Spatial populations with seed-bank
Oomen, M.

CHAPTER 5

Proofs long-time behaviour $N < \infty$

In this chapter we prove Theorems 4.3.2–4.3.3. The integral criterion for $\rho = \infty$ in (4.50) is explained in Section 5.1. Theorem 4.3.2 is proved in Section 5.2 and Theorem 4.3.3 in Section 5.3.

§5.1 Explanation of clustering criterion for infinite seed-bank

Recall Fig. 4.6. Suppose that $g = dg_{\mathrm{FW}}$, so that we have a dual. We will show that the integral criterion in (4.50) determines whether or not two dual lineages coalesce with probability 1. Since two lineages in the dual can only coalesce when they are active at the same site, we need to keep track of the probabilities that the lineages are active at a given time. Because the lineages can only migrate when they are active, we also need to keep track of the total time they are active up to a given time.

Recall the renewal interpretation of the dual process (see Remark 4.2.9). We argue heuristically as follows. If $\rho = \infty$, then the activity times σ_k are much smaller than the sleeping times τ_k, and we may assume that $\tau_k + \sigma_k \asymp \tau_k$, $k \to \infty$. Discretising time, we can use the results from [1] for the intersection of two independent renewal processes. Then the integral criterion in (4.50) can be interpreted as follows:

- If $\gamma \in (0,1)$, then the probability for each of the lineages to be active at time s decays like $\asymp \varphi(s)^{-1} s^{-(1-\gamma)}$ [1]. Hence the total time they are active up to time s is $\asymp \varphi(s)^{-1} s^{\gamma}$. Because the lineages only move when they are active, the probability that the two lineages meet at time s is $\asymp a^{(N)}_{\varphi(s)^{-1}s^{\gamma}}(0,0)$. Hence the total hazard is $\asymp \int_1^\infty \mathrm{d}s\, [\varphi(s)^{-1} s^{-(1-\gamma)}]^2\, a^{(N)}_{\varphi(s)^{-1}s^{\gamma}}(0,0)$. After the transformation $t = t(s) = \varphi(s)^{-1} s^{\gamma}$, the latter turns into the integral in (4.50), modulo a constant. When carrying out this transformation, we need that $s\varphi'(s)/\varphi(s) \to 0$, which follows from (4.49), and $\varphi(t(s))/\varphi(s) \asymp 1$, which follows from the bound we imposed on ψ in (4.49) together with the fact that $\log \varphi(s)/\log s \to 0$. This computation is spelled out in Appendix B.1.

- If $\gamma = 1$, then the probability for each of the lineages to be active at time s decays like $\hat{\varphi}(s)^{-1}$ [1], and so the total time they are active up to time s is $\asymp s\hat{\varphi}(s)^{-1}$. Recall from (4.48) that $\hat{\varphi}(t) = \mathbb{E}[\tau \wedge t]$ is also slowly varying.) Hence the total hazard is $\asymp \int_1^\infty \mathrm{d}s\, [\hat{\varphi}(s)^{-1}]^2\, a^{(N)}_{\hat{\varphi}(s)^{-1}s}(0,0)$. After the transformation

$t = t(s) = \hat{\varphi}(s)^{-1}s$ (for which we can use the same type of computation as in Appendix B.1), the latter turns into the integral in (4.50), modulo a constant.

§5.2 Scaling of wake-up time and migration kernel for infinite seed-bank

We can prove Theorem 4.3.2 by direct computation via assumptions (4.52)–(4.53). We start by computing γ. Afterwards we compute $\hat{\varphi}(t)$ and $a_t^{\Omega_N}(0,0)$.

Computation of γ. Recall (4.40), which reads

$$\mathbb{P}(\tau > t) = \frac{1}{\chi} \sum_{m \in \mathbb{N}_0} K_m \frac{e_m}{N^m}\, \mathrm{e}^{-(e_m/N^m)t}. \tag{5.1}$$

Since we are interested in the asymptotic behaviour of $\mathbb{P}(\tau > t)$ as $t \to \infty$, we need to consider only large values of t. For large values of t, only large values of m (for which $\frac{e_m}{N^m}$ is small) contribute to the sum in (5.1). Hence we can estimate the latter by an integral and insert the assumptions made in (4.52)–(4.53). Subsequently, using the change of variable $s = \frac{e_m}{N^m}$ and taking the logarithm to express m in terms of s, we obtain the following values of γ after extracting the t-dependence:

$$\begin{aligned} &(4.52) \quad \Longrightarrow \quad \gamma = 1, \quad \varphi(t) \asymp (\log t)^{-\alpha}, \\ &(4.53) \quad \Longrightarrow \quad \gamma = \gamma_{N,K,e} = \frac{\log(N/Ke)}{\log(N/e)}, \quad \varphi(t) \asymp 1. \end{aligned} \tag{5.2}$$

In order to guarantee that $\rho = \infty$, we must require that $\alpha \in (-\infty, 1]$, respectively, $K \in [1, \infty)$ (while β, respectively, e play no role). Subject to (4.52),

$$\hat{\varphi}(t) \asymp \begin{cases} (\log t)^{1-\alpha}, & \alpha \in (-\infty, 1), \\ \log\log t, & \alpha = 1, \end{cases} \tag{5.3}$$

while subject to (4.53),

$$\hat{\varphi}(t) \asymp \begin{cases} 1, & K \in (1, \infty), \\ \log t, & K = 1. \end{cases} \tag{5.4}$$

Computation of $a_t^{\Omega_N}(0,0)$. To compute $a_t^{\Omega_N}(0,0)$, we first rewrite the migration kernel $a^{\Omega_N}(\cdot,\cdot)$ in (4.6) as

$$a^{\Omega_N}(0,\eta) = \frac{r_{\|\eta\|}}{N^{\|\eta\|-1}(N-1)} \tag{5.5}$$

with

$$r_{\|\eta\|} = \frac{1}{D(N)} \frac{N-1}{N} \sum_{l \geq \|\eta\|} \frac{c_{l-1}}{N^{l-1}} \frac{1}{N^{l-\|\eta\|}}, \tag{5.6}$$

where $D(N)$ is a renormalisation constant such that $\sum_{j\in\mathbb{N}} r_j = 1$. For transition kernels of the form (5.5), the time-t transition kernel $a_t^{\Omega_N}(\cdot,\cdot)$ was computed in [35] with the help of Fourier analysis, see also [19]. Namely,

$$a_t^{\Omega_N}(0,\eta) = \sum_{j\geq k} K_{jk}(N)\,\frac{\exp[-h_j(N)t]}{N^j}, \qquad t\geq 0, \quad \eta\in\Omega_N\colon\ d_{\Omega_N}(0,\eta)=k\in\mathbb{N}_0, \tag{5.7}$$

where

$$K_{jk}(N) = \begin{cases} 0, & j=k=0,\\ -1, & j=k>0,\\ N-1, & \text{otherwise}, \end{cases} \qquad j,k\in\mathbb{N}_0, \tag{5.8}$$

and

$$h_j(N) = \frac{N}{N-1}\,r_j(N) + \sum_{i>j} r_i(N), \qquad j\in\mathbb{N}. \tag{5.9}$$

The expressions in (5.6)–(5.9) simplify considerably in the limit as $N\to\infty$, namely, the term with $i=j$ dominates and

$$h_j(N) \sim r_j(N) \sim \frac{c_{j-1}}{D(N)N^{j-1}}, \quad j\in\mathbb{N}, \qquad D(N)\sim c_0. \tag{5.10}$$

We show why this is true for $h_j(N)$ (the argument for $r_j(N)$ and $D(N)$ is similar). Write

$$\begin{aligned} h_j(N) &= \frac{N}{N-1}\,r_j(N) + \sum_{i>j} r_i(N)\\ &= \frac{1}{D(N)}\left(\sum_{l\geq j}\frac{c_{l-1}}{N^{l-1}}\frac{1}{N^{l-j}} + \frac{N-1}{N}\sum_{l>j}\frac{c_{l-1}}{N^{l-1}}\sum_{i<j\leq l}\frac{1}{N^{l-i}}\right)\\ &= \frac{1}{D(N)}\frac{c_{j-1}}{N^{j-1}}\left(1+\left[1+O\left(\frac{1}{N}\right)\right]\left(\frac{c_{j-1}}{N^{j-1}}\right)^{-1}\sum_{l>j}\frac{c_{l-1}}{N^{l-1}}\right). \end{aligned} \tag{5.11}$$

Hence it suffices to show that

$$\limsup_{N\to\infty}\left(\frac{c_{j-1}}{N^{j-1}}\right)^{-1}\sum_{l>j}\frac{c_{l-1}}{N^{l-1}} = 0, \qquad j\in\mathbb{N}. \tag{5.12}$$

To do so, note that, since $\limsup_{k\to\infty}\frac{1}{k}\log c_k < \log N$ by (4.7), for N large enough we have

$$\sup_{k\in\mathbb{N}_0} c_k^{1/k} < N. \tag{5.13}$$

Let $\bar{N} = \inf\{N\in\mathbb{N}\colon\ \sup_{k\in\mathbb{N}_0} c_k^{1/k} < N\}$. Then

$$\begin{aligned} &\limsup_{N\to\infty}\left(\frac{c_{j-1}}{N^{j-1}}\right)^{-1}\sum_{l>j}\frac{c_{l-1}}{N^{l-1}} \leq \limsup_{N\to\infty}\frac{1}{c_{j-1}}\sum_{l>j}\frac{\bar{N}^{l-1}}{N^{l-1}}N^{j-1}\\ &= \frac{\bar{N}^{j-1}}{c_{j-1}}\limsup_{N\to\infty}\frac{\frac{\bar{N}}{N}}{1-\frac{\bar{N}}{N}} = 0, \qquad j\in\mathbb{N}, \end{aligned} \tag{5.14}$$

which settles (5.12).

To understand what (5.9) gives for finite N, note that for asymptotically polynomial coefficients (recall (4.52))

$$\begin{aligned}\left(\frac{c_{j-1}}{N^{j-1}}\right)^{-1}\sum_{l>j}\frac{c_{l-1}}{N^{l-1}} &= [1+o(1)]\,\frac{N^{j-1}}{F(j-1)^{-\phi}}\sum_{l>j}\frac{F(l-1)^{-\phi}}{N^{l-1}}\\ &= [1+o(1)]\sum_{l>j}\frac{(l-1)^{-\phi}}{(j-1)^{-\phi}}N^{-(l-j)} \\ &= [1+o(1)]\sum_{k\geq 1}\left(1+\frac{k}{j-1}\right)^{-\phi}N^{-k}, \qquad j\in\mathbb{N}.\end{aligned} \tag{5.15}$$

For $\phi \geq 0$ the right-hand side is bounded from above by $\sum_{k\geq 1} N^{-k} = \frac{1}{N-1}$ and for $\phi < 0$ by $N^{-1}\sum_{k\geq 1}(1+k)^{-\phi}N^{-(k-1)} \leq N^{-1}C_\phi$. On the other hand, for pure exponential coefficients (recall (4.53)),

$$\left(\frac{c_{j-1}}{N^{j-1}}\right)^{-1}\sum_{l>j}\frac{c_{l-1}}{N^{l-1}} = \sum_{k\geq 1}\left(\frac{c}{N}\right)^{-k} = \frac{c}{N-c}. \tag{5.16}$$

Hence, for both choices of coefficients we have the following:

> For $N\to\infty$ the quantities $h_j(N), r_j(N)$ are bounded from above and below by positive finite constants times the right-hand side of (5.10) uniformly in $j\in\mathbb{N}$. (5.17)

Picking $\eta = 0$ ($k = 0$) in (5.7), we obtain

$$a_t^{\Omega_N}(0,0) = \sum_{j\in\mathbb{N}}(N-1)\frac{\exp[-h_j(N)t]}{N^j}. \tag{5.18}$$

Since we are interested in the asymptotic behaviour of $a_t^{\Omega_N}(0,0)$, only large values of j are relevant and we can estimate the sum in (5.18) by an integral. To do so, we change variables by putting $s = h_j(N)$ and exploit (5.17). Take the logarithm to express j in terms of s, compute $\mathrm{d}s/\mathrm{d}j$, and extract the t-dependence. This gives

$$\begin{aligned}(4.52) \quad &\Longrightarrow \quad a_t^{\Omega_N}(0,0) \asymp t^{-1}\log^{\phi} t,\\ (4.53) \quad &\Longrightarrow \quad a_t^{\Omega_N}(0,0) \asymp t^{-1-\delta_{N,c}},\end{aligned} \tag{5.19}$$

where

$$\delta_{N,c} = \frac{\log c}{\log(N/c)}. \tag{5.20}$$

§5.3 Hierarchical clustering

In this section we prove Theorem 4.3.3 by substituting the results of Theorem 4.3.2 into the clustering criterion in (4.50).

Combining (4.51), (5.2)–(5.4) and (5.19)–(5.20), we find the following clustering criterion for *fixed* N and infinite seed-bank:

- Subject to (4.52), clustering prevails if and only if

$$-\phi \leq \alpha \leq 1. \tag{5.21}$$

- Subject to (4.53), clustering prevails if and only

$$\delta_{N,c} \leq -\frac{1-\gamma_{N,K,e}}{\gamma_{N,K,e}}. \tag{5.22}$$

In view of (5.2) and (4.56), the condition in (5.22) amounts to

$$\log N \times \log(Kc) \leq \log c \times \log(K^2 e), \tag{5.23}$$

where we use that $c < N$ and $Ke < N$ (recall (4.7) and (4.12)). The condition in (5.23) holds for all N when

$$Kc = 1 \text{ with } \begin{cases} c = 1, & K^2 e \in (0, \infty), \\ c > 1, & K^2 e \geq 1, \\ c < 1, & K^2 e \leq 1. \end{cases} \tag{5.24}$$

It also holds for N large enough when $Kc < 1$ and fails for N large enough when $Kc > 1$. Thus, for infinite seed-bank, clustering prevails for N large enough if and only if

$$Kc \leq 1 \leq K, \tag{5.25}$$

which is the analogue of (5.21).