x) = x_{l'}\sigma_{l'\mathbf{i}}$, and $\bar{\lambda}_{\mathbf{i},\mathbf{i}-e_{l'}}^{(2)} = i_{l'}(\varphi_{2l'} + \varphi_{3l'})$. The transition rates $\bar{\lambda}_{\mathbf{i}ll'}^{(3)}$ and $\lambda_{\mathbf{i}ll'}^{(3)}(x)$ relate to events that change $\mathcal{A}$'s variety from l to l'; it is tacitly assumed that no other changes take place when this happens, but more general possibilities could have been allowed. The rates $\bar{\lambda}_{\mathbf{i}ll'}^{(4)}$ and $\lambda_{\mathbf{i}ll'}^{(4)}(x)$ relate to births of migrants as offspring of an l-animal. The rates $\bar{\delta}_{\mathbf{i}l}' \geq \bar{\delta}_{\mathbf{i}}$ and $\delta_{\mathbf{i}l}'(x) \geq \delta_{\mathbf{i}}(x)$ include a contribution from the mortality rate of an animal of variety l in a patch of type $\mathbf{i}$, in addition to the rate of destruction of the patch itself. As for the single patch dynamics, the rates for the process Z_N at time t are obtained by replacing x with $x^N(t)$ in the expressions (2.4).

6

These constructions immediately suggest approximating the processes Y_N and Z_N by random processes Y and Z, in which the transition rates at time t are obtained by replacing x by $x(t)$ in (2.3) and (2.4). Consider first the processes Y_N and Y. Suppose, for some $\delta > 0$, that the functions $\lambda_{\mathbf{ij}}$, $\gamma_{\mathbf{i}l}$, $\sigma_{l\mathbf{i}}$ and $\delta_{\mathbf{i}}$ are all of uniformly bounded Lipschitz μ-norm, for x in a set $B_{T,\delta} := \{x \in \mathcal{X}' : \inf_{0 \le t \le T} \|x - x(t)\|_\mu \le \delta\}$ of points close to the deterministic trajectory $(x(t), 0 \le t \le T)$. Then, in view of (2.3), the jump rates of Y_N and Y at any time $t \in [0, T]$ differ only by a small amount, on the event that $\sup_{0 \le t \le T} \|x^N(t) - x(t)\|_\mu \le \varepsilon_N$, provided that N is large enough that $\varepsilon_N \le \delta$. Indeed, defining $f^* := \sup_{x \in B_{T,\delta}} |f(x)|$ for any $f \colon \mathcal{X} \to \mathbb{R}$, and setting

$$|Df|(x) := \limsup_{\varepsilon \to 0} \sup_{0 < \|y - x\|_\mu < \varepsilon} \{|f(y) - f(x)| / \|y - x\|_\mu\},$$

it follows that, if $|x - x(t)| \le \varepsilon < \delta$ and $0 \le t \le T$, then the sum of the differences of the transition rates out of x and $x(t)$ is bounded by

$$\sup_{\mathbf{i} \in \mathcal{Z}_1} \left\{ \sum_{\mathbf{j} \in \mathcal{Z}_1} |\lambda_{\mathbf{ij}}(x) - \lambda_{\mathbf{ij}}(x(t))| + \sum_{l=1}^d |\gamma_{\mathbf{i}l}(x) - \gamma_{\mathbf{i}l}(x(t))| \right.$$
$$\left. + \sum_{l=1}^d |x_l \sigma_{l\mathbf{i}}(x) - x_l(t)\sigma_{l\mathbf{i}}(x(t))| + |\delta_{\mathbf{i}}(x) - \delta_{\mathbf{i}}(x(t))| \right\} \le \varepsilon D_Y(T, \delta),$$

where, writing $\hat{\sigma}_{l\mathbf{i}}(x) := x_l \sigma_{l\mathbf{i}}(x)$, we define

$$D_Y(T, \delta) := \sup_{\mathbf{i} \in \mathcal{Z}_1} \left\{ \sum_{\mathbf{j} \in \mathcal{Z}_1} |D\lambda_{\mathbf{ij}}|^* + \sum_{l=1}^d \{|D\gamma_{\mathbf{i}l}|^* + |D\hat{\sigma}_{l\mathbf{i}}|^*\} + |D\delta_{\mathbf{i}}|^* \right\}.$$

Thus, until the time at which first $\|x^N(t) - x(t)\|_\mu > \varepsilon_N$, the aggregate difference between the jump rates of the processes Y_N and Y is bounded by $\varepsilon_N D_Y(T, \delta)$, if also $t \le T$. This immediately leads to the following theorem.

Theorem 2.1 *Suppose that (2.2) holds, and that $D_Y(T, \delta) < \infty$ for some $\delta > 0$. Then, if $Y_N(0) = Y(0)$ and $\varepsilon_N \le \delta$, the processes Y_N and Y can be constructed on the same probability space in such a way that*

$$\mathbb{P}[Y_N(t) = Y(t) \text{ for all } 0 \le t \le T] \ge 1 - \{T\varepsilon_N D_Y(T, \delta) + P_T(N, \varepsilon_N)\}.$$

Proof. Let Y_1 and Y_2 be time-inhomogeneous Markov processes on a countable state space $\mathcal{Y}$, with transition rates $q_1(t, y, y')$ and $q_2(t, y, y')$ respectively. Starting with $Y_1(0) = Y_2(0) = y_0$, the processes can be coupled by representing them as the marginals of a joint process $((Y_1(t), Y_2(t)), t \geq 0)$, whose transition rates at points on the diagonal are given by

$$\begin{aligned}
q(t, (y, y), (y', y')) &:= \min\{q_1(t, y, y'), q_2(t, y, y')\}; \\
q(t, (y, y), (y, y')) &:= \{q_2(t, y, y') - q_1(t, y, y')\}_+; \\
q(t, (y, y), (y', y)) &:= \{q_1(t, y, y') - q_2(t, y, y')\}_+,
\end{aligned}$$

and with the components evolving independently when off the diagonal. Let $\tau := \inf\{t \geq 0 \colon Y_1(t) \neq Y_2(t)\}$, and let E_t^η denote the event $\{Q(s, Y_1(s)) \leq \eta$ for all $0 \leq s \leq t\}$, where

$$Q(t, y) := \sum_{y' \in \mathcal{Y}} |q_2(t, y, y') - q_1(t, y, y')|.$$

Then the one-jump process $(I[\{\tau \leq t\} \cap E_t^\eta], t \geq 0)$ has compensator

$$A_t := \int_0^{t \wedge \tau} Q(s, Y_1(s)) I[E_s^\eta] \, ds \leq \eta t.$$

This implies that, for any $T > 0$,

$$\mathbb{P}[\{\tau \leq T\} \cap E_T^\eta] = \mathbb{E}\{I[\{\tau \leq T\} \cap E_T^\eta]\} = \mathbb{E}A_T \leq \eta T,$$

from which it follows that $\mathbb{P}[\tau \leq T] \leq \eta T + \mathbb{P}[(E_T^\eta)^c]$. Thus this construction realizes Y_1 and Y_2 on the same probability space, in such a way that the two remain identical up to time T with probability at least $1 - (\eta T + \mathbb{P}[(E_T^\eta)^c])$.

Now, taking Y_N for Y_1 and Y for Y_2, and setting $\eta = \varepsilon_N D_Y(T, \delta)$, the theorem follows from (2.2). $\blacksquare$

Since all the transitions in (2.3) involve a single patch, the theorem generalizes easily to any group of K patches. The transition rates for the process $(Y_N^{[1]}, Y_N^{[2]}, \ldots, Y_N^{[K]})$ at time t from a state $(\mathbf{i}^{(1)}, \ldots, \mathbf{i}^{(K)})$ to one in which $\mathbf{i}^{(k)}$ is replaced by $\mathbf{i}^{(k')}$, with $\mathbf{i}^{(k')}$ either of the form $\mathbf{i}^{(k)} + \mathbf{j}$, $\mathbf{j} \in \mathbb{Z}^d$, or Δ, are given by the formulae in (2.3) with $\mathbf{i}^{(k)}$ for $\mathbf{i}$, and with $x^N(t)$ for x. The rates for a vector of independent processes $Y^{[k]}$, $1 \leq k \leq K$, each distributed as Y, with $Y^{[k]}(0) = \mathbf{i}^{(k)}$, are the corresponding rates with $x(t)$ for x. This leads to the following corollary.

Corollary 2.2 *Under the conditions of Theorem 2.1,*

$$\mathbb{P}[(Y_N^{[1]}(t), \ldots, Y_N^{[K]}(t)) = (Y^{[1]}(t), \ldots, Y^{[K]}(t)) \text{ for all } 0 \le t \le T]$$
$$\ge 1 - \{KT\varepsilon_N D_Y(T, \delta) + P_T(N, \varepsilon_N)\}.$$

Thus the joint distribution of K_N patches is asymptotically close to that of K_N independently evolving patches over any fixed interval $[0, T]$, as $N \to \infty$, if $K_N \varepsilon_N \to 0$, $P_T(N, \varepsilon_N) \to 0$ and $D_Y(T, \delta) < \infty$ for some $\delta > 0$.

For the life history process of an animal, the argument for a single individual is very similar. We consider the differences in the transition rates (2.4) with arguments $x^N(t)$ and $x(t)$; defining

$$D_Z(T, \delta) := \max_{1 \le l \le d} \left(\sup_{\mathbf{i} \in \mathcal{Z}_1} \Big\{ \sum_{s \in \mathbb{Z}_+^d} |D\lambda_{\mathbf{i}ls}^{(1)}|^* + \sum_{\mathbf{j} \in \mathcal{Z}_1} |D\lambda_{\mathbf{ij}}^{(2)}|^* + \sum_{l'=1}^{d} |D\lambda_{\mathbf{i}ll'}^{(3)}|^* \right.$$
$$\left. + \sum_{l'=1}^{d} |D\lambda_{\mathbf{i}ll'}^{(4)}|^* + |D\delta_{\mathbf{i}}'|^* + |D\gamma_{\mathbf{i}l}|^* + |D\hat{\sigma}_{l\mathbf{i}}'|^* \Big\} + |D\zeta_l|^* \right),$$

where $\hat{\sigma}_{l\mathbf{i}}'(x) := x_{\mathbf{i}} \sigma_{l\mathbf{i}}(x)$, this gives the following result.

Theorem 2.3 *Suppose that (2.2) holds, and that $D_Z(T, \delta) < \infty$ for some $\delta > 0$. Then, if $\varepsilon_N \le \delta$ and $Z_N(0) = Z(0)$, the processes Z_N and Z can be constructed on the same probability space in such a way that*

$$\mathbb{P}[Z_N(t) = Z(t) \text{ for all } 0 \le t \le T] \ge 1 - \{T\varepsilon_N D_Z(T, \delta) + P_T(N, \varepsilon_N)\}.$$

For the joint distribution of a group of K animals, asymptotic independence is not quite as straightforward, since all but the fourth and the last transitions in (2.4) simultaneously change the state of any other animal in the same patch. Hence it is necessary to begin with all animals in different patches, and the simple coupling breaks down once two of them are to be found in the same patch. This can only occur when a migrant enters a patch that already contains another of the K animals. For a given animal of variety l, an upper bound for the maximum rate at which it can enter such a patch is $N^{-1}(K-1)\sup_{\mathbf{i}} |\sigma_{l\mathbf{i}}|^*$, because the $(K-1)$ other animals of the group can be in at most $K-1$ distinct patches, and $\sigma_{l\mathbf{i}}(x) \le |\sigma_{l\mathbf{i}}|^*$; and there are K animals that could migrate into such a patch. Hence the event that no two of the K animals are in the same patch during the interval $[0, T]$ has probability bounded by $K^2 N^{-1} \sigma^+$, where $\sigma^+ := \sup_{\mathbf{i} \in \mathcal{Z}_1} \max_{1 \le l \le d} |\sigma_{l\mathbf{i}}|^*$. This leads to the following corollary.

Corollary 2.4 *Suppose that (2.2) holds, and that $D_Z(T, \delta) < \infty$ for some $\delta > 0$. Then, if $\varepsilon_N \leq \delta$ and the K individuals are initially all in distinct patches, we have*

$$\mathbb{P}[(Z_N^{[1]}(t), \ldots, Z_N^{[K]}(t)) = (Z^{[1]}(t), \ldots, Z^{[K]}(t)) \text{ for all } 0 \leq t \leq T]$$
$$\geq 1 - \{KT\varepsilon_N D_Z(T, \delta) + TK^2 N^{-1}\sigma^+ + P_T(N, \varepsilon_N)\},$$

where the $Z^{[k]}$, $1 \leq k \leq K$, are independent copies of Z with $Z^{[k]}(0) = Z_N^{[k]}(0)$.

Thus, if (2.2) holds and $D_Z(T, \delta) < \infty$ for some $\delta > 0$, any group of K_N animals that are initially in different patches behaves asymptotically as a group of independent individuals, under the same asymptotic scenario as before, if also $N^{-1}K_N^2 \to 0$ as $N \to \infty$.

The model in Arrigoni (2003) does not conform to our general prescription, because migration is assumed to take place instantaneously, rather than by way of an intermediate migration state. However, the state dependent elements of its transition rates are locally uniformly Lipschitz, and (2.2) holds, so that analogous theorems hold for this model as well. We do not include instantaneous migration in our general formulation, partly because it seems unrealistic, but mainly because, for the methods in [BL] (2012a) to be applied, only rather restrictive choices can be allowed for the migration transitions. For instance, in the Arrigoni model, it is important that the migration rate $\bar{\gamma}_i$ out of patches with i individuals is given by $\bar{\gamma}_i = \gamma i$; variants in which $i^{-1}\bar{\gamma}_i$ increases with i would not lead to a locally Lipschitz drift F in (3.14) below.

3 Establishing the law of large numbers

We now need to prove that (2.2) holds. For this, we need to find conditions on the transition rates in I–VI that allow us to apply the results of [BL] (2012a) to the process X^N. First, we need to make some small modifications to the setting in the previous section. We start by augmenting the type space $\mathcal{Z}$ to $\widetilde{\mathcal{Z}}$, by substituting $\widetilde{\mathcal{Z}}_2 := \{1, 2, \ldots, d\} \times \{0, 1\}$ for $\mathcal{Z}_2$, where the type $(l, 1)$ replaces the previous type l in representing an individual of variety l in migration, and type $(l, 0)$ is to be thought of as an unused place available for a migrant of variety l. Then, in transitions IV and IV′, $e(l)$ is replaced by $e(l, 1) - e(l, 0)$ and, in transitions V and VI, $-e(l)$ is replaced by $e(l, 0) - e(l, 1)$ and X_l by X_{l1}. The number X_{l0} of patches of type $e(l, 0)$ can be deduced

from the number X_{l1} of $e(l, 1)$ patches, since the sum $X_{l1} + X_{l0}$ remains constant in all transitions, and is therefore always the same as its initial value. However, to prevent the number of type $(l, 0)$ patches becoming negative, the process X^N has to be stopped at the time $\tau_{0,N} := \inf\{t \geq 0 : \min_{1 \leq l \leq d} X_{l0}^N = 0\}$. So that this has little effect on the process, $X^N(0)$ is chosen with $X_{l0}^N \geq Nh_l$, $1 \leq l \leq d$, with the h_l so large that, for fixed T, the event $\{\tau_{0,N} \leq T\}$ has asymptotically small probability as $N \to \infty$. The reason for introducing the empty migration patches will emerge shortly.

3.1   A priori bounds

We now introduce a measure ν of the size of a patch, defining $\nu(l, 0) = \nu(l, 1) := 1$ for $1 \leq l \leq d$, and $\nu(\mathbf{i}) := \|\mathbf{i}\|_1 + 1$, one more than the number of individuals in a type $\mathbf{i}$ patch. More flexible choices for ν are allowed in [BL] (2012a), but this suffices here. It is then necessary to make assumptions ensuring that, for enough values of $r \in \mathbb{Z}_+$, the empirical moments $S_r(x^N(t)) := \sum_{z \in \widetilde{\mathcal{Z}}} \nu(z)^r x_z^N(t)$ remain bounded with high probability as N increases, if they are initially bounded. Let J denote a finite linear combination of coordinate vectors in $\widetilde{\mathcal{Z}}$. Let $\mathcal{J}$ denote the jumps J that appear in the transitions I–VI, with the above modification replacing $e(l)$ by $e(l, 1) - e(l, 0)$, and let the associated transition rates be denoted by $N\alpha_J(x)$. Note that we can suppose that $x \in \mathcal{X}'$, if the l coordinates in $\mathcal{Z}$ are identified with the $(l, 1)$ coordinates in $\widetilde{\mathcal{Z}}$, since the values $x_{(l,0)}$ do not appear in the expressions for the transition rates I–VI. For $J := \sum_{k=1}^{K} a_k e(\mathbf{j}^{(k)}) \in \mathcal{J}$, write

$$\nu_r^+(J) := \sum_{k=1}^{K} a_k \{\nu(\mathbf{j}^{(k)})\}^r, \tag{3.1}$$

and, for $r \in \mathbb{Z}_+$, define

$$U_r(x) := \sum_{J \in \mathcal{J}} \alpha_J(x)\nu_r^+(J); \quad V_r(x) := \sum_{J \in \mathcal{J}} \alpha_J(x)\{\nu_r^+(J)\}^2. \tag{3.2}$$

Then, in order to be able to apply the theorems of [BL] (2012a), we assume that, for some $r^{(1)} \geq 1$ and for all $0 \leq r \leq r^{(1)}$,

$$\sum_{J \in \mathcal{J}} \alpha_J(N^{-1}X)|\nu_r^+(J)| < \infty \text{ for each } X \in \mathcal{X}, \tag{3.3}$$

11

and that, for suitable constants k_{rl} and all $x \in \mathcal{X}'$,

$$
\begin{aligned}
U_0(x) &\leq k_{01}S_0(x) + k_{04}; \\
U_1(x) &\leq k_{11}S_1(x) + k_{14}; \\
U_r(x) &\leq \{k_{r1} + k_{r2}S_0(x)\}S_r(x) + k_{r4}, \qquad 2 \leq r \leq r^{(1)},
\end{aligned}
\tag{3.4}
$$

and, for some $r^{(2)} \geq 1$,

$$
\begin{aligned}
V_0(x) &\leq k_{03}S_1(x) + k_{05}; \\
V_r(x) &\leq k_{r3}S_{p(r)}(x) + k_{r5}, \qquad 1 \leq r \leq r^{(2)},
\end{aligned}
\tag{3.5}
$$

are satisfied, where $1 \leq p(r) \leq r^{(1)}$ for $1 \leq r \leq r^{(2)}$.

In our setting, satisfying the condition (3.3) is straightforward except for the transitions of the form II, since, for $X \in \mathcal{X}$, only finitely many of the $X_{\mathbf{i}}$ are non-zero; and transitions of the form II are also the only ones that make positive contributions to $U_0(x)$. One plausible assumption, covering these and later conditions, is to require that

$$
\beta_{\mathbf{j}}(x) \leq c'_{\mathbf{j}}(\|x\|_1 + 1), \quad \text{where} \quad \sum_{\mathbf{j} \in \mathcal{Z}_1} c'_{\mathbf{j}}\{\nu(\mathbf{j})\}^r < \infty \quad \text{for each } r \in \mathbb{Z}_+. \tag{3.6}
$$

Here, and in what follows, c and c' are used to denote generic constants. If the types $(l, 0)$ had not been introduced, there would also be positive contributions of $\sum_{\mathbf{j} \in \mathcal{Z}_1} X_{\mathbf{j}} \bar{\gamma}_{\mathbf{j}l}$ to $U_0(x)$ from transitions IV, and the most natural assumption for the value of $\bar{\gamma}_{\mathbf{j}l}$ is $\gamma_l j_l$, for some constant γ_l, corresponding to a constant *per capita* migration rate for l-individuals. Thus $\sum_{\mathbf{j} \in \mathcal{Z}_1} X_{\mathbf{j}} \bar{\gamma}_{\mathbf{j}l}$ would be bounded by a multiple of $S_1(x)$, rather than by a multiple of $S_0(x)$, and so would not have come within the scope of [BL] (2012a). For the remaining conditions concerning $U_r(x)$, $r \geq 1$, it is enough to assume that, for $\mathbf{i} \in \mathcal{Z}_1$ and for all $x \in \mathcal{X}'$,

$$
\sum_{\mathbf{j} \in \mathcal{Z}_1} \bar{\lambda}_{\mathbf{ij}} + \sum_{\mathbf{j} \in \mathcal{Z}_1} (\bar{\lambda}_{\mathbf{ij}} + \lambda_{\mathbf{ij}}(x))\{\nu(\mathbf{j}) - \nu(\mathbf{i})\}_+ \leq c\nu(\mathbf{i}); \tag{3.7}
$$

$$
\sum_{\mathbf{j} \in \mathcal{Z}_1} (\bar{\lambda}_{\mathbf{ij}} + \lambda_{\mathbf{ij}}(x))(\{\nu(\mathbf{j})\}^r - \{\nu(\mathbf{i})\}^r)_+ \leq c\{\nu(\mathbf{i})\}^r(\|x\|_1 + 1), \tag{3.8}
$$

and that, for $1 \leq l \leq d$ and for all $x \in \mathcal{X}'$,

$$
\sigma_{l\mathbf{i}}(x) \leq c; \quad \sum_{\mathbf{j} \in \mathcal{Z}_1} x_{\mathbf{j}} \sigma_{l\mathbf{j}}(x) \leq c. \tag{3.9}
$$

For the conditions concerning $V_r(x)$, $r \geq 0$, with $p(r) = 2r + 1$ as in [BL] (2012), we assume further that, for $\mathbf{i} \in \mathcal{Z}_1$ and $1 \leq l \leq d$ and for all $x \in \mathcal{X}'$,

$$\bar{\delta}_\mathbf{i} + \delta_\mathbf{i}(x) \leq c\nu(\mathbf{i}), \quad \bar{\gamma}_{\mathbf{j}l} + \gamma_{\mathbf{i}l}(x) \leq c\nu(\mathbf{i}) \quad \text{and} \quad \bar{\gamma}'_{\mathbf{i}l} + \gamma'_{\mathbf{i}l}(x) \leq c\nu(\mathbf{i}), \quad (3.10)$$

and that

$$\sum_{\mathbf{j} \in \mathcal{Z}_1} (\bar{\lambda}_{\mathbf{ij}} + \lambda_{\mathbf{ij}}(x))(\{\nu(\mathbf{j})\}^r - \{\nu(\mathbf{i})\}^r)^2 \leq c\{\nu(\mathbf{i})\}^{2r+1}. \quad (3.11)$$

3.2 The deterministic equation

The process $x^N := N^{-1}X^N$ has infinitesimal drift $F_0(x)$, $x \in \mathcal{X}'$, whose components are formally given by

$$F_{0;\mathbf{i}}(x) := \sum_{\mathbf{j} \in \mathcal{Z}_1} x_\mathbf{j}\{\bar{\lambda}_{\mathbf{ji}} + \lambda_{\mathbf{ji}}(x)\} - x_\mathbf{i}\sum_{\mathbf{j} \in \mathcal{Z}_1}\{\bar{\lambda}_{\mathbf{ij}} + \lambda_{\mathbf{ij}}(x)\} - x_\mathbf{i}\{\bar{\delta}_\mathbf{i} + \delta_\mathbf{i}(x)\}$$

$$+ \beta_\mathbf{i}(x) + \sum_{l=1}^{d} x_{\mathbf{i}+e_l}\{\bar{\gamma}_{\mathbf{i}+e_l,l} + \gamma_{\mathbf{i}+e_l,l}(x)\} - x_\mathbf{i}\sum_{l=1}^{d}\{\bar{\gamma}_{\mathbf{i}l} + \gamma_{\mathbf{i}l}(x)\}$$

$$+ \sum_{l=1}^{d} x_{l1}\{x_{\mathbf{i}-e_l}\sigma_{l,i-e_l}(x) - x_\mathbf{i}\sigma_{l\mathbf{i}}(x)\}, \quad (3.12)$$

for $\mathbf{i} \in \mathcal{Z}_1$, and, for $1 \leq l \leq d$,

$$F_{0;l1}(x) := \sum_{\mathbf{j} \in \mathcal{Z}_1} x_\mathbf{j}\{\bar{\gamma}_{\mathbf{j}l} + \gamma_{\mathbf{j}l}(x) + \bar{\gamma}'_{\mathbf{j}l} + \gamma'_{\mathbf{j}l}(x)\}$$

$$- x_{l1}\sum_{\mathbf{j} \in \mathcal{Z}_1} x_\mathbf{j}\sigma_{l\mathbf{j}}(x) - x_{l1}\{\bar{\zeta}_l + \zeta_l(x)\}; \quad (3.13)$$

these expressions only make sense if the $\mathbf{j}$-sums are all finite. The drift in the $(l, 0)$ coordinate is given by $-F_{0;l1}(x)$, but we do not use it explicitly. Thus, for $x \in \mathcal{X}'$ such that $F(x)$ exists, we can write

$$F_0(x) := Ax + F(x), \quad (3.14)$$

to be interpreted as an element of $\mathbb{R}_+^{\mathcal{Z}}$, where

$$
\begin{aligned}
A_{\mathbf{ij}} &:= \bar{\lambda}_{\mathbf{ji}} + \sum_{l=1}^{d} \mathbf{1}_{\{\mathbf{j}=\mathbf{i}+e_l\}}\bar{\gamma}_{\mathbf{j}l}, \quad \mathbf{i} \neq \mathbf{j} \in \mathcal{Z}_1; \\
A_{\mathbf{ii}} &:= -\sum_{\mathbf{j} \in \mathcal{Z}_1} \bar{\lambda}_{\mathbf{ij}} - \bar{\delta}_{\mathbf{i}} - \sum_{l=1}^{d} \bar{\gamma}_{\mathbf{i}l}, \quad \mathbf{i} \in \mathcal{Z}_1; \\
A_{\mathbf{i}l} &:= 0, \; A_{l\mathbf{i}} := \bar{\gamma}_{\mathbf{i}l} + \bar{\gamma}_{\mathbf{i}l}', \; A_{ll} := -\bar{\zeta}_l, \; A_{ll'} := 0, \quad \mathbf{i} \in \mathcal{Z}_1, \, 1 \leq l, l' \leq d,
\end{aligned}
$$
$$(3.15)$$

with l in the indices of A as shorthand for $(l,1)$; and where

$$
\begin{aligned}
F_{\mathbf{i}}(x) &:= \sum_{\mathbf{j} \in \mathcal{Z}_1} x_{\mathbf{j}} \lambda_{\mathbf{ji}}(x) - x_{\mathbf{i}} \sum_{\mathbf{j} \in \mathcal{Z}_1} \lambda_{\mathbf{ij}}(x) + \beta_{\mathbf{i}}(x) - x_{\mathbf{i}} \delta_{\mathbf{i}}(x) + \sum_{l=1}^{d} x_{\mathbf{i}+e_l} \gamma_{\mathbf{i}+e_l, l}(x) \\
&\quad - x_{\mathbf{i}} \sum_{l=1}^{d} \gamma_{\mathbf{i}l}(x) + \sum_{l=1}^{d} x_{l1}\{x_{\mathbf{i}-e_l}\sigma_{l,\mathbf{i}-e_l}(x) - x_{\mathbf{i}}\sigma_{l\mathbf{i}}(x)\},
\end{aligned}
$$
$$(3.16)$$

for $\mathbf{i} \in \mathcal{Z}_1$, and, for $1 \leq l \leq d$,

$$
F_{l1}(x) := \sum_{\mathbf{j} \in \mathcal{Z}_1} x_{\mathbf{j}}\{\gamma_{\mathbf{j}l}(x) + \gamma_{\mathbf{j}l}'(x)\} - x_{l1} \sum_{\mathbf{j} \in \mathcal{Z}_1} x_{\mathbf{j}}\sigma_{l\mathbf{j}}(x) - x_{l1}\zeta_l(x). \quad (3.17)
$$

The reason for splitting the drift as above is to treat models in which the transition rates are not bounded as $\nu(\mathbf{i})$ increases — migration, birth and death rates proportional to the numbers of individuals in a patch are very natural — enabling the theory of perturbed linear operators to be applied.

We first assume that there is a real $\mu \in [1,\infty)^{\mathcal{Z}}$ such that, for some $w \geq 0$,

$$
A^T \mu \leq w\mu, \quad (3.18)
$$

and use it to define the μ-norm

$$
\|x\|_\mu := \sum_{z \in \mathcal{Z}} \mu(z)|x_z| \quad \text{on} \quad \mathcal{X}_\mu' := \{x \in \mathbb{R}^{\mathcal{Z}}: \|x\|_\mu < \infty\}, \quad (3.19)
$$

with x_l identified with x_{l1} as before. Note that, if (3.18) is assumed, we must have $\sum_{z \in \mathcal{Z}} \bar{\lambda}_{\mathbf{i}z}\mu(z) < \infty$ for each $\mathbf{i}$. Then, as in [BL] (2012a, Theorem 3.1), there exists a μ-strongly continuous semigroup $\{R(t), t \geq 0\}$ with

14

elementwise derivative $R'(0) = A$. Furthermore, if $F : \mathcal{X}'_\mu \to \mathcal{X}'_\mu$ is locally μ-Lipschitz and $\|x(0)\|_\mu < \infty$, the integral equation

$$x(t) = R(t)x(0) + \int_0^t R(t-u)F(x(u))\,du \qquad (3.20)$$

has a unique, μ-continuous solution on $[0, T]$ for any $0 < T < t_{\max}$, for some $t_{\max} \le \infty$. This x is the deterministic curve that approximates $x^N(t)$ when $x^N(0)$ is μ-close enough to $x(0)$.

From now on, we take $\mu(\mathbf{j}) := \|\mathbf{j}\|_1 + 1$ for $\mathbf{j} \in \mathcal{Z}_1$ and $\mu(l) := 1$ for $1 \le l \le d$. Inequality (3.18) is then satisfied if

$$\sum_{\mathbf{j} \in \mathcal{Z}_1} \bar{\lambda}_{\mathbf{ij}}(\mu(\mathbf{j}) - \mu(\mathbf{i})) + \sum_{l=1}^d (\bar{\gamma}_{\mathbf{i}-e_l,l} - \bar{\gamma}_{\mathbf{i}l})\mu(\mathbf{i} - e_l) + \sum_{l=1}^d \bar{\gamma}'_{\mathbf{i}l} \le w\mu(\mathbf{i}) \quad (3.21)$$

for all $\mathbf{i} \in \mathcal{Z}_1$. In order then to deduce that $F \colon \mathcal{X}'_\mu \to \mathcal{X}'_\mu$ is locally μ-Lipschitz, sufficient conditions are that, for $1 \le l \le d$ and $\mathbf{i} \in \mathcal{Z}_1$, and for any $R > 0$,

$\sigma_{\mathbf{i}l}(x), \delta_\mathbf{i}(x), \gamma_{\mathbf{i}l}(x), \gamma'_{\mathbf{i}l}(x), \zeta_l(x)$ and $\displaystyle\sum_{\mathbf{j} \in \mathcal{Z}_1} \lambda_{\mathbf{ij}}(x)$ are uniformly bounded, and

$\delta_\mathbf{i}(x), \gamma_{\mathbf{i}l}(x), \gamma'_{\mathbf{i}l}(x), \sigma_{\mathbf{i}l}(x)$ and $\zeta_l(x)$ are μ-uniformly Lipschitz, in $x \in B_R$;

$$\sum_{\mathbf{j} \in \mathcal{Z}_1} |\lambda_{\mathbf{ij}}(x) - \lambda_{\mathbf{ij}}(y)| \le c\|x - y\|_\mu, \quad \sum_{\mathbf{j} \in \mathcal{Z}_1} |\beta_\mathbf{j}(x) - \beta_\mathbf{j}(y)|\mu(\mathbf{j}) \le c\|x - y\|_\mu,$$

$$\sum_{\mathbf{j} \in \mathcal{Z}_1} |\lambda_{\mathbf{ij}}(x) - \lambda_{\mathbf{ij}}(y)|\mu(\mathbf{j}) \le c\mu(\mathbf{i})\|x - y\|_\mu \quad \text{and} \quad \sum_{\mathbf{j} \in \mathcal{Z}_1} \lambda_{\mathbf{ij}}(x)\mu(\mathbf{j}) \le c\mu(\mathbf{i}),$$

uniformly in $x, y \in B_R$,

$$(3.22)$$

for suitable constants $c = c_R$, where B_R is the ball of radius R in $\mathcal{X}'_\mu$.

3.3 The law of large numbers approximation

In order to apply the results of [BL] (2012a), we still need to check that their Assumption 4.2 is satisfied. Part (1) is satisfied with $r_\mu = 1$, because $\mu(z) = \nu(z)$ for all $z \in \widetilde{\mathcal{Z}}$. For Part (2), we define $\zeta(\mathbf{i}) := (\|\mathbf{i}\|_1 + 1)^{2d+5}$ for $\mathbf{i} \in \mathcal{Z}_1$ and $\zeta(l, 1) := \zeta(l, 0) = 1$ for $1 \le l \le d$, and observe that then, using

conditions (3.7) and (3.10), the sum

$$Z := \sum_{\mathbf{j} \in \mathcal{Z}_1} \frac{\mu(\mathbf{j})(A_{\mathbf{jj}} + 1)}{\sqrt{\zeta(\mathbf{j})}} = O\left(\sum_{j \geq 0} j^{(d-1)+2-(d+5/2)}\right) < \infty.$$

This implies that [BL] (2012a, Assumption 4.2(2)) is satisfied, provided that ζ satisfies [BL] (2012a, Assumption (2.25)). Defining $f(J) := \sum_{k=1}^{K} |a_k| \zeta(\mathbf{j}^{(k)})$ when $J := \sum_{k=1}^{K} a_k \mathbf{j}^{(k)}$, this in turn requires that

$$\sum_{J \in \mathcal{J}} \alpha_J(x) f(J) \leq \{k_1 S_r(x) + k_2\}, \quad x \in \mathcal{X}', \tag{3.23}$$

for some constants k_1 and k_2 and for some $r \leq r^{(2)}$. However, this also follows from conditions (3.7)–(3.10), if $r = 2d + 6$. Hence it is necessary to have $r^{(2)} \geq 2d + 6$ in (3.5) and thus $r^{(1)} \geq 4d + 13$ in (3.4).

Suppose now that the assumptions (3.6)–(3.11) of Section 3.1, and (3.18), (3.21) and (3.22) of Section 3.2, are all satisfied. Then it follows from [BL] (2012a, Theorem 4.7) that, for a sequence of initial conditions satisfying

$$x_N(0) \in \mathcal{X}', \ N \geq 1; \quad S_{2d+6}(x_N(0)) \leq C_* \text{ for some } C_* < \infty, \tag{3.24}$$

and

$$\|x_N(0) - x(0)\|_\mu = O(N^{-1/2}\sqrt{\log N}) \quad \text{for some } x(0) \in \mathcal{X}'_\mu, \tag{3.25}$$

the deterministic approximation (2.2) holds for any T, with

$$\varepsilon_N = k_T N^{-1/2} \sqrt{\log N} \quad \text{and} \quad P_T(N, \varepsilon_N) = k'_T N^{-1} \log N,$$

for suitably chosen constants k_T and k'_T. Note that equation (3.20) remains the same, whatever the values h_l, $1 \leq l \leq d$, chosen as lower bounds for x_{l0}^N. Hence, in view of this approximation, it follows that the event $\{\tau_{0,N} \leq T\}$ has probability at most $P_T(N, \varepsilon_N)$ if the h_l are chosen to satisfy $h_l \geq \sup_{0 \leq t \leq T} x_{l1} + \delta$ for each l, for some $\delta > 0$, whenever N is so large that $\varepsilon_N < \delta$. Thus, under the above conditions on the rates for the transitions I–VI, the results of Section 2 all hold, with the above values of ε_N and $P_T(N, \varepsilon_N)$. In particular, groups of patches or of animals of sizes $K_N = O(N^\alpha)$, for any $\alpha < 1/2$, behave asymptotically independently.

Remark 3.1 The assumptions concerning the transition rates are rather general, and cover many biologically useful models. They can be extended somewhat, as far as the permissible variation with x is concerned, by noting that the inequality (3.5), for $r \geq 1$, could be replaced by

$$V_r(x) \leq k_{r3}S_{p(r)}(x)(1 + S_0(x)) + k_{r5};$$

this would require only minor modification to the proof of [BL] (2012a, Theorem 2.4). For our purposes, the bounds in (3.10) and (3.11) could then be relaxed by multiplying their right hand sides by a factor $(\|x\|_1 + 1)$. However, it is not obvious that the inequality in (3.7) can be relaxed in this way, and this restricts the freedom for $\lambda_{\mathbf{ij}}(x)$ to vary with x.

4 Examples

4.1 Example 1: The finite patch size models of Metz & Gyllenberg (2001)

The first model, with N patches and just one variety of animal, has transitions of the form I–VI, with index set $\mathbb{Z}_+ \cup \{D\}$, where D is used here as index for the migrants (Metz & Gyllenberg use D to denote our x_D). In their notation, in a patch with i occupants, the birth rate is $\bar{\lambda}_{i,i+1} := i\lambda_i(1 - d_i)$, the death rate $\bar{\lambda}_{i,i-1} := i\mu_i$, the catastrophe rate $\bar{\lambda}_{i,0} := \gamma_i$ and the birth rate of (juvenile) migrants $\bar{\gamma}'_{iD} := i\lambda_i d_i$; here, $0 \leq d_i \leq 1$ for all i. The arrival rate of a migrant into an i-patch is $\sigma_{Di}(x) := \alpha s_i$, where $0 \leq s_i \leq 1$ for all i, and the death rate of a migrant is $\zeta_D := \mu_D$. All other transition rates are zero; in particular, there is none of the explicit dependence on x that would be allowed in our formulation, for functions such as $\lambda_{ij}(x)$.

We take $\nu(i) = \mu(i) = i + 1$, $i \in \mathbb{Z}_+$, and $\nu(D) = \mu(D) = 1$. Then assumption (3.6) is trivially satisfied, and (3.7) and (3.8) require λ_i to be bounded (so, as is reasonable, the *per capita* birth rate of an animal is to be bounded), in which case (3.10) is also satisfied. For (3.9), we require s_i to be bounded, which is satified since s_i are assumed to be probabilities. Condition (3.11) also involves γ_i and μ_i, and is satisfied if, in addition, μ_i and $i^{-1}\gamma_i$ are bounded in $i \geq 1$. The conditions (3.22) are trivially satisfied, and (3.21) is satisfied for

$$w := \sup_{i \geq 1}\{\lambda_i(1 - d_i) - \mu_i - i^{-1}\gamma_i + ((i-1)\lambda_{i-1}d_{i-1} - i\lambda_i d_i)\},$$

finite if also $u_i := (i-1)\lambda_{i-1}d_{i-1} - i\lambda_i d_i$ is bounded above in $i \geq 1$. The quantity u_i is the amount by which the total migration from a patch declines, when the number of individuals in the patch increases from $i-1$ to i, and for this to be bounded is again an entirely reasonable hypothesis. Finally, the quantities $D_Y(T, \delta)$ and $D_Z(T, \delta)$ are bounded, since the s_i are bounded. Hence, assuming that

$$\lambda_i, \mu_i, i^{-1}\gamma_i, \text{ and } u_i \text{ are bounded,} \tag{4.1}$$

our theorems apply to the initial model of Metz & Gyllenberg (2001), for initial conditions $x^N(0)$ satisfying (3.24) and (3.25). As it happens, the authors restricted their model by imposing a maximal number of animals per patch 'to make life easy', so that (4.1) is trivially satisfied in their context; but such a restriction is unnatural, and we have shown that it can be replaced by (4.1). Metz & Gyllenberg use the deterministic approximation $x := \{x(t), t \geq 0\}$ as the basis for their analysis, and this is justified over any fixed finite time interval $[0, T]$ by the discussion in Section 3, provided that N is large enough.

The results of Section 2 now show, in addition, that small groups of individuals behave almost independently of each other, according to time inhomogeneous Markov jump processes whose transition rates are determined by x. For a chosen patch $\mathcal{P}$, the Markov process has transition rates at time t given by

$$
\begin{array}{llll}
i & \to & i+1 & \text{at rate} \quad i\lambda_i(1 - d_i) + x_D(t)\alpha s_i, \quad & i \geq 0; \\
i & \to & i-1 & \text{at rate} \quad i\mu_i, & i \geq 2; \\
i & \to & 0 & \text{at rate} \quad \gamma_i + \mu_1\mathbf{1}_{\{1\}}(i), & i \geq 1.
\end{array} \tag{4.2}
$$

Any particular animal $\mathcal{A}$ is born either as a migrant, or in a patch. Once in a patch, it never migrates again. Its Markov process has transition rates at time t given by

$$
\begin{array}{llll}
(i, m) & \to & (i+1, m+1) & \text{at rate} \quad \lambda_i(1 - d_i); & i \geq 1 \\
(i, m) & \to & (i+1, m) & \text{at rate} \quad (i-1)\lambda_i(1 - d_i) + x_D(t)\alpha s_i; & i \geq 2 \\
(i, m) & \to & (i-1, m) & \text{at rate} \quad (i-1)\mu_i; & i \geq 2 \\
(i, m) & \to & (i, m+1) & \text{at rate} \quad \lambda_i d_i; & i \geq 1 \\
(i, m) & \to & (\Delta, m) & \text{at rate} \quad \mu_i + \gamma_i; & i \geq 1 \\
(D, 0) & \to & (i, 0) & \text{at rate} \quad \alpha x_{i-1}(t)s_{i-1}; & i \geq 1 \\
(D, 0) & \to & (\Delta, 0) & \text{at rate} \quad \mu_D.
\end{array} \tag{4.3}
$$

In either case, the process depends on $x(t)$ only through the arrival rates of migrants into patches.

The second model of Metz & Gyllenberg (2001) has animals of two different varieties, that interact through living in common patches, in that their *per capita* birth and death rates λ and μ and their migration parameters d and s vary with the entire composition (i_1, i_2) of the populations of the two varieties in a patch. Under assumptions analogous to (4.1), the deterministic process $\{x(t),\, t \geq 0\}$ with index set $Z_+^2 \cup \{D_1, D_2\}$ again acts as a good approximation to the random process x^N, and small groups of individuals and patches behave asymptotically almost independently. Sufficient conditions for this are bounded *per capita* birth, death, catastrophe and migrant arrival rates, together with u_{i_1, i_2} being bounded in $i_1, i_2 \geq 0$, where

$$u_{i,j} := (i-1)\lambda_{i-1,j}d_{i-1,j} - i\lambda_{ij}d_{ij} + (j-1)\lambda_{i,j-1}^* d_{i,j-1}^* - j\lambda_{ij}^* d_{ij}^*;$$

here, the starred quantities are those for the second variety, and the unstarred those for the first.

However, Metz & Gyllenberg are interested in using the approximation when just a small number of animals of the second variety have been introduced into a resident metapopulation consisting only of the first variety. Under such circumstances, the development of the introduced variety has an essentially random component — it may die out by chance, even if at a theoretical advantage — making it more reasonable to treat it as a small group of individuals, of a different variety, evolving at random among a resident population. The following discussion represents a theoretical justification for the analysis in Metz & Gyllenberg (2001, Section 2(d)).

We begin by choosing $x^N(0) = \tilde{x}^N(0) + N^{-1}K_N e_{D_2}$, where $\tilde{x}^N(0)$ is an initial composition consisting only of individuals of the first variety, and $\|\tilde{x}^N(0) - \tilde{x}(0)\|_\mu = O(N^{-1/2}\sqrt{\log N})$ for some fixed $\tilde{x}(0) \in \mathcal{X}'_\mu$, which thus also consists only of 1-individuals. Then, in the transition rates for any Markov process approximating individual dynamics, the argument $x(t)$ can be taken to be $\tilde{x}(t)$, where $\tilde{x}$ denotes the solution of (3.20) starting at $\tilde{x}(0)$, provided that $K_N = O(N^\beta)$ for any $\beta < 1/2$, because then $\|x^N(0) - \tilde{x}(0)\|_\mu = O(N^{-1/2}\sqrt{\log N})$ also. But since $\tilde{x}(0)$ consists only of 1-individuals, so does $\tilde{x}(t)$ for all $t > 0$, and $\tilde{x}(t)$ is the solution to the deterministic equation for the initial model of Metz & Gyllenberg (2001), with the parameters of the resident population.

Since a 2-juvenile, once arrived in a patch, never leaves it, the development of the introduced species is best described in terms of the evolution

of the patches that 2-juveniles reach. Each such patch can be treated as an 'individual', and the 2-migrants that leave it as its offspring, up to the time at which the patch contains no more 2-individuals. This patch process, of a 'p-individual', can thus be interpreted as a life history process W, beginning with the juvenile 2-migrant, whose offspring are the 2-migrants that leave its chosen patch. The entire process begins with a group of K_N juvenile 2-migrants, and the 2-migrant offspring of the resulting p-individuals in turn initiate new W-processes, so that the entire process, if the bound deduced from Corollary 2.4 is small, can be approximated by a Crump–Mode–Jagers (CMJ) branching process (Crump & Mode (1968a,b), Jagers (1968); see also Jagers (1975, Chapter 6)).

Let $W(t) = ((i, j), m)$ indicate that, at time t, the patch contains i 1-individuals and j 2-individuals, and that m 2-migrants have left the patch up to time t; if (i, j) is replaced by Δ, this indicates that the initial juvenile and all of its offspring that did not migrate, if there were any, have died, and D_2 is used when the state consists of the single juvenile 2-migrant, before it reaches a patch. The transition rates of W at time t can then be expressed as

$$
\begin{array}{llll}
((i,j),m) & \to & ((i,j+1),m) & \text{at rate} \quad j\lambda_{ij}^*(1 - d_{ij}^*); & i \geq 0,\ j \geq 1 \\
((i,j),m) & \to & ((i+1,j),m) & & \\
& & \quad \text{at rate} \quad i\lambda_{ij}(1 - d_{ij}) + \tilde{x}_D(t)\alpha s_{ij}; & i \geq 1,\ j \geq 1 \\
((i,j),m) & \to & ((i-1,j),m) & \text{at rate} \quad i\mu_{ij}; & i \geq 1,\ j \geq 1 \\
((i,j),m) & \to & ((i,j),m+1) & \text{at rate} \quad j\lambda_{ij}^*d_{ij}^*; & i \geq 0,\ j \geq 1 \\
((i,j),m) & \to & ((i,j-1),m) & \text{at rate} \quad j\mu_{ij}^*; & i \geq 0,\ j \geq 2 \\
((i,j),m) & \to & (\Delta, m) & \text{at rate} \quad \mu_1^*\mathbf{1}_{\{1\}}(j) + \gamma_{ij}; & i \geq 0,\ j \geq 1 \\
(D_2, 0) & \to & ((i,1),0) & \text{at rate} \quad \alpha\tilde{x}_i(t)s_{i0}^*; & i \geq 0 \\
(D_2, 0) & \to & (\Delta, 0) & \text{at rate} \quad \mu_D^*. &
\end{array}
$$

(4.4)

In particular, if the resident population started at an equilibrium of the deterministic equations, so that $\tilde{x}(t) = \tilde{x}(0)$ for all t, then these transition rates are time homogeneous. Note also that, since the *per capita* birth rate of the second variety is uniformly bounded over all patch compositions, comparison with a linear pure birth process shows that the expectation of the square of the number of 2-individuals that were ever alive during $[0, T]$ is bounded by $c_T K_N^2$, for a suitable $c_T < \infty$. Hence the probability that any 2-migrant, whenever it was born, arrives during $[0, T]$ in a patch which has already been

20

visited by individuals of the second variety is of order $O(N^{-1}K_N^2)$, and this is asymptotically small if $K_N = O(N^\beta)$ for any $\beta < 1/2$.

Thus, in view of Corollary 2.4, the evolution of the introduced species over any finite time interval $[0, T]$, measured in terms of the number of juvenile migrants, is the same as that of a CMJ–branching process, with probability of order $O(N^{-1+2\beta})$. The individual life history consists of a period of migration, followed either by death (with probability μ_D^*/S, where $S := \mu_D^* + \sum_{i \geq 0} \alpha \tilde{x}_i(0) s_{i0}^*$) or arrival in a patch (of type $(i, 0)$ with probability $\alpha \tilde{x}_i(0) s_{i0}^*/S$), after which its subsequent life history follows that of the Markov process with rates (4.4), started in the state $((i, 1), 0)$. In particular, each transition of this process in which the third component increases corresponds to the birth of a new juvenile migrant. If $P(i, j, t)$ denotes the probability $\mathbb{P}[(W_1(t), W_2(t)) = (i, j) \mid W(0) = (D_2, 0)]$, then the mean intensity of the offspring process is $m(t) := \sum_{i \geq 0} \sum_{j \geq 1} P(i, j, t) j \lambda_{ij}^* d_{ij}^* \, dt$, and the mean number of offspring is $\bar{m} := \int_0^\infty m(t) \, dt \leq \infty$.

The approximation using a branching process gives a lot of insight into the development of the introduced species. In particular, if the equation $\int_0^\infty e^{-\rho t} m(t) \, dt = 1$ has a solution $\rho > 0$ (which has to be the case if $1 < \bar{m} < \infty$), then the introduced species, if it becomes established, grows exponentially with rate ρ, and the probability that it becomes established from an initial population of K juvenile migrants is $1 - q^K$, where q is the extinction probability of the Galton–Watson process, starting with a single individual, whose offspring distribution is the distribution of the total number of offspring in the CMJ–process. If $\bar{m} \leq 1$, the introduced species dies out with probability one. However, the current theorems only guarantee this approximation to be valid over a fixed time interval $[0, T]$, and then for N sufficiently large. In Barbour, Hamza, Kaspi & Klebaner (2013), the development of an introduced species, including the branching approximation, is considered over much longer time intervals, but in the context of *finite dimensional* Markov population processes. It would be interesting to establish analogous results in the current context.

Metz & Gyllenberg (2001) made the (intuitively obvious) conjecture that, if the introduced species has exactly the same parameters as the original, and is introduced in equilibrium, then $\bar{m} = 1$. This is equivalent to saying that, in equilibrium, each migrant generates a process that results in an average of exactly one new migrant. They were, however, unable to give a proof of this. If the random process for finite N were ergodic, it would be natural to use

21

arguments based on long term time averages as the basis of a proof. However, the finite N process is eventually absorbed in the zero population extinction state, so such arguments cannot be used. However, we sketch a proof of the conjecture, under assumptions that include those of Metz & Gyllenberg, in the appendix.

4.2 Example 2: Kretzschmar's (1993) model

In Kretzschmar's (1993) model of parasitic infection, N denotes the initial number of hosts, these playing the role of patches. The index $i \in \mathbb{Z}_+$ denotes the number of parasites living in the host. The model has transitions of the form I–VI, with $\lambda_{i,i-1} := i\mu$, $\lambda_{i,i+1} := \lambda\varphi(x)$, $\beta_0(x) := \beta \sum_{i \geq 0} x_i \theta^i$ and $\delta_i := \kappa + i\alpha$, all other transition rates being zero; here, $0 \leq \theta \leq 1$, and $\varphi(x) := \sum_{j \geq 1} j x_j / (c + \|x\|_1)$ for some $c > 0$. It is shown in [BL] (2012a, Example 5.1) that, if the initial conditions satisfy (3.24) and (3.25), then the law of large numbers approximation (2.2) holds with $\varepsilon_N = k_T N^{-1/2}\sqrt{\log N}$ and $P_T(N, \varepsilon_N) = k_T' N^{-1} \log N$, for suitably chosen constants k_T, k_T', where, as usual, $\mu(i) = i + 1$. It is also easy to check that $D_Y(T, \delta) < \infty$ for all T and δ. The patch process Y on $\mathbb{Z}_+ \cup \Delta$ has transition rates at time t given by

$$
\begin{array}{llll}
i & \to & i+1 \quad \text{at rate} \quad \lambda\varphi(x(t)), & i \geq 0; \\
i & \to & i-1 \quad \text{at rate} \quad i\mu, & i \geq 1; \\
i & \to & \Delta \quad\;\; \text{at rate} \quad \kappa + i\alpha, & i \geq 0.
\end{array}
\tag{4.5}
$$

One way of looking at this process is as a superposition of Poisson processes. Each parasite on arrival decides independently either to die or to kill the host, with probabilities $\mu/(\mu + \alpha)$ and $\alpha/(\mu + \alpha)$ respectively. The time of this event is exponentially distributed with mean $1/(\mu + \alpha)$. Independently, the host is killed after an exponentially distributed time with mean $1/\kappa$. Because of the independence of marked Poisson streams, given that the host is alive at time T, the number of parasites living in it has a Poisson distribution with mean

$$
\int_0^T \lambda\varphi(x(t))e^{-(\mu+\alpha)(T-t)} \, dt.
$$

Thus a cohort consisting of K_N hosts of given age T would exhibit an approximately Poisson distribution of parasites per host, if $K_N = O(N^\gamma)$ for some $\gamma < 1/2$. Thus, within age classes, Kretzschmar's model does not generate over-dispersed distributions of parasites per host, though mixing over

age classes in a sample may be expected to do so. Even then, if α and κ are much smaller than μ, and x is in equilibrium, the departure from Poisson may not be very noticeable, unless there are many young hosts (with ages comparable to $1/\mu$) in the sample.

Appendix

In this section, we establish the conjecture of Metz & Gyllenberg (2001) discussed above. For this purpose, we can take their single type model, since all individuals behave in the same way. Let Z denote the CMJ-branching process associated with the process W of Example 1, when the underlying process x is in equilibrium. Suppose first that its mean $\bar{m}$ exceeds 1, so that its extinction probability q is less than 1. In this case, given any $M > 0$, there exists a finite time T_M such that

$$\mathbb{P}_1[Z(T_M) > M] \; > \; (1 - q)/2,$$

where $\mathbb{P}_1$ denotes probability starting from a single migrant. Starting the x_N-process close to the equilibrium $\bar{x}$, there are $d_N \approx N\bar{x}_D$ migrants at time 0. We assume that $\bar{x}_D > 0$, which is true, for instance, under the irreducibility condition introduced below. Let Z_N^j denote the process of migrant descendants of the j-th of them. As noted above, it has distribution close to that of Z for large N, by Theorem 2.3. Set $I_j := I[Z_N^j(T_M) > M]$, and let N be so large that $\mathbb{E}I_j =: p_N > (1 - q)/2$. Then, because any two of the processes Z_N^j and Z_N^k, $k \neq j$, are asymptotically independent as $N \to \infty$, by Corollary 2.4, it follows that $\mathbb{E}(I_j I_k) = p_N^2 + o(1)$ as $N \to \infty$, implying in turn that $S_N := \sum_{j=1}^{d_N} I_j$ has $\mathbb{E}S_N > N\bar{x}_D(1 - q)/2$ for all N large enough, and that $\operatorname{Var} S_N = o(N^2)$. Thus, by Chebyshev's inequality, $\mathbb{P}[MS_N \geq MN\bar{x}_D(1-q)/4] \to 1$ as $N \to \infty$. But this contradicts (2.2) if M is chosen such that $M(1 - q)/4 > 1$, because $MS_N \leq Nx_D^N(T_M)$, and (2.2) implies that $\mathbb{P}[Nx_D^N(T_M) \leq N\bar{x}_D(1 + \varepsilon)] \to 1$ for any $\varepsilon > 0$.

The proof of contradiction if $\bar{m} < 1$ is more involved. Recall that $m(\cdot)$ denotes the mean offspring measure of the CMJ-branching process $W =: W_0$ starting with $W_0(0) = (D_2, 0)$. Let $m_i(\cdot)$ denote the mean offspring measure for the initial individual in the process W_i, starting with $W_i(0) = ((i, 1), 0)$. All of its migrant children have mean offspring measure m, but the initial individual in general has a different measure. Let $n_i(t)$ denote the mean number of migrants alive at time t in the process W_i. Then $n_i(t) =$

$\int_0^t m_i(dv)n_0(t-v)$. The assumption $\bar{m} < 1$ implies that $\lim_{t\to\infty} n_0(t) = 0$, and so $\lim_{t\to\infty} \sum_{i\geq0} i\bar{x}_i n_i(t) = 0$ also, if $\sum_{i\geq0} i\bar{x}_i \int_0^\infty m_i(dv) < \infty$, by dominated convergence. The latter is true, if $\sum_{i\geq0} i\bar{x}_i < \infty$ and if $\sup_i \int_0^\infty m_i(dv) = m^* < \infty$.

We now make four assumptions. The first three are that $0 < \sum_{i\geq0} i\bar{x}_i < \infty$, that $\lambda^* := \sup_i \lambda_i < \infty$, and that, for some $\varepsilon > 0$, there exists i_0 such that $\mu_i + \gamma_i \geq \lambda_i\{1 - (1-\varepsilon)d_i\}$ for all $i \geq i_0$. The fourth is an irreducibility assumption: we require that the birth, death and catastrophe rates are such that a patch with $i \geq 1$ occupants can evolve into a patch with $i' \geq 0$ occupants, for $i \neq i' \leq i^*$, where i^* is the maximum possible number of occupants of a patch (infinity, if there is no maximum); that $s_0 > 0$; and that $\lambda_i d_i > 0$ for some $i \geq 1$.

The second of the assumptions ensures that mean proportion of the contribution to $\mathbb{E}X_D^N(t)$ arising from individuals in $X^N(0)$ whose family trees do *not* remain coupled to the corresponding branching process over any fixed interval $[0,T]$ is asymptotically negligible as $N \to \infty$, for T fixed: the worst contribution from any such individual is $\exp\{\lambda^*T\}$, and the proportion of them is asymptotically negligible as $N \to \infty$, by Theorem 2.3. The fourth assumption, together with $\bar{m} < 1$, ensures that $\int_0^\infty m_i(dv) < \infty$ for each i, since there is then a positive probability that a migrant is at some time in a patch with $i - 1$ other occupants, and its total mean number of migrant offspring is finite. The third assumption ensures that $m^* < \infty$. This can be proved by analyzing a system of recurrence equations satisfied by the quantities $\int_0^\infty m_i(dv)$, showing that, in $i \geq i_0$, $\int_0^\infty m_i(dv)$ is uniformly bounded by a quantity of the form $c_1 + c_2 \int_0^\infty m_{i_0}(dv)$. This, combined with the first assumption, shows that the contribution to $N^{-1}\mathbb{E}X_D^N(t)$ arising from individuals for which the coupling is maintained over $[0,T]$ is asymptotically close to $\sum_{i\geq0} i\bar{x}_i n_i(T)$ as $N \to \infty$, which can be made as small as desired by choosing T large enough. Furthermore, because $\lambda^* < \infty$, the variance of the contribution to $X_D^N(T)$ from any individual is uniformly bounded in i, and the correlation between the contributions from pairs of different individuals is asymptotically small in N, by Corollary 2.4. Hence, with ever higher probability as $N \to \infty$, $N^{-1}X_D^N(T)$ stays close to its (small) expectation. However, for x_N in equilibrium, it has also to be asymptotically close to the fixed value $\bar{x}_D$, by (2.2), and this is a contradiction, if $\bar{x}_D > 0$; and this is the case, because of the fourth assumption.

Metz & Gyllenberg (2001) actually assume that there is a largest index

$i^* < \infty$. In this case the conditions are typically satisfied, if i_0 is taken equal to the largest index i^* in the third assumption. However, there are some trivial possibilities where their conjecture is not true. For instance, if $i^* = 1$ and $\mu_1 + \gamma_1 = 0$ and $\lambda_1 > 0$ (in which case, from the definition of i^*, $d_1 = 1$, and also $\sigma_1 = 0$), and if $\lambda_0 > 0$, one would have $\bar{x}_0 = 0$, $\bar{x}_1 = 1$ and $\bar{m} = 0$, but $\bar{x}_D = \lambda_1/\mu_D > 0$. Of course, this is a biologically implausible scenario, and it violates both the third and fourth assumptions.

Acknowledgement

References

[1] F. ARRIGONI (2003) Deterministic approximation of a stochastic meta-population model. *Adv. Appl. Probab.* **35**, 691–720. MR1990610

[2] A. D. BARBOUR, K. HAMZA, H. KASPI & F. C. KLEBANER (2013) Escape from the boundary in Markov population processes. `arXiv:1312.5788`

[3] A. D. BARBOUR & M. KAFETZAKI (1993) A host–parasite model yielding heterogeneous parasite loads. *J. Math. Biol.* **31**, 157–176.

[4] A. D. BARBOUR & M. J. LUCZAK (2008) Laws of large numbers for epidemic models with countably many types. *Ann. Appl. Probab.* **18**, 2208–2238. MR2473655

[5] A. D. BARBOUR & M. J. LUCZAK (2012a) A law of large numbers approximation for Markov population processes with countably many types. *Prob. Theory Rel. Fields* **153**, 727–757. MR2948691

[6] A. D. Barbour & M. J. Luczak (2012b) Central limit approximations for Markov population processes with countably many types. *Electr. J. Probab.* **17**, no. 90, 16pp.

[7] K. S. Crump & C. J. Mode (1968a) A general age-dependent branching process, I. *J. Math. Anal. Appl.* **24**, 494–508.

[8] K. S. Crump & C. J. Mode (1968b) A general age-dependent branching process, II. *J. Math. Anal. Appl.* **25**, 8–17.

[9] I. Hanski & M. Gilpin (1991) Metapopulation dynamics: brief history and conceptual domain. *Biol. J. Linnean Soc.* **42**, 3–16.

[10] P. Jagers (1968) Age-dependent branching processes allowing immigration. *Teor. Verojatnost. i Primenen* **13**, 230–242.

[11] P. Jagers (1975) *Branching processes with biological applications.* Wiley, New York.

[12] M. Kretzschmar (1993) Comparison of an infinite dimensional model for parasitic diseases with a related 2-dimensional system. *J. Math. Analysis Applics* **176**, 235–260. MR1222167

[13] T. G. Kurtz (1970) Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Probab.* **7**, 49–58. MR0254917

[14] T. G. Kurtz (1971) Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J. Appl. Probab.* **8**, 344–356. MR0287609

[15] C. Léonard (1990) Some epidemic systems are long range interacting particle systems. In: *Stochastic Processes in Epidemic Theory*, Eds J. P. Gabriel, C. Lefèvre and P. Picard, Lecture Notes in Biomathematics 86, pp. 170–183. Springer, New York.

[16] R. Levins (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bull. Entomol. Soc. Amer.* **15**, 237–240.

[17] C. J. Luchsinger (2001a) Stochastic models of a parasitic infection, exhibiting three basic reproduction ratios. *J. Math. Biol.* **42**, 532–554. MR1845591

[18] C. J. Luchsinger (2001b) Approximating the long term behaviour of a model for parasitic infection. *J. Math. Biol.* **42**, 555–581. MR1845592

[19] J. A. J. Metz & M. Gyllenberg (2001) How should we define fitness in structured metapopulation models? Including an application to the calculation of evolutionarily stable dispersal strategies. *Proc. Roy. Soc.* B **268**, 499–508.