Probability Surveys
Vol. 7 (2010) 53–83
ISSN: 1549-5787
DOI: 10.1214/10-PS158

Limit theorems for discrete-time metapopulation models[*]

F.M. Buckley

e-mail: fbuckley@maths.uq.edu.au

and

P.K. Pollett

e-mail: pkp@maths.uq.edu.au

Abstract: We describe a class of one-dimensional chain binomial models of use in studying metapopulations (population networks). Limit theorems are established for time-inhomogeneous Markov chains that share the salient features of these models. We prove a law of large numbers, which can be used to identify an approximating deterministic trajectory, and a central limit theorem, which establishes that the scaled fluctuations about this trajectory have an approximating autoregressive structure.

AMS 2000 subject classifications: Primary 60J10, 92B05; secondary 60J80.

Received January 2010.

1. Metapopulations

A metapopulation is a population confined to a network of geographically separated habitat patches that may suffer extinction locally and be recolonized through dispersal of individuals from other patches. The term was coined by Levins [41], but the idea goes back much further, to MacArthur and Wilson [42], Andrewartha and Birch [4] and Wright [67], and has been refined more recently by Hanski and others (see for example [26] and [27]). Levins [40] was the first to provide a succinct mathematical description of a metapopulation, proposing that the number n_t of occupied patches at time t in a group of N patches should follow the law of motion

$$\frac{dn}{dt} = \frac{c}{N} n(N-n) - en, \tag{1}$$

with c being the colonization rate and e being the local extinction rate. This is Verhulst's model [63] for population growth and Levins used Pearl's rationale [51, 52, 53] to derive it. Furthermore, Levins was able to divine an explicit solution to (1) in the case where both c and e are time dependent, and he derived a diffusion approximation for n_t (surprisingly, the time-inhomogeneous

---

[*] This is an original survey paper.

version of Levins' model has had no traction in the ecology literature, despite the obvious implications for biological control in varying environments).

The natural stochastic analogue of (1) is a birth-death process on $S = \{0, 1, \ldots, N\}$ with birth rates $\lambda_n = \lambda n(N - n)$, where $\lambda = c/N$, and death rates $\mu_n = en$ (in the usual notation); indeed (1) can be viewed as its large-N mean field limit [54]. Known as the SIS (susceptible-infectious-susceptible) model in the epidemiology literature [64], it is an example of Feller's stochastic logistic model [23], which includes variants of the SIS model which are also of use in ecology, for example, models that account for colonisation from an external source [2, 57]. A significant drawback of the SIS model is that patches are assumed to be identical [50], although it does account for patch proximity (λ is inversely proportional to N). A more immediate drawback is that seasonal variation is not taken into account. Seasonal variation could certainly be accommodated using a time-inhomogeneous version, allowing c and e to be time dependent, but a simpler approach is to suppose that the number of occupied patches follows a Markov chain in discrete time. Akçakaya and Ginzburg [1] and Day and Possingham [22] developed a Markov chain model that assumes colonisation and extinction events occur in distinct successive phases; one might envisage an annual cycle, with local populations being susceptible to extinction during winter, while new populations establish in empty patches during the spring. It is assumed that a census takes place either at the end of successive colonisation phases (EC model) or at the end of successive extinction phases (CE model) and the state of the Markov chain is the state of the population at a census time. This approach has become predominant in the applied metapopulation literature, because it provides a vehicle for parameter estimation [46] and permits control mechanisms to be investigated using simple optimisation tools such as dynamic programming [60, 62]. Indeed, discrete-time Markov chain models predominate in the ecology literature (even in cases where they are not faithful to population dynamics), perhaps due in part to a widespread misconception that a discrete time model is needed if populations are observed (and controlled) at discrete time points [59, 58]. Numerical methods and simulation are generally used to analyse discrete time metapopulation models, typically the EC case only [29, 34, 68], and until recently there have been few analytical studies [18, 45].

Our purpose here is to present limit theorems that can be brought to bear in the study of chain-binomial metapopulation models (and indeed any discrete-time Markovian model that shares their salient features). We present a law of large numbers, which is used to identify an approximating (discrete-time) deterministic trajectory, and a central limit theorem, which establishes that the scaled fluctuations about this trajectory have an approximating autoregressive structure. Limit theorems of this kind are standard fare in the context of continuous-time Markovian models (see Kurtz [35, 36, 37, 38, 39] and Barbour [11, 12, 13, 14] and, more recently, Darling and Norris [20]); an approximating continuous-time deterministic trajectory is identified, and the fluctuations about this trajectory are approximated by a Gaussian diffusion. These methods have been exploited in the analysis of continuous-time metapopulation models [2, 5, 6, 7, 15, 56, 57, 54]. Diffusion approximations can also be realised for discrete-time Markovian models [49], but these approximations do not respect the discrete-time structure.

We begin by describing a class of one-dimensional discrete-time metapopulation models, called homogeneous SPOMs (stochastic patch occupancy models) in the ecology literature—*homogeneous* because patches are assumed to be identical. Section 3 contains the basic limit theorems. Our major source is the work of Klebaner and Nerman [33] (see also Klebaner [32]), who studied a generalisation of the Galton-Watson process where the offspring distribution was allowed to depend on the current population size measured as a proportion of some threshold. We explain how their results can be extended to accommodate time-inhomogeneous discrete-time Markov chains with density dependent transition probabilities. These results are applied to our metapopulation models in Sections 4 and 5. We conclude by exploring the relationship between discrete-time metapopulation models and their counterparts in continuous time.

2. Discrete-time metapopulation models

Extinction and colonisation are assumed to occur in alternating Markovian phases. In the extinction phase occupied patches are assumed to go extinct independently, each with the same probability e $(0 < e < 1)$, while in the colonisation phase empty patches are colonised independently, each with a probability $c(x)$ that depends on the proportion x of occupied patches at the start of this phase. For example, we might have $c(x) = cx$, where c $(0 < c \leq 1)$ is the (hypothetical) probability that a single unoccupied patch would be colonised by the fully occupied network, or $c(x) = c_0 + cx$ $(c \geq 0,\ c_0 > 0,\ c_0 + c \leq 1)$ if we wish to account for an external source of colonisation. Or, we might have $c(x) = 1 - \exp(-\beta x)$ $(\beta > 0)$, which effectively assumes (see [34, 29]) that colonising individuals propagate from each occupied patch at rate β. However, we will assume only that c is continuous, increasing and concave, with $c(0) \geq 0$ and $c(x) \leq 1$.

A census is assumed to take place either at the end of successive colonisation phases (the EC model) or at the end of successive extinction phases (the CE model); we consider both scenarios. If N is the total number of patches and n_t the number occupied at census $t \in \{0, 1, \dots\}$, then $(n_t,\ t \geq 0)$ is a Markov chain taking values in $S_N = \{0, 1, \dots, N\}$ with the following chain-binomial structure:

$$n_{t+1} = \tilde{n}_t + \mathrm{Bin}(N - \tilde{n}_t, c(\tilde{n}_t/N)) \qquad \tilde{n}_t = n_t - \mathrm{Bin}(n_t, e) \qquad \text{(EC model)}$$
$$n_{t+1} = \tilde{n}_t - \mathrm{Bin}(\tilde{n}_t, e) \qquad \tilde{n}_t = n_t + \mathrm{Bin}(N - n_t, c(n_t/N)). \qquad \text{(CE model)}$$

(We adopt the convention that $\mathrm{Ber}(p)$, $\mathrm{Bin}(n,p)$, $\mathrm{Poi}(\mu)$ and $\mathrm{N}(\mu, \sigma^2)$ denote random variables with the corresponding Bernoulli, Binomial, Poisson and Gaussian distribution.) Under the conditions we have imposed, S_N is irreducible unless $c(0) = 0$, in which case there is a single absorbing state 0, corresponding to total extinction of the population, with the remaining states forming an irreducible transient class $E_N = \{1, 2, \dots, N\}$ from which 0 is accessible.

Models with a state-independent colonisation probability $c(x) = c_0$ (> 0) have appeared in different guise in the epidemiology literature (see Section 4.4 of [19]). They were studied by us in detail in [18]. We found that n_t has the same distribution as the sum of two *independent* binomial random variables, $\mathrm{Bin}(n_0, p_t)$ and $\mathrm{Bin}(N - n_0, q_t)$, whose success probabilities (p_t) and (q_t) could be exhibited explicitly. Indeed $n_\bullet$ is equivalent (in law) to the urn model $n_{t+1} = \mathrm{Bin}(n_t, p_1) + \mathrm{Bin}(N - n_t, q_1)$, so that p_1 and q_1 can be interpreted as 'effective' survival and colonisation probabilities. This special property of equivalent independent phases carries over to some extent in the general case: for the CE model (only) it is easy to show that n_{t+1} has the same distribution as the sum of two independent binomial random variables, $\mathrm{Bin}(n_t, 1 - e)$ and $\mathrm{Bin}(N - n_t, (1 - e)c(n_t/N))$, so that $(1 - e)c(x)$ is the effective colonisation probability when the proportion of occupied patches is x. In [18] we also examined the proportion $X_t^N = n_t/N$ of occupied patches and considered what happens for fixed t as N gets large, assuming $X_0^N \to x_0$. We proved a law of large numbers that established existence of a limiting deterministic trajectory $x_\bullet$ with initial value x_0, which could be exhibited explicitly. The corresponding central limit law for the scaled fluctuations $Z_t^N = \sqrt{N}(X_t^N - x_t)$ about this trajectory was also proved, assuming $Z_0^N \to z_0$, the limiting Gaussian distribution having a variance that could be exhibited explicitly, and we mooted that the process $Z_\bullet^N$ might converge (in the sense of finite-dimensional distributions) to a Gaussian Markov chain $Z_\bullet$ with initial value z_0.

These results can be extended to accommodate general discrete-time metapopulation models with state-dependent colonization probabilities. We will see that, under mild conditions, $X_t^N \overset{\mathrm{P}}{\to} x_t$ for all $t \geq 0$, where $x_\bullet$ is determined by $x_{t+1} = f(x_t)$ $(t \geq 0)$ with f specific to the model. It is also possible to identify conditions which ensure that if $Z_0^N \overset{\mathrm{D}}{\to} z_0$, then $Z_\bullet^N$ converges weakly (in the usual product topology) to the Gaussian Markov chain $Z_\bullet$ defined by $Z_{t+1} = f'(x_t)Z_t + E_t$ $(Z_0 = z_0)$, with the 'errors' (E_t) being independent with $E_t \sim \mathrm{N}(0, v(x_t))$, where v is specific to the model.

We will also examine the following infinite-patch models, with $n_t \in S = \{0, 1, \dots\}$:

$$n_{t+1} = \tilde{n}_t + \mathrm{Poi}(m(\tilde{n}_t)) \qquad \tilde{n}_t = n_t - \mathrm{Bin}(n_t, e) \qquad \text{(EC model)}$$

$$n_{t+1} = \tilde{n}_t - \mathrm{Bin}(\tilde{n}_t, e) \qquad \tilde{n}_t = n_t + \mathrm{Poi}(m(n_t)), \qquad \text{(CE model)}$$

where $m(n)$ (> 0) is the expected number of patches colonised during any one colonisation phase when the number occupied at the start of that phase is n. Even though there is now no ceiling on the number of occupied patches, the dependence of m on n_t allows for regulation of the colonisation process. The case $m(n) = mn$ (where the constant $m > 0$ is the expected number of colonisations by any one occupied patch) is a natural analogue of the N-patch models described above, for if $c(0) = 0$ and c has a continuous second derivative near 0, then, for fixed n, $\mathrm{Bin}(N - n, c(n/N)) \overset{\mathrm{D}}{\to} \mathrm{Poi}(mn)$ as $N \to \infty$, where $m = c'(0)$.

3. General structure: Density dependence

Let $(n_t^N,\, t \geq 0)$ be a family of Markov chains indexed by N, each taking values in a subset set S_N of $\mathbb{Z}_+$. Suppose that the family is *density dependent* in that there are sequences of non-negative functions (f_t) and (v_t) such that $\mathbb{E}(n_{t+1}^N | n_t^N) = N f_t(n_t^N/N)$ and $\mathrm{Var}\,(n_{t+1}^N | n_t^N) = N v_t(n_t^N/N)$. Then, the 'density process' $(X_t^N,\, t \geq 0)$, defined by $X_t^N = n_t^N/N$, will have $\mathbb{E}(X_{t+1}^N | X_t^N) = f_t(X_t^N)$ and $N\,\mathrm{Var}\,(X_{t+1}^N | X_t^N) = v_t(X_t^N)$. Our first result is a law of large numbers that establishes convergence of the density process to a deterministic trajectory.

Theorem 1. *Suppose that, for all $t \geq 0$, $f_t(x)$ and $v_t(x)$ are continuous in x and such that $f_t(X_t^N)$ and $v_t(X_t^N)$ are a.s. uniformly bounded. Then, if $X_0^N \xrightarrow{\mathrm{P}} x_0$ (a constant), $X_t^N \xrightarrow{\mathrm{P}} x_t$ for all $t \geq 1$, where $x_\bullet$ is determined by $x_{t+1} = f_t(x_t)$ $(t \geq 0)$.*

Proof. We will use mathematical induction. Suppose $X_t^N \xrightarrow{\mathrm{P}} x_t$ for some $t \geq 0$. Then, since f_t is continuous, $\mathbb{E}(X_{t+1}^N | X_t^N) = f_t(X_t^N) \xrightarrow{\mathrm{P}} f_t(x_t)$. But, $f_t(X_t^N)$ is a.s. uniformly bounded, and so $f_t(X_t^N) \xrightarrow{r} f_t(x_t)$ (convergence in r-th mean) for all $r \geq 1$, which entails, in particular, that $\mathbb{E}X_{t+1}^N = \mathbb{E}f_t(X_t^N) \to f_t(x_t)$ and $\mathrm{Var}\,f_t(X_t^N) \to 0$. Similarly, $\mathbb{E}\,v_t(X_t^N) \to v_t(x_t)$ because v_t is continuous and $v_t(X_t^N)$ is a.s. uniformly bounded. Therefore,

$$\begin{aligned}\mathrm{Var}\,X_{t+1}^N &= \mathbb{E}\,\mathrm{Var}\,(X_{t+1}^N | X_t^N) + \mathrm{Var}\,\mathbb{E}(X_{t+1}^N | X_t^N)\\ &= \frac{1}{N}\,\mathbb{E}\,v_t(X_t^N) + \mathrm{Var}\,f_t(X_t^N) \to 0.\end{aligned}$$

But, for all $\epsilon > 0$,

$$\begin{aligned}\Pr(|X_{t+1}^N - f_t(x_t)| \geq \epsilon) &\leq \frac{1}{\epsilon^2}\mathbb{E}(X_{t+1}^N - f_t(x_t))^2\\ &= \frac{1}{\epsilon^2}\left(\mathrm{Var}\,X_{t+1}^N + (\mathbb{E}X_{t+1}^N - f_t(x_t))^2\right) \to 0,\end{aligned}$$

that is, $X_{t+1}^N \xrightarrow{\mathrm{P}} x_{t+1}$, and the proof is complete. □

We anticipate applying Theorem 1 in cases where X_t^N itself is bounded, for example, when X_t^N is a *proportion* (as it is in our N-patch models). To accommodated cases where X_t^N is unbounded (as it can be in our infinite-patch models) we relax uniform boundedness in favour of a Lipschitz condition, but at the expense of requiring a more stringent initial condition, that X_0^N converges to x_0 in mean square.

Theorem 2. *Suppose that, for all $t \geq 0$, $f_t(x)$ and $v_t(x)$ are Lipschitz continuous in x. If $X_0^N \xrightarrow{2} x_0$ (a constant), then $X_t^N \xrightarrow{2} x_t$ (and hence $X_t^N \xrightarrow{\mathrm{P}} x_t$) for all $t \geq 1$, where $x_\bullet$ is determined by $x_{t+1} = f_t(x_t)$ $(t \geq 0)$.*

Proof. We will again use mathematical induction. Suppose $X_t^N \xrightarrow{2} x_t$ for some $t \geq 0$. Since $f_t(x)$ is Lipschitz continuous, $|f_t(X_t^N) - f_t(x_t)| \leq \kappa_t |X_t^N - x_t|$, and hence $(f_t(X_t^N) - f_t(x_t))^2 \leq \kappa_t^2(X_t^N - x_t)^2$, for some positive constant κ_t. On taking expectations, we see that $f_t(X_t^N) \overset{2}{\to} f_t(x_t)$, which implies in particular that (i) $\operatorname{Var} f_t(X_t^N) \to 0$ and (ii) $\mathbb{E} f_t(X_t^N) \to f_t(x_t)$, that is, $\mathbb{E} X_{t+1}^N \to x_{t+1}$. Similarly, since $v_t(x)$ is Lipschitz continuous, $\mathbb{E} v_t(X_t^N) \to v_t(x_t)$. We deduce that $\operatorname{Var} X_{t+1}^N \to 0$, for, as noted earlier, $\operatorname{Var} X_{t+1}^N = \mathbb{E}\, v_t(X_t^N)/N + \operatorname{Var} f_t(X_t^N)$. But,

$$\mathbb{E}(X_{t+1}^N - x_{t+1})^2 = \operatorname{Var} X_{t+1}^N + (\mathbb{E} X_{t+1}^N - x_{t+1})^2,$$

and so $X_{t+1}^N \overset{2}{\to} x_{t+1}$. □

Having established convergence in probability to a limiting deterministic trajectory $x_\bullet$, we next consider the 'fluctuations process' (Z_t^N) defined by $Z_t^N = \sqrt{N}(X_t^N - x_t)$. Assuming now that $Z_0^N \overset{\mathrm{D}}{\to} z_0$, we aim to identify conditions under which $Z_\bullet^N$ converges weakly to a Gaussian Markov chain $Z_\bullet$. Additional structure is needed. We will assume that

$$n_{t+1}^N = g_t^N + \textstyle\sum_{j=1}^{r_t^N} \xi_{jt}^N \qquad (t \geq 0), \tag{2}$$

where $r_t^N = N r_t(n_t^N/N)$ and $g_t^N = N g_t(n_t^N/N)$ with $r_t(x)$ and $g_t(x)$ being continuous in x, and ξ_{jt}^N $(j = 1, \ldots, r_t^N)$ are iid having a distribution that depends only on t and on n_t^N/N, and which has bounded third moment. In particular, we assume that there are functions $m_t(x)$ and $\sigma_t^2(x)$ such that $\mathbb{E}\, \xi_{jt}^N = m_t(n_t^N/N)$ and $\operatorname{Var}(\xi_{jt}^N) = \sigma_t^2(n_t^N/N)$, and a function $b_t(x)$ such that $\mathbb{E}(\xi_{jt}^N - m_t(x))^3 = b_t(n_t^N/N)$, which is bounded in x. Of course all of our ingredients must be such that r_t^N and g_t^N are positive integers and, then, that $n_{t+1}^N \in S_N$. Notice that $\mathbb{E}(n_{t+1}^N | n_t^N) = N f_t(n_t^N/N)$, where $f_t(x) = g_t(x) + r_t(x) m_t(x)$, and $\operatorname{Var}(n_{t+1}^N | n_t^N) = N v_t(n_t^N/N)$, where $v_t(x) = r_t(x)\sigma_t^2(x)$.

This setup is similar to that of Klebaner and Nerman [33] (see also Klebaner [32]) who studied a generalisation of the Galton-Watson process where the offspring distribution was allowed to depend on the current population size measured as a proportion of some threshold N. Their model was time homogeneous and had $g_t \equiv 0$ and $r_t(x) = x$. None the less, many of the results in Section 3 of their paper carry over to the present case with only minor changes. We content ourselves with the following central limit law.

Theorem 3. *Suppose that, for all $t \geq 0$, $f_t(x)$ is twice continuously differentiable in x with bounded second derivative and that $X_t^N \overset{\mathrm{P}}{\to} x_t$, where $x_\bullet$ satisfies $x_{t+1} = f_t(x_t)$ $(t \geq 0)$. If $Z_0^N \overset{\mathrm{D}}{\to} z_0$ (a constant), then $Z_\bullet^N$ converges weakly to the Gaussian Markov chain $Z_\bullet$ defined by*

$$Z_{t+1} = f_t'(x_t) Z_t + E_t \qquad (Z_0 = z_0), \tag{3}$$

where (E_t) are independent with $E_t \sim \mathrm{N}(0, v_t(x_t))$.

Proof. First observe that we may rewrite (2) as

$$X_{t+1}^N = f_t(X_t^N) + \frac{1}{\sqrt{N}}\, \eta_t^N(X_t^N), \text{ where } \eta_t^N(x) = \frac{1}{\sqrt{N}} \sum_{j=1}^{[N r_t(x)]} (\xi_{jt}^N - m_t(x)) \tag{4}$$

($[\,\cdot\,]$ denotes integer part), noting that, for fixed x, $\eta_t^N(x)$ is independent of X_t^N. Next, for fixed t and x, $\eta_t^N(x) \xrightarrow{\text{D}} \text{N}(0, v_t(x))$ as $N \to \infty$. To see this, apply the Feller-Lindeberg Theorem (see for example Theorems 27.2 and 27.3 of Billingsley [16]) to the zero-mean triangular array $(W_{Nj},\ j = 1, \dots, r_N)$, where $W_{Nj} = \xi_{jt}^N - m_t(x)$ and $r_N = [N r_t(x)]$, noting that $\mathbb{E}(W_{Nj}^2) = \sigma_t^2(x)$ and the Lyapounov condition, that

$$\frac{\mathbb{E}|W_{N1}|^{2+\delta}}{r_N^{\delta/2}(\mathbb{E}\, W_{N1}^2)^{(1+\delta/2)}} \to 0 \quad (\text{for some } \delta > 0),$$

is satisfied here with $\delta = 1$ because $b_t(x)$, the third centred moment of ξ_{jt}^N, is bounded in x.

We have assumed that $X_t^N \xrightarrow{\text{P}} x_t$ for all $t \geq 1$, where $x_\bullet$ is determined by $x_{t+1} = f_t(x_t)$ $(t \geq 0)$. But, since $f_t(x)$ is twice continuously differentiable in x, we also have, by Taylor's theorem, that

$$f_t(X_t^N) = f_t(x_t) + f_t'(x_t)(X_t^N - x_t) + \tfrac{1}{2} f_t''(\theta_t^N)(X_t^N - x_t)^2,$$

for some θ_t^N between X_t^N and x_t, and so, from (4),

$$Z_{t+1}^N = \eta_t^N(X_t^N) + f_t'(x_t) Z_t^N + \tfrac{1}{2\sqrt{N}} f_t''(\theta_t^N)(Z_t^N)^2.$$

Since $f_t''(x)$ is bounded in x, we may thus write

$$Z_{t+1}^N = f_t'(x_t) Z_t^N + \eta_t^N(X_t^N) + o_t^N(1), \tag{5}$$

where $o_t^N(1) \xrightarrow{\text{P}} 0$ as $N \to \infty$.

To establish weak convergence of $Z_\bullet^N$ to $Z_\bullet$ it is sufficient to establish convergence of the finite-dimensional distributions. To this end, consider the characteristic function $\phi_t^N(\omega_t, \dots, \omega_0) = \mathbb{E} \exp\left(\mathrm{i}\left(\omega_t Z_t^N + \cdots + \omega_0 Z_0^N\right)\right)$ of $(Z_0^N, \dots, Z_t^N)$. Then, from (5),

$$\begin{aligned}\phi_{t+1}^N(\omega_{t+1}, \omega_t, \dots, \omega_0) \\ = \mathbb{E} \exp\Big(&\mathrm{i}\,\omega_{t+1}\eta_t^N(X_t^N) + \mathrm{i}\,(\omega_t + \omega_{t+1} f_t'(x_t)) Z_t^N \\ &+ \mathrm{i}\big(\omega_{t-1} Z_{t-1}^N + \cdots + \omega_0 Z_0^N\big) + o_t^N(1)\Big).\end{aligned}$$

Since, for fixed t and x, $\eta_t^N(x) \xrightarrow{\text{D}} \text{N}(0, v_t(x))$ as $N \to \infty$, it follows, from the Markov property and our premise $Z_0^N \xrightarrow{\text{D}} z_0$, that $\lim_{N\to\infty} \phi_t^N = \phi_t$ exists for all $t \geq 1$ and satisfies

$$\phi_{t+1}(\omega_{t+1}, \omega_t, \dots \omega_0) = \exp(-\tfrac{1}{2}\omega_{t+1}^2 v_t(x_t))\, \phi_t(\omega_t + \omega_{t+1} f_t'(x_t), \omega_{t-1}, \dots \omega_0),$$

with $\phi_0(\omega_0) = \exp(\mathrm{i}\,\omega_0 z_0)$ being the characteristic function of $Z_0 = z_0$. But, this iteration defines the characteristic function of $(Z_0, \dots, Z_t)$, where $Z_\bullet$ is the proposed limiting Gaussian process. This completes the proof. $\square$

Remarks. (i) An alternative approach to proving Theorems 1 and 3 would be to adapt the results of Karr [30] to a time inhomogeneous setting. He considered a sequence of *time-homogeneous* Markov chains $(X_t^N, t \geq 0)$ on a general state space $(E, \mathcal{E})^{\mathbb{N}}$ with transition kernels $(K^N(x, A),\ x \in E, A \in \mathcal{E})$ and initial distributions $(\pi^N(A),\ A \in \mathcal{E})$. He proved that if $\pi^N \Rightarrow \pi$ and $K^N(x^N, \cdot) \Rightarrow K(x, \cdot)$ when $x^N \to x$ in E, then the corresponding probability measures $(\mathbb{P}_{\pi^N}^N)$ on $(E, \mathcal{E})^{\mathbb{N}}$ also converge: $\mathbb{P}_{\pi^N}^N \Rightarrow \mathbb{P}_\pi$. Karr's main result (Theorem 1 of [30]) remains true, with obvious modifications, for a time *inhomogeneous* Markov chain. Applying this to our density process (X_t^N), and then to the two-dimensional process (X_t^N, Z_t^N), would establish the required convergence, and the form of the limiting processes would become apparent once the limiting kernels were evaluated.

(ii) The chain binomial models described in Section 2 are time homogeneous, yet we need the present level of generality to accommodate them because it is not always possible to establish density dependence, even on occasions when we anticipate the kind of asymptotic (large N) behaviour exhibited in Theorems 1 and 3. For our models (or, more generally, for Markov chains with two alternating phases), it is natural to construct a time-inhomogeneous Markov chain $(n_t^N, t \geq 0)$ by setting $n_{2t}^N = n_t$ and $n_{2t+1}^N = \tilde{n}_t$. Then, density dependence in the phases is often enough to establish the required asymptotic behaviour. Of course, this programme extends to models with more than two density-dependent phases. For example, our methods can easily accommodate the (three-phase) extinction-reproduction-settlement model of Klok and De Roos [34].

(iii) Klebaner and Nerman (Theorem 3 via Theorem 6 of [33]) anticipated a special case of Theorems 1 and 3: when the time homogeneous version of (4) is in force with $r_t(x) = x$.

(iv) The mean and covariance function of $Z_\bullet$ are easy to evaluate by iterating (3):

$$\mu_t := \mathbb{E}\, Z_t = z_0\, \Pi_{0,\,t} \quad (t \geq 1) \tag{6}$$

and

$$c_{t,\,s} := \mathrm{Cov}\,(Z_t, Z_s) = V_t\, \Pi_{t,\,s} \quad (s \geq t \geq 1), \tag{7}$$

where

$$\Pi_{u,\,v} = \prod_{w=u}^{v-1} f'_w(x_w) \quad (v > u) \tag{8}$$

and

$$V_t := \mathrm{Var}\,(Z_t) = \sum_{s=0}^{t-1} v_s(x_s)\Pi_{s+1,\,t}^2 \quad (t \geq 1). \tag{9}$$

(Here and henceforth empty products are to be interpreted as being equal to 1.) Furthermore, it is clear that, for any $t \geq 1$, $Z_t^N \xrightarrow{\mathrm{D}} \mathrm{N}(\mu_t, V_t)$, and so these formulae can be used to approximate the mean and covariance function of $n_\bullet^N$. Indeed, the joint distribution of $n_{t_1}^N, \ldots, n_{t_n}^N$, where $t_1, \ldots, t_n$ is any finite set of times, can be approximated by an n-dimensional Gaussian distribution with $\mathbb{E}n^N_{t_i} \simeq Nx_{t_i} + \sqrt{N}\mu_{t_i}$ and $\mathrm{Cov}\,(n^N_{t_i}, n^N_{t_j}) \simeq Nc_{t_i,\,t_j}$.

(v) Even when $n^N_{\bullet}$ is time homogeneous, and hence the approximating deterministic model has the form $x_{t+1} = f(x_t)$, the full range of long-term behaviour, including chaos, is possible. For example, if f has a stable fixed point x^* and $X^N_0 \overset{\mathrm{P}}{\to} x^*$, then, assuming f and v are continuous, and $f(X^N_t)$ and $v(X^N_t)$ are a.s. uniformly bounded, we will have $X^N_t \overset{\mathrm{P}}{\to} x^*$ for all t. Furthermore, $Z^N_t = \sqrt{N}(X^N_t - x^*)$ and, if f is twice continuously differentiable with bounded second derivative, then, assuming $Z^N_0 \overset{\mathrm{D}}{\to} z_0$, the limit process $Z_{\bullet}$ will be an *autoregressive* (AR-1) process with the representation $Z_{t+1} = f'(x^*)Z_t + E_t$, where (E_t) are independent and $E_t \sim \mathrm{N}(0, v(x^*))$. In this case $Z^N_t \overset{\mathrm{D}}{\to} \mathrm{N}(\mu_t, V_t)$, where $\mu_t = z_0 a^t$ $(= \mathbb{E}Z_t)$ and $c_{t,\,s} = V_t\, a^{|s-t|}$ $(= \mathrm{Cov}\,(Z_t, Z_s))$, where $a = f'(x^*)$, and, $V_t = v(x^*)t$ if $|a| = 1$ and $V_t = v(x^*)(1 - a^{2t})/(1 - a^2)$ otherwise. Furthermore, if $Z^N_0 \overset{\mathrm{D}}{\to} z_0$, and $|a| < 1$, then there will be a sequence of times (t_N) such that $Z^N_{t_N} \overset{\mathrm{D}}{\to} \mathrm{N}(0, V^*)$, where $V^* = v(x^*)/(1 - a^2)$.

More generally, if f admits a stable limit cycle, $x^*_0, x^*_1, \ldots, x^*_{d-1}$, and $X^N_0 \overset{\mathrm{P}}{\to} x^*_0$, we will have $Z^N_{nd+j} = \sqrt{N}(X^N_{nd+j} - x^*_j)$ $(n \geq 0,\ j = 0, \ldots, d-1)$ and, assuming $Z^N_0 \overset{\mathrm{D}}{\to} z_0$, the limit process $Z_{\bullet}$ will have the following representation: $(\boldsymbol{Y}_n,\, n \geq 0)$, where $\boldsymbol{Y}_n = (Z_{nd}, Z_{nd+1}, \ldots, Z_{(n+1)d-1})^\top$ with $Z_0 = z_0$, is a d-variate AR-1 process of the form $\boldsymbol{Y}_{n+1} = A\boldsymbol{Y}_n + \boldsymbol{E}_n$, where $(\boldsymbol{E}_n)$ are independent and $\boldsymbol{E}_n \sim \mathrm{N}(\boldsymbol{0}, \Sigma)$. The distribution of $\boldsymbol{Y}_0$, and both the coefficient matrix A and the covariance matrix Σ, are determined using (6)–(9) (obtained by iterating (3)) with $x_{nd+j} = x^*_j$ $(n \geq 0,\ j = 0, \ldots, d-1)$. This was done explicitly by Klebaner and Nerman (Theorem 4 of [33]) for the population dependent branching process, and observed to be true more generally (Theorem 6 of [33]). Indeed, their Theorems 4 and 5 hold word for word, but with obvious adjustments in the definitions of f and v, in the present more general context of a time-homogeneous density dependent family.

4. *N*-patch models

Here we explain how the results of Section 3 can be used to study our N-patch metapopulation models. We first identify an approximating deterministic model and describe the structure of the approximating Gaussian process, and then give a detailed account of long-term behaviour.

4.1. Limit theorems for the proportion of occupied patches

Let $X^N_t = n_t/N$ be the proportion of occupied patches at census t. We first prove a law of large numbers for the process $X^N_{\bullet}$, thus establishing the existence of an approximating deterministic trajectory, and a central limit law for the fluctuations about this trajectory. For the N-patch CE model, we may evaluate $\mathbb{E}(n_{t+1}|n_t)$ and $\mathrm{Var}\,(n_{t+1}|n_t)$ explicitly, because $\mathbb{E}(n_{t+1}|\tilde{n}_t)$ and $\mathrm{Var}\,(n_{t+1}|\tilde{n}_t)$

are both linear functions of $\tilde{n}_t$:

$$\begin{aligned}\mathbb{E}(n_{t+1}|n_t) &= \mathbb{E}\big(\mathbb{E}(n_{t+1}|\tilde{n}_t,\, n_t)|n_t\big) = \mathbb{E}\big(\mathbb{E}(n_{t+1}|\tilde{n}_t)|n_t\big) = (1-e)\mathbb{E}\big(\tilde{n}_t|n_t\big)\\ &= (1-e)(n_t + (N-n_t)c(n_t/N)),\end{aligned}$$

implying that $f_t(x) := f(x) = (1-e)(x + (1-x)c(x))$, and

$$\begin{aligned}\operatorname{Var}(n_{t+1}|n_t) &= \mathbb{E}\big(\operatorname{Var}(n_{t+1}|\tilde{n}_t,\, n_t)|n_t\big) + \operatorname{Var}\big(\mathbb{E}(n_{t+1}|\tilde{n}_t,\, n_t)|n_t\big)\\ &= \mathbb{E}\big(\operatorname{Var}(n_{t+1}|\tilde{n}_t)|n_t\big) + \operatorname{Var}\big(\mathbb{E}(n_{t+1}|\tilde{n}_t)|n_t\big)\\ &= e(1-e)\mathbb{E}(\tilde{n}_t|n_t) + (1-e)^2\operatorname{Var}(\tilde{n}_t|n_t)\\ &= e\,\mathbb{E}(n_{t+1}|n_t) + (1-e)^2(N-n_t)c(n_t/N)(1-c(n_t/N)),\end{aligned}$$

implying that

$$\begin{aligned}v_t(x) := v(x) &= ef(x) + (1-e)^2(1-x)c(x)(1-c(x))\\ &= (1-e)\left(ex + (1-x)c(x)(1-(1-e)c(x))\right).\end{aligned}$$

We may apply Theorem 1 directly since both f and v are continuous, and $f(X_t^N)$ and $v(X_t^N)$ are bounded because $0 \le X_t^N \le 1$ and $c(x) \le 1$. For the N-patch EC model we cannot evaluate the conditional mean and variance explicitly, and it will be clear that the model is not density dependent unless we impose further restrictions on c. For the N-patch EC model (and indeed, alternatively, for the CE model) we apply Theorem 1 to the time-inhomogeneous Markov chain $(n_t^N,\, t \ge 0)$ obtained by setting $n_{2t}^N = n_t$ and $n_{2t+1}^N = \tilde{n}_t$. Then, for the EC model, $f_{2t}(x) = f_0(x) = (1-e)x$ and $f_{2t+1}(x) = f_1(x) = x + (1-x)c(x)$, and, $v_{2t}(x) = v_0(x) = e(1-e)x$ and $v_{2t+1}(x) = v_1(x) = (1-x)c(x)(1-c(x))$. All of these functions are continuous, and $f_t(X_t^N)$ and $v_t(X_t^N)$ are uniformly bounded because $0 \le X_t^N \le 1$ and $c(x) \le 1$. The limiting trajectory satisfies (in particular) $x_{2(t+1)} = f(x_{2t})$, where $f = f_1 \circ f_0$. Thus we have the following simple result.

Theorem 4. *For the N-patch metapopulation models with parameters e and $c(x)$, let X_t^N be the proportion of occupied patches at census t. If $X_0^N \xrightarrow{\text{P}} x_0$ (a constant), then $X_t^N \xrightarrow{\text{P}} x_t$ for all $t \ge 1$, where $x_\bullet$ is determined by $x_{t+1} = f(x_t)$ $(t \ge 0)$ with*

$$f(x) = (1-e)x + (1-(1-e)x)c((1-e)x) \qquad \text{(EC model)}$$
$$f(x) = (1-e)(x + (1-x)c(x)). \qquad \text{(CE model)}$$

To obtain the corresponding central limit law for $Z_t^N = \sqrt{N}(X_t^N - x_t)$, first observe that our N-patch models can be represented as

$$n_{t+1} = \tilde{n}_t + \textstyle\sum_{j=1}^{N-\tilde{n}_t}\mathrm{Ber}_j(c(\tilde{n}_t/N)) \qquad \tilde{n}_t = n_t - \textstyle\sum_{j=1}^{n_t}\mathrm{Ber}_j(e) \qquad \text{(EC model)}$$
$$n_{t+1} = \tilde{n}_t - \textstyle\sum_{j=1}^{\tilde{n}_t}\mathrm{Ber}_j(e) \qquad \tilde{n}_t = n_t + \textstyle\sum_{j=1}^{N-n_t}\mathrm{Ber}_j(c(n_t/N)), \qquad \text{(CE model)}$$

where the $(\mathrm{Ber}_j(p))$ are collections of iid Bernoulli random variables with success probability p. We may thus apply Theorem 3 to the time-inhomogeneous Markov chain $(n_t^N,\, t \geq 0)$ obtained (as above) by setting $n_{2t}^N = n_t$ and $n_{2t+1}^N = \tilde{n}_t$, because this chain will have the form (2) with the $(\pm\,\xi_{jt}^N)$ being appropriate sequences of iid Bernoulli random variables. For both models, $g_{2t}(x) = g_{2t+1}(x) = x$. For the EC model, $r_{2t}(x) = x$, $r_{2t+1}(x) = 1 - x$, $m_{2t}(x) = -e$, $\sigma_{2t}^2(x) = e(1-e)$, $m_{2t+1}(x) = c(x)$ and $\sigma_{2t+1}^2(x) = c(x)(1-c(x))$, leading to

$$f_{2t}(x) = f_0(x) = (1-e)x \qquad f_{2t+1}(x) = f_1(x) = x + (1-x)c(x) \tag{10}$$

$$v_{2t}(x) = v_0(x) = e(1-e)x \quad v_{2t+1}(x) = v_1(x) = (1-x)c(x)(1-c(x)), \tag{11}$$

and, $b_{2t}(x) = e(1-e)(1-2e)$ and $b_{2t+1}(x) = c(x)(1-c(x))(1-2c(x))$. For the CE model, $r_{2t}(x) = 1-x$, $r_{2t+1}(x) = x$, $m_{2t}(x) = c(x)$, $\sigma_{2t}^2(x) = c(x)(1-c(x))$, $m_{2t+1}(x) = -e$ and $\sigma_{2t+1}^2(x) = e(1-e)$, leading to

$$\begin{aligned} f_{2t}(x) &= x + (1-x)c(x) \qquad & f_{2t+1}(x) &= (1-e)x \\ v_{2t}(x) &= (1-x)c(x)(1-c(x)) \qquad & v_{2t+1}(x) &= e(1-e)x, \end{aligned}$$

and, $b_{2t}(x) = c(x)(1-c(x))(1-2c(x))$ and $b_{2t+1}(x) = e(1-e)(1-2e)$. Thus, for the CE model the roles of f_0 and f_1, and v_0 and v_1, are reversed. If c is twice continuously differentiable, then (remembering that c is increasing and concave with $c(0) \geq 0$ and $c(x) \leq 1$) in both cases $f_t(x)$ will be twice continuously differentiable in x with bounded second derivative, $v_t(x)$ will be continuous in x, and $b_t(x)$ will be bounded in x.

We have already seen that our deterministic trajectory satisfies $x_{2(t+1)} = f(x_{2t})$, where $f = f_1 \circ f_0$ for the EC model and $f = f_0 \circ f_1$ for the CE model, with f_0 and f_1 as given in (10). Similarly, it is clear that our limiting Gaussian Markov chain $Z_\bullet$ must satisfy

$$Z_{2(t+1)} = f'(x_{2t})Z_{2t} + \hat{E}_{2t}, \qquad \text{with} \quad \hat{E}_{2t} \sim \mathrm{N}(0, v(x_{2t})),$$

where $v = v_1 \circ f_0 + (f_1' \circ f_0)^2 v_0$ for the EC model and $v = v_0 \circ f_1 + (f_0' \circ f_1)^2 v_1$ for the CE model, with v_0 and v_1 as given in (11), noting that $f_0'(x) = 1 - e$ and $f_1'(x) = 1 - c(x) + (1-x)c'(x)$. Thus we arrive at the following result.

Theorem 5. *For the N-patch metapopulation models with parameters e and $c(x)$, suppose that c is twice continuously differentiable. Let $Z_t^N = \sqrt{N}(X_t^N - x_t)$, where X_t^N is the proportion of occupied patches at census t and where $x_\bullet$ is determined by $x_{t+1} = f(x_t)$ $(t \geq 0)$ with f given as in Theorem 4. Then, if $Z_0^N \xrightarrow{\mathrm{D}} z_0$, $Z_\bullet^N$ converges weakly to the Gaussian Markov chain $Z_\bullet$ defined by $Z_{t+1} = f'(x_t)Z_t + E_t$ $(Z_0 = z_0)$, with (E_t) independent and $E_t \sim \mathrm{N}(0, v(x_t))$, where*

$$\begin{aligned} v(x) &= (1-(1-e)x)c((1-e)x)(1-c((1-e)x)) \\ &\qquad + e(1-e)x\big[1 - c((1-e)x) + (1-(1-e)x)c'((1-e)x)\big]^2 \qquad \text{(EC model)} \\ v(x) &= (1-e)\big[ex + (1-x)c(x)\big(1-(1-e)c(x)\big)\big]. \qquad \text{(CE model)} \end{aligned}$$

Since our limiting Gaussian Markov chain $Z_\bullet$ satisfies (3), now with f given as in Theorem 4, an immediate consequence of Theorem 5 is that $Z_t^N \xrightarrow{\mathrm{D}} \mathrm{N}(\mu_t, V_t)$, where

$$\mu_t = z_0 \prod_{s=0}^{t-1} f'(x_s) \quad \text{and} \quad V_t = \sum_{s=0}^{t-1} v(x_s) \prod_{u=s+1}^{t-1} f'(x_u)^2$$

(even though the constructed Markov chain $n_{\bullet}^{N}$ is time-inhomogeneous). Also, the joint distribution of numbers of occupied patches, observed at census times $t_1, \ldots, t_n$, can be approximated by an n-dimensional Gaussian distribution with means $Nx_{t_i} + \sqrt{N}\mu_{t_i}$ and covariances Nc_{t_i,t_j}, where $c_{t,s} := \mathrm{Cov}(Z_t, Z_s) = V_t \prod_{u=t}^{s-1} f'(x_u)$ $(s \geq t)$.

4.2. Long-term behaviour

Next we look at stationarity/quasi stationarity, and begin by examining the long-term $(t \to \infty)$ behaviour of our deterministic models. First notice that, in an obvious notation, $f_{\mathrm{CE}}((1-e)x) = (1-e)f_{\mathrm{EC}}(x)$, and so the fixed points of f_{CE} and f_{EC} are related by $x^*_{\mathrm{CE}} = (1-e)x^*_{\mathrm{EC}}$ (again in an obvious notation). Furthermore, x^*_{CE} and x^*_{EC} will have the same stability properties, because $f'_{\mathrm{CE}}((1-e)x) = f'_{\mathrm{EC}}(x)$, implying that $f'_{\mathrm{CE}}(x^*_{\mathrm{CE}}) = f'_{\mathrm{EC}}(x^*_{\mathrm{EC}})$, and $(1-e)f''_{\mathrm{CE}}((1-e)x) = f''_{\mathrm{EC}}(x)$, implying that $(1-e)f''_{\mathrm{CE}}(x^*_{\mathrm{CE}}) = f''_{\mathrm{EC}}(x^*_{\mathrm{EC}})$. Now, x^* is a fixed point of f_{CE} if and only if $c(x^*) = r(x^*)$, where $r(x) = \rho x/(1-x)$ and $\rho = e/(1-e)$, the function r having slope ρ at $x = 0$ and increasing strictly from 0 to ∞ as x increases from 0 to 1. But, recall that c is strictly increasing from $c(0) \geq 0$ and concave with $c(x) \leq 1$. Therefore, we always have precisely one stable fixed point and x_t approaches this point monotonically. We have the following three cases.

(i) *Stationarity*: $c(0) > 0$. There is a unique fixed point x^* in $[0,1]$ and this satisfies $0 < x^* < 1$. Moreover it is stable, because clearly $c'(x^*_{\mathrm{CE}}) < r'(x^*_{\mathrm{CE}})$, and a simple calculation shows that this entails $f'_{\mathrm{CE}}(x^*_{\mathrm{CE}}) < 1$.
(ii) *Evanescence*: $c(0) = 0$ and $c'(0) \leq \rho$. Now 0 is the unique fixed point in $[0,1]$. It is stable because if $c'(0) < \rho$, then $f'_{\mathrm{CE}}(0) = (1-e)(1+c'(0)) < 1$, while if $c'(0) = \rho$, then $f'_{\mathrm{CE}}(0) = 1$, but $f''_{\mathrm{CE}}(0) = (1-e)(c''(0) - 2c'(0)) < 0$.
(iii) *Quasi stationarity*: $c(0) = 0$ and $c'(0) > \rho$. There are two fixed points in $[0,1]$, 0 and $x^* \in (0,1)$; 0 is unstable because now $f'_{\mathrm{CE}}(0) = (1-e)(1 + c'(0)) > 1$ and x^* is stable by the same argument as used in (i).

Thus, in Case (i) we expect the unique stationary distribution of the process $n_{\bullet}^{N}$ to be centred near Nx^*. In Cases (ii) and (iii) there is an absorbing state 0 which is reached in a finite time. However, in Case (ii) we expect the process to be absorbed quickly (even for N quite large), while in Case (iii) a quasi-equilibrium will be reached; we expect the unique quasi-stationary distribution of $n_{\bullet}^{N}$ (being the limiting conditional state probabilities, conditional on non-extinction) to be centred near Nx^*.

Notice that $x^*_{\mathrm{CE}} < x^*_{\mathrm{EC}}$. Indeed, since $f_{\mathrm{CE}}(x) < f_{\mathrm{EC}}(x)$, the deterministic trajectory $x_{\bullet}$ will be uniformly smaller for the CE model than for the EC model.

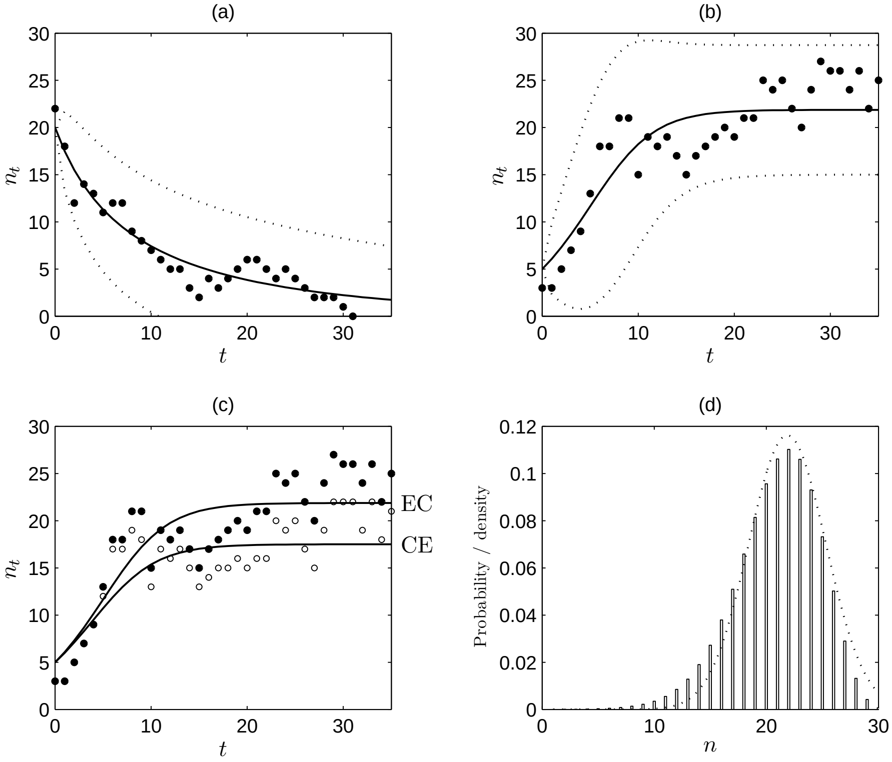

Fig 1. *N-patch metapopulation model with $N = 30$, $e = 0.2$ and $c(x) = cx$. Simulation (solid circles) of the EC Model with (a) $c = 0.2$ (evanescence) and (b) $c = 0.6$ (quasi stationarity); deterministic trajectories are shown (solid), together with ± 2 standard deviations of the Gaussian approximation (dotted). (c) Simulation of the EC (solid circles) and CE (open circles) models with $c = 0.6$; both deterministic trajectories shown (solid). (d) Quasi-stationary distribution (bars) of the EC Model with $c = 0.6$ and the stationary Gaussian pdf (dotted).*

This is to be expected, for our models differ only in when the census is taken; numbers observed following an extinction phase would likely be smaller than numbers observed following a colonisation phase. These remarks are supported by illustrations in Figure 1. Simulations are depicted for both the EC and CE models with $c(x) = cx$ (see Example 1 below), as well as the corresponding deterministic trajectories and quantities relating to the limiting Gaussian processes. The quasi-stationary distribution, $p^N = (p_i^N,\, i \in E_N)$, of $n_\bullet^N$ was evaluated as the normalized left eigenvector of the transition matrix restricted to E_N corresponding to its Perron-Frobenius eigenvalue (see Darroch and Seneta [21]), and this was compared with the approximating Gaussian pdf with mean Nx^* and variance NV^*, where $V^* = v(x^*)/(1 - f'(x^*)^2)$ (see Corollary 1 below). We note that, whilst the quasi-stationary distributions for our two models cannot be exhibited explicitly, they are always related by $\pi_{\mathrm{EC}} = \pi_{\mathrm{CE}}\bar{C}$ (in an obvious notation) with common Perron-Frobenius eigenvalue, where $\bar{C}$ denotes the matrix C (the colonisation transition matrix) restricted to E_N.

Our next result is obtained from Theorems 4 and 5 on setting $x_0 = x^*$. It established that in the stationary and quasi-stationary cases, where there is a positive stable deterministic equilibrium x^*, the fluctuations $Z_t^N = \sqrt{N}(X_t^N - x^*)$ of $X_\bullet^N$ about x^* can be approximated by an AR-1 process whose parameters can be exhibited explicitly.

Corollary 1. *For the N-patch metapopulation models with parameters e and $c(x)$, let X_t^N be the proportion of occupied patches at census t and let $Z_t^N = \sqrt{N}(X_t^N - x^*)$. Suppose that, in addition to c being twice continuously differentiable, $c(0) > 0$ or $c(0) = 0$ and $c'(0) > e/(1-e)$, and let x^* be the stable fixed point of f (f given as in Theorem 4). Then, $X_0^N \xrightarrow{\mathrm{P}} x^*$ implies that $X_t^N \xrightarrow{\mathrm{P}} x^*$ for all $t \geq 1$, in which case if $Z_0^N \xrightarrow{\mathrm{D}} z_0$, then $Z_\bullet^N$ converges weakly to the AR-1 process $Z_\bullet$ defined by $Z_{t+1} = f'(x^*)Z_t + E_t$ ($Z_0 = z_0$), with iid errors $E_t \sim \mathrm{N}(0, v(x^*))$, where v is given as in Theorem 5.*

One consequence of the corollary is that, for all $t \geq 1$, $Z_t^N \xrightarrow{\mathrm{D}} \mathrm{N}(z_0 a^t, V_t)$, where $a = f'(x^*)$ and now $V_t = v(x^*)(1-a^{2t})/(1-a^2)$. Another is that, if $Z_0^N \xrightarrow{\mathrm{D}} z_0$, then there will be a sequence of times (t_N) such that $Z_{t_N}^N \xrightarrow{\mathrm{D}} \mathrm{N}(0, V^*)$, where $V^* = v(x^*)/(1-a^2)$. Also, we expect that if the process has reached equilibrium/quasi equilibrium then the joint distribution of the numbers of occupied patches, observed at census times $t_1, \ldots, t_n$, can be approximated by an n-dimensional Gaussian distribution with means $Nx_{t_i} + \sqrt{N}\mu_{t_i}$ and covariances Nc_{t_i, t_j}, where $c_{t,s} := \mathrm{Cov}\,(Z_t, Z_s) = V_t\, a^{|s-t|}$. It would be of interest to determine how closely, for how long, and over what ranges, $X_\bullet^N$ is faithfully approximated. To this end, we might look at the time $\tau_N = \inf\{t \geq 1 : |Z_t^N| \geq e_N\}$ of first exit of $X_\bullet^N$ from an interval containing x^*, where $e_N \to \infty$. Based on Theorem 1 of Barbour [12] (who considered this problem for the continuous time analogue—a limiting Ornstein-Uhlenbeck process), we conjecture that if e_N does not grow too quickly, $X_\bullet^N$ is asymptotically equally likely to leave to the right of the interval as to the left, and, conditional on (say) leaving to the right, the exit time is asymptotically geometrically distributed.

We have already noted the simple relationship between the deterministic equilibria of our two models, $x^*_{\mathrm{CE}} = (1-e)x^*_{\mathrm{EC}}$, and that the decay rates are the same: $a = f'_{\mathrm{CE}}(x^*_{\mathrm{CE}}) = f'_{\mathrm{EC}}(x^*_{\mathrm{EC}})$. The stationary variances of the approximating AR-1 processes are also related. First, because $c(x^*) = r(x^*)$, it is easy to prove that

$$v_{\mathrm{CE}}(x^*_{\mathrm{CE}}) = e x^*_{\mathrm{CE}}\left(2 - e/(1 - x^*_{\mathrm{CE}})\right). \tag{12}$$

And, since it can also be shown that

$$(1-e)^2 v_{\mathrm{EC}}(x) = v_{\mathrm{CE}}((1-e)x) + e\big((1-e)x(f'_{\mathrm{CE}}((1-e)x))^2 - f_{\mathrm{CE}}((1-e)x)\big),$$

we have

$$(1-e)^2 v_{\mathrm{EC}}(x^*_{\mathrm{EC}}) = v_{\mathrm{CE}}(x^*_{\mathrm{CE}}) - e x^*_{\mathrm{CE}}(1-a^2), \tag{13}$$

and therefore $(1-e)^2 V^*_{\mathrm{EC}} = V^*_{\mathrm{CE}} - e x^*_{\mathrm{CE}}$.

4.3. Examples

We now illustrate these results by looking at particular instances of $c(x)$.

Example 1. Suppose that $c(x) = cx$ $(0 < c \leq 1)$. In this case we may write $f(x) = x(1 + r(1 - x/x^*))$, where $r = c(1 - e) - e$ for both models and x^* is the appropriate equilibrium: $x^*_{\mathrm{EC}} = r/(c(1 - e)^2)$ or $x^*_{\mathrm{CE}} = r/(c(1 - e))$ (both being strictly positive, and then stable, if and only if $c > e/(1 - e)$). Thus, our limiting deterministic trajectory follows the discrete logistic model (see for example Section 3.2 of Renshaw [55]), with r being the 'natural growth rate' and x^* being the 'carrying capacity' (expressed as a proportion of the ceiling N). Of course the logistic model is well known to exhibit a wide range of dynamic behaviour, but we emphasise that here $0 < 1 + r = (1 - e)(1 + c) < 2$. Our limiting Gaussian Markov chain has error variance

$$\begin{aligned} v(x) &= (1 - e)x\big[c(1 - (1 - e)x)(1 - c(1 - e)x) \\ &\qquad\qquad + e\big(1 + c - 2c(1 - e)x\big)^2\big] && \text{(EC model)} \\ v(x) &= (1 - e)x\big[e + c(1 - x)(1 - c(1 - e)x)\big]. && \text{(CE model)} \end{aligned}$$

In the quasi-equilibrium case $(r > 0)$, the limiting AR-1 process is defined by $Z_{t+1} = aZ_t + E_t$, where $a = 1 - r$ $(0 < a < 1)$, with $E_t \sim \mathrm{N}(0, v^*)$, where, from (12) and (13), $v^*_{\mathrm{CE}} = er(1 - e + a)/(e + r)$ and $(1 - e)^2 v^*_{\mathrm{EC}} = er(a(1 + a) - e)/(e + r)$. The stationary variance of $Z_\bullet$ is $V^* = v^*/(1 - a^2)$. Note that this decreases with a, so the faster the decay in the mean, the smaller the stationary variance.

Example 2. Suppose that $c(x) = c_0$, where $0 < c_0 \leq 1$. This case was studied in detail by us in [18]. As mentioned earlier, the metapopulation behaves as if, at every census, each occupied patch remains occupied with probability p, and, *independently*, each unoccupied patch is colonised with probability q, where

$$\begin{aligned} p &= 1 - e(1 - c_0) & q &= c_0 && \text{(EC model)} \\ p &= 1 - e & q &= (1 - e)c_0. && \text{(CE model)} \end{aligned}$$

Notice that for the EC model the 'effective' extinction probability is $e(1 - c_0)$. This accords with Hanski's [25] interpretation of the 'rescue effect', a term coined by Brown and Kodric-Brown [17] to describe the deleterious effect of colonisation on extinction when colonisation is frequent; Hanski argued that the extinction probability should be $(1 - c_0)e$, where e is the extinction probability in the absence of migration.

In [18] we proved that, for all $t \geq 1$, n^N_t has the same distribution as the sum of two independent random variables, $\mathrm{Bin}(n^N_0, p_t)$ and $\mathrm{Bin}(N - n^N_0, q_t)$, with success probabilities $q_t = q^*(1 - a^t)$ and $p_t = q_t + a^t$ $(t \geq 0)$, where $a = p - q = (1 - e)(1 - c_0)$ (the same for both EC and CE) and $q^* = q/(1 - a)$. The proportion X^N_t of occupied patches at time t has mean and variance given by $\mathbb{E}X^N_t = x_t(X^N_0)$ and $N\mathrm{Var}(X^N_t) = V_t(X^N_0)$, where

$$x_t(x_0) = x_0 p_t + (1 - x_0)q_t \quad \text{and} \quad V_t(x_0) = x_0 p_t(1 - p_t) + (1 - x_0)q_t(1 - q_t).$$

So, on the one hand, as $t \to \infty$, $\mathbb{E}X_t^N$ converges to q^* at geometric rate a (note that $0 < a < 1$) and $N\mathrm{Var}\,(X_t^N) \to q^*(1-q^*)$ (indeed, $n_\bullet^N$ has a $\mathrm{Bin}(N, q^*)$ stationary distribution). On the other, letting $N \to \infty$ with t fixed, $\mathbb{E}X_t^N \to x_t(x_0)$ and $N\mathrm{Var}\,(X_t^N) \to V_t(x_0)$ whenever $X_0^N \xrightarrow{\mathrm{P}} x_0$. Furthermore, because n_t^N is the sum of two independent binomial random variables, it is clear that $Z_t^N := \sqrt{N}(X_t^N - x_t)$ and $Y_t^N := \sqrt{N}(X_t^N - q^*)$ will converge in distribution to Gaussian random variables if their initial values converge. Theorem 5 and Corollary 1 provide more detailed information. Since $f(x) = px + q(1-x)$, and hence $f'(x) = a$, and $v(x) = p(1-p)x + q(1-q)(1-x)$, we deduce that if $Z_0^N \xrightarrow{\mathrm{D}} z_0$, then $Z_\bullet^N$ converges weakly to a Gaussian Markov chain $Z_\bullet$ with $\mathbb{E}Z_t = a^t z_0$ and $\mathrm{Cov}\,(Z_t, Z_s) = V_t(x_0)\, a^{|s-t|}$, while if $Y_0^N \xrightarrow{\mathrm{D}} y_0$, then $Y_\bullet^N$ converges weakly to an AR-1 process $Y_\bullet$ with $\mathbb{E}Y_t = a^t y_0$ and $\mathrm{Cov}\,(Y_t, Y_s) = q^*(1-q^*)a^{|s-t|}(1-a^{2t})$; the error variance here is $v(q^*) = q^*(1-q^*)(1-a^2)$.

Finally, we remark that if n_0^N follows the stationary $\mathrm{Bin}(N, q^*)$ law, our representation $n_{t+1} = \mathrm{Bin}(n_t, p_1) + \mathrm{Bin}(N - n_t, q_1)$ is termed *binomial autoregressive* [43, 44, 65, 66]. Our results establish a connection with standard autoregressive processes.

Example 3. Suppose that $c(x) = c_0 + cx$, where $c_0 > 0$, $c > 0$ and $c_0 + c \le 1$. Now we may write $f(x) = \nu + x(1 + r(1 - x/K))$, where $\nu = c_0(1-e)$ and $r = (c - c_0)(1-e) - e$ for both models, and K depends on which model: $K_{\mathrm{EC}} = r/(c(1-e)^2)$ or $K_{\mathrm{CE}} = r/(c(1-e))$ (in an obvious notation). Since $c(0) = c_0 > 0$, we have unique stable equilibria $x^*_{\mathrm{EC}} = x^*_{\mathrm{CE}}/(1-e)$ with x^*_{CE} being the unique positive solution to $c(1-e)x^2 - rx - \nu = 0$. The common decay rate is $a = 1 + r(1 - 2x^*/K)$. The error variance can be evaluated, but omitted for brevity's sake. The limiting AR-1 process has stationary variance $v^*/(1-a^2)$, where v^* is given by (12) or (13).

Example 4. Suppose that $c(x) = 1 - \exp(-\beta x)$, where $\beta > 0$ is the propagation rate. Since $c(0) = 0$ and $c'(0) = \beta > 0$, we have evanescence if $\beta \le e/(1-e)$ and quasi stationarity if $\beta > e/(1-e)$. The limiting Gaussian Markov chain has error variance

$$v(x) = \mathrm{e}^{-\beta(1-e)x}\Big((1 - (1-e)x)(1 - \mathrm{e}^{-\beta(1-e)x}) + e(1-e)x\big[1 + (1 - (1-e)x)\beta\big]^2 \mathrm{e}^{-\beta(1-e)x}\Big) \qquad \text{(EC model)}$$

$$v(x) = (1-e)\big[ex + (1-x)(1 - \mathrm{e}^{-\beta x})\big(e + (1-e)\mathrm{e}^{-\beta x}\big)\big]. \qquad \text{(CE model)}$$

In the quasi-stationary case the deterministic equilibria cannot be exhibited explicitly, but can be evaluated numerically by iterating the map $f_{\mathrm{CE}}(x) = (1-e)(1 - (1-x)\exp(-\beta x))$, remembering that $x^*_{\mathrm{EC}} = x^*_{\mathrm{CE}}/(1-e)$. The limiting AR-1 process has stationary variance $v^*/(1-a^2)$, where v^* is evaluated using (12) or (13). A simple calculation reveals that

$$a = \frac{(1 + \beta(1 - x^*_{\mathrm{CE}}))(1 - e - x^*_{\mathrm{CE}})}{(1-e)(1 - x^*_{\mathrm{CE}})}.$$

5. Infinite-patch models

As previously, let n_t be the number of occupied patches at time t, but suppose now that $(n_t,\, t \geq 0)$ is a Markov chain taking values in $S = \{0, 1, \dots\}$ that evolves as follows:

$$
\begin{aligned}
n_{t+1} &= \tilde{n}_t + \mathrm{Poi}(m(\tilde{n}_t)) \qquad & \tilde{n}_t &= n_t - \mathrm{Bin}(n_t, e) \qquad & &\text{(EC model)}\\
n_{t+1} &= \tilde{n}_t - \mathrm{Bin}(\tilde{n}_t, e) \qquad & \tilde{n}_t &= n_t + \mathrm{Poi}(m(n_t)), \qquad & &\text{(CE model)}
\end{aligned}
$$

where $m(n) \geq 0$. So, as before, extinction and colonisation occur in alternating phases, and occupied patches go extinct independently with probability e $(0 < e < 1)$, but now the number of colonisations follows a Poisson law and the expected number of colonisations is a function of the number of patches presently occupied.

Before embarking on the general case, let us examine the important special case where the expected number of colonisations is a linear function of the number of patches presently occupied.

5.1. The infinite-patch branching model

Suppose that $m(n) = mn$, where $m > 0$. The parameter m can be interpreted as the expected number of colonisations by *any one* occupied patch. As we noted earlier, this is the natural analogue of our N-patch models, for recall that, if $c(0) = 0$ and c has a continuous second derivative near 0, then $\mathrm{Bin}(N - n, c(n/N)) \overset{\mathrm{D}}{\to} \mathrm{Poi}(mn)$ as $N \to \infty$, where $m = c'(0)$. But, this infinite-patch scheme has a simplifying feature, namely *branching*, which makes both models much simpler to analyse. Notice that if there are n occupied patches at the beginning of any given phase, then the number occupied at the end of that phase has the same distribution as the sum of n independent copies of either $B = \mathrm{Ber}(1-e)$ (extinction phase) or $P_+ := 1 + \mathrm{Poi}(m)$ (colonisation phase). And, since the phases are conditionally independent, the net effect is that n_{t+1} will have the same distribution as the sum of n_t independent copies of Y, where Y is either B independent copies of P_+ (the EC model) or P_+ independent copies of B (the CE model). We therefore make the following simple observation.

Proposition 1. *The process $(n_t,\, t \geq 0)$ is a Galton-Watson process whose offspring distribution has pgf $G(z)$ given by*

$$
\begin{aligned}
G(z) &= e + (1-e)z\mathrm{e}^{-m(1-z)} \qquad & &\text{(EC model)}\\
G(z) &= (e + (1-e)z)\mathrm{e}^{-m(1-e)(1-z)}. \qquad & &\text{(CE model)}
\end{aligned}
$$

Thus, we think of the census times as marking the 'generations' of our branching process, the 'particles' being the occupied patches, and the 'offspring' being the occupied patches that they notionally replace in the succeeding generation. For the EC model, it is as if each occupied patch becomes empty with probability e, or otherwise colonises $\mathrm{Poi}(m)$ patches, while, for the CE model, it is as if each occupied patch survives with probability $1-e$ or becomes empty with probability e, but, whatever happens, it colonises $\mathrm{Poi}(m(1-e))$ patches.

We may now invoke the encylopaedic theory of branching processes [8, 9, 10, 28] to prove results for this important special case of the model; it is just a matter of which questions are of interest. For example, it is easy to prove that offspring distribution has mean $\mu = (1+m)(1-e)$ (the same for both models), and so $\mathbb{E}(n_t|n_0) = n_0\mu^t$ $(t \geq 1)$. Also, our branching process is *subcritical*, *critical* or *supercritical* according as m is *less than*, *equal to* or *greater than* the critical value $\rho = e/(1-e)$. This accords immaculately with our earlier criteria for evanescence ($c'(0) \leq \rho$) versus quasi stationarity ($c'(0) > \rho$) of our N-patch models with $c(0) = 0$. We also have the following simple result concerning the probability that the metapopulation becomes extinct (totally extinct), starting with n_0 patches occupied.

Corollary 2. *For the infinite-patch branching model, total extinction occurs with probability* 1 *if and only if* $m \leq \rho$*; otherwise total extinction occurs with probability* η^{n_0}*, where* η *is the unique fixed point of* G *on the interval* $(0,1)$*, with* G *given as in Proposition 1.*

The extinction probability η cannot be exhibited explicitly, but can of course be obtained numerically by iterating the map G.

Finally, the variance of the offspring distribution is $\sigma^2 = \sigma^2_{\mathrm{EC}} := (1-e)((1+c)^2e + c)$ or $\sigma^2 = \sigma^2_{\mathrm{CE}} := (e+c)(1-e)$, depending on which model, and so

$$\mathrm{Var}\,(n_t|n_0) = \begin{cases} n_0\sigma^2 t & \text{if } \mu = 1 \quad (\, m = \rho\,) \\ n_0\sigma^2(\mu^t - 1)\mu^{t-1}/(\mu - 1) & \text{if } \mu \neq 1 \quad (\, m \neq \rho\,), \end{cases}$$

for all $t \geq 1$.

A simple extension of the branching model is obtained by setting $m(n) = m_0 + mn$, where now $m \geq 0$, and the new parameter m_0 (> 0) is to be interpreted as the expected number of colonisations from an external source. It can be derived from our N-patch models if they are modified so that colonisation probability $c(n/N)$ is replaced by $m_0/N + c(n/N)$ (imagine that an external colonization potential m_0 is apportioned equally among all N patches), for then $\mathrm{Bin}(N-n, m_0/N + c(n/N)) \overset{\mathrm{D}}{\to} \mathrm{Poi}(m_0 + c'(0)n)$. The effect is to introduce an additional (independent) $\mathrm{Poi}(m_0)$ number of colonisations in each colonisation phase. It is easy to see that the resulting process $(n_t,\, t \geq 0)$ will be the Galton-Watson process identified in Proposition 1, but modified so that there are $\mathrm{Poi}(d)$ immigrant particles in each generation, where $d = m_0$ for the EC model and $d = (1-e)m_0$ for the CE model. Again we can invoke general theory. For example, on applying Theorem VI.7.2 of [10], we learn that $n_\bullet$ has a proper limiting distribution if and only if $m < \rho$.

5.2. The infinite-patch model with regulated colonisation

Returning now to the general case, where the expected number of colonisations depends arbitrarily on the number of occupied patches, let us consider what happens when the initial number of occupied patches becomes large. We will suppose that there is an index N such that $m(n) = N\mu(n/N)$, where μ is continuous with bounded first derivative. We may take N to be simply n_0 or, more generally, following Klebaner [32], we may interpret N as being a 'threshold' with the property that $n_0/N \to x_0$ as $N \to \infty$. By choosing μ appropriately, we may allow for a degree of regulation in the colonisation process; for example, $\mu(x)$ might be of the form $\mu(x) = rx(a - x)$ $(0 \le x \le a)$ (logistic growth), $\mu(x) = xe^{r(1-x)}$ $(x \ge 0)$ (Ricker growth dynamics) or $\mu(x) = \lambda x/(1 + ax)^b$ $(x \ge 0)$ (Hassell growth dynamics) (see [55]). Under these conditions we can establish a law of large numbers for $X_t^N = n_t/N$, the number of occupied patches at census t measured *relative to* the threshold.

The CE model is always density dependent because

$$\mathbb{E}(n_{t+1}|n_t) = (1-e)(n_t + m(n_t)) = (1-e)(n_t + N\mu(n_t/N)),$$

implying that $f_t(x) := f(x) = (1-e)(x + \mu(x))$, and

$$\begin{aligned}\operatorname{Var}(n_{t+1}|n_t) &= e(1-e)\mathbb{E}(\tilde{n}_t|n_t) + (1-e)^2 \operatorname{Var}(\tilde{n}_t|n_t)\\ &= e(1-e)(n_t + m(n_t)) + (1-e)^2 m(n_t)\\ &= e(1-e)(n_t + N\mu(n_t/N)) + (1-e)^2 N\mu(n_t/N),\end{aligned}$$

implying that $v_t(x) := v(x) = (1-e)(ex + \mu(x))$, but, whilst f and v are both continuous, we cannot apply Theorem 1 because X_t^N is not necessarily bounded. However, Theorem 2 can be used; since μ is continuous with bounded first derivative, it is Lipschitz continuous and hence so too are f and v. The EC model is not always density dependent, but we may work with the phases separately, applying Theorem 2 to the time-inhomogeneous Markov chain $(n_t^N, t \ge 0)$ obtained by setting $n_{2t}^N = n_t$ and $n_{2t+1}^N = \tilde{n}_t$. We let $f_{2t}(x) = f_0(x) = (1-e)x$ and $f_{2t+1}(x) = f_1(x) = x + \mu(x)$, and, $v_{2t}(x) = v_0(x) = e(1-e)x$ and $v_{2t+1}(x) = v_1(x) = \mu(x)$, noting that all are Lipschitz continuous. Thus, $X_0^N \xrightarrow{2} x_0$ as $N \to \infty$ is sufficient for convergence of $X_\bullet^N$ to a limiting deterministic trajectory $x_\bullet$, which satisfies (in particular) $x_{2(t+1)} = f(x_{2t})$, where $f = f_1 \circ f_0$. We may summarise these observations as follows.

Theorem 6. *For the infinite-patch metapopulation models with parameters e and $\mu(x)$, let $X_t^N = n_t/N$ be the number of occupied patches at census t relative to the threshold N. Suppose that μ is continuous with bounded first derivative. If $X_0^N \xrightarrow{2} x_0$ as $N \to \infty$, then $X_t^N \xrightarrow{2} x_t$ (and hence $X_t^N \xrightarrow{\mathrm{P}} x_t$) for all $t \ge 1$, where $x_\bullet$ is determined by $x_{t+1} = f(x_t)$ $(t \ge 0)$ with*

$$\begin{aligned} f(x) &= (1-e)x + \mu((1-e)x) && \text{(EC model)}\\ f(x) &= (1-e)(x + \mu(x)). && \text{(CE model)}\end{aligned}$$

Having established that $X_t^N \xrightarrow{\mathrm{P}} x_t$ for all $t \ge 0$, we can also prove a central limit law for the scaled fluctuations $Z_t^N = \sqrt{N}(X_t^N - x_t)$ under the stronger condition that μ is twice continuously differentiable with bounded second derivative.

First observe that our infinite-patch models have the equivalent representation

$$n_{t+1} = \tilde{n}_t + \textstyle\sum_{j=1}^{N} \mathrm{Poi}_j(\mu(\tilde{n}_t/N)) \qquad \tilde{n}_t = n_t - \textstyle\sum_{j=1}^{n_t} \mathrm{Ber}_j(e) \qquad \text{(EC model)}$$

$$n_{t+1} = \tilde{n}_t - \textstyle\sum_{j=1}^{\tilde{n}_t} \mathrm{Ber}_j(e) \qquad \tilde{n}_t = n_t + \textstyle\sum_{j=1}^{N} \mathrm{Poi}_j(\mu(n_t/N)), \qquad \text{(CE model)}$$

where the $(\mathrm{Poi}_j(\,\cdot\,))$ are collections of iid Poisson random variables with mean $\mu(n_t/N)$. Thus, mirroring our argument leading to Theorem 5 for our N-patch models, we may apply Theorem 3 to the time-inhomogeneous Markov chain $(n_t^N,\, t \geq 0)$ obtained by setting $n_{2t}^N = n_t$ and $n_{2t+1}^N = \tilde{n}_t$. This chain also has the form (2), but with the $(\pm\xi_{jt}^N)$ now being appropriate sequences of iid Poisson random variables. For both models, $g_{2t}(x) = g_{2t+1}(x) = x$. For the EC model, $r_{2t}(x) = x$, $r_{2t+1}(x) = 1$, $m_{2t}(x) = -e$, $\sigma_{2t}^2(x) = e(1-e)$ and $m_{2t+1}(x) = \sigma_{2t+1}^2(x) = \mu(x)$, leading to

$$f_{2t}(x) = f_0(x) = (1-e)x \qquad f_{2t+1}(x) = f_1(x) = x + \mu(x) \tag{14}$$

$$v_{2t}(x) = v_0(x) = e(1-e)x \qquad v_{2t+1}(x) = v_1(x) = \mu(x), \tag{15}$$

and, $b_{2t}(x) = b_0(x) = e(1-e)(1-2e)$ and $b_{2t+1}(x) = b_1(x) = \mu(x)$. For the CE model, $r_{2t}(x) = 1$ and $r_{2t+1}(x) = x$, $m_{2t}(x) = \sigma_{2t}^2(x) = \mu(x)$, $m_{2t+1}(x) = -e$ and $\sigma_{2t+1}^2(x) = e(1-e)$, leading to the same expressions for f_t, v_t and b_t, but with the roles f_0 and f_1, v_0 and v_1, and b_0 and b_1, reversed. Since μ is twice continuously differentiable with bounded second derivative, in both cases $f_t(x)$ will be twice continuously differentiable in x with bounded second derivative, $v_t(x)$ will be continuous in x, and $b_t(x)$ will be bounded in x.

We have just seen that the limiting deterministic trajectory satisfies $x_{2(t+1)} = f(x_{2t})$, where $f = f_1 \circ f_0$ for the EC model and $f = f_0 \circ f_1$ for the CE model, with f_0 and f_1 as given in (14). Similarly, it is clear that our limiting Gaussian Markov chain $Z_\bullet$ should take the form

$$Z_{2(t+1)} = f'(x_{2t}) Z_{2t} + \hat{E}_{2t}, \qquad \text{with} \quad \hat{E}_{2t} \sim \mathrm{N}(0, v(x_{2t})),$$

where $v = v_1 \circ f_0 + (f_1' \circ f_0)^2 v_0$ for the EC model and $v = v_0 \circ f_1 + (f_0' \circ f_1)^2 v_1$ for the CE model, with v_0 and v_1 as given in (15), noting that $f_0'(x) = 1 - e$ and $f_1'(x) = 1 + \mu'(x)$. Thus we arrive at the following result.

Theorem 7. *For the infinite-patch metapopulation models with parameters e and $\mu(x)$, suppose that μ is twice continuously differentiable with bounded second derivative. Let X_t^N be the proportion of occupied patches at census t, and suppose that $X_0^N \stackrel{2}{\to} x_0$, so that $X_t^N \stackrel{\mathrm{P}}{\to} x_t$ for all $t \geq 0$, where $x_\bullet$ is determined by $x_{t+1} = f(x_t)$ $(t \geq 0)$ with f given as in Theorem 6. Let $Z_t^N = \sqrt{N}(X_t^N - x_t)$ and suppose that $Z_0^N \stackrel{\mathrm{D}}{\to} z_0$. Then, $Z_\bullet^N$ converges weakly to the Gaussian Markov chain $Z_\bullet$ defined by $Z_{t+1} = f'(x_t) Z_t + E_t$ $(Z_0 = z_0)$, with (E_t) independent and $E_t \sim \mathrm{N}(0, v(x_t))$, where*

$$v(x) = \mu((1-e)x) + e(1-e)x\,(1 + \mu'((1-e)x)) \qquad \text{(EC model)}$$

$$v(x) = (1-e)(ex + \mu(x)). \qquad \text{(CE model)}$$

Equilibrium behaviour is richer and more interesting than for the earlier N-patch models, because now the limiting deterministic models can exhibit the full range of long-term behaviour, and we cannot be as precise in classifying this behaviour as we were earlier. First notice that, in the notation adopted earlier, $f_{\mathrm{CE}}((1-e)x) = (1-e)f_{\mathrm{EC}}(x)$, and so fixed points of f_{CE} and f_{EC} are related by $x^*_{\mathrm{CE}} = (1-e)x^*_{\mathrm{EC}}$, and, as before, $f'_{\mathrm{CE}}(x^*_{\mathrm{CE}}) = f'_{\mathrm{EC}}(x^*_{\mathrm{EC}})$ and $(1-e)f''_{\mathrm{CE}}(x^*_{\mathrm{CE}}) = f''_{\mathrm{EC}}(x^*_{\mathrm{EC}})$, implying that x^*_{CE} and x^*_{EC} have the same stability properties. Notice also that x^*_{CE} will be a fixed point of f_{CE} if and only if $\mu(x^*_{\mathrm{CE}}) = \rho x^*_{\mathrm{CE}}$, where $\rho = e/(1-e)$. So, if $\mu(0) = 0$ then 0 is a fixed point; it is stable if $\mu'(0) < 1$ and unstable if $\mu'(0) > 1$ (if $\mu'(0) = 1$ its stability is determined by higher derivatives of μ near $x = 0$). However, even when $\mu(0) = 0$, there might be other (conceivably many) fixed points; our conditions on μ do not preclude this. Certainly if there *is* a unique positive fixed point x^*, it will be stable if $\mu'(x^*) < 1$ and unstable if $\mu'(x^*) > 1$ (again we need to consider higher derivatives when $\mu'(x^*) = 1$). Finally, notice that the d-th iterates of our maps are also related by $f^{(d)}_{\mathrm{CE}}((1-e)x) = (1-e)f^{(d)}_{\mathrm{EC}}(x)$, which means that if $x^*_0, x^*_1, \dots, x^*_{d-1}$ is a limit cycle for the deterministic EC model, then $(1-e)x^*_0, (1-e)x^*_1, \dots, (1-e)x^*_{d-1}$ is a limit cycle for the deterministic CE model.

To illustrate, we look at the case where the expected number of colonisations relative to the threshold obeys a Ricker law. For this model the full range of behaviour is exhibited, and we *can* be precise in classifying this behaviour.

Ricker growth dynamics. Suppose that $\mu(x) = x\exp(r(1-x))$, where $r > 0$, so that the colonisation potential of the occupied patches is greatest when their number is close to N/r; the parameter r can be interpreted as the growth rate and N the carrying capacity of the metapopulation *in the absence of extinction*. The fixed points of f_{CE} are 0 and $x^*_{\mathrm{CE}} = 1 - r_0/r$, where $r_0 = \log(\rho)$. Notice that $f'_{\mathrm{CE}}(x) = (1-e)(1 + (1-rx)e^{r(1-x)})$, implying that $f'_{\mathrm{CE}}(0) = (1-e)(1+e^r)$, and $f''_{\mathrm{CE}}(x) = -(1-e)(2-rx)re^{r(1-x)}$, implying that $f''_{\mathrm{CE}}(0) = -2(1-e)re^r$. Therefore, if $r \le r_0$, 0 is the unique non-negative fixed point, and it is stable. If $r > r_0$, then x^*_{CE} is an additional positive fixed point; it is stable because $f'_{\mathrm{CE}}(x^*_{\mathrm{CE}}) = 1 - e(r - r_0) < 1$ (and 0 is unstable). However, if r is sufficiently large, we get limiting cycles with period doubling towards chaos, as illustrated in Figure 2.

Figure 3 illustrates some of the range of behaviour exhibited by the stochastic CE model with Ricker growth dynamics. Four cases are depicted. *Evanescence* (a): for r between 0 and r_0 ($\simeq 0.8473$), corresponding to 0 being the unique stable fixed point, the process dies out quickly. *Quasi stationarity* (b): for r between r_0 and $r_1 \simeq 0.7$, corresponding to x^*_{CE} ($\simeq 1 - 0.8473/r$) being the unique stable fixed point, the process exhibits (quasi-) equilibrium behaviour; r_1 is the point of first period-doubling (we have not been able to determine r_1 analytically). *Oscillation* (c) and (d): for r bigger than r_1 the process 'tracks' the limit cycles of the deterministic model (period 2 and 4, respectively). To be sure, the theory predicts the kind of behaviour shown in Figure 3(c), but it is no less remarkable when one witnesses it; one could not mistake it for, say, bistability.

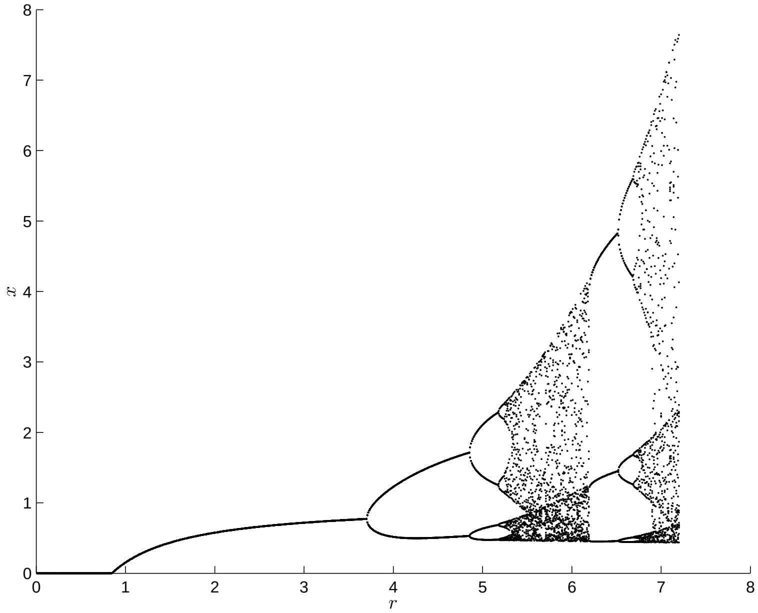

FIG 2. *Bifurcation diagram for the infinite-patch deterministic CE model with Ricker growth dynamics:* $x_{n+1} = (1-e)x_n(1+e^{r(1-x_n)})$. *Here* $e = 0.7$ *and* r *ranges from* 0 *to* 7.2.

Returning now to generality, we complete the picture by presenting two results concerning the fluctuations of $X_\bullet^N$ about a positive stable equilibrium x^*, assuming $X_0^N \overset{2}{\to} x^*$, or a stable limit cycle $x_0^*, x_1^*, \ldots, x_{d-1}^*$, assuming $X_0^N \overset{2}{\to} x_0^*$. Corollary 3 follows from Theorem 7 on setting $x_0 = x^*$, so that then $x_t = x^*$ for all $t \geq 0$, and evaluating the stationary error variance $v(x^*)$ in both cases. Corollary 4 follows from Theorem 7 on setting $x_0 = x_0^*$, so that then $x_\bullet$ tracks the limit cycle, that is, $x_{nd+j} = x_j^*$ $(n \geq 0,\ j = 0, \ldots, d-1)$. The representation of $Z_\bullet$ as a d-variate AR-1 process $\boldsymbol{Y}_\bullet$, and in particular the form of the coefficient matrix A and the error covariance matrix Σ_d, follow by iterating $Z_{t+1} = f'(x_t)Z_t + E_t$ $(Z_0 = z_0)$, with (E_t) independent $\mathrm{N}(0, v(x_t))$ random variables: using expressions (6) to (9) with $f_t = f$ and $v_t = v$, noting that $\Pi_{i,\,j} = \prod_{k=i}^{j-1} f'(x_k^*) = a_j/a_i$ for $1 \leq i \leq j \leq d$, we obtain a representation of Z_{nd+j} $(j = 1, \ldots, d)$ in terms of Z_{nd} $(n \geq 0)$, as well as the stationary covariance matrix V. Notice that, because $Z_\bullet$ is Markovian, only $Z_{(n+1)d-1}$ contributes to the drift in $\boldsymbol{Y}_n$.

Corollary 3. *Suppose that* f *given in Theorem 6 admits a unique positive fixed point* x^* *satisfying* $\mu'(x^*) < 1$. *Then, if* $X_0^N \overset{2}{\to} x^*$, $x_t = x^*$ *for all* t *and, assuming* $Z_0^N \to z_0$, *the limit process* $Z_\bullet$ *determined by Theorem 7 is AR-1 process* $Z_\bullet$ *defined by* $Z_{t+1} = aZ_t + E_t$ $(Z_0 = z_0)$, *where* $a = (1-e)(1+\mu'(x_{\mathrm{CE}}^*))$ *(being the same for both models), and with iid errors* $E_t \sim \mathrm{N}(0, v)$, *where* $v = e(1+a)x^*$ *(EC model) or* $v = e(2-e)x^*$ *(CE model).*

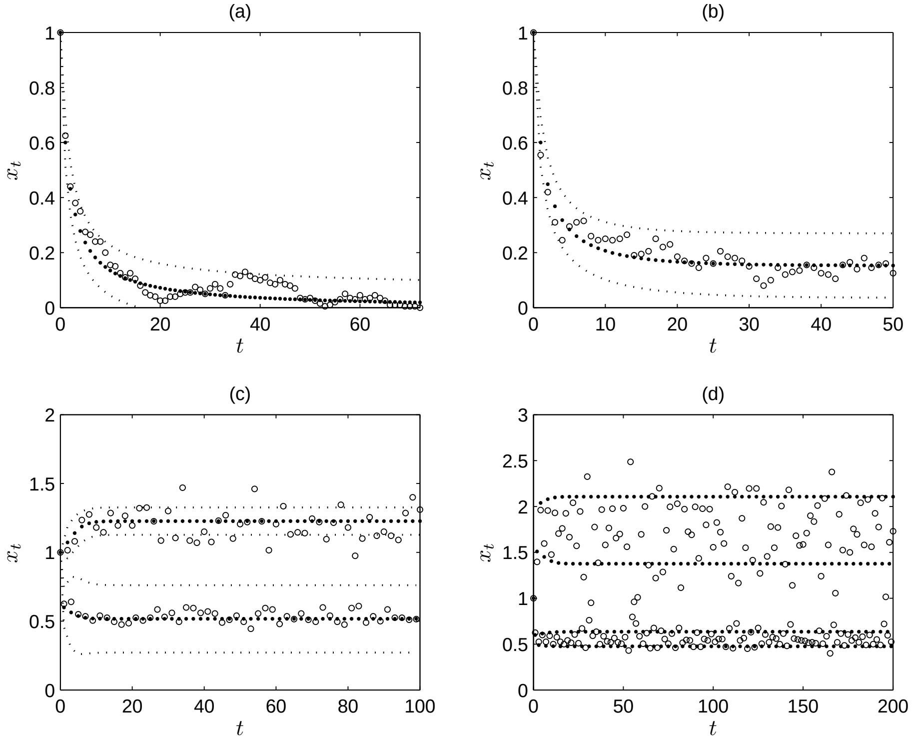

Fig 3. *Simulation (open circles) of the infinite-patch CE model with Ricker growth dynamics, together with the corresponding limiting deterministic trajectories (small solid circles). Here* $e = 0.7$ *and* $N = 200$*, and, (a)* $r = 0.84$*, (b)* $r = 1$ *(c)* $r = 4$ *and (d)* $r = 5$*. In (a), (b) and (c), the dotted lines indicate* ± 2 *standard deviations of the Gaussian approximation (in (c) every* second *point is joined to indicate variation about each of the two limit cycle values).*

Corollary 4. *Suppose that* f *given in Theorem 6 admits a stable limit cycle* $x_0^*, x_1^*, \ldots, x_{d-1}^*$ *with* $X_0^N \xrightarrow{2} x_0^*$*. Then,* $x_{nd+j} = x_j^*$ $(n \geq 0,\ j = 0, \ldots, d-1)$ *and, assuming* $Z_0^N \to z_0$*, the limit process* $Z_\bullet$ *determined by Theorem 7 has the following representation:* $(\boldsymbol{Y}_n,\ n \geq 0)$*, where* $\boldsymbol{Y}_n = (Z_{nd}, Z_{nd+1}, \ldots, Z_{(n+1)d-1})^\top$ *with* $Z_0 = z_0$*, is a* d*-variate AR-1 process of the form* $\boldsymbol{Y}_{n+1} = A\boldsymbol{Y}_n + \boldsymbol{E}_n$*, where* $(\boldsymbol{E}_n)$ *are independent and* $\boldsymbol{E}_n \sim \mathrm{N}(\mathbf{0}, \Sigma_d)$*; here* A *is the* $d \times d$ *matrix*

$$A = \begin{pmatrix} 0 & 0 & \cdots & a_1 \\ 0 & 0 & \cdots & a_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_d \end{pmatrix},$$

where $a_j = \prod_{i=0}^{j-1} f'(x_i^*)$*,* $\Sigma_d = (\sigma_{ij})$ *is the* $d \times d$ *symmetric matrix with entries*

$$\sigma_{ij} = a_i a_j \sum_{k=0}^{i-1} v(x_k^*)/a_{k+1}^2 \qquad (1 \leq i \leq j \leq d),$$

where v *is given as in Theorem 7, and the random entries,* $(Z_1, \ldots, Z_{d-1})$, *of* $\boldsymbol{Y}_0$ *have a Gaussian* $\mathrm{N}(\boldsymbol{a} z_0, \Sigma_{d-1})$ *distribution, where* $\boldsymbol{a} = (a_1, \ldots, a_{d-1})$. *Furthermore,* $\boldsymbol{Y}_{\bullet}$ *has a Gaussian* $\mathrm{N}(\boldsymbol{0}, V)$ *stationary distribution, where* $V = (v_{ij})$ *has entries given by*

$$v_{ij} = \frac{a_i a_j}{1 - a_d^2} \sum_{k=0}^{d-1} v(x_k^*)/a_{k+1}^2 \qquad (1 \le i \le j \le d).$$

6. Continuous-time analogues

When extinction and colonisation events happen in random order, rather than in alternating phases, it is natural to take $(n_t,\, t \ge 0)$, where n_t is the number occupied patches at time t, to be a Markov chain in continuous time. But, what is the most appropriate model?

If the probability of colonisation in a small time interval were independent of the number of occupied patches, then the SIS model would seem to be the most appropriate N-patch model, and the Levins model (1) could be used, in much the same way as above, to draw conclusions about its long-term behaviour. Evanescence and quasi stationarity could be distinguished by examining the stability of the equilibrium points, 0 and $n^* = N(1 - e/c)$; if $c \le e$, 0 would be stable and the population would have genuine evanescent character, while if $c > e$, n^* would be strictly positive and stable, and the population would persist (in this latter case the quasi-stationary distribution would be centred near n^* [47, 48]).

If, as envisaged here, the probability of colonisation were to depend on the proportion of patches currently occupied, the natural N-patch continuous-time model would be a birth-death process on $S = \{0, 1, \ldots, N\}$ with birth rates $\lambda_n = c(n/N)(N - n)$ and death rates $\mu_n = en$, where $c(x)$ is as above (assumed to be continuous, increasing and concave, with $c(0) \ge 0$ and $c(x) \le 1$). To see this, suppose that an occupied patch becomes empty in a time interval of length h with probability $eh + \mathrm{o}(h)$, or remains occupied with probability $1 - eh + \mathrm{o}(h)$, while if there are n occupied patches at time t, then any given unoccupied patch becomes occupied in the interval $(t, t + h]$ with probability $c(n/N)h + \mathrm{o}(h)$, or remains unoccupied with probability $1 - c(n/N)h + \mathrm{o}(h)$. Suppose also that the chance of two or more events of either kind happening in $(t, t + h]$ is $\mathrm{o}(h)$. Then, a transition from n to $n + 1$ in time h is effected by having *exactly one* colonization and *no* extinctions (there are other ways, but all have probability $\mathrm{o}(h)$), and the chance of this happening in time h is

$$\begin{aligned}
\Pr(n_{t+h} &= n + 1 | n_t = n) \\
&= (N - n)(c(n/N)h + \mathrm{o}(h))(1 - c(n/N)h + \mathrm{o}(h))^{N-n-1} \\
&\qquad \times (1 - eh + \mathrm{o}(h))^n + \mathrm{o}(h) \\
&= (N - n)c(n/N)h + \mathrm{o}(h) = \lambda_n h + \mathrm{o}(h).
\end{aligned}$$

Similarly,

$$\begin{aligned}\Pr(n_{t+h} &= n-1 | n_t = n) \\ &= n(eh+\mathrm{o}(h))(1-eh+\mathrm{o}(h))^{n-1}(1-c(n/N)h+\mathrm{o}(h))^{N-n}+\mathrm{o}(h) \\ &= enh+\mathrm{o}(h) = \mu_n h+\mathrm{o}(h)\end{aligned}$$

and

$$\begin{aligned}\Pr(n_{t+h} &= n | n_t = n) \\ &= (1-eh+\mathrm{o}(h))^{n}(1-c(n/N)h+\mathrm{o}(h))^{N-n}+\mathrm{o}(h) \\ &= 1-((N-n)c(n/N)+en)\,h+\mathrm{o}(h) = 1-(\lambda_n+\mu_n)h+\mathrm{o}(h),\end{aligned}$$

as well as $\Pr(n_{t+h} = m | n_t = n) = \mathrm{o}(h)$ when $|m-n| \geq 2$.

So, analogous to Examples 1, 2 and 3 above, if $c(x) = cx$ we get the stochastic SIS model, if $c(x) = c_0$ we get the continuous-time Ehrenfest model (see Section 1.4 of [31]), while if $c(x) = c_0 + cx$ we get the mainland-island model of Alonso and McKane [2] (see also [57]), all being instances of Feller's stochastic logistic model [23]. Whatever the form of $c(x)$, we can obtain continuous-time analogues of Theorems 4 and 5 and Corollary 1, because our birth-death model is *density dependent* in the sense of Kurtz [35, 36]. It follows immediately from Theorem 3.1 of [35] that the proportion $X_t^N = n_t/N$ of occupied patches at time t converges (uniformly in probability over finite time intervals) to a deterministic trajectory $(x_t,\, t \geq 0)$ satisfying the law of motion

$$\frac{dx}{dt} = F(x) \quad (t \geq 0), \quad \text{where} \quad F(x) = c(x)(1-x) - ex \quad (0 \leq x \leq 1), \tag{16}$$

assuming of course that $X_0^N \to x_0$ (our conditions on c imply that F is Lipschitz continuous on $[0,1]$). Furthermore, if we let $Z_t^N = \sqrt{N}\left(X_t^N - x_t\right)$, then, assuming $Z_0^N \to z_0$, Theorem 3.5 of [36] can be used to show that process $(Z_t^N,\, t \geq 0)$ converges weakly in $\mathcal{D}[0,t]$ (the space of right-continuous left-limits functions on $[0,t]$) to a Gaussian diffusion $(Z_t,\, t \geq 0)$ with initial value $Z_0 = z_0$ and with mean $\mathbb{E}\, Z_t = M_t z_0$, where $M_t = \exp(\int_0^t B_u\, du)$ and $B_t := F'(x_t)$, and variance $V_t = M_t^2 \int_0^t M_u^{-2} G(x_u)\, du$, where $G(x) = F(x) + 2ex$. In the important special case where x_0 is taken to be an equilibrium point x^* of (16), usually a stable equilibrium, the approximating diffusion is an *Ornstein-Uhlenbeck process*, and more precise results are available [61]. For example, if $F(x^*) = 0$ and $B := F'(x^*) < 0$, then $\mathbb{E}\, Z_t = e^{-at} z_0$, where $a = -B$, and $V_t = V^*(1 - e^{-2at})$, where $V^* = G(x^*)/(-2B) = ex^*/a$ (the stationary variance).

There is a one to one correspondence between the equilibria of (16) and the fixed points f_{CE}; we simply replace ρ $(= e/(1-e))$ above by e, because $F(x) = 0$ if and only if $c(x) = ex/(1-x)$. There is the same correspondence in their classification (again simply replace ρ by e), because $F'(x) = c'(x)(1-x) - c(x) - e$ and $F''(x) = c''(x)(1-x) - 2c'(x)$, and, since

$$f_{\text{CE}}(x) - x = (1-e)(x + (1-x)c(x)) - x = (1-e)(c(x)(1-x) - \rho x),$$

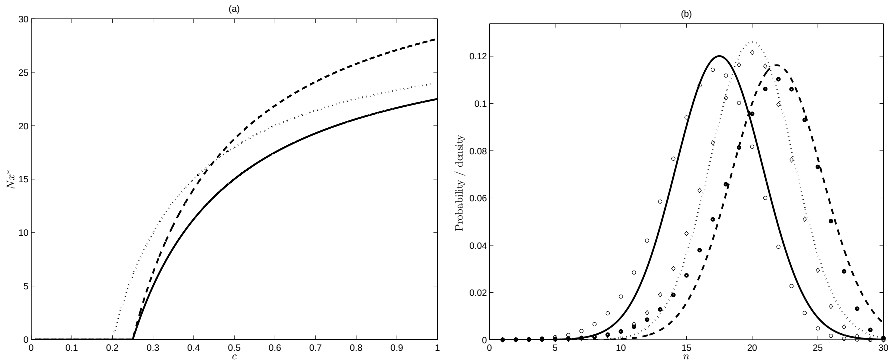

Fig 4. *N-patch metapopulation model with* $N = 30$, $e = 0.2$ *and* $c(x) = cx$. *(a)* x^*_{CE} *(solid),* x^*_{SIS} *(dotted) and* x^*_{EC} *(dashed) versus* c. *(b) pdf of the stationary distribution of the limiting Gaussian process for the CE model (solid), the SIS model (dotted) and the EC model (dashed) with* $c = 0.6$, *together with the corresponding quasi-stationary distributions (open circles, diamonds and solid circles).*

we have that $f'_{\mathrm{CE}}(x) - 1 = (1-e)(c'(x)(1-x) - c(x) - \rho)$ and $f''_{\mathrm{CE}}(x) - 1 = (1-e)(c''(x)(1-x) - 2c'(x))$. Thus, stationarity happens when $c(0) > 0$, evanescence when $c(0) = 0$ and $c'(0) \leq e$, and quasi stationarity when $c(0) = 0$ and $c'(0) > e$. One might have expected that, in stationary and quasi-stationary regimes, the stable equilibrium would lie between x^*_{EC} and x^*_{EC}, but this turns out not to be the case.

To illustrate this, and compare other equilibrium characteristics, suppose that $c(x) = cx$. We will assume that $\rho < c \leq 1$, so that also $c > e$. Then, $x^*_{\mathrm{CE}} = 1 - e/(c(1-e))$, $x^*_{\mathrm{EC}} = x^*_{\mathrm{CE}}/(1-e)$ and $x^*_{\mathrm{SIS}} = 1 - e/c$ (in an obvious notation) are all positive and stable. Thus, we always have $x^*_{\mathrm{SIS}} > x^*_{\mathrm{CE}}$, but $x^*_{\mathrm{EC}} > x^*_{\mathrm{SIS}}$ *only if* $c > \rho + e$. Figure 4(a) shows the equilibria plotted against c for fixed e ($e = 0.2$); note that here $\rho = 0.25$ and $\rho + e = 0.45$.

Figure 4(b) shows the quasi-stationary distributions of the three models, CE, SIS and EC, and the pdfs of the corresponding $\mathrm{N}(0, V^*)$ stationary distributions of the limiting Gaussian processes; $V^*_{\mathrm{CE}} = er(1-e+a)/((e+r)(1-a^2))$, where $r = c(1-e) - e$ and $a = 1 - r$, $V^*_{\mathrm{EC}} = er(a(1+a) - e)/((e+r)(1-e)^2(1-a^2))$ and $V^*_{\mathrm{SIS}} = e/c$ (in an obvious notation).

The natural continuous-time analogue of our infinite-patch models is the birth-death process on $S = \{0, 1, \dots\}$ with birth rates $\lambda_n = m(n)$ and death rates $\mu_n = en$. When $m(n) = mn$, where $m > 0$, we get the simple immigration-death process (see Section 6.1 of [24]), which can be interpreted as a Markov branching process with binary splitting. Even with the inclusion of an immigration term, $m(n) = m_0 + mn$, the model is completely tractable (explicit results are catalogued in Section 3.2 of [3]). If, more generally, $m(n) = N\mu(n/N)$, for some threshold N, where μ is continuous with bounded first derivative (implying that F is Lipschitz continuous), Theorem 3.1 of [35] guarantees that if the number $X^N_t = n_t/N$ of occupied patches at time t, measured relative to the threshold, converges (uniformly in probability over finite time intervals) to a deterministic trajectory $(x_t,\, t \geq 0)$ satisfying

$$\frac{dx}{dt} = F(x) \quad (t \geq 0), \quad \text{where} \quad F(x) = \mu(x) - ex \quad (x \geq 0), \tag{17}$$

whenever $X_0^N \to x_0$. Furthermore, Theorem 3.5 of [36] guarantees that $(Z_t^N,\, t \geq 0)$, where $Z_t^N = \sqrt{N}\left(X_t^N - x_t\right)$, converges weakly in $\mathcal{D}[0,t]$ to a Gaussian diffusion $(Z_t,\, t \geq 0)$ with initial value $Z_0 = z_0$, assuming $Z_0^N \to z_0$. Its mean is $\mathbb{E}\, Z_t = M_t z_0$, where $M_t = \exp(\int_0^t B_u\, du)$ and $B_t := F'(x_t) = \mu'(x_t) - e$, and its variance is $V_t = M_t^2 \int_0^t M_u^{-2} G(x_u)\, du$, where $G(x) = F(x) + 2ex$. When x_0 is taken to be a stable equilibrium point x^* of (17) $Z_\bullet$ is an Ornstein-Uhlenbeck process with $\mathbb{E}\, Z_t = e^{-at} z_0$, where $a = e - \mu'(x^*)$, and $V_t = V^*(1 - e^{-2at})$, where $V^* = ex^*/a$ is the stationary variance. It will be clear that the equilibria of (17), and their classification, will be the same as that laid out in Section 5.2 for the discrete-time models, *but with ρ replaced by e.*

Acknowledgements

Our thanks go to Servet Martinez for pointing out the connection between our finite-patch state-independent colonisation models and the urn model described in Section 2, and to Ross McVinish for comments on an early draft of this manuscript and for making us aware of the connection with binomial autoregressive processes. We are also grateful to the referee and the Editor for helpful comments and suggestions. This research was supported by the Australian Research Council Centre of Excellence for Mathematics and Statistics of Complex Systems.

References

[1] Akçakaya, H. and Ginzburg, L. (1991). Ecological risk analysis for single and multiple populations. In *Species Conservation: A Population Biological Approach*, A. Seitz and V. Loescheke, Eds., Birkhauser, Basel, 78–87.

[2] Alonso, D. and McKane, A. (2002). Extinction dynamics in mainland-island metapopulations: an N-patch stochastic model. *Bull. Math. Biol. 64*, 913–958.

[3] Anderson, W. (1991). *Continuous-Time Markov Chains: An Applications-Oriented Approach*. Springer-Verlag, New York. MR1118840

[4] Andrewartha, H. and Birch, L. (1954). *The Distribution and Abundance of Animals*. University of Chicago Press, Chicago, USA.

[5] Arrigoni, F. (2001). Approssimazione deterministica di modelli stocastici di metapopulazione. *Bollettino dell'Unione Matematica Italiana, Serie VIII IV-A*, 387–390.

[6] Arrigoni, F. (2003). A deterministic approximation of a stochastic metapopulation model. *Adv. Appl. Probab. 3*, 691–720. MR1990610

[7] Arrigoni, F. and Pugliese, A. (2002). Limits of a multi-patch SIS epidemic model. *J. Math. Biol. 45*, 419–440. MR1942373

[8] Asmussen, S. and Hering, H. (1983). *Branching Processes*. Birkhauser, Boston. MR0701538

[9] Athreya, K. and Jagers, P. (1997). *Classical and Modern Branching Processes*. The IMA Volumes in Mathematics and its Applications, Vol. **84**. Springer-Verlag, New York. Papers from the IMA Workshop held at the University of Minnesota, Minneapolis, MN, June 13–17, 1994. MR1601681

[10] Athreya, K. and Ney, P. (1972). *Branching Processes*. Springer-Verlag, Berlin. MR0373040

[11] Barbour, A. (1974). On a functional central limit theorem for Markov population processes. *Adv. Appl. Probab. 6*, 21–39. MR0341610

[12] Barbour, A. (1976). Quasi-stationary distributions in Markov population processes. *Adv. Appl. Probab. 8*, 296–314. MR0410960

[13] Barbour, A. (1980a). Density dependent Markov population processes. In *Biological growth and spread: mathematical theories and applications (Proc. Conf., Heidelberg, 1979), Lecture Notes in Biomathematics*, W. Jager, H. Rost, and P. Tautu, Eds. Vol. **38**. Springer, Berlin, 36–49. MR0609344

[14] Barbour, A. (1980b). Equilibrium distributions for Markov population processes. *Adv. Appl. Probab. 12*, 591–614. MR0578839

[15] Barbour, A. and Pugliese, A. (2005). Asymptotic behaviour of a metapopulation model. *Ann. Appl. Probab. 15*, 1306–1338. MR2134105

[16] Billingsley, P. (1995). *Probability and Measure*, 3rd ed. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York. MR1324786

[17] Brown, J. and Kodric-Brown, A. (1977). Turnover rates in insular biogeography: effect of immigration on extinction. *Ecology 58*, 445–449.

[18] Buckley, F. and Pollett, P. (2009). Analytical methods for a stochastic mainland-island metapopulation model. In *Proceedings of the 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation*, R. Anderssen, R. Braddock, and L. Newham, Eds. Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, Canberra, Australia, 1767–1773.

[19] Daley, D. and Gani, J. (1999). *Epidemic Modelling: an Introduction*. Cambridge Studies in Mathematical Biology, Vol. **15**. Cambridge University Press, Cambridge. MR1688203

[20] Darling, R. and Norris, J. (2008). Differential equation approximations for Markov chains. *Probab. Surv. 5*, 37–79. MR2395153

[21] Darroch, J. and Seneta, E. (1965). On quasi-stationary distributions in absorbing discrete-time Markov chains. *J. Appl. Probab. 2*, 88–100. MR0179842

[22] Day, J. and Possingham, H. (1995). A stochastic metapopulation model with variable patch size and position. *Theoret. Pop. Biol. 48*, 333–360.

[23] Feller, W. (1939). Die grundlagen der volterraschen theorie des kampfes ums dasein in wahrscheinlichkeitsteoretischer behandlung. *Acta Biotheoretica 5*, 11–40. MR0690284

[24] Grimmett, G. and Stirzaker, D. (2001). *Probability and Random Processes*, 3rd ed. Oxford University Press, New York. MR2059709

[25] Hanski, I. (1994). A practical model for metapopulation dynamics. *J. Animal Ecol. 63*, 151–162.

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

[27] Hanski, I. and Simberloff, D. (1997). The metapopulation approach, its history, conceptual domain, and application to conservation. In *Metapopulation Biology*, I. Hanski and M. Gilpin, Eds. Academic Press, San Diego, CA, USA, 5–26.

[28] Harris, T. (1963). *The Theory of Branching Processes*. Springer-Verlag, Berlin. MR0163361

[29] Hill, M. and Caswell, H. (2001). The effects of habitat destruction in finite landscapes: a chain-binomial metapopulation model. *OIKOS 93*, 321–331.

[30] Karr, A. (1975). Weak convergence of a sequence of Markov chains. *Probability Theory and Related Fields 33*, 41–48. MR0394882

[31] Kelly, F. (1979). *Reversibility and Stochastic Networks*. Wiley, Chichester. MR0554920

[32] Klebaner, F. (1993). Population-dependent branching processes with a threshold. *Stochastic Process. Appl. 46*, 115–127. MR1217689

[33] Klebaner, F. and Nerman, O. (1994). Autoregressive approximation in branching processes with a threshold. *Stochastic Process. Appl. 51*, 1–7. MR1288280

[34] Klok, C. and De Roos, A. (1998). Effects of habitat size and quality on equilibrium density and extinction time of *Sorex araneus* populations. *J. Animal Ecol. 67*, 195–209.

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

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

[37] Kurtz, T. (1972). The relationship between stochastic and deterministic models in chemical reactions. *J. Chem. Phys. 57*, 2976–2978.

[38] Kurtz, T. (1976). Limit theorems and diffusion approximations for density dependent Markov chains. *Math. Prog. Study 5*, 67–78. MR0445626

[39] Kurtz, T. (1978). Strong approximation theorems for density dependent Markov chains. *Stochastic Process. Appl. 6*, 223–240. MR0464414

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

[41] Levins, R. (1970). Extinction. In *Some Mathematical Questions in Biology*, M. Gerstenhaber, Ed. American Mathematical Society, Providence, RI, USA, 75–107. MR0286512

[42] MacArthur, R. and Wilson, E. (1967). *The Theory of Island Biogeography*. Princeton University Press, Princeton, NJ, USA.

[43] McKenzie, E. (1985). Some simple models for discrete variate time series. *Water Resources Bull. 21*, 645–650.

[44] McKenzie, E. (2003). Discrete variate time series. In *Handbook of Statistics*, C. Rao and D. Shanbhag, Eds., Elsevier Science, Amsterdam, 573–606. MR1973555

[45] McVinish, R. and Pollett, P. (2009). Limits of large metapopulations with patch dependent extinction probabilities. Submitted for publication.

[46] Moilanen, A. (1999). Patch occupancy models of metapopulation dynamics: efficient parameter estimation using implicit statistical inference. *Ecology 80*, 1031–1043.

[47] Nåsell, I. (1996). The quasi-stationary distribution of the closed endemic SIS model. *Adv. Appl. Probab. 28*, 895–932. MR1404315

[48] Nåsell, I. (1999). On the quasi-stationary distribution of the stochastic logistic epidemic. *Math. Biosci. 156*, 21–40. MR1686454

[49] Norman, M. (1972). *Markov Processes and Learning Models*. Academic Press, New York. MR0423546

[50] Ovaskainen, O. and Hanski, I. (2004). Metapopulation dynamics in highly fragmented landscapes. In *Ecology, Genetics, and Evolution in Metapopulations*, I. Hanski and O. Gaggiotti, Eds. Academic Press, San Diego, CA, USA, 73–103.

[51] Pearl, R. (1925). *The Biology of Population Growth*. Alfred A. Knopf, New York.

[52] Pearl, R. (1927). The growth of populations. *Quart. Rev. Biol. 2*, 532–548.

[53] Pearl, R. and Reed, L. (1920). On the rate of growth of population of the united states since 1790 and its mathematical representation. *Proc. Nat. Academy Sci. 6*, 275–288.

[54] Pollett, P. (2001). Diffusion approximations for ecological models. In *Proceedings of the International Congress on Modelling and Simulation*, F. Ghassemi, Ed. Vol. **2**. Modelling and Simulation Society of Australia and New Zealand, Canberra, Australia, 843–848.

[55] Renshaw, E. (1991). *Modelling Biological Populations in Space and Time*. Cambridge Studies in Mathematical Biology, Vol. **11**. Cambridge University Press, Cambridge. MR1130616

[56] Ross, J. (2006a). A stochastic metapopulation model accounting for habitat dynamics. *J. Math. Biol. 52*, 788–806. MR2235528

[57] Ross, J. (2006b). Stochastic models for mainland-island metapopulations in static and dynamic landscapes. *Bull. Math. Biol. 68*, 417–449. MR2224776

[58] Ross, J., Pagendam, D., and Pollett, P. (2009). On parameter estimation in population models II: multi-dimensional processes and transient dynamics. *Theoret. Pop. Biol. 75*, 123–132.

[59] Ross, J., Taimre, T., and Pollett, P. (2006). On parameter estimation in population models. *Theoret. Pop. Biol. 70*, 498–510.

[60] Rout, T., Hauser, C., and Possingham, H. (2007). Minimise long-term loss or maximise short-term gain? Optimal translocation strategies for threatened species. *Ecological Modelling 201*, 67–74.

[61] Schach, S. (1971). Weak convergence results for a class of multivariate Markov processes. *Ann. Math. Statist. 42*, 451–465. MR0295441

[62] Tenhumberg, B., Tyre, A., Shea, K., and Possingham, H. (2004). Linking wild and captive populations to maximise species persistence: optimal translocation strategies. *Conservation Biology 18*, 1304–1314.

[63] Verhulst, P. (1838). Notice sur la loi que la population suit dans son accroisement. *Corr. Math. et Phys. X*, 113–121.

[64] Weiss, G. and Dishon, M. (1971). On the asymptotic behaviour of the stochastic and deterministic models of an epidemic. *Math. Biosci. 11*, 261–265. MR0295451

[65] Weiss, C. (2009a) A new class of autoregressive models for time series of binomial counts. *Comm. Statist. Theory Meth. 38*, 447–460. MR2510797

[66] Weiss, C. (2009b) Monitoring correlated processes with binomial marginals. *J. Appl. Statist. 36*, 399–414.

[67] Wright, S. (1940). Breeding structure of populations in relation to speciation. *American Naturalist 74*, 232–248.

[68] Zhang, Y., Liu, L., and Xu, R. (2009). The effect of migration on the viability, dynamics and structure of two coexisting metapopulations. *Ecological Modelling 220*, 272–282.