Electronic Journal of Probability

Electron. J. Probab. **17** (2012), no. 90, 1–16.
ISSN: 1083-6489 DOI: 10.1214/EJP.v17-1760

Central limit approximations for Markov population processes with countably many types*

A. D. Barbour† M. J. Luczak‡

Abstract

When modelling metapopulation dynamics, the influence of a single patch on the metapopulation depends on the number of individuals in the patch. Since there is usually no obvious natural upper limit on the number of individuals in a patch, this leads to systems in which there are countably infinitely many possible types of entity. Analogous considerations apply in the transmission of parasitic diseases. In this paper, we prove central limit theorems for quite general systems of this kind, together with bounds on the rate of convergence in an appropriately chosen weighted ℓ_1 norm.

Keywords: Epidemic models ; metapopulation processes ; countably many types ; central limit approximation ; Markov population processes.
AMS MSC 2010: 92D30 ; 60J27 ; 60B12.
Submitted to EJP on January 26, 2012, final version accepted on September 5, 2012.

1 Introduction

Metapopulations, introduced by Levins (1969), are used to describe the evolution of the population of a species in a fragmented habitat. The metapopulation consists of a number of distinct patches, together with (a summary of) the population present in each patch, and its development over time is governed by specified within and between patch dynamics. In the Markovian structured mean-field metapopulation model of Arrigoni (2003), the state of the system consists simply of the numbers of individuals in each patch. Individuals reproduce within patches and migrate between patches, and each patch is subject to random catastrophes, which reduce its population to zero. Letting N be the total number of patches, thought of as being large, and letting $X_t^{N,i}$ denote the number of patches with i individuals at time t, the transitions out of state $X \in \mathbb{Z}_+^{\mathbb{Z}_+}$

*ADB was supported in part by Schweizerischer Nationalfonds Projekt Nr. 20–107935/1 and by Australian Research Council Grants Nos DP120102728 and DP120102398. MJL was supported in part by an EPSRC Leadership Fellowship and by Australian Research Council Grant No DP120102398.
†Universität Zürich, Switzerland. E-mail: a.d.barbour@math.uzh.ch
‡University of Sheffield, England. E-mail: m.luczak@qmul.ac.uk

to states $X + J$ in her model are as follows:

$$\begin{aligned}
J &= e^{(i-1)} - e^{(i)} && \text{at rate } Nix^i(d_i + \gamma(1-\rho)), && i \geq 2;\\
J &= e^{(0)} - e^{(1)} && \text{at rate } Nx^1(d_1 + \gamma(1-\rho) + \kappa); && \\
J &= e^{(i+1)} - e^{(i)} && \text{at rate } Nix^i b_i, && i \geq 1;\\
J &= e^{(0)} - e^{(i)} && \text{at rate } Nx^i\kappa, && i \geq 2;\\
J &= e^{(k+1)} - e^{(k)} + e^{(i-1)} - e^{(i)} && \text{at rate } Nix^i x^k \rho\gamma, \quad k \geq 0, && i \geq 1;
\end{aligned}$$

here, $x := N^{-1}X$, and $e^{(i)}$ denotes the unit vector along the i coordinate. The total number $N := \sum_{j\geq 0} X^{N,j}$ of patches remains constant throughout, and each transition involves at most 4 components of X. The *per capita* death and birth rates $(d_i), (b_i)$ within each patch are allowed to vary with the current population size i, but in the same way in all patches. They would usually be chosen to correspond to one of the traditional single species demographic models. We shall always assume that the b_i are bounded above, which is a biologically realistic requirement, and excludes the possibility that the population in a patch could become infinite in finite time. The *per capita* migration rate γ is taken to be the same for all individuals, as is the probability ρ that a migration is successful, and a successful migrant chooses its new patch uniformly at random. Each patch is independently subject to catastrophes at the same rate κ.

If there were an absolute upper limit for the number of individuals in each patch, the model would be a finite dimensional Markov population process. The behaviour of these finite dimensional models can be approximated using the methods pioneered by Kurtz (1970, 1971), who was able to establish a law of large numbers approximation, in the form of a system of ordinary differential equations, and a corresponding diffusion approximation. However, there are no upper limits on population number in the usual single population models, and it is the stochastic evolution according to the rules of the model that dictates the region in which population numbers typically lie. Thus it seems unnatural to introduce an *a priori* upper limit in the system above, just because more than one population is being modelled. Similar considerations surface in a number of other population models, including the epidemic models of Luchsinger (2001a,b) and Kretzschmar (1993), and models involving the copy number of genes in cells in Kimmel & Axelrod (2002, Chapter 7). Instead, it makes sense to consider Markov population processes with a countably infinite number of dimensions as models in their own right.

A law of large numbers in a general setting of this kind was first established by Eibeck & Wagner (2003). Under appropriate conditions, Barbour & Luczak (2008, 2011) strengthened the law of large numbers by providing an error bound, in a weighted ℓ_1 norm, that is close to optimal order in N. In this paper, these latter results are complemented by a central limit approximation, together with a corresponding error estimate.

Our general setting, as in Barbour & Luczak (2011) [BL], is that of families of Markov population processes $X^N := (X_t^N,\, t \geq 0)$, $N \geq 1$, taking values in the countable space $\mathcal{X}_+ := \{X \in \mathbb{Z}_+^{\mathbb{Z}_+};\, \sum_{m\geq 0} X^m < \infty\}$. The component $X_t^{N,j}$ of X_t^N represents the number of individuals of type j that are present at time t, and there are countably many types possible; however, at any given time, there are only finitely many individuals in the system. The process evolves as a Markov process with state–dependent transitions

$$X \;\to\; X + J \quad \text{at rate} \quad N\alpha_J(N^{-1}X), \qquad X \in \mathcal{X}_+,\; J \in \mathcal{J}, \tag{1.1}$$

where each jump is of bounded influence, in the sense that

$$\mathcal{J} \subset \Big\{X \in \mathbb{Z}^{\mathbb{Z}_+};\, \sum_{m\geq 0} |X^m| \leq J_* < \infty\Big\}, \quad \text{for some fixed} \quad J_* < \infty, \tag{1.2}$$

so that the number of individuals affected at each transition is uniformly bounded. Density dependence is reflected in the fact that the arguments of the functions α_J are counts normalised by the 'typical size' N. Write $\mathcal{R} \subset R_+^{\mathbb{Z}_+}$ for the domain of the functions α_J, which are assumed to satisfy

$$\sum_{J \in \mathcal{J}} \alpha_J(\xi) \;<\; \infty, \qquad \xi \in \mathcal{R}_0, \tag{1.3}$$

where $\mathcal{R}_0 := \{\xi \in \mathcal{R}\colon\ \xi_i = 0 \text{ for all but finitely many } i\}$; this assumption implies that the processes X^N are indeed pure jump processes, at least for some (possibly random) non-zero length of time. To prevent the paths leaving $\mathcal{X}_+$, we also assume that $J^l \geq -1$ for each l, and that $\alpha_J(\xi) = 0$ if $\xi^l = 0$ for any $J \in \mathcal{J}$ such that $J^l = -1$.

In the finite dimensional case, the law of large numbers is expressed in terms of the system of *deterministic equations*

$$\frac{d\xi}{dt} \;=\; \sum_{J \in \mathcal{J}} J\alpha_J(\xi). \tag{1.4}$$

In [BL], it is assumed that

$$\sum_{J \in \mathcal{J}} J\alpha_J(\xi) \;=\; A\xi + F(\xi), \tag{1.5}$$

where A is a constant $\mathbb{Z}_+ \times \mathbb{Z}_+$ matrix, and (1.4) is then treated as a perturbed linear system (Pazy 1983, Chapter 6). Under suitable assumptions on A, there exists a measure μ on $\mathbb{Z}_+$, defining a weighted ℓ_1 norm $\|\cdot\|_\mu$ on $\mathcal{R}$, and a strongly $\|\cdot\|_\mu$–continuous semigroup $\{R(t),\, t \geq 0\}$ of transition matrices having pointwise derivative $R'(0) = A$. If F is locally $\|\cdot\|_\mu$–Lipschitz, the solution x of the integral equation

$$x_t \;=\; R(t)x_0 + \int_0^t R(t-s)F(x_s)\,ds\,, \tag{1.6}$$

for $\|x_0\|_\mu < \infty$, replaces that of (1.4) as an approximation to $x^N := N^{-1}X^N$.

Under suitable conditions, it is shown in [BL, Theorem 4.7] that

$$\sup_{0 \leq t \leq T} \|x_t^N - x_t\|_\mu \;=\; O(N^{-1/2}\sqrt{\log N}),$$

except on an event of probability of order $O(N^{-1}\log N)$, provided that $\|x_0^N - x_0\|_\mu = O(N^{-1/2}\sqrt{\log N})$. The conditions under which this approximation holds can be divided into three categories: growth conditions on the transition rates, so that the *a priori* bounds, which have the character of moment bounds, can be established; conditions on the matrix A, sufficient to limit the growth of the semigroup R, and (together with the properties of F) to determine the weights defining the metric in which the approximation is to be carried out; and conditions on the initial state of the system. The conditions are described in the next section. They are all needed in the current paper, too, in which we investigate the difference $x_t^N - x_t$ in greater detail.

Our main result, Theorem 6.1, shows that, under some extra conditions, it is possible to construct a diffusion process Y on the same probability space as X^N in such a way that

$$\sup_{0 \leq t \leq T} \|N^{1/2}(x_t^N - x_t) - Y_t\|_\mu \;=\; O(N^{-b_1}), \tag{1.7}$$

except on an event of probability of order $O(N^{-b_2})$, for specific values of b_1 and b_2. With the best possible control of moments, as for the model of Arrigoni (2003) mentioned above, one can take any $b_1 < 1/4$ and any $b_2 < 1$, provided that the initial conditions are appropriately chosen. The process Y can be interpreted as the infinite dimensional

analogue of the diffusion approximation in Kurtz (1971), satisfying the formal stochastic differential equation

$$dY_t \;=\; \{A + DF(x_t)\}Y_t\,dt + dW_t. \tag{1.8}$$

Here, dW is a time–inhomogeneous white noise process with infinitesimal covariance matrix $\sigma^2(t) := \sum_{J\in\mathcal{J}} JJ^T\alpha_J(x_t)$, and Y has time–inhomogeneous linear drift with coefficient matrix $A + DF(x_t)$. In particular, if $\bar{x}$ is an equilibrium of the deterministic equations, satisfying $A\bar{x} + F(\bar{x}) = 0$, then Y is an infinite dimensional Ornstein–Uhlenbeck process, with constant drift coefficient matrix $A + DF(\bar{x})$ and infinitesimal covariance matrix $\sum_{J\in\mathcal{J}} JJ^T\alpha_J(\bar{x})$.

Basic approach

The structure of the argument is as follows. It is shown in [BL, (4.8)] that, under suitable conditions, the process x^N satisfies an equation very similar to (1.6):

$$x_t^N \;=\; R(t)x_0^N + \int_0^t R(t-s)F(x_s^N)\,ds + \widetilde{m}_t^N, \tag{1.9}$$

where

$$\widetilde{m}_t^N \;:=\; \int_0^t R(t-s)\,dm_s^N \;=\; m_t^N + \int_0^t R(t-s)Am_s^N\,ds, \tag{1.10}$$

and

$$m_t^N \;:=\; x_t^N - x_0^N - \int_0^t \{Ax_s^N + F(x_s^N)\}\,ds \tag{1.11}$$

is a local martingale. Taking the difference between (1.9) and (1.6), and multiplying by $\sqrt{N}$, gives

$$U_t^N \;=\; R(t)U_0^N + \int_0^t R(t-s)\,DF(x_s)[U_s^N]\,ds + \eta_t^N + N^{1/2}\widetilde{m}_t^N, \tag{1.12}$$

with $U_t^N := N^{1/2}\{x_t^N - x_t\}$ and

$$\eta_t^N \;:=\; N^{1/2}\int_0^t R(t-s)\{F(x_s^N) - F(x_s) - DF(x_s)[N^{-1/2}U_s^N]\}\,ds. \tag{1.13}$$

Starting with this representation of U^N, the first step is to show that η^N is uniformly small with high probability, so that the randomness in U^N is driven principally by the process $N^{1/2}\widetilde{m}^N$. This quantity is in turn determined, through (1.10), by the local martingale $N^{1/2}m^N$. The next step is to show that $N^{1/2}m^N$ is close to a diffusion W, formally expressible as

$$W_t \;:=\; \sum_{J\in\mathcal{J}} JW_J(A_J(t)), \tag{1.14}$$

where the $\{W_J,\, J\in\mathcal{J}\}$ are independent standard Brownian motions, and we define $A_J(t) := \int_0^t \alpha_J(x_s)\,ds$: this is the diffusion appearing in (1.8). Analogously to (1.10), we then show that we can define a process $\widetilde{W}$ such that

$$\widetilde{W}_t \;:=\; W_t + \int_0^t R(t-s)AW_s\,ds, \tag{1.15}$$

and that $\widetilde{W}$ is close to $N^{1/2}\widetilde{m}^N$. Finally, returning to (1.12), we show that U^N is close to the solution Y to the analogous equation

$$Y_t \;=\; R(t)Y_0 + \int_0^t R(t-s)\,DF(x_s)[Y_s]\,ds + \widetilde{W}_t, \tag{1.16}$$

which in turn can be shown to exist and be unique. The random process solving (1.16) at first sight seems rather mysterious. However, partial integration represents $\widetilde{W}_t$ in the form $\int_0^t R(t-s)\,dW_s$, and so the expression for Y can indeed be interpreted as the variation of constants representation of the solution to the formal stochastic differential equation (1.8).

In the remaining sections, this programme is carried out in detail. Section 2 is concerned with specifying the conditions under which the main theorem is true, and with recalling the results from [BL] that are needed here. In the subsequent sections, the steps sketched above are examined in turn.

2 Assumptions and preliminaries

We assume henceforth that (1.2) and (1.3) are satisfied. Since the index $j \in \mathbb{Z}_+$ is symbolic in nature, we fix an $\nu \in \mathcal{R}$, such that $\nu(j)$ reflects in some sense the 'size' of j:

$$\nu(j) \geq 1 \text{ for all } j \geq 0 \quad \text{and} \quad \lim_{j\to\infty} \nu(j) = \infty. \tag{2.1}$$

We then assume that most indices are large and that most transitions involve some large indices, in the sense that, for $\mathcal{T}_M := \{j\colon\ \nu(j) \leq M\}$ and $\mathcal{J}_M := \{J\colon\ J^j = 0 \text{ for all } j \notin \mathcal{T}_M\}$, we have

$$|\mathcal{T}_M| \leq n_1 M^{\beta_1}; \qquad |\mathcal{J}_M| \leq n_2 M^{\beta_2}, \qquad M \geq 1, \tag{2.2}$$

for some n_1, n_2 and $\beta_1 \leq \beta_2$; note that in fact $\beta_2 \leq 2\beta_1 J_*$ also. As a consequence of these assumptions, for any $s > \beta_1$, there exists $K_s < \infty$ such that

$$\sum_{j\geq 0} \nu(j)^{-s} \ < \ \infty; \qquad \sum_{j\notin\mathcal{T}_M} \nu(j)^{-s} \ < \ K_s M^{-s+\beta_1}; \tag{2.3}$$

moreover, if $\nu(J) := \max_{\{j\colon\ J^j\neq 0\}} \nu(j)$, then, for any $s > \beta_2$,

$$\sum_{J\in\mathcal{J}} \nu(J)^{-s} \ < \ \infty; \qquad \sum_{J\notin\mathcal{J}_M} \nu(J)^{-s} \ < \ K'_s M^{-s+\beta_2}, \tag{2.4}$$

for some $K'_s < \infty$.

Moment assumptions

In the proofs that follow, it is important to be able to show that x^N is largely concentrated on indices j with $\nu(j)$ not too large. This is shown to be the case in [BL, Section 2], under the following 'moment' assumptions. Defining $S_r(x) := \sum_{j\geq 0} x^j \{\nu(j)\}^r$, $x \in \mathcal{R}_0,\ r \geq 0$, and then

$$U_r(x) \ := \ \sum_{J\in\mathcal{J}} \alpha_J(x)\Big(\sum_{j\geq 0} J^j\{\nu(j)\}^r\Big); \quad V_r(x) \ := \ \sum_{J\in\mathcal{J}} \alpha_J(x)\Big(\sum_{j\geq 0} J^j\{\nu(j)\}^r\Big)^2, \tag{2.5}$$

$x \in \mathcal{R}$, the assumptions that we need are as follows.

Assumption 2.1. *For ν as above, assume that there exist $r^{(1)}_{\max}, r^{(2)}_{\max} \geq 1$ such that, for all $\xi \in \mathcal{R}_0$,*

$$\sum_{J\in\mathcal{J}} \alpha_J(\xi) \left|\sum_{j\geq 0} J^j\{\nu(j)\}^r\right| \ < \ \infty, \qquad 0 \leq r \leq r^{(1)}_{\max}, \tag{2.6}$$

and also that, for some non-negative constants k_{rl}, the inequalities

$$\begin{aligned} U_0(x) &\leq k_{01}S_0(x) + k_{04}, \\ U_1(x) &\leq k_{11}S_1(x) + k_{14}, \qquad\qquad 2 \leq r \leq r^{(1)}_{\max}, \\ U_r(x) &\leq \{k_{r1} + k_{r2}S_0(x)\}S_r(x) + k_{r4}, \end{aligned} \tag{2.7}$$

and

$$\begin{aligned} V_0(x) &\le k_{03}S_1(x) + k_{05}, \\ V_r(x) &\le k_{r3}S_{p(r)}(x) + k_{r5}, \end{aligned} \qquad 1 \le r \le r^{(2)}_{\max}, \tag{2.8}$$

are satisfied, where $1 \le p(r) \le r^{(1)}_{\max}$ *for* $1 \le r \le r^{(2)}_{\max}$.

As a result of these assumptions, it is shown in [BL, Lemma 2.3 and Theorem 2.4] that, if r is such that $1 \le r \le r^{(2)}_{\max}$ and if $\max\{S_r(x_0^N), S_{p(r)}(x_0^N)\} \le C$ for some C, then there are constants C_1 and C_2, depending on C, r and T, such that

$$\mathbf{P}\Big[\sup_{0\le t\le T} S_r(x_t^N) > C_1\Big] \;\le\; N^{-1}C_2. \tag{2.9}$$

Semigroup assumptions

In order to make sense of (1.6), we need some assumptions about A. We assume that

$$A_{ij} \ge 0 \text{ for all } i \ne j \ge 0; \qquad \sum_{j\ne i} A_{ji} < \infty \text{ for all } i \ge 0, \tag{2.10}$$

and that, for some $\mu \in R_+^{\mathbb{Z}_+}$ such that $\mu(m) \ge 1$ for each $m \ge 0$, and for some $w \ge 0$,

$$A^T\mu \;\le\; w\mu. \tag{2.11}$$

We then use μ to define the μ-norm

$$\|\xi\|_\mu \;:=\; \sum_{m\ge 0} \mu(m)|\xi^m| \quad \text{on} \quad \mathcal{R}_\mu \;:=\; \{\xi \in \mathcal{R}\colon\; \|\xi\|_\mu < \infty\}, \tag{2.12}$$

and, under these assumptions, the transition semigroup R is well defined [BL, Section 3], and

$$\sum_{i\ge 0} \mu(i)R_{ij}(t) \;\le\; \mu(j)e^{wt} \qquad \text{for all } j \text{ and } t. \tag{2.13}$$

Note that there may be many possible choices for μ, but that we also require that F is locally Lipschitz in the μ-norm, in order to ensure that (1.6) has a μ-continuous solution: we assume that, for any $z > 0$,

$$\sup_{x\ne y\colon\; \|x\|_\mu, \|y\|_\mu \le z} \|F(x) - F(y)\|_\mu / \|x - y\|_\mu \;\le\; K(\mu, F; z) \;<\; \infty, \tag{2.14}$$

and this should be borne in mind when choosing μ. We further assume that, for some β_3, β_4,

$$\mu(j) \;\le\; \nu(j)^{\beta_3} \quad \text{and} \quad |A_{jj}| \;\le\; \nu(j)^{\beta_4}. \tag{2.15}$$

Transition rate assumptions

We need to ensure that the sum of the transition rates, even when weighted by largish powers of $\nu(j)$, remains bounded. To ensure this, we assume that, for some r_0 large enough, there exist $r_1 \le r^{(2)}_{\max}$, $b \ge 1$ and $k_1, k_2 > 0$ such that

$$\sum_{J\in\mathcal{J}} \alpha_J(x) \sum_{j\ge 0} |J^j| \{\nu(j)\}^{r_0} \;\le\; \{k_1 S_{r_1}(x) + k_2\}^b; \tag{2.16}$$

this assumption is a specialized version of [BL, (2.25)]. In view of (2.9), this implies that, if $\max\{S_{r_1}(x_0^N), S_{p(r_1)}(x_0^N)\} \le C$, then there are constants C_1 and C_2 depending on C and T, such that

$$\mathbf{P}\Big[\sup_{0\le t\le T} \sum_{J\in\mathcal{J}} \alpha_J(x_t^N) \sum_{j\ge 0} |J^j| \{\nu(j)\}^{r_0} > C_1\Big] \;\le\; N^{-1}C_2. \tag{2.17}$$

We shall therefore assume that the initial condition needed for (2.17) is indeed satisfied: that, for some $C < \infty$,

$$\max\{S_{r_1}(x_0^N), S_{p(r_1)}(x_0^N)\} \le C. \tag{2.18}$$

It can be seen from the statement of Theorem 6.1 that the larger we can take r_0 in (2.16), the sharper the approximation bound that we get in (1.7), in that ζ can be taken smaller for a given value of the product ζr_0, resulting in larger values of $b_1(\zeta)$.

Since it is immediate that

$$\sum_{J\in\mathcal{J}}\sum_{j\ge 0}|J^j|\{\nu(j)\}^r\alpha_J(x) \;\ge\; \sum_{J\in\mathcal{J}}\{\nu(J)\}^r\alpha_J(x),$$

it follows that, for any $r, s \ge 0$,

$$\begin{aligned}\sum_{J\notin\mathcal{J}_M}\{\nu(J)\}^r\alpha_J(x) \;&\le\; M^{-s}\sum_{J\notin\mathcal{J}_M}\{\nu(J)\}^{r+s}\alpha_J(x)\\ &\le\; M^{-s}\sum_{J\in\mathcal{J}}\sum_{j\ge 0}|J^j|\{\nu(j)\}^{r+s}\alpha_J(x),\end{aligned}$$

so that, if $r + s \le r_0$, (2.17) implies that

$$\sup_{0\le t\le T}\sum_{J\notin\mathcal{J}_M}\{\nu(J)\}^r\alpha_J(x_t^N) \;\le\; C_1M^{-s}, \tag{2.19}$$

except on an event of probability at most C_2N^{-1}.

Smoothness assumptions

We need some smoothness conditions on the rates near the deterministic path x. First, for some $\delta > 0$, we assume that F has second order partial derivatives in the tube

$$B(t,x,\delta) \;:=\; \{z\in\mathcal{R}\colon\; \|z-x_s\|_\mu \le \delta \text{ for some } 0\le s\le t\}, \tag{2.20}$$

where x solves (1.6), and that, for any j, k, l,

$$\sup_{z\in B(x,t,\delta)}\left|D_{kl}F^j(z)\right| \;\le\; v_{jkl}, \tag{2.21}$$

where the v_{jkl} are such that

$$\sum_{j\ge 0}\mu(j)v_{jkl} \;\le\; K_{F2}\mu(k)\mu(l), \tag{2.22}$$

for some $K_{F2} < \infty$. Note that (2.22) is satisfied if

$$v_{jkl} \;\le\; v^j\mu(k)\mu(l), \qquad \text{where} \quad \|v\|_\mu < \infty. \tag{2.23}$$

It is also true under the following condition: that, for each k, there exists $N(k) \subset \mathbb{Z}_+$ with $|N(k)| \le n_0$ and $\max_{j\in N(k)}\mu(j) \le K_0\mu(k)$ such that

$$v_{jkl} \;\le\; K_1\{\mu(l)\mathbf{1}_{\{N(k)\}}(j) + \mu(k)\mathbf{1}_{\{N(l)\}}(j)\}; \qquad v_{jkl} \;=\; 0 \quad \text{otherwise}, \tag{2.24}$$

for suitable n_0, K_0 and K_1, all finite. The function F has already been assumed to be μ-Lipschitz in (2.14); with the assumption (2.21), F becomes continuously μ-differentiable in the tube, so that, for some constant K_{F1},

$$\sup_{0\le t\le T}\sum_{j\ge 0}\mu(j)|D_kF^j(x_t)| \;\le\; K_{F1}\mu(k) \quad \text{for all} \quad k. \tag{2.25}$$

We also assume that the individual transition rates α_J are uniformly μ-Lipschitz in $B(T,x,\delta)$, with

$$\sup_{z_1,z_2\in B(T,x,\delta)} |\alpha_J(z_1)-\alpha_J(z_2)|/\|z_1-z_2\|_\mu \;\le\; K_\alpha\{\nu(J)\}^{\beta_5} \tag{2.26}$$

for some $K_\alpha, \beta_5 > 0$. This assumption, and those on the second derivatives of F, go beyond what is required for the law of large numbers in [BL]; the same is true of the assumptions (2.2) and (2.15).

Preliminary conclusions

We now assume, in addition, that we can take

$$r_0 \;>\; 2(\beta_1+\beta_3+\beta_4) \tag{2.27}$$

in (2.16). Then, under the assumptions of this section, it follows from [BL, Theorem 4.7], with $\zeta(j) := \{\nu(j)\}^{r_0}$, that the following result holds: for each $T>0$, there exist constants $K_T^{(1)}, K_T^{(2)}$ and $K_T^{(3)}$ such that, for all N large enough, if

$$\|x_0^N - x_0\|_\mu \le K_T^{(1)}\sqrt{\frac{\log N}{N}}, \tag{2.28}$$

then

$$\mathbf{P}\Big(\sup_{0\le t\le T}\|x_t^N - x_t\|_\mu > K_T^{(2)}\sqrt{\frac{\log N}{N}}\Big) \le K_T^{(3)}\frac{\log N}{N}. \tag{2.29}$$

We shall from now on also assume that (2.18) holds with x_0 for x_0^N. Since then x can be represented as a limit of processes x^M satisfying (2.18), because of (2.29), it follows in view of (2.17) and (2.26) that we also have

$$\sup_{0\le t\le T}\sum_{J\in\mathcal{J}}\alpha_J(x_t)\sum_{j\ge 0}|J^j|\{\nu(j)\}^{r_0} \;\le\; C_1, \tag{2.30}$$

and therefore, as for (2.19), for $r+s\le r_0$,

$$\sup_{0\le t\le T}\sum_{J\notin\mathcal{J}_M}\{\nu(J)\}^r\alpha_J(x_t) \;\le\; C_1M^{-s}. \tag{2.31}$$

When approximating x^N by a deterministic path x, it is natural to choose their initial values to be close, as in (2.28). The impact of also assuming (2.18) is to force the initial values of both paths to put very little weight on components whose indices j have $\nu(j)$ large.

Example

In the model of Arrigoni (2003) presented in the introduction, we can take $\nu(j) = j+1$, in which case $\beta_1 = 1$ and $\beta_2 = 2$, the latter because of the migration transition. Calculation shows that (2.7) is satisfied for all r, in particular because the b_i are assumed to be uniformly bounded. (2.8) also holds, if the d_i are at most polynomially growing in i; if the d_i grow like i^{m_d}, we can take $p(r) = \max\{2r, 2r-1+m_d\}$, and then $r_{\max}^{(1)} = r_{\max}^{(2)} = \infty$. Furthermore, (2.16) is then satisfied for any r_0, with $r_1 = r_0+1+m_d$. The quantities A and F are given by

$$\begin{aligned} A_{ii} &= -\{\kappa + i(b_i+d_i+\gamma)\};\; A^T_{i,i-1} \;=\; i(d_i+\gamma);\; A^T_{i,i+1} \;=\; ib_i, \quad i\ge 1;\\ A_{00} &= -\kappa, \end{aligned}$$

with all other elements of A equal to zero, and, writing $s(x) := \sum_{j\ge 1} jx^j$,

$$F^i(x) \;=\; \rho\gamma(x^{i-1}-x^i)s(x), \quad i\ge 1; \qquad F^0(x) \;=\; -\rho\gamma x^0 s(x)+\kappa,$$

where we have used the fact that $\sum_{j\geq 0} x^j = 1$. Hence Assumption (2.10) is immediate, and Assumption (2.11) holds for $\mu(j) = j+1$ (so that $\beta_3 = 1$), with $w = \max_i(b_i - d_i - \gamma - \kappa)_+$; the value of β_4 is $m_d + 1$. For instance, the stochastic version of Ricker's (1954) model has both the b_i and the d_i uniformly bounded, in which case we can take $\beta_4 = 1$. However, in the stochastic analogue of Verhulst's (1838) logistic model, the d_i grow linearly with i, and then one needs $\beta_4 = 2$.

With the above choice of μ, F can easily be seen to be locally Lipschitz in the μ-norm, with $K(\mu, F; z) \leq 4\rho\gamma z$. The partial derivatives of F are given by

$$\begin{aligned} D_k F^i(x) &= \rho\gamma k(x^{i-1} - x^i) + \rho\gamma s(x)\{\mathbf{1}_{\{k\}}(i-1) - \mathbf{1}_{\{k\}}(i)\}; \\ D_{kl} F^i(x) &= \rho\gamma\{k[\mathbf{1}_{\{l\}}(i-1) - \mathbf{1}_{\{l\}}(i)] + l[\mathbf{1}_{\{k\}}(i-1) - \mathbf{1}_{\{k\}}(i)]\}, \end{aligned}$$

for any $i, k, l \geq 0$ (we take $x^{-1} = 0$). From this, it follows (using the elementary bound $j+1 \leq 2j$ if $j \geq 1$) that we can take $K_{F1} = 6\rho\gamma \sup_{0\leq t\leq T} \|x_t\|_\mu$ in (2.25), and that (2.24) is satisfied with $n_0 = 2$, $K_0 = 2$ and $K_1 = \rho\gamma$, so that (2.22) is also satisfied (one can in fact take $K_{F2} = 4\rho\gamma$). Finally, (2.26) is satisfied, with $\beta_5 = \max\{1, m_d\}$ and with K_α of the form $K'(1 + \sup_{0\leq t\leq T} \|x_t\|_\mu)$.

The models of Kimmel & Axelrod (2002, p.157) and Kretzschmar (1993), and that of Luchsinger (2001a,b) with Poisson infection distribution, can also be treated in much the same way.

3 Controlling η^N

From now on, we assume that all the assumptions of Section 2 are in force. We first show that the effect of the perturbation η^N is negligible. For η^N_t, from (1.13), we need to consider the difference

$$\int_0^t R(t-s)\{F(x^N_s) - F(x_s) - DF(x_s)[x^N_s - x_s]\}\, ds.$$

We note first that, if $\|h\|_\mu \leq \delta$ for δ as in Condition (2.20), then, from (2.22),

$$\|F(x_s + h) - F(x_s) - DF(x_s)[h]\|_\mu \;\leq\; \sum_{j\geq 0} \mu(j) \sum_{k\geq 0}\sum_{l\geq 0} |h_k h_l| v_{jkl} \;\leq\; K_{F2}\|h\|^2_\mu.$$

Hence, from (2.29) and from (2.13), for all N large enough so that $K^{(2)}_T N^{-1/2}\sqrt{\log N} \leq \delta$, we have

$$\sup_{0\leq s\leq t} \|\eta^N_s\|_\mu \;\leq\; K_{F2} t \{K^{(2)}_T\}^2 e^{wt} N^{-1/2} \log N, \tag{3.1}$$

for all $0 < t \leq T$, except on a set of probability at most $K^{(3)}_T N^{-1} \log N$.

4 Discrete to diffusion

We now show that $N^{1/2} m^N$ is close in the μ-norm to the diffusion W, given by

$$W_t \;:=\; \sum_{J\in\mathcal{J}} J W_J(A_J(t)),$$

as in (1.14). We first need to show that this W indeed has paths in $\mathcal{R}_\mu$. For this, it is enough to show that

$$\sum_{j\geq 0} \mu(j) \sum_{J\in\mathcal{J}} |J^j| |W_J(A_J(t))| \;<\; \infty \tag{4.1}$$

for all t.

We begin by noting that, using the reflection principle, if B is standard Brownian motion, then there exists a constant $\gamma < \infty$ such that, for all $a > 0$,

$$\mathbf{P}[\sup_{0\le x\le 1} |B(x)| > a\gamma] \;\le\; e^{1-a^2}. \tag{4.2}$$

Thus, from (4.2), for any $C > 1$ and $p, T > 0$, we have

$$|W_J(A_J(t))| \;=_d\; A_J(T)^{1/2}|B(A_J(t)/A_J(T))| \;\le\; A_J(T)^{1/2}\gamma\sqrt{p\log(C\nu(J))},$$

for all $0 \le t \le T$, except on a set of probability at most $e\{C\nu(J)\}^{-p}$. Hence it follows that, for all $0 \le t \le T$ and for all $J \in \mathcal{J}$, by (2.4),

$$|W_J(A_J(t))| \;\le\; A_J(T)^{1/2}\gamma\sqrt{p\log(C\nu(J))}, \tag{4.3}$$

except on a set of probability at most $eC^{-p}\sum_{J\in\mathcal{J}}\{\nu(J)\}^{-p} < \infty$, if $p > \beta_2$.

For $x, y \ge 0$, one has $\sqrt{x+y} \le (1+\sqrt{x})(1+\sqrt{y})$. Substituting from (4.3) into (4.1) shows that $\|W_t\|_\mu < \infty$ a.s. for all $0 \le t \le T$, provided that

$$\sum_{j\ge 0}\mu(j)\sum_{J\in\mathcal{J}}|J^j|\{1+\sqrt{\log\nu(J)}\}A_J(T)^{1/2} \;<\; \infty,$$

since C can be chosen arbitrarily in (4.3). However, by (2.15) and recalling the definition of J_*, we have, for any $\varepsilon > 0$,

$$\begin{aligned}
&\sum_{j\ge 0}\mu(j)\sum_{J\in\mathcal{J}}|J^j|\nu(J)^{\varepsilon}A_J(t)^{1/2}\\
&\quad\le\quad \sum_{J\in\mathcal{J}}\sum_{j\ge 0}|J^j|\nu(J)^{-\beta_2/2-\varepsilon}\,\nu(J)^{\beta_3+2\varepsilon+\beta_2/2}A_J(t)^{1/2}\\
&\quad\le\quad J_*\left\{\left(\sum_{J\in\mathcal{J}}\nu(J)^{-\beta_2-2\varepsilon}\right)^{1/2}\left(\sum_{J\in\mathcal{J}}\int_0^T\nu(J)^{\beta_2+2\beta_3+4\varepsilon}\alpha_J(x_s)\,ds\right)^{1/2}\right\},
\end{aligned} \tag{4.4}$$

and both sums in the final expression are finite, by (2.4) and (2.30), provided that ε is small enough and that $\beta_2 + 2\beta_3 < r_0$.

Having established that W indeed has paths in $\mathcal{R}_\mu$, we now need to show that it is close to $N^{1/2}m^N$ in the μ-norm, if the Brownian motions W_J are suitably chosen. The relationship between $N^{1/2}m^N$ and W arises because $N^{1/2}m^N$ can be represented in the form

$$\begin{aligned}
Nm_t^N \;&:=\; N\left\{x_t^N - x_0^N - \int_0^t\sum_{J\in\mathcal{J}}J\alpha_J(x_s^N)\,ds\right\}\\
&=\; \sum_{J\in\mathcal{J}}J\{P_J(NA_J^N(t)) - NA_J^N(t)\},
\end{aligned} \tag{4.5}$$

where $A_J^N(t) := \int_0^t\alpha_J(x_s^N)\,ds$, and the P_J's are independent unit rate Poisson processes. Now $\{N^{-1/2}(P_J(Nt) - Nt),\, t \ge 0\}$ can be well approximated by a Brownian motion, and $A_J^N(t)$ is close to $A_J(t) := \int_0^t\alpha_J(x_s)\,ds$, by (2.26) and (2.29).

We thus wish to show that the W_J can be chosen in such a way that

$$\sup_{0\le t\le T}\sum_{j\ge 0}\mu(j)\left|\sum_{J\in\mathcal{J}}J^jZ_J^N(A_J^N(t)) - \sum_{J\in\mathcal{J}}J^jW_J(A_J(t))\right| \tag{4.6}$$

is small, where we define

$$Z_J^N(s) \;:=\; N^{-1/2}\{P_J(Ns) - Ns\}. \tag{4.7}$$

For use in the next section, in bounding (5.2), we prove somewhat more: that, under appropriate conditions, we can replace $\mu(j)$ in (4.6) by the larger quantity $\nu_*(j) := \{\nu(j)\}^{\beta_3+\beta_4}$, and still obtain something that is small. To do so, we begin by bounding the sum by $T_1(t) + T_2(t) + T_3(t)$, where

$$\begin{aligned} T_1(t) &:= \sum_{J\in\mathcal{J}_M}\sum_{j\geq 0}|J^j|\,\nu_*(j)\,|Z_J^N(A_J^N(t)) - W_J(A_J(t))|; \\ T_2(t) &:= \sum_{J\notin\mathcal{J}_M}\sum_{j\geq 0}|J^j|\,\nu_*(j)\,|Z_J^N(A_J^N(t))|; \\ T_3(t) &:= \sum_{J\notin\mathcal{J}_M}\sum_{j\geq 0}|J^j|\,\nu_*(j)\,|W_J(A_J(t))|. \end{aligned} \tag{4.8}$$

Here, M is to be chosen later as N^ζ, for some suitable small $\zeta > 0$.

We begin with $T_2(t)$, which we deal with by showing that, for suitable choice of M, $N\sum_{J\notin\mathcal{J}_M} A_J^N(T)$ is small. Indeed,

$$N\sum_{J\notin\mathcal{J}_M} A_J^N(T) \;=\; N\int_0^T \sum_{J\notin\mathcal{J}_M}\alpha_J(x_u^N)\,du,$$

which is bounded by using (2.19) with $r = 0$ and $s \leq r_0$, together with (2.17); the quantity is of order NM^{-s} for any $s \leq r_0$, except on an event of probability of order $O(N^{-1})$. Thus, except on an event with probability of order $O(N^{-1} + NM^{-r_0})$, $N^{-1/2}P_J(NA_J^N(T)) = 0$ for all $J \notin \mathcal{J}_M$. Furthermore, the contribution from the compensators is bounded by

$$\begin{aligned} N^{1/2}\sum_{J\notin\mathcal{J}_M}\nu_*(j)|J^j|A_J^N(T) &\leq N^{1/2}T\sup_{0\leq t\leq T}\sum_{J\notin\mathcal{J}_M}\nu(j)^{\beta_3+\beta_4}|J^j|\alpha_J(x_t^N) \\ &= O(N^{1/2}M^{-s'}), \end{aligned} \tag{4.9}$$

by (2.19), if $\beta_3 + \beta_4 + s' \leq r_0$, except on an event whose probability is of order $O(N^{-1})$. Recalling (2.27), this proves that, for any $s' \leq r_0 - \beta_3 - \beta_4$,

$$\sup_{0\leq t\leq T} T_2(t) = O(N^{1/2}M^{-s'}), \tag{4.10}$$

except on an event of probability of order $O(N^{-1} + NM^{-r_0})$.

For $T_3(t)$, we use (4.2) to give

$$\mathbf{P}\Big[\sup_{0\leq t\leq T}|W_J(A_J(t))| > \{A_J(T)\}^{1/2}\gamma\sqrt{p\log\nu(J)}\Big] \;\leq\; e\{\nu(J)\}^{-p}, \tag{4.11}$$

for any $p > 0$. Hence, for any $p > \beta_2$, it follows that $|W_J(A_J(t))| \leq \{A_J(T)\}^{1/2}\gamma\sqrt{p\log\nu(J)}$ for all $J \notin \mathcal{J}_M$ and for all $0 \leq t \leq T$, except on an event E_3 of probability of order $O(M^{-p+\beta_2})$, from (2.4). But then, except on E_3, for all $0 \leq t \leq T$,

$$\begin{aligned} \sum_{J\notin\mathcal{J}_M}\sum_{j\geq 0}|J^j|\nu_*(j)|W_J(A_J(t))| &\leq \sum_{J\notin\mathcal{J}_M}\sum_{j\geq 0}|J^j|\nu_*(j)\{A_J(T)\}^{1/2}\gamma\sqrt{p\log\nu(J)} \\ &\leq K_\varepsilon\sqrt{p}\sum_{J\notin\mathcal{J}_M} J_*\nu(J)^{\beta_3+\beta_4+\varepsilon}\{A_J(T)\}^{1/2}, \end{aligned}$$

for any $\varepsilon > 0$, with suitable choice of K_ε. But now, for any $r > 0$,

$$\begin{aligned}
&\sum_{J \notin \mathcal{J}_M} \nu(J)^{\beta_3+\beta_4+\varepsilon} \{A_J(T)\}^{1/2} \\
&\quad \le \left\{ \sum_{J \notin \mathcal{J}_M} \nu(J)^{-2(r-\beta_3-\beta_4-\varepsilon)} \right\}^{1/2} \left\{ \sum_{J \notin \mathcal{J}_M} \nu(J)^{2r} A_J(T) \right\}^{1/2} \\
&\quad \le \{K'_{2(r-\beta_3-\beta_4-\varepsilon)}\}^{1/2} \left\{ M^{-2(r-\beta_3-\beta_4-\varepsilon)+\beta_2} \right\}^{1/2} \left\{ T M^{-2r'} \right\}^{1/2}, \qquad (4.12)
\end{aligned}$$

by (2.4) and (2.31), so long as $r > \beta_3 + \beta_4 + \varepsilon + \beta_2/2$ and $2(r + r') \le r_0$. Hence, if $r_0 > \beta_2 + 2(\beta_3 + \beta_4)$, then for any $r' < (r_0 - \beta_2)/2 - (\beta_3 + \beta_4)$ we have

$$p^{-1/2} \sup_{0 \le t \le T} T_3(t) = O(M^{-r'}), \qquad (4.13)$$

with an implied constant uniform for all $p > \beta_2$, except on an event with probability of order $O(M^{-p+\beta_2})$.

So far, the bounds have been achieved without any specific choice of the Brownian motions W_J, but, for $T_1(t)$, we need to be more precise. We treat each J separately, since the underlying Poisson processes P_J are independent, and match the centred and normalized Poisson process Z_J^N to a Brownian motion W_J using the KMT construction. We need only to do this over a limited time interval, since, from (2.17),

$$\sup_{0 \le t \le T} \sum_{J \in \mathcal{J}} \alpha_J(x_t^N) \ \le \ C_1,$$

except on an event E_0 of probability of order $O(N^{-1})$, so that, off E_0,

$$A_J^N(T) \ \le \ TC_1 \quad \text{for all } J. \qquad (4.14)$$

We use Komlós, Major & Tusnády (1975, Theorem 1 (ii)), together with (4.2) to interpolate between integer time points, applied to the centred unit rate Poisson process. This implies that, for any $p > 0$, we can choose W_J in such a way that

$$\sup_{0 \le t \le TC_1} \{N^{-1/2}(P_J(Nt) - Nt) - W_J(t)\} \ \le \ k_p N^{-1/2} \log N,$$

for a constant k_p, except on an event $\widetilde{E}_{Jp}$ of probability of order $O(N^{-p})$. Thus the same bound holds for all $J \in \mathcal{J}_M$ except on an event $\widetilde{E}_p$ of probability of order $O(M^{\beta_2} N^{-p})$. Hence, except on $E_0 \cup \widetilde{E}_p$, an event of probability of order $O(N^{-1} + M^{\beta_2} N^{-p})$, we have

$$\begin{aligned}
T_{11}(t) \ &:= \ \sum_{J \in \mathcal{J}_M} \sum_{j \ge 0} |J^j| \, \nu_*(j) \, |Z_J^N(A_J^N(t)) - W_J(A_J^N(t))| \\
&\le \ J_* \sum_{J \in \mathcal{J}_M} \{\nu(J)\}^{\beta_3+\beta_4} \, k_p N^{-1/2} \log N \\
&= \ O(M^{\beta_2+\beta_3+\beta_4} N^{-1/2} \log N), \qquad (4.15)
\end{aligned}$$

for all $0 \le t \le T$. To bound $T_1(t)$ from (4.8), it thus remains to consider

$$T_{12}(t) \ := \ \sum_{J \in \mathcal{J}_M} \sum_{j \ge 0} |J^j| \, \nu_*(j) \, |W_J(A_J^N(t)) - W_J(A_J(t))|. \qquad (4.16)$$

First, note that, by (2.26),

$$\sup_{0 \le t \le T} |A_J^N(t) - A_J(t)| \ \le \ TK_\alpha \{\nu(J)\}^{\beta_5} \sup_{0 \le t \le T} \|x_t^N - x_t\|_\mu, \qquad (4.17)$$

and that

$$\sup_{0\le t\le T} \|x_t^N - x_t\|_\mu \;\le\; K_T^{(2)} N^{-1/2}\sqrt{\log N}, \tag{4.18}$$

by (2.29), except on an event with probability of order $O(N^{-1}\log N)$. Furthermore, by (2.30), $A_J(T) \le TC_1$ for all J. Then, for a Brownian motion W and for $0<\delta<1$, by a standard argument based on (4.2),

$$\mathbf{P}\left[\sup_{0\le u\le A, |s|\le\delta} |W(u+s)-W(u)| > 3\gamma\delta^{1/2}\sqrt{r\log(1/\delta)}\right] \;\le\; (A\delta^{-1}+2)e\delta^r, \tag{4.19}$$

for any $r>0$, with γ chosen as for (4.2). Hence, taking $W=W_J$ and, in view of (4.17) and (4.18), taking

$$\delta \;=\; \delta_J \;=\; TK_\alpha K_T^{(2)}\{\nu(J)\}^{\beta_5} N^{-1/2}\sqrt{\log N}$$

and $A=TC_1$ in (4.19), it follows that, for any $\varepsilon, r'>0$, there is a $K_{r'\varepsilon}<\infty$ such that

$$\sup_{0\le t\le T} |W_J(A_J^N(t)) - W_J(A_J(t))| \;\le\; K_{r'\varepsilon} M^{\beta_5/2} N^{-1/4+\varepsilon} \quad \text{for all} \quad J\in\mathcal{J}_M,$$

except on an event of probability of order $O(N^{-1}\log N + M^{\beta_2}\{M^{2\beta_5}/N\}^{r'})$. Off the exceptional event, we have

$$\begin{aligned}\sup_{0\le t\le T} T_{12}(t) \;&\le\; J_* \sum_{J\in\mathcal{J}_M} \nu_*(J) K_{r'\varepsilon} M^{\beta_5/2} N^{-1/4+\varepsilon}\\ &=\; O(M^{\beta_2+\beta_3+\beta_4+\beta_5/2} N^{-1/4+\varepsilon}),\end{aligned} \tag{4.20}$$

for any $\varepsilon>0$, and, by choosing r' large enough, the exceptional event can be made to have probability of order $O(N^{-1}\log N)$, provided that M is bounded by a small enough power of N.

Combining (4.10), (4.13), (4.15) and (4.20), and choosing $M=N^\zeta$, we see that we have no useful bound unless $\zeta r_0>1$ (because of the exceptional event in (4.10)) and also $\zeta<1/\{4(\beta_2+\beta_3+\beta_4)+2\beta_5\}$ (in view of (4.20)), so that $r_0>4(\beta_2+\beta_3+\beta_4)+2\beta_5$ is a minimal requirement. Note that this assumption on r_0 is more restrictive than that assumed in Section 2. The error bound in (4.15) is always smaller than that in (4.20), and with $\zeta r_0>1$, the error bound in (4.10) is smaller than that in (4.13). This translates into the following conclusion: if $r_0>4(\beta_2+\beta_3+\beta_4)+2\beta_5$, then for any $1/r_0<\zeta<1/\{4(\beta_2+\beta_3+\beta_4)+2\beta_5\}$ we have

$$\begin{aligned}&\sup_{0\le t\le T} \|N^{1/2} m_t^N - W_t\|_\mu\\ &\quad\le \sup_{0\le t\le T} \sum_{j\ge 0} \{\nu(j)\}^{\beta_3+\beta_4} \left|\sum_{J\in\mathcal{J}} J^j Z_J^N(A_J^N(t)) - \sum_{J\in\mathcal{J}} J^j W_J(A_J(t))\right|\\ &\quad=\; O(N^{-b_1})\end{aligned} \tag{4.21}$$

except on an event of probability of order $O(N^{-b_2})$, for any

$$\begin{aligned} b_1 \;&<\; b_1(\zeta)\\ &:=\; \min\{\tfrac14 - \zeta(\beta_2+\beta_3+\beta_4+\beta_5/2), \tfrac12\zeta(r_0-\beta_2-2(\beta_3+\beta_4))\};\\ b_2 \;&<\; b_2(\zeta) \;:=\; \min\{\zeta r_0 - 1, 1\}.\end{aligned} \tag{4.22}$$

5 The existence of $\widetilde{W}$, and its approximation

The next step in the argument is to show that the process $\widetilde{W}$ in (1.15), related to W exactly as $\widetilde{m}^N$ is related to m^N through (1.10), is well defined, and that it is indeed the limiting analogue of the process $N^{1/2}\widetilde{m}_t^N$. For its existence, recalling (1.15), it is enough to show that $R(t-s)AW_s$ exists for each s,t, and belongs to $\mathcal{R}_\mu$. For this, it is enough to show that

$$\sum_{i\geq 0}\mu(i)\sum_{j\geq 0}R_{ij}(t-s)\sum_{k\geq 0}|A_{jk}|\sum_{J\in\mathcal{J}}|J_k|\,|W_J(A_J(s))|$$

is a.s. bounded, uniformly for $s\in[0,t]$. Now, in view of (2.11), (2.13) and (2.15) and recalling that the off-diagonal entries of A are non-negative, this will be the case if we can bound

$$\sum_{k\geq 0}\mu(k)(1+|A_{kk}|)\sum_{J\in\mathcal{J}}|J_k||W_J(A_J(s))| \;\leq\; J_*\sum_{J\in\mathcal{J}}\{\nu(J)\}^{\beta_3+\beta_4}|W_J(A_J(s))| \tag{5.1}$$

uniformly in s. Now, once again from (4.2), for any $C,r>0$,

$$\mathbf{P}\left[\sup_{0\leq t\leq T}|W_J(A_J(t))|>\{A_J(T)\}^{1/2}\gamma\sqrt{r\log(C\nu(J))}\right] \;\leq\; e\{C\nu(J)\}^{-r},$$

so that, in view of (2.4), taking any $r>\beta_2$, there is a (random) C such that

$$\sup_{0\leq t\leq T}|W_J(A_J(t))| \;\leq\; \gamma\sqrt{r}\{A_J(T)\}^{1/2}\sqrt{\log(C\nu(J))}$$

a.s. for all J. But now, returning to (5.1), we just need to show that the quantity

$$\sum_{J\in\mathcal{J}}\{\nu(J)\}^{\beta_3+\beta_4+\varepsilon}\{A_J(T)\}^{1/2}$$

is finite for some $\varepsilon>0$, and this is achieved as in (4.12), if $r_0>\beta_2+2(\beta_3+\beta_4)$.

To show that $\widetilde{W}$ is a good approximation to $N^{1/2}\widetilde{m}^N$, we begin with the result proved in (4.21) above: that $\sup_{0\leq t\leq T}\|N^{1/2}m_t^N-W_t\|_\mu=O(N^{-b_1})$ except on an event of probability of order $O(N^{-b_2})$, for any $b_1<b_1(\zeta)$ and $b_2<b_2(\zeta)$. This quantity is one element of $\|N^{1/2}\widetilde{m}_t^N-\widetilde{W}_t\|_\mu$; the other is

$$\int_0^t R(t-s)A(N^{1/2}m_s^N-W_s)\,ds.$$

Arguing much as in the previous paragraph, we need to bound

$$\sup_{0\leq t\leq T}\sum_{J\in\mathcal{J}}\{\nu(J)\}^{\beta_3+\beta_4}|W_J(A_J(t))-Z_J^N(A_J^N(t))|. \tag{5.2}$$

But this is exactly what we achieved in (4.21). Hence, for any $b_1<b_1(\zeta)$ and $b_2<b_2(\zeta)$, where ζ is such that $1/r_0<\zeta<1/\{4(\beta_2+\beta_3+\beta_4)+2\beta_5\}$, we have

$$\sup_{0\leq t\leq T}\|\widetilde{W}_t-N^{1/2}\widetilde{m}_t^N\|_\mu \;=\; O(N^{-b_1}), \tag{5.3}$$

except on an event of probability of order $O(N^{-b_2})$.

6 The final approximation

The final step in the argument is to compare the solution U^N to (1.12) with the solution Y to (1.16). Both satisfy the general equation

$$Z_t \;=\; R(t)Z_0 + \int_0^t R(t-s)DF(x_s)[Z_s] + z_t, \tag{6.1}$$

but with different initial conditions Z_0 and forcing functions z; and their difference $Y - U^N$ also satisfies (6.1), with initial conditions and forcing functions subtracted. Now, for U^N, the forcing function $\eta^N + N^{1/2}\widetilde{m}^N$ is close to the forcing function $\widetilde{W}$ for Y, because of (3.1) and (5.3), and we shall assume that U_0^N and Y_0 are also close to one another, so that both differences are small. We now show that this implies that the difference between Y and U^N is also small.

First, the assumption (2.25) implies that, for $w \in \mathcal{R}_\mu$, $\|DF(x_s)[w]\|_\mu \le K_{F1}\|w\|_\mu$. It is then immediate from (2.13) that

$$\left\| \int_0^t R(t-s)DF(x_s)[Z_s]\,ds \right\|_\mu \;\le\; K_{F1}e^{wt}\int_0^t e^{-ws}\|Z_s\|_\mu\,ds$$

and that $\|R(t)Z_0\|_\mu \le \|Z_0\|_\mu e^{wt}$. Hence, for $0 \le t \le T$, it follows that

$$e^{-wt}\|Z_t\|_\mu \;\le\; \|Z_0\|_\mu + e^{-wt}\sup_{0\le s\le T}\|z_s\|_\mu + K_{F1}\int_0^t e^{-ws}\|Z_s\|_\mu\,ds,$$

and hence that

$$\|Z_t\|_\mu \;\le\; \{\|Z_0\|_\mu + \sup_{0\le s\le T}\|z_s\|_\mu\}e^{Ct}, \tag{6.2}$$

with $C = K_{F1} + w$. Applying (6.2) to $Z = Y - U^N$, and using the bounds in (3.1) and (5.3), it follows that, except on an event of probability of order $O(N^{-b_2})$,

$$\sup_{0\le t\le T}\|Y_t - U_t^N\|_\mu \;=\; O(N^{-b_1}),$$

where $U^N := N^{1/2}(x^N - x)$ and Y is the solution to (1.16), provided that $\|Y_0 - U_0^N\|_\mu = O(N^{-b_1})$ also. This proves the main theorem of the paper:

Theorem 6.1. *Suppose that the assumptions of Section 2 are satisfied, and that we can take $r_0 > 4(\beta_2+\beta_3+\beta_4) + 2\beta_5$ in (2.16). For any ζ such that $1/r_0 < \zeta < 1/\{4(\beta_2+\beta_3+\beta_4)+2\beta_5\}$, define*

$$b_1(\zeta) \;:=\; \min\{\tfrac{1}{4} - \zeta(\beta_2+\beta_3+\beta_4+\beta_5/2), \tfrac{1}{2}\zeta(r_0 - \beta_2 - 2(\beta_3+\beta_4))\},$$

and $b_2(\zeta) \;:=\; \min\{\zeta r_0 - 1, 1\}$. Suppose that $\|Y_0 - U_0^N\|_\mu \le KN^{-b_1(\zeta)}$. Then, for any $b_1 < b_1(\zeta)$, $b_2 < b_2(\zeta)$, we can construct copies of Y and U^N on the same probability space, in such a way that

$$\sup_{0\le t\le T}\|Y_t - U_t^N\|_\mu \;=\; O(N^{-b_1}),$$

except on an event whose probability is of order $O(N^{-b_2})$.

So, for example, in the model of Arrigoni (2003), we can take r_0 as big as we wish, and then $\zeta r_0 = 2$, allowing $b_1 = 1/4 - \varepsilon$ and $b_2 = 1 - \varepsilon$ for any $\varepsilon > 0$. However, these rates can only be attained for correspondingly well controlled initial conditions: in addition to (2.28), it is necessary to ensure that (2.18) is satisfied, so that $S_{2(r_0+1)}(x_0^N) \le C$ for some $C < \infty$ and for all N, and that $S_{2(r_0+1)}(x_0) \le C$ also. For stochastic logistic dynamics within the patches, with $b_i = b$ and $d_i = d + ci$, we need to take r_0 to exceed $4(\beta_2+\beta_3+\beta_4) + 2\beta_5 = 22$ to yield an error bound that converges to zero with N, and thus require the initial conditions to have uniformly bounded $(46+\delta)$-th moments for some $\delta > 0$.

References

[1] Arrigoni, F.: Deterministic approximation of a stochastic metapopulation model. *Adv. Appl. Probab.* **35**, (2003), 691–720. MR-1990610

[2] Barbour, A. D. and Luczak, M. J.: Laws of large numbers for epidemic models with countably many types. *Ann. Appl. Probab.* **18**, (2008), 2208–2238. MR-2473655

[3] Barbour, A. D. and Luczak, M. J.: A law of large numbers approximation for Markov population processes with countably many types. *Prob. Theory Rel. Fields* **153**, (2012), 727–757. MR-2948691

[4] Eibeck, A. and Wagner, W.: Stochastic interacting particle systems and non-linear kinetic equations. *Ann. Appl. Probab.* **13**, (2003), 845–889. MR-1994039

[5] Kimmel, M. and Axelrod, D. E.: *Branching processes in biology.* Springer, Berlin, 2003. MR-1903571

[6] Komlós, J., Major, P. and Tusnády, G.: An approximation of partial sums of independent RV'-s, and the sample DF. I *Z. Wahrscheinlichkeitstheorie verw. Geb.* **32**, (1975), 111–131. MR-0375412

[7] Kretzschmar, M.: Comparison of an infinite dimensional model for parasitic diseases with a related 2-dimensional system. *J. Math. Analysis Applics* **176**, (1993), 235–260. MR-1222167

[8] Kurtz, T. G.: Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Probab.* **7**, (1970), 49–58. MR-0254917

[9] Kurtz, T. G.: Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J. Appl. Probab.* **8**, (1971), 344–356. MR-0287609

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

[11] Luchsinger, C. J.: Stochastic models of a parasitic infection, exhibiting three basic reproduction ratios. *J. Math. Biol.* **42**, (2001a), 532–554. MR-1845591

[12] Luchsinger, C. J.: Approximating the long term behaviour of a model for parasitic infection. *J. Math. Biol.* **42**, (2001b), 555–581. MR-1845592

[13] Pazy, A.: *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Springer, Berlin, 1983. MR-0710486

[14] Ricker, W. E.: Stock and Recruitment. *J. Fisheries Res. Board Canada* **11**, (1954), 559–623.

[15] Verhulst, P.-F.: Notice sur la loi que la population poursuit dans son accroissement. *Correspondance Mathématique et Physique* **10**, (1838), 113–121.

Acknowledgments. The authors express their warm thanks to two referees, for their detailed and thoughtful reading of the original version of the paper. They also wish to thank the Institute for Mathematical Sciences of the National University of Singapore and the University of Melbourne for providing welcoming environments while part of this work was accomplished. MJL also thanks the University of Zürich and ADB Monash University for their hospitality on a number of visits.