3 Limits

Add some words here to link from previous chapter once that is written.

3.1 Encoding genealogies

3.1.1 The genealogical process

Before we can analyse genealogies, we need a way to encode them. The encoding will only include the information relevant to the sample genealogy, namely which lineages coalesce at which times. Information about particle positions and "killed" particles is ignored.

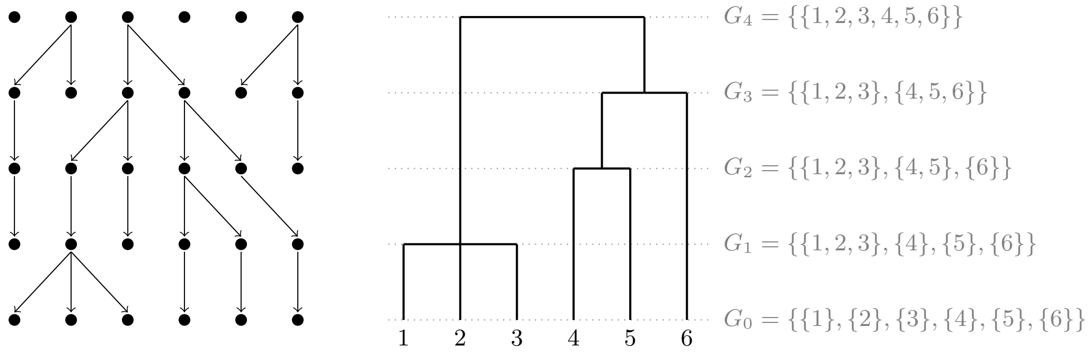
Let $\mathcal{P}_n$ be the space of partitions on $\{1, \ldots, n\}$. For convenience, we now label time in reverse, so the terminal particles are at time 0, their parents are at time 1, and so on. Consider a randomly chosen sample of n terminal particles among a total of N particles, and label the sampled particles $1, \ldots, n$. The genealogical process $(G_t^{(n,N)})_{t \in \mathbb{N}_0}$ for this sample is the $\mathcal{P}_n$-valued stochastic process such that labels i and j are in the same block of the partition $G_t^{(n,N)}$ if and only if terminal particles i and j have a common ancestor at time t (i.e. t generations back).

A formulation where $G_t^{(n,N)}$ takes values in the space of equivalence relations from $[n]$ to $[n]$ is sometimes used (e.g. Möhle 1999); interpreting partition blocks as equivalence classes, this formulation is equivalent to ours.

The initial value of the process is the partition of singletons $G_0^{(n,N)} = \{\{1\}, \ldots, \{n\}\}$, since all of the terminal particles are in separate lineages. The only possible non-identity transitions are those that merge some blocks of the partition, encoding the coalescence of the corresponding lineages. The trivial partition $\{\{1, \ldots, n\}\}$ is therefore an absorbing state, corresponding to all lineages in the sample having coalesced (i.e. the MRCA has been reached). The construction of the genealogical process from the resampling relationships (i.e. the vector of parental indices at each generation) is illustrated in Figure 3.1.

3.1.2 Time scale

In order to have a well-defined limit for the genealogical process as $N \to \infty$, we must scale time by a suitable function $\tau_N(\cdot)$. In the population genetics literature the time scale function is typically deterministic (Section 2.2.3), but in our case τ_N depends on the offspring counts and is therefore random. To define the time scale we first define the pair



(a) Resampling relationships

(b) Genealogy of terminal particles

Figure 3.1: Illustration of how the sample genealogy is encoded. (a) Relationships induced by resampling in a sample of $n = 6$ particles over four iterations. (b) The genealogy of these six particles, labelled with the value of the genealogical process G_t at each time.

merger rate

$$c_N(t) := \frac{1}{(N)_2} \sum_{i=1}^{N} (\nu_t^{(i)})_2. \tag{3.1}$$

This is the probability, conditional on $\nu_t^{(1:N)}$, that a randomly chosen pair of lineages in generation t merges exactly one generation back. To achieve a limiting pair merger rate of 1, as in the n-coalescent, we rescale time by the generalised inverse

$$\tau_N(t) := \inf \left\{ s \in \mathbb{N} : \sum_{r=1}^{s} c_N(r) \geq t \right\}. \tag{3.2}$$

The function τ_N maps continuous to discrete time, providing the link between the discrete-time SMC dynamics and the continuous-time Kingman limit. We will also need the following quantity, which is an upper bound on the conditional probability of a multiple merger (three or more lineages merging, or two or more simultaneous pairwise mergers):

$$D_N(t) := \frac{1}{N(N)_2} \sum_{i=1}^{N} (\nu_t^{(i)})_2 \left\{ \nu_t^{(i)} + \frac{1}{N} \sum_{j \neq i} (\nu_t^{(j)})^2 \right\}. \tag{3.3}$$

This will be used to control the rate of multiple mergers, which must be dominated by the pair-merger rate as $N \to \infty$ if we are to recover a Kingman limit (in which almost surely the only non-identity transitions are pair mergers). Some basic properties of c_N, D_N and τ_N are stated in Proposition 3.1.

Proposition 3.1. *For all* $t \in \mathbb{N}$, $t' > s' > 0$,

(a) $c_N(t), D_N(t) \in [0,1]$

(b) $D_N(t) \leq c_N(t)$

(c) $c_N(t)^2 \leq c_N(t)$

(d) $t' \leq \sum_{r=1}^{\tau_N(t')} c_N(r) \leq t' + 1.$

(e) $t' - s' - 1 \leq \sum_{r=\tau_N(s')+1}^{\tau_N(t')} c_N(r) \leq t' - s' + 1.$

(f) $\tau_N(t') \geq t'.$

Proof. **(a)** $c_N(t)$ and $D_N(t)$ are clearly non-negative. Both are maximised when one of the offspring counts is equal to N and the rest are zero, in which case $c_N(t) = D_N(t) = 1$.
(b) As outlined in Koskela et al. (2018, p.9),

$$\begin{aligned}
D_N(t) &:= \frac{1}{(N)_2} \sum_{i=1}^{N} (\nu_t^{(i)})_2 \frac{1}{N} \left\{ \nu_t^{(i)} + \frac{1}{N} \sum_{j \neq i}^{N} (\nu_t^{(j)})^2 \right\} \\
&\leq \frac{1}{(N)_2} \sum_{i=1}^{N} (\nu_t^{(i)})_2 \frac{1}{N} \left\{ \nu_t^{(i)} + \frac{1}{N} \sum_{j \neq i}^{N} N \nu_t^{(j)} \right\} \\
&= \frac{1}{(N)_2} \sum_{i=1}^{N} (\nu_t^{(i)})_2 \frac{1}{N} \left\{ \sum_{j=1}^{N} \nu_t^{(j)} \right\} = \frac{1}{(N)_2} \sum_{i=1}^{N} (\nu_t^{(i)})_2 = c_N(t).
\end{aligned}$$

(c) is immediate given (a).
(d) follows directly from the definition of τ_N in (3.2).
(e) Writing

$$\sum_{r=\tau_N(s')+1}^{\tau_N(t')} c_N(r) = \sum_{r=1}^{\tau_N(t')} c_N(r) - \sum_{r=1}^{\tau_N(s')} c_N(r),$$

the result follows by applying (d) to both sums.
(f) follows from (a) and the definition of τ_N in (3.2). ■

Another useful property is the following, based on Koskela et al. (2018, Lemma 2). There the special case $f(r) \equiv c_N(r)$ is proved, but the authors remark that the result also holds for other choices of f. Here we state the result in full generality.

Lemma 3.2. *Fix $t > 0$. Let $(\mathcal{F}_r)$ be the backwards-in-time filtration generated by the offspring counts $\nu_r^{(1:N)}$ at each generation r. Let $f(r)$ be any deterministic function of $\nu_r^{(1:N)}$ such that for all r there exists $B < \infty$ for which $0 \le f(r) \le B$. Then*

$$\mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} f(r)\right] = \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{E}[f(r) \mid \mathcal{F}_{r-1}]\right].$$

Proof. Define

$$M_s := \sum_{r=1}^{s} \{f(r) - \mathbb{E}[f(r) \mid \mathcal{F}_{r-1}]\}.$$

It is easy to establish that (M_s) is a martingale with respect to $(\mathcal{F}_s)$, and $M_0 = 0$. Now fix $K \ge 1$ and note that $\tau_N(t) \wedge K$ is a bounded $\mathcal{F}_t$-stopping time. Hence we can apply the optional stopping theorem:

$$\begin{aligned}
\mathbb{E}[M_{\tau_N(t)\wedge K}] &= \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)\wedge K} \{f(r) - \mathbb{E}[f(r) \mid \mathcal{F}_{r-1}]\}\right] \\
&= \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)\wedge K} f(r)\right] - \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)\wedge K} \mathbb{E}[f(r) \mid \mathcal{F}_{r-1}]\right] = 0.
\end{aligned}$$

Since this holds for all $K \ge 1$,

$$\lim_{K\to\infty} \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)\wedge K} f(r)\right] = \lim_{K\to\infty} \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)\wedge K} \mathbb{E}[f(r) \mid \mathcal{F}_{r-1}]\right].$$

The monotone convergence theorem allows the limit to pass inside the expectation on each side (since increasing K can only increase each sum, by possibly adding some non-negative terms). Hence

$$\begin{aligned}
\mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} f(r)\right] &= \mathbb{E}\left[\lim_{K\to\infty} \sum_{r=1}^{\tau_N(t)\wedge K} f(r)\right] = \mathbb{E}\left[\lim_{K\to\infty} \sum_{r=1}^{\tau_N(t)\wedge K} \mathbb{E}[f(r) \mid \mathcal{F}_{r-1}]\right] \\
&= \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{E}[f(r) \mid \mathcal{F}_{r-1}]\right],
\end{aligned}$$

which concludes the proof. ■

3.1.3 Transition probabilities

Recall that $\mathcal{P}_n$ denotes the space of partitions of $\{1, \dots, n\}$. For any $\xi, \eta \in \mathcal{P}_n$ and $t \in \mathbb{N}$, let $p_{\xi\eta}(t)$ denote the conditional transition probabilities of the genealogical process given $\nu_t^{(1:N)}$. The transition probability $p_{\xi\eta}(t)$ can only be non-zero when η is obtained from ξ

by merging some blocks of ξ (i.e. some lineages coalescing). Ordering the blocks by their least element, denote by b_i the number of blocks of ξ that merge to form block i in η, for each $i \in \{1, \dots, |\eta|\}$. Hence $b_1 + \cdots + b_{|\eta|} = |\xi|$. Then the transition probability is given by

$$p_{\xi\eta}(t) := \frac{1}{(N)_{|\xi|}} \sum_{\substack{i_1 \neq \cdots \neq i_{|\eta|} \\ =1}}^{N} (\nu_t^{(i_1)})_{b_1} \cdots (\nu_t^{(i_{|\eta|})})_{b_{|\eta|}}. \tag{3.4}$$

We will only need to work directly with the identity transition probabilities $p_{\xi\xi}(t)$. Upper and lower bounds on these probabilities are presented in Propositions 3.3 and 3.4.

Proposition 3.3. *Let $\xi \in \mathcal{P}_n$, $N > 2$. Then*

$$p_{\xi\xi}(t) \geq 1 - \binom{|\xi|}{2} \frac{N^{|\xi|-2}}{(N-2)_{|\xi|-2}} \left[c_N(t) + B_{|\xi|} D_N(t) \right]$$

where

$$B_{|\xi|} = K(|\xi|-1)!(|\xi|-2)\exp(2\sqrt{2(|\xi|-2)})$$

for some $K > 0$ that does not depend on $|\xi|$.

Proof. We have the following expression for $p_{\xi\xi}(t)$, by subtracting all possible non-identity transitions (the omitted $k = |\xi|$ term would count identity transitions):

$$p_{\xi\xi}(t) = 1 - \frac{1}{(N)_{|\xi|}} \sum_{k=1}^{|\xi|-1} \sum_{\substack{b_1 \geq \ldots \geq b_k = 1 \\ b_1 + \ldots + b_k = |\xi|}}^{|\xi|} \frac{|\xi|!}{\prod_{j=1}^{|\xi|} (j!)^{\kappa_j} \kappa_j!} \sum_{\substack{i_1 \neq \ldots \neq i_k = 1 \\ \text{all distinct}}}^{N} (\nu_t^{(i_1)})_{b_1} \ldots (\nu_t^{(i_k)})_{b_k},$$

where $\kappa_i = |\{j : b_j = i\}|$ is the multiplicity of mergers of size i (κ_1 counts non-merger events, and we have the identity $\kappa_1 + 2\kappa_2 + \cdots + |\xi|\kappa_{|\xi|} = |\xi|$). The combinatorial factor is the number of partitions of a sequence of length $|\xi|$ having κ_j subsequences of length j for each j (Fu 2006, Equation (11)).

We separate the $k = |\xi|-1$ term (which counts single pair mergers), for which $(b_1, b_2, \ldots, b_{|\xi|-1}) = (2, 1, \ldots, 1)$ and

$$\frac{|\xi|!}{\prod_{j=1}^{|\xi|} (j!)^{\kappa_j} \kappa_j!} = \binom{|\xi|}{2}.$$

For the remaining terms we use

$$\frac{|\xi|!}{\prod_{j=1}^{|\xi|} (j!)^{\kappa_j} \kappa_j!} \leq |\xi|!.$$

Thus

$$p_{\xi\xi}(t) \geq 1 - \frac{1}{(N)_{|\xi|}} \binom{|\xi|}{2} \sum_{\substack{i_1 \neq \ldots \neq i_{|\xi|-1}=1 \\ \text{all distinct}}}^{N} (\nu_t^{(i_1)})_2 \nu_t^{(i_2)} \ldots \nu_t^{(i_{|\xi|-1})}$$
$$- \frac{1}{(N)_{|\xi|}} \sum_{k=1}^{|\xi|-1} \sum_{\substack{b_1 \geq \ldots \geq b_k = 1 \\ b_1 + \ldots + b_k = |\xi|}}^{|\xi|} |\xi|! \sum_{\substack{i_1 \neq \ldots \neq i_k = 1 \\ \text{all distinct}}}^{N} (\nu_t^{(i_1)})_{b_1} \ldots (\nu_t^{(i_k)})_{b_k} \tag{3.5}$$

Now, for the $k = |\xi| - 1$ term we use the bound

$$\sum_{i_1 \neq \ldots \neq i_{|\xi|-1}=1}^{N} (\nu_t^{(i_1)})_2 \nu_t^{(i_2)} \ldots \nu_t^{(i_{|\xi|-1})} \leq N^{|\xi|-2} \sum_{i=1}^{N} (\nu_t^{(i)})_2$$

while for the other terms we have (similarly to Koskela et al. 2018, Lemma 1 Case 3)

$$\sum_{\substack{i_1 \neq \ldots \neq i_k = 1 \\ \text{all distinct}}}^{N} (\nu_t^{(i_1)})_{b_1} \ldots (\nu_t^{(i_k)})_{b_k} \leq \sum_{i=1}^{N} (\nu_t^{(i)})_2 \left(N^{|\xi|-2} - \sum_{\substack{j_1 \neq \ldots \neq j_{|\xi|-2}=1 \\ \text{all distinct and } \neq i}}^{N} \nu_t^{(j_1)} \ldots \nu_t^{(j_{|\xi|-2})} \right)$$
$$\leq \sum_{i=1}^{N} (\nu_t^{(i)})_2 \left\{ N^{|\xi|-2} - (N - \nu_t^{(i)})^{|\xi|-2} + \binom{|\xi|-2}{2} \sum_{j \neq i} (\nu_t^{(j)})^2 \left(\sum_{k \neq i} \nu_t^{(k)} \right)^{|\xi|-4} \right\}$$
$$\leq \sum_{i=1}^{N} (\nu_t^{(i)})_2 \left\{ (|\xi|-2) \nu_t^{(i)} N^{|\xi|-3} + \binom{|\xi|-2}{2} \sum_{j \neq i} (\nu_t^{(j)})^2 N^{|\xi|-4} \right\},$$

where the last step uses $(N-x)^b \geq N^b - bxN^{b-1}$ for $x \leq N$, $b \geq 0$. Hence

$$p_{\xi\xi}(t) \geq 1 - \frac{1}{(N)_{|\xi|}} \binom{|\xi|}{2} N^{|\xi|-2} \sum_{i=1}^{N} (\nu_t^{(i)})_2$$
$$- \frac{N^{|\xi|-3}}{(N)_{|\xi|}} |\xi|! \sum_{k=1}^{|\xi|-1} \sum_{\substack{b_1 \geq \ldots \geq b_k = 1 \\ b_1 + \ldots + b_k = |\xi|}}^{|\xi|} \sum_{i=1}^{N} (\nu_t^{(i)})_2 \left\{ (|\xi|-2) \nu_t^{(i)} + \binom{|\xi|-2}{2} \frac{1}{N} \sum_{j \neq i} (\nu_t^{(j)})^2 \right\}.$$

The summands in the last line are independent of $k, b_1, \ldots, b_k$, and the number of terms in the sums over k and $b_1, \ldots, b_k$ is bounded by $\gamma_{|\xi|-2}(|\xi|-2)$, where γ_n is the number of integer partitions of n. By Hardy and Ramanujan (1918, Section 2), $\gamma_n < Ke^{2\sqrt{2n}}/n$ for a constant $K > 0$ independent of n. Thus, for $|\xi| > 2$,

$$
\begin{aligned}
p_{\xi\xi}(t) \geq 1 &- \frac{N^{|\xi|-2}}{(N-2)_{|\xi|-2}} \binom{|\xi|}{2} c_N(t) \\
&- \frac{N^{|\xi|-2}}{(N-2)_{|\xi|-2}} K \exp(2\sqrt{2(|\xi|-2)})|\xi|! \frac{1}{N(N)_2} \\
&\qquad \sum_{i=1}^{N} (\nu_t^{(i)})_2 \left\{ (|\xi|-2)\nu_t^{(i)} + \binom{|\xi|-2}{2} \frac{1}{N} \sum_{j \neq i} (\nu_t^{(j)})^2 \right\} \\
\geq 1 &- \frac{N^{|\xi|-2}}{(N-2)_{|\xi|-2}} \binom{|\xi|}{2} c_N(t) \\
&- \frac{N^{|\xi|-2}}{(N-2)_{|\xi|-2}} K \exp(2\sqrt{2(|\xi|-2)})|\xi|! \binom{|\xi|-1}{2} D_N(t) \\
\geq 1 &- \frac{N^{|\xi|-2}}{(N-2)_{|\xi|-2}} \binom{|\xi|}{2} \left[c_N(t) + B_{|\xi|} D_N(t) \right]
\end{aligned}
$$

where

$$
\begin{aligned}
B_{|\xi|} &= \binom{|\xi|}{2}^{-1} K \exp(2\sqrt{2(|\xi|-2)})|\xi|! \binom{|\xi|-1}{2} \\
&= K(|\xi|-1)!(|\xi|-2) \exp(2\sqrt{2(|\xi|-2)}).
\end{aligned}
$$

When $|\xi| = 2$, (3.5) becomes

$$p_{\xi\xi}(t) \geq 1 - c_N(t)$$

and when $|\xi| = 1$, (3.5) becomes

$$p_{\xi\xi}(t) \geq 1;$$

in both cases the result is immediate. ■

Proposition 3.4. *Let $\xi \in \mathcal{P}_n$. Then, for N sufficiently large,*

$$p_{\xi\xi}(t) \leq 1 - \binom{|\xi|}{2} \{1 + O(N^{-1})\} \left[c_N(t) - B'_{|\xi|} D_N(t) \right]$$

where $B'_{|\xi|} = \binom{|\xi|-1}{2}$.

A proof is given in Koskela et al. (2018, Lemma 1 Case 1). refer to the erratum once available, which is more explicit about this proof.

3.2 An existing limit theorem

Under the assumption (A1) stated below, it is sufficient for our purposes to consider only offspring counts $\nu_t^{(1:N)} = (\nu_t^{(1)}, \ldots, \nu_t^{(N)})$, where $\nu_t^{(i)} = |\{j : a_t^{(j)} = i\}|$, rather than the parental indices $a_t^{(1:N)}$ which are generally more informative.

(A1) The conditional distribution of parental indices $a_t^{(1:N)}$ given offspring counts $\nu_t^{(1:N)}$ is uniform over all assignments such that $|\{j : a_t^{(j)} = i\}| = \nu_t^{(i)}$ for all i.

As we saw in Section 2.2, the n-coalescent is *exchangeable*, so for instance the pair of lineages merging at each event is chosen uniformly. (A1) is a weaker condition than exchangeability of the particles within a generation which is sufficient to admit an exchangeable process in the limit. Exchangeability of the particles would imply neutrality, an unreasonable assumption in the setting of SMC. In contrast, (A1) can easily be enforced upon any SMC algorithm by applying a random permutation to the offspring indices immediately after resampling.

Koskela et al. (2018) proved the following theorem which gives sufficient conditions under which sampled genealogies of (non-neutral) interacting particle systems converge to the n-coalescent as $N \to \infty$. Naturally, such a result can only be expected to hold for genealogies of finite samples ($n << N$), and not for the entire genealogy of the N particles. For instance the genealogies arising in SMC algorithms are not restricted to single pair mergers only, although within a sparse sample we may, under mild conditions, see only single pair mergers. That is to say, there is not an extension of this result whereby the whole-population genealogy converges to the Kingman coalescent as $N \to \infty$, unless very restrictive conditions are imposed.

Theorem 3.5 (Koskela et al. 2018)**.** *Fix $n \leq N$ as the observed number of particles from the output of an interacting particle system with N particles which satisfies (A1). Suppose that for any $0 \leq s < t < \infty$, we have*

$$\lim_{N\to\infty} \mathbb{E}\left[\sum_{r=\tau_N(s)+1}^{\tau_N(t)} D_N(r) \right] = 0, \tag{3.6}$$

$$\lim_{N\to\infty} \mathbb{E}[c_N(t)] = 0, \tag{3.7}$$

$$\lim_{N\to\infty} \mathbb{E}\left[\sum_{r=\tau_N(s)+1}^{\tau_N(t)} c_N(r)^2 \right] = 0, \tag{3.8}$$

and
$$\mathbb{E}[\tau_N(t) - \tau_N(s)] \leq C_{t,s} N, \tag{3.9}$$

for some constant $C_{t,s} > 0$ that is independent of N. Then $(G^{(n,N)}_{\tau_N(t)})_{t\geq 0}$ converges to the Kingman n-coalescent in the sense of finite-dimensional distributions as $N \to \infty$.

To ensure samples of size n have Kingman genealogies in the limit, with pair mergers only, we require that multiple mergers (that is, where more than two lineages merge into one, or where two or more mergers happen simultaneously) occur on a slower time scale than pair mergers. This is the role of condition (3.6).

Conditions (3.7) and (3.8) ensure that the limiting process is continuous and has the required unit pair merger rate. For (3.7) to fail to hold, the expected number of mergers at some generation would have to be $O(N^2)$. This can only happen if the resampling scheme is very bad (e.g. star resampling) or the weights are particularly badly-behaved. The latter is ruled out in the corollaries of Chapter 5 by imposing bounds on the potential functions; this is discussed further in Section 5.1.

Condition (3.9) specifies that the time scale must be $O(N)$. As we saw in Section 2.2.3, this is the correct time scale for the Wright-Fisher model, but for instance the Moran model has time scale $O(N^2)$ and hence violates this condition. Since we know that the neutral Moran model also has Kingman genealogies in the limit, condition (3.9) clearly is not necessary. The simplified statement in Theorem 3.6 does not impose any such condition on the time scale.

The proof of Koskela et al. (2018) does not explicitly use (3.7) but rather the similar condition

$$\lim_{N\to\infty} \mathbb{E}[c_N(\tau_N(t))] = 0. \tag{3.10}$$

However, as we will see in the next section (Lemmata 3.8 and 3.9), both (3.7) and (3.10) are implied by (3.6), so the theorem is correct. Such redundancies in the statement of Theorem 3.5 are removed in Theorem 3.6.

The proof of Theorem 3.5 (i.e. Koskela et al. 2018, Theorem 1) proceeds in three parts. The first is a vanishing upper bound on finite-dimensional distributions of the genealogical process when the path of the process involves any multiple mergers. The second is showing that, when the path of the genealogy consists of only single pair mergers, the finite-dimensional distributions of the n-coalescent upper bound those of the genealogical process in the limit $N \to \infty$. The final piece is a similar lower bound, which together with the upper bound establishes convergence of the finite-dimensional distributions.

3.3 A new limit theorem

We now present a related theorem, having the same conclusion but with conditions that are more tractable and remove some redundancies in the statement of Theorem 3.5. While we do not prove that this is a strict generalisation, there are certainly systems which satisfy the conditions of Theorem 3.6 but not of Theorem 3.5.

Theorem 3.6. *Let $\nu_t^{(1:N)}$ denote the offspring numbers in an interacting particle system satisfying (A1) such that, for any N sufficiently large, $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all finite t. Suppose that there exists a deterministic sequence $(b_N)_{N\geq 1}$ such that $\lim_{N\to\infty} b_N = 0$ and*

$$\frac{1}{(N)_3} \sum_{i=1}^{N} \mathbb{E}_t[(\nu_t^{(i)})_3] \leq b_N \frac{1}{(N)_2} \sum_{i=1}^{N} \mathbb{E}_t[(\nu_t^{(i)})_2] \tag{3.11}$$

for all N, uniformly in $t \geq 1$. Fix $n \leq N$ and consider a randomly chosen sample of n terminal particles. Then the resulting rescaled genealogical process $(G^{(n,N)}_{\tau_N(t)})_{t\geq 0}$ converges in the sense of finite-dimensional distributions to Kingman's n-coalescent as $N \to \infty$.

On the RHS of (3.11) is the filtered expectation of $c_N(t)$, i.e. the expected pair merger rate, and the LHS is the corresponding rate of triple mergers. Intuitively, (3.11) says that pair mergers dominate triple mergers, the expected rate of which vanishes as $N \to \infty$. As we will see, this implies that pair mergers also dominate all other larger mergers, such as simultaneous pair mergers.

Our result improves on Theorem 3.5 by eliminating the restrictive condition (3.9), which we know is unnecessary. This allows our result to apply to some models not previously included; for example the neutral Moran model. Although we do not prove that Theorem 3.6 is a true generalisation of Theorem 3.5, Möhle and Sagitov (2003, Theorem 5.4) showed that in neutral models the straightforward analogue of (3.11) is both necessary and sufficient, suggesting that in general this condition is not significantly stronger than (3.6)–(3.8) combined.

Our conditions are also significantly easier to verify than those of Theorem 3.5. Not only are four conditions replaced with one, but the condition (3.11) only involves marginal moments of the offspring counts, whereas (3.6) and (3.8) involve mixed moments. As we will see in Chapter 4, once we move beyond conditionally independent resampling schemes such as multinomial resampling, the joint distributions of offspring counts become complex and it may only be feasible to calculate their moments marginally. As such, we are able to verify the conditions of Theorem 3.6 in several cases, including for resampling schemes that induce strong correlations between offspring counts, whereas Koskela et al. (2018) apply their theorem only to multinomial resampling.

Our condition on the time scale, $\mathbb{P}[\tau_N(t) = \infty] = 0$, is not very restrictive. Essentially, it rules out systems in which coalescences occur at only finitely many generations. This condition is not actually necessary for Theorem 3.6 to hold, as such, but if it is violated then the limiting object is an n-coalescent under an infinite time-scaling, that is n lineages never coalescing. This would constitute a qualitatively different result and one that is of little interest for SMC, so we follow Möhle (1998) in excluding it.

3.3.1 Proof of theorem

First we prove that (3.10) and the assumptions (3.6)–(3.8) of Theorem 3.5 all follow from (3.11). Figure 3.2 illustrates how the following Lemmata 3.7–3.10 fit together. The argument differs slightly from that presented in Brown et al. (2021) in that we will here show (3.11) ⇒ (3.6) ⇒ (3.7) rather than (3.11) ⇒ (3.6) and (3.11) ⇒ (3.7). This highlights the redundancy in Theorem 3.5, where condition (3.6) directly implies two of the other stated conditions.

The second step in the proof is to show that condition (3.9) is not necessary. In particular, the parts of the proof of Koskela et al. (2018) which relied on (3.9) are rewritten using Proposition 3.3 instead. Proposition 3.3 is a lower bound on the probability of an identity transition, which holds in general without the need for further conditions, so we really are removing condition (3.9) rather than substituting it for a different condition.

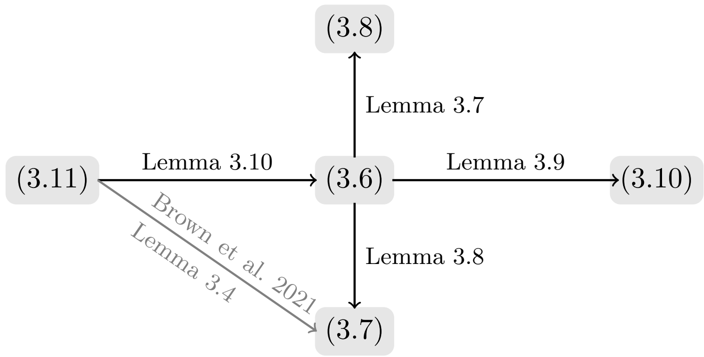


Figure 3.2: Dependencies between conditions of Theorems 3.5 and 3.6. Arrows represent logical implication; labels on arrows indicate the lemma in which the implication is stated. In Brown et al. (2021, Lemma 3.4) the direct implication (3.11) ⇒ (3.7) was proved, but here we will instead show that (3.6) ⇒ (3.7).

Lemma 3.7. *If for all* $0 \le s < t < \infty$

$$\lim_{N\to\infty} \mathbb{E}\left[\sum_{r=\tau_N(s)+1}^{\tau_N(t)} D_N(r)\right] = 0$$

then for all $0 \le s < t < \infty$

$$\lim_{N\to\infty} \mathbb{E}\left[\sum_{r=\tau_N(s)+1}^{\tau_N(t)} c_N(r)^2\right] = 0.$$

Proof. We have

$$\begin{aligned}
c_N(t)^2 &= \frac{1}{N(N-1)(N)_2}\sum_{i=1}^{N}(\nu_t^{(i)})_2\left\{\nu_t^{(i)}(\nu_t^{(i)}-1)+\sum_{\substack{j=1\\j\neq i}}^{N}(\nu_t^{(j)})_2\right\} \\
&= \frac{1}{N(N)_2}\sum_{i=1}^{N}(\nu_t^{(i)})_2\left\{\frac{\nu_t^{(i)}(\nu_t^{(i)}-1)}{N-1}+\frac{1}{N-1}\sum_{\substack{j=1\\j\neq i}}^{N}(\nu_t^{(j)})_2\right\} \\
&\leq \frac{1}{N(N)_2}\sum_{i=1}^{N}(\nu_t^{(i)})_2\left\{\nu_t^{(i)}+\frac{1}{N-1}\sum_{\substack{j=1\\j\neq i}}^{N}(\nu_t^{(j)})_2\right\} \\
&\leq \frac{1}{N(N)_2}\sum_{i=1}^{N}(\nu_t^{(i)})_2\left\{\nu_t^{(i)}+\frac{N/(N-1)}{N}\sum_{\substack{j=1\\j\neq i}}^{N}(\nu_t^{(j)})^2\right\} \\
&\leq \frac{N/(N-1)}{N(N)_2}\sum_{i=1}^{N}(\nu_t^{(i)})_2\left\{\nu_t^{(i)}+\frac{1}{N}\sum_{\substack{j=1\\j\neq i}}^{N}(\nu_t^{(j)})^2\right\} = \frac{N}{N-1}D_N(t)
\end{aligned}$$

which is sufficient for the result. ■

Lemma 3.8. *If for all* $0 \leq s < t < \infty$

$$\lim_{N\to\infty}\mathbb{E}\left[\sum_{r=\tau_N(s)+1}^{\tau_N(t)} D_N(r)\right] = 0$$

then for all $t \in \mathbb{N}$

$$\lim_{N\to\infty}\mathbb{E}[c_N(t)] = 0.$$

Proof. Fix $\epsilon > 0$, and assume $N > 2/\epsilon$. Following Möhle and Sagitov (2003), define the events

$$A_i(t) := \{\nu_t^{(i)} \leq N\epsilon\}. \tag{3.12}$$

Then

$$\begin{aligned}
c_N(t) &= \frac{1}{(N)_2}\sum_{i=1}^{N}(\nu_t^{(i)})_2[\mathbb{1}_{A_i(t)}+\mathbb{1}_{A_i(t)^c}] \\
&\leq \frac{N\epsilon}{(N)_2}\sum_{i=1}^{N}\nu_t^{(i)}+\sum_{i=1}^{N}\mathbb{1}_{A_i(t)^c} \\
&= \frac{N\epsilon}{N-1}+\sum_{i=1}^{N}\mathbb{1}_{A_i(t)^c}.
\end{aligned}$$

Taking expectations and applying the generalised Markov inequality,

$$\begin{aligned}
\mathbb{E}[c_N(t)] &\leq \epsilon 1_N + \sum_{i=1}^{N} \mathbb{P}[\nu_t^{(i)} > N\epsilon] \\
&\leq \epsilon 1_N + \sum_{i=1}^{N} \frac{\mathbb{E}[(\nu_t^{(i)})_3]}{(N\epsilon)_3} \\
&\leq \epsilon 1_N + \frac{N(N)_2}{(N\epsilon)_3} \mathbb{E}[D_N(t)] \\
&= \epsilon 1_N + \epsilon^{-3} 1_N \mathbb{E}[D_N(t)] \\
&\leq \epsilon 1_N + \epsilon^{-3} 1_N \mathbb{E}\left[\sum_{r=1}^{t} D_N(r)\right] \\
&\leq \epsilon 1_N + \epsilon^{-3} 1_N \mathbb{E}\left[\sum_{r=\tau_N(0)+1}^{\tau_N(t)} D_N(r)\right].
\end{aligned}$$

Taking limits,

$$\lim_{N\to\infty} \mathbb{E}[c_N(t)] \leq \epsilon.$$

Since ϵ was arbitrary this concludes the proof. ∎

Lemma 3.9. *If for all* $0 \leq s < t < \infty$

$$\lim_{N\to\infty} \mathbb{E}\left[\sum_{r=\tau_N(s)+1}^{\tau_N(t)} D_N(r)\right] = 0$$

then for all $0 < t < \infty$

$$\lim_{N\to\infty} \mathbb{E}[c_N(\tau_N(t))] = 0.$$

Proof. Analogously to the proof of Lemma 3.8, we find

$$\begin{aligned}
\mathbb{E}[c_N(\tau_N(t))] &\leq \epsilon 1_N + \sum_{i=1}^{N} \mathbb{P}[\nu_{\tau_N(t)}^{(i)} > N\epsilon] \\
&\leq \epsilon 1_N + \epsilon^{-3} 1_N \mathbb{E}\left[D_N(\tau_N(t))\right] \\
&\leq \epsilon 1_N + \epsilon^{-3} 1_N \mathbb{E}\left[\sum_{r=\tau_N(0)+1}^{\tau_N(t)} D_N(r)\right] \\
&\xrightarrow[N\to\infty]{} \epsilon
\end{aligned}$$

which concludes the proof. ∎

Lemma 3.10. *If there exists a deterministic sequence $(b_N)_{N\geq 1}$ such that $\lim_{N\to\infty} b_N = 0$ and*

$$\frac{1}{(N)_3}\sum_{i=1}^{N}\mathbb{E}_t[(\nu_t^{(i)})_3] \leq b_N \frac{1}{(N)_2}\sum_{i=1}^{N}\mathbb{E}_t[(\nu_t^{(i)})_2]$$

for all N, uniformly in $t \in \mathbb{N}$, then for all $0 \leq s < t < \infty$

$$\lim_{N\to\infty}\mathbb{E}\left[\sum_{r=\tau_N(s)+1}^{\tau_N(t)} D_N(r)\right] = 0.$$

Proof. We decompose $D_N(t)$ as the sum of two terms and consider their filtered expectations. The first is

$$\begin{aligned}
\frac{1}{N(N)_2}\sum_{i=1}^{N}\mathbb{E}_t[(\nu_t^{(i)})_2\nu_t^{(i)}] &= \frac{1}{N(N)_2}\sum_{i=1}^{N}\mathbb{E}_t[2(\nu_t^{(i)})_2 + (\nu_t^{(i)})_3] \\
&\leq \frac{2}{N}\mathbb{E}_t[c_N(t)] + \frac{1}{(N)_3}\sum_{i=1}^{N}\mathbb{E}_t[(\nu_t^{(i)})_3] \\
&\leq \left(\frac{2}{N} + b_N\right)\mathbb{E}_t[c_N(t)]. \qquad (3.13)
\end{aligned}$$

The second is

$$\begin{aligned}
\frac{1}{N^2(N)_2}\sum_{j=1}^{N}\sum_{i\neq j}\mathbb{E}_t[(\nu_t^{(i)})_2(\nu_t^{(j)})^2] &= \frac{1}{N^2(N)_2}\sum_{j=1}^{N}\sum_{i\neq j}\mathbb{E}_t[(\nu_t^{(i)})_2(\nu_t^{(j)})_2 + (\nu_t^{(i)})_2\nu_t^{(j)}] \\
&\leq \frac{1}{N^2(N)_2}\sum_{j=1}^{N}\sum_{i\neq j}\mathbb{E}_t[(\nu_t^{(i)})_2(\nu_t^{(j)})_2] + \frac{\mathbb{E}_t[c_N(t)]}{N}. \qquad (3.14)
\end{aligned}$$

Now, with the events $A_i(t)$ defined as in (3.12),

$$\begin{aligned}
\sum_{j=1}^{N}\sum_{i\neq j}\mathbb{E}_t\{(\nu_t^{(i)})_2(\nu_t^{(j)})_2\} &= \sum_{j=1}^{N}\sum_{i\neq j}\left\{\mathbb{E}_t[(\nu_t^{(i)})_2(\nu_t^{(j)})_2\mathbb{1}_{A_i(t)}] + \mathbb{E}_t[(\nu_t^{(i)})_2(\nu_t^{(j)})_2\mathbb{1}_{A_i(t)^c}]\right\} \\
&\leq N\epsilon\sum_{j=1}^{N}\sum_{i\neq j}\mathbb{E}_t[\nu_t^{(i)}(\nu_t^{(j)})_2\mathbb{1}_{A_i(t)}] + N^3\sum_{j=1}^{N}\sum_{i\neq j}\mathbb{E}_t[\nu_t^{(j)}\mathbb{1}_{A_i(t)^c}] \\
&\leq N^2(N)_2\epsilon\mathbb{E}_t[c_N(t)] + N^4\sum_{i=1}^{N}\mathbb{P}[\nu_t^{(i)} > N\epsilon \mid \mathcal{F}_{t-1}]. \qquad (3.15)
\end{aligned}$$

For $N \geq 3/\epsilon$, by the generalised Markov inequality,

$$\sum_{i=1}^{N} \mathbb{P}(\nu_t^{(i)} > N\epsilon \mid \mathcal{F}_{t-1}) \leq \frac{1}{(N\epsilon)_3} \sum_{i=1}^{N} \mathbb{E}_t\{(\nu_t^{(i)})_3\} = \frac{\{1 + O(N^{-1})\}}{\epsilon^3 (N)_3} \sum_{i=1}^{N} \mathbb{E}_t\{(\nu_t^{(i)})_3\}$$
$$\leq \{1 + O(N^{-1})\} \frac{b_N}{\epsilon^3} \mathbb{E}_t\{c_N(t)\}. \tag{3.16}$$

Substituting (3.16) into (3.15) gives

$$\sum_{j=1}^{N} \sum_{i \neq j} \mathbb{E}_t[(\nu_t^{(i)})_2 (\nu_t^{(j)})_2] \leq N^4 (1 + O(N^{-1})) \left(\epsilon + \frac{b_N}{\epsilon^3} \right) \mathbb{E}_t[c_N(t)] \tag{3.17}$$

and substituting (3.17) into (3.14) gives

$$\frac{1}{N^2 (N)_2} \sum_{j=1}^{N} \sum_{i \neq j} \mathbb{E}_t[(\nu_t^{(i)})_2 (\nu_t^{(j)})^2] \leq \left[(1 + O(N^{-1})) \left(\epsilon + \frac{b_N}{\epsilon^3} \right) + \frac{1}{N} \right] \mathbb{E}_t[c_N(t)]. \tag{3.18}$$

Combining (3.13) and (3.18), we have that

$$\mathbb{E}_t[D_N(t)] = \left[(1 + O(N^{-1})) \left(\epsilon + \frac{b_N}{\epsilon^3} \right) + \frac{3}{N} + b_N \right] \mathbb{E}_t[c_N(t)].$$

Finally, invoking Lemma 3.2 twice gives

$$\begin{aligned}
\mathbb{E}\left[\sum_{r=\tau_N(s)+1}^{\tau_N(t)} D_N(r) \right] &= \mathbb{E}\left[\sum_{r=\tau_N(s)+1}^{\tau_N(t)} \mathbb{E}_r[D_N(r)] \right] \\
&\leq \left\{ (1 + O(N^{-1})) \left(\epsilon + \frac{b_N}{\epsilon^3} \right) + \frac{3}{N} + b_N \right\} \mathbb{E}\left[\sum_{r=\tau_N(s)+1}^{\tau_N(t)} c_N(r) \right] \\
&\leq \left\{ (1 + O(N^{-1})) \left(\epsilon + \frac{b_N}{\epsilon^3} \right) + \frac{3}{N} + b_N \right\} (t - s + 1) \\
&\underset{N \to \infty}{\longrightarrow} \epsilon (t - s + 1),
\end{aligned}$$

and recalling that $\epsilon > 0$ was arbitrary concludes the proof. ∎

Update KJJS references in the following to point to relevant places in the erratum. The proof is similar to `Dropbox/SMC_Genealogies/Asymptotic_genealogies_of_interacting_particle_` `Post_acceptance_correction/Erratum/AOS/Round_3/Indicators/full_redraft.pdf`, from page 22. We need to refer to the KJJS erratum (once published), rather than the arXiv version which is not correct (lacking some indicators) and doesn't contain all the required results, then update equation and lemma numbers etc. as required.

To complete the proof of Theorem 3.6 it remains to show that condition (3.9) is unnecessary. We will show that Proposition 3.3 can be used instead of (3.9) to obtain the same result. The only part of Koskela et al. (2018, Proof of Theorem 1) making use of condition

(3.9) is the lower bound on finite-dimensional distributions of the genealogical process for paths involving single pair mergers only. A slight modification of the argument allows a similar lower bound to be obtained via Proposition 3.3 such that as $N \to \infty$ the bound coincides with the corresponding finite-dimensional distributions of the n-coalescent, as required. The modified section of the proof is presented below, using the notation of Koskela et al. (2018) for ease of comparison. Maybe I should actually write out a full proof of the new theorem?

Proof. Let $\chi_d^\star$ be the conditional transition probability of a transition from state η_{d-1} to state η_d at times $\tau_N(t_{d-1})$ and $\tau_N(t_d)$ respectively, conditional on the offspring counts between those times $\nu^{(1:N)}_{\tau_N(d-1)+1}, \dots, \nu^{(1:N)}_{\tau_N(d)}$. This transition can happen via any valid path of merger events, but we restrict to paths involving binary mergers only, and denote by χ_d the conditional transition probability subject to this restriction. Compared to Koskela et al. (2018, Proof of Theorem 1), the derivation of an upper bound on χ_d holds without modification, while the first step in the derivation of a lower bound (Koskela et al. 2018, p.14) involves the application of Koskela et al. (2018, Lemma 1 Case 1) to bound χ_d from below and the subsequent application of (3.9). Instead, we apply Proposition 3.3 to obtain, for sufficiently large N,

$$\chi_d \geq \sum_{\substack{s_1 < \ldots < s_\alpha \\ = \tau_N(t_{d-1})+1}}^{\tau_N(t_d)} (\tilde{Q}^\alpha)_{\eta_{d-1}\eta_d} \left(\prod_{r=1}^{\alpha} \mathbb{1}_{\{c_N(s_r) > \binom{n-2}{2} D_N(s_r)\}} \left[c_N(s_r) - \binom{n-2}{2} 1_N D_N(s_r) \right] \right)$$
$$\times \prod_{\substack{r=\tau_N(t_{d-1})+1 \\ r \neq s_1, \ldots, r \neq s_\alpha}}^{\tau_N(t_d)} \left[1 - \tilde{B}_n 1_N D_N(r) - \binom{|\eta_{d-1}| - |\{i : s_i < r\}|}{2} 1_N c_N(r) \right]$$
$$\times \mathbb{1}_{\{c_N(r) < (\tilde{B}_n + \binom{n}{2})^{-1}\}}.$$

Here $\tilde{Q}$ is the matrix obtained from the generator Q of Kingman's n-coalescent (see Definition 2.1) by setting the diagonal entries to 0. The number of pair-merger steps required to transition from η_{d-1} to η_d is $\alpha = |\eta_{d-1}| - |\eta_d|$. The sequences $s_1, \dots, s_\alpha$ denote the times at which these pair-mergers happen. At the remaining times r the partition is unchanged, and the bound of Proposition 3.3 has been applied to the one-step transition probabilities corresponding to these identity transitions. The constant is $\tilde{B}_n := B_n \binom{n}{2}$ where B_n is the constant defined in Proposition 3.3, and we have replaced $|\eta_d|$ by its upper bound n.

The rest of the proof proceeds as in Koskela et al. (2018), albeit from this modified initial lower bound. A multinomial expansion of the product on the second line, noting that $(1_N)^a = 1_N$ for any $a \in \mathbb{R}$, yields

$$\chi_d \geq \left(\prod_{r=\tau_N(t_{d-1})+1}^{\tau_N(t_d)} \mathbb{1}_{\{c_N(r) > \binom{n-2}{2} D_N(r)\}} \mathbb{1}_{\{c_N(r) < (\tilde{B}_n + \binom{n}{2}))^{-1}\}} \right)$$
$$\times \sum_{\beta=0}^{\tau_N(t_d)-\tau_N(t_{d-1})-\alpha} (\tilde{Q}^{\alpha})_{\eta_{d-1}\eta_d} \sum_{\substack{(\lambda,\mu)\in\Pi_2([\alpha+\beta]):\\ |\lambda|=\alpha}} 1_N$$
$$\times \sum_{\substack{s_1 < \ldots < s_{\alpha+\beta}\\ =\tau_N(t_{d-1})+1}}^{\tau_N(t_d)} \left(\prod_{r\in\lambda} \left[c_N(s_r) - \binom{n-2}{2} 1_N D_N(s_r) \right] \right)$$
$$\times \prod_{r\in\mu} \left\{ -\binom{|\eta_{d-1}| - |\{i \in \lambda : i < r\}|}{2} c_N(s_r) - \tilde{B}_n D_N(s_r) \right\}$$

where $\Pi_i([n])$ denotes the set of partitions of $\{1, \ldots, n\}$ into exactly i blocks. Expanding the product over λ gives

$$\chi_d \geq \left(\prod_{r=\tau_N(t_{d-1})+1}^{\tau_N(t_d)} \mathbb{1}_{\{c_N(r) > \binom{n-2}{2} D_N(r)\}} \mathbb{1}_{\{c_N(r) < (\tilde{B}_n + \binom{n}{2}))^{-1}\}} \right)$$
$$\times \sum_{\beta=0}^{\tau_N(t_d)-\tau_N(t_{d-1})-\alpha} (\tilde{Q}^{\alpha})_{\eta_{d-1}\eta_d} \sum_{\substack{(\lambda,\mu,\pi)\in\Pi_3([\alpha+\beta]):\\ |\mu|=\beta}} \binom{n-2}{2}^{|\pi|} (-1)^{|\pi|} 1_N$$
$$\times \sum_{\substack{s_1 < \ldots < s_{\alpha+\beta}\\ =\tau_N(t_{d-1})+1}}^{\tau_N(t_d)} \left\{ \prod_{r\in\lambda} c_N(s_r) \right\} \left\{ \prod_{r\in\pi} D_N(s_r) \right\}$$
$$\times \prod_{r\in\mu} \left\{ -\binom{|\eta_{d-1}| - |\{i \in \lambda \cup \pi : i < r\}|}{2} c_N(s_r) - \tilde{B}_n D_N(s_r) \right\}$$

and expanding the product over μ results in

$$\chi_d \geq \left(\prod_{r=\tau_N(t_{d-1})+1}^{\tau_N(t_d)} \mathbb{1}_{\{c_N(r) > \binom{n-2}{2} D_N(r)\}} \mathbb{1}_{\{c_N(r) < (\tilde{B}_n + \binom{n}{2}))^{-1}\}} \right)$$
$$\times \sum_{\beta=0}^{\tau_N(t_d)-\tau_N(t_{d-1})-\alpha} (\tilde{Q}^{\alpha})_{\eta_{d-1}\eta_d} \sum_{\substack{(\lambda,\mu,\pi,\sigma)\in\Pi_4([\alpha+\beta]):\\ |\mu|+|\sigma|=\beta}} \tilde{B}_n^{|\sigma|} \binom{n-2}{2}^{|\pi|} (-1)^{|\pi|+|\sigma|}$$
$$\times 1_N \left\{ \prod_{r\in\mu} -\binom{|\eta_{d-1}| - |\{i \in \lambda \cup \pi : i < r\}|}{2} \right\}$$
$$\times \sum_{\substack{s_1 < \ldots < s_{\alpha+\beta}\\ =\tau_N(t_{d-1})+1}}^{\tau_N(t_d)} \left\{ \prod_{r\in\lambda\cup\mu} c_N(s_r) \right\} \prod_{r\in\pi\cup\sigma} D_N(s_r).$$

Via a further multinomial expansion, the lower bound for the k-step transition probability can be written as

$$\lim_{N\to\infty}\mathbb{E}\left[\prod_{d=1}^{k}\chi_d\right] \geq \lim_{N\to\infty}\mathbb{E}\Bigg[\left(\prod_{r=\tau_N(t_0)+1}^{\tau_N(t_k)} \mathbb{1}_{\{c_N(r)>\binom{n-2}{2}D_N(r)\}}\mathbb{1}_{\{c_N(r)<(\tilde{B}_n+\binom{n}{2})^{-1}\}}\right)$$
$$\times \sum_{\beta_1=0}^{\infty}\cdots\sum_{\beta_k=0}^{\infty}\sum_{\substack{(\lambda_1,\mu_1,\pi_1,\sigma_1)\in\Pi_4([\alpha_1+\beta_1]):\\|\mu_1|+|\sigma_1|=\beta_1}}\cdots\sum_{\substack{(\lambda_k,\mu_k,\pi_k,\sigma_k)\in\Pi_4([\alpha_k+\beta_k]):\\|\mu_k|+|\sigma_k|=\beta_k}}$$
$$\tilde{B}_n^{\sum_{d=1}^{k}|\sigma_d|}\binom{n-2}{2}^{\sum_{d=1}^{k}|\pi_d|}(-1)^{\sum_{d=1}^{k}|\pi_d|+|\sigma_d|}1_N$$
$$\times\left\{\prod_{d=1}^{k}(\tilde{Q}^{\alpha_d})_{\eta_{d-1}\eta_d}\prod_{r\in\mu_d}-\binom{|\eta_{d-1}|-|\{i\in\lambda_d\cup\pi_d : i<r\}|}{2}\right\}$$
$$\times\sum_{\substack{s_1^{(1)}<\ldots<s^{(1)}_{\alpha_1+\beta_1}\\=\tau_N(t_0)+1}}^{\tau_N(t_1)}\cdots\sum_{\substack{s_1^{(k)}<\ldots<s^{(k)}_{\alpha_k+\beta_k}\\=\tau_N(t_{k-1})+1}}^{\tau_N(t_k)}$$
$$\prod_{d=1}^{k}\mathbb{1}_{\{\tau_N(t_d)-\tau_N(t_{d-1})\geq\alpha_d+\beta_d\}}\left\{\prod_{r\in\lambda_d\cup\mu_d}c_N(s_r^{(d)})\right\}\prod_{r\in\pi_d\cup\sigma_d}D_N(s_r^{(d)})\Bigg].$$

An argument completely analogous to that in Koskela et al. (2018, Appendix) shows that passing the expectation and the limit through the infinite sums is justified, whereupon the contribution of terms with $\sum_{d=1}^{k}(|\pi_d|+|\sigma_d|)>0$ vanishes. To see why, follow the argument used to show that the contribution of multiple merger trajectories vanishes in the corresponding upper bound in Koskela et al. (2018). That leaves

$$\lim_{N\to\infty}\mathbb{E}\left[\prod_{d=1}^{k}\chi_d\right] \geq \sum_{\beta_1=0}^{\infty}\cdots\sum_{\beta_k=0}^{\infty}\sum_{\substack{(\lambda_1,\mu_1)\in\Pi_2([\alpha_1+\beta_1]):\\|\mu_1|=\beta_1}}\cdots\sum_{\substack{(\lambda_k,\mu_k)\in\Pi_2([\alpha_k+\beta_k]):\\|\mu_k|=\beta_k}}$$
$$\left\{\prod_{d=1}^{k}(\tilde{Q}^{\alpha_d})_{\eta_{d-1}\eta_d}\prod_{r\in\mu_d}-\binom{|\eta_{d-1}|-|\{i\in\lambda_d\cup\pi_d : i<r\}|}{2}\right\}$$
$$\times\lim_{N\to\infty}\mathbb{E}\Bigg[\left(\prod_{r=\tau_N(t_{d-1})+1}^{\tau_N(t_d)}\mathbb{1}_{\{c_N(r)>\binom{n-2}{2}D_N(r)\}}\mathbb{1}_{\{c_N(r)<(\tilde{B}_n+\binom{n}{2})^{-1}\}}\right)$$
$$\times\sum_{\substack{s_1^{(1)}<\ldots<s^{(1)}_{\alpha_1+\beta_1}\\=\tau_N(t_0)+1}}^{\tau_N(t_1)}\cdots\sum_{\substack{s_1^{(k)}<\ldots<s^{(k)}_{\alpha_k+\beta_k}\\=\tau_N(t_{k-1})+1}}^{\tau_N(t_k)}$$
$$\prod_{d=1}^{k}\mathbb{1}_{\{\tau_N(t_d)-\tau_N(t_{d-1})\geq\alpha_d+\beta_d\}}\left\{\prod_{r\in\lambda_d\cup\mu_d}c_N(s_r^{(d)})\right\}\Bigg]. \tag{3.19}$$

Recall (Koskela et al. 2018, Eq (11)):

$$\sum_{\substack{(\lambda,\mu)\in\Pi_2([\alpha+\beta]):\\ |\mu|=\beta}} (\tilde{Q}^{\alpha})_{\eta_{d-1}\eta_d} \prod_{r\in\mu} -\binom{|\eta_{d-1}| - |\{i\in\lambda\cup\pi : i<r\}|}{2} = (Q^{\alpha+\beta})_{\eta_{d-1}\eta_d}.$$

Applying this k times in (3.19) yields

$$\begin{aligned}
\lim_{N\to\infty} \mathbb{E}\left[\prod_{d=1}^{k} \chi_d\right] \geq \sum_{\beta_1=0}^{\infty}\cdots\sum_{\beta_k=0}^{\infty} & \left\{\prod_{d=1}^{k} (Q^{\alpha_d+\beta_d})_{\eta_{d-1}\eta_d}\right\} \\
& \times \lim_{N\to\infty} \mathbb{E}\Bigg\{\left(\prod_{r=\tau_N(t_{d-1})+1}^{\tau_N(t_d)} \mathbb{1}_{\{c_N(r)>\binom{n-2}{2}D_N(r)\}}\mathbb{1}_{\{c_N(r)<(\tilde{B}_n+\binom{n}{2})^{-1}\}}\right) \\
& \qquad \times \left(\prod_{d=1}^{k} \mathbb{1}_{\{\tau_N(t_d)-\tau_N(t_{d-1})\geq\alpha_d+\beta_d\}}\right) \\
& \qquad \times \sum_{\substack{s_1^{(1)}<\ldots<s^{(1)}_{\alpha_1+\beta_1}\\ =\tau_N(t_0)+1}}^{\tau_N(t_1)} \cdots \sum_{\substack{s_1^{(k)}<\ldots<s^{(k)}_{\alpha_k+\beta_k}\\ =\tau_N(t_{k-1})+1}}^{\tau_N(t_k)} \prod_{d=1}^{k}\prod_{r\in\lambda_d\cup\mu_d} c_N(s_r^{(d)})\Bigg\}.
\end{aligned}$$

We now apply equations (14) and (15), respectively, of Koskela et al. (2018), to those terms with a negative ($|\boldsymbol{\beta}|$ odd) and positive ($|\boldsymbol{\beta}|$ even) sign, respectively, to obtain

$$\begin{aligned}
\lim_{N\to\infty} \mathbb{E}\left[\prod_{d=1}^{k} \chi_d\right] &\geq \sum_{\beta_1=0}^{\infty}\cdots\sum_{\beta_k=0}^{\infty} \left\{\prod_{d=1}^{k} (Q^{\alpha_d+\beta_d})_{\eta_{d-1}\eta_d} \frac{(t_d-t_{d-1})^{\alpha_d+\beta_d}}{(\alpha_d+\beta_d)!}\right\} \\
& \qquad \times \lim_{N\to\infty} \mathbb{E}\Bigg[\left(\prod_{r=\tau_N(t_{d-1})+1}^{\tau_N(t_d)} \mathbb{1}_{\{c_N(r)>\binom{n-2}{2}D_N(r)\}}\mathbb{1}_{\{c_N(r)<(\tilde{B}_n+\binom{n}{2})^{-1}\}}\right) \\
& \qquad\qquad \times \left(\prod_{d=1}^{k} \mathbb{1}_{\{\tau_N(t_d)-\tau_N(t_{d-1})\geq\alpha_d+\beta_d\}}\right)\Bigg] \\
&\geq \sum_{\beta_1=0}^{\infty}\cdots\sum_{\beta_k=0}^{\infty} \left\{\prod_{d=1}^{k} (Q^{\alpha_d+\beta_d})_{\eta_{d-1}\eta_d} \frac{(t_d-t_{d-1})^{\alpha_d+\beta_d}}{(\alpha_d+\beta_d)!}\right\}
\end{aligned}$$

where the expectation of the indicators converges to 1 due to Koskela et al. (2018, Equation (16)) and Lemma 4.12 and Lemma 4.11. Or refer to Koskela et al. (2018, Equation (16)) and Lemma 4 in the appendix of `full_redraft.pdf`. ■

4 Weak Convergence

In this chapter we present a weak convergence result which is identical to Theorem 3.6 except that the mode of convergence is strengthened from convergence of the finite-dimensional distributions to weak convergence. Weak convergence is desirable because it implies convergence of a strictly larger class of functions of genealogies, granting access to the distributions of statistics such as the time to the sample MRCA, the total branch length, and the probability that the MRCA of a subsample is equal to the sample MRCA okay, technically if this one is going to be a “statistic”, I'm talking about the indicator on this event.

The extension from Theorem 3.6 to weak convergence requires an additional tightness argument. The proof is rather long-winded since we do not make such strong simplifying assumptions on the dynamics of the interacting particle system as are seen for example in Möhle (1999) and others...?. The proof is broken down into a series of technical results which culminate in Theorem 4.1. The overall structure of the proof is depicted graphically in Figure 4.1.

We start by defining a suitable metric space. Let $\mathcal{P}_n$ be the space of partitions of $\{1, \dots, n\}$. Denote by $\mathcal{X}$ the set of all functions mapping $[0, \infty)$ to $\mathcal{P}_n$ that are right-continuous with left limits. (Our rescaled genealogical process $(\mathcal{G}^{(n,N)}_{\tau_N(t)})_{t \geq 0}$ and our encoding of the n-coalescent are piecewise-constant functions mapping time $t \in [0, \infty)$ to partitions, and thus live in the space $\mathcal{X}$.) Finally, equip the space $\mathcal{P}_n$ with the discrete metric,

$$\rho(\xi, \eta) = 1 - \delta_{\xi\eta} := \begin{cases} 0 & \text{if } \xi = \eta \\ 1 & \text{otherwise} \end{cases}$$

for any $\xi, \eta \in \mathcal{P}_n$.

Theorem 4.1. *Let $\nu_t^{(1:N)}$ denote the offspring numbers in an interacting particle system satisfying (A1) and such that, for any N sufficiently large, for all finite t, $\mathbb{P}[\tau_N(t) = \infty] = 0$. Suppose that there exists a deterministic sequence $(b_N)_{N \in \mathbb{N}}$ such that $\lim_{N\to\infty} b_N = 0$ and*

$$\frac{1}{(N)_3} \sum_{i=1}^{N} \mathbb{E}_t \left[(\nu_t^{(i)})_3 \right] \leq b_N \frac{1}{(N)_2} \sum_{i=1}^{N} \mathbb{E}_t \left[(\nu_t^{(i)})_2 \right] \tag{4.1}$$

almost surely for all N, uniformly in $t \geq 1$. Then the rescaled genealogical process $(G^{(n,N)}_{\tau_N(t)})_{t \geq 0}$ converges weakly in $(\mathcal{X}, \rho)$ to Kingman's n-coalescent as $N \to \infty$.

Proof of Theorem 4.1. The structure of the proof follows Möhle (1999), albeit with considerable technical complication due to the dependence between generations (non-neutrality) in our model. Is this the main/only source of complication? To make it digestible, the proof is broken down into a number of results which are organised into sections; the relationships between these are shown in Figure 4.1.

Since we already have convergence of the finite-dimensional distributions (Theorem 3.6), strengthening this to weak convergence requires relative compactness of the sequence of processes $\{(G^{(n,N)}_{\tau_N(t)})_{t\geq 0}\}_{N\in\mathbb{N}}$.

Ethier and Kurtz (2009, Chapter 3, Corollary 7.4) provide a necessary and sufficient condition for relative compactness: $\mathcal{P}_n$ is finite and therefore complete and separable, and the sample paths of $(G^{(n,N)}_{\tau_N(t)})_{t\geq 0}$ live in $\mathcal{X}$, so the conditions of their corollary are satisfied. The corollary states that the sequence of processes $\{(G^{(n,N)}_{\tau_N(t)})_{t\geq 0}\}_{N\in\mathbb{N}}$ is relatively compact if and only if the following two conditions hold:

1. For every $\epsilon > 0$, $t \geq 0$ there exists a compact set $\Gamma \subseteq \mathcal{P}_n$ such that

$$\liminf_{N\to\infty} \mathbb{P}[G^{(n,N)}_{\tau_N(t)} \in \Gamma] \geq 1 - \epsilon$$

2. For every $\epsilon > 0$, $t > 0$ there exists $\delta > 0$ such that

$$\liminf_{N\to\infty} \mathbb{P}[\omega(G^{(n,N)}_{\tau_N(\cdot)}, \delta, t) < \epsilon] \geq 1 - \epsilon$$

where ω is the modified modulus of continuity:

$$\omega(G^{(n,N)}_{\tau_N(\cdot)}, \delta, t) := \inf \max_{i\in[K]} \sup_{u,v\in[T_{i-1},T_i)} \rho\left(G^{(n,N)}_{\tau_N(u)}, G^{(n,N)}_{\tau_N(v)}\right)$$

with the infimum taken over all partitions of the form $0 = T_0 < T_1 < \cdots < T_{K-1} < t \leq T_K$ (for some K) such that $\min_{i\in[K]}(T_i - T_{i-1}) > \delta$.

In our case, Condition 1 is satisfied automatically with $\Gamma = \mathcal{P}_n$, since $\mathcal{P}_n$ is finite and hence compact. Intuitively, Condition 2 ensures that the jumps of the process are well-separated. In our case where ρ is the zero-one metric, we see that $\rho(G^{(n,N)}_{\tau_N(u)}, G^{(n,N)}_{\tau_N(v)})$ is equal to 1 if there is a jump between times u and v, and 0 otherwise. Taking the supremum and maximum then indicates whether there is a jump inside any of the intervals of the given partition; this can only be equal to zero if all of the jumps up to time t occur exactly at the times $T_0, \ldots, T_K$. The infimum over all allowed partitions, then, can only be equal to zero if no two jumps occur less than δ (unscaled) time apart, because of the restriction we placed on these partitions.

The proof is concentrated on proving Condition 2. To do this, we use a coupling with another process that contains all of the jumps of the genealogical process, with the addition of some extra jumps. This process is constructed in such a way that it can be shown to satisfy Condition 2, and hence so does the genealogical process.

Define $p_t := \max_{\xi \in \mathcal{P}_n}\{1 - p_{\xi\xi}(t)\} = 1 - p_{\Delta\Delta}(t)$, where Δ denotes the trivial partition of singletons $\{\{1\}, \ldots, \{n\}\}$. For a proof that the maximum is attained at $\xi = \Delta$, see Lemma 4.2. Following Möhle (1999), we now construct the two-dimensional Markov process $(Z_t, S_t)_{t\in\mathbb{N}_0}$ on $\mathbb{N}_0 \times \mathcal{P}_n$ with transition probabilities

$$\begin{aligned}\mathbb{P}[Z_t = j, S_t = \eta \mid Z_{t-1} = i, S_{t-1} = \xi, \mathcal{F}_\infty]& \\ = \begin{cases} 1 - p_t & \text{if } j = i \text{ and } \xi = \eta \\ p_{\xi\xi}(t) + p_t - 1 & \text{if } j = i+1 \text{ and } \xi = \eta \\ p_{\xi\eta}(t) & \text{if } j = i+1 \text{ and } \xi \neq \eta \\ 0 & \text{otherwise} \end{cases}&\end{aligned} \tag{4.2}$$

and initial state $Z_0 = 0$, $S_0 = \Delta$. Unlike the corresponding process in Möhle (1999), in our case the transition probabilities depend on offspring counts, thus the process is only Markovian conditional on $\mathcal{F}_\infty$. It can be thought of as a Markov process in a random environment.

The construction is such that the marginal (S_t) has the same distribution as the genealogical process of interest, and (Z_t) has jumps at all the times (S_t) does plus some extra jumps. The definition of p_t ensures that the probability in the second case of (4.2) is non-negative, attaining the value zero when $\xi = \Delta$. And the transition probabilities (jump times) of Z do not depend on the current state.

Denote by $0 = T_0^{(N)} < T_1^{(N)} < \ldots$ the jump times of the rescaled process $(Z_{\tau_N(t)})_{t\ge0}$, and by $\varpi_i^{(N)} := T_i^{(N)} - T_{i-1}^{(N)}$ the corresponding holding times.

Suppose that for some fixed $\varpi_1^{(N)}, \ldots, \varpi_m^{(N)}$ and $t > 0$, there exists $m \in \mathbb{N}$ and $\delta > 0$ such that $\varpi_i^{(N)} > \delta$ for all $i \in \{1, \ldots, m\}$, and $T_m^{(N)} \ge t$. Then $K_N := \min\{i : T_i^{(N)} \ge t\}$ is well-defined with $1 \le K_N \le m$, and $T_1^{(N)}, \ldots, T_{K_N}^{(N)}$ form a partition of the form required for Condition 2. Indeed $(Z_{\tau_N(\cdot)})$ is constant on every interval $[T_{i-1}^{(N)}, T_i^{(N)})$ by construction, so $\omega((Z_{\tau_N(\cdot)}), \delta, t) = 0$. We therefore have that for each $m \in \mathbb{N}$ and $\delta > 0$,

$$\mathbb{P}[\omega((Z_{\tau_N(\cdot)}), \delta, t) < \epsilon] \ge \mathbb{P}[T_m^{(N)} \ge t, \varpi_i^{(N)} > \delta\, \forall i \in \{1, \ldots, m\}].$$

Thus a sufficient condition for Condition 2 is: for any $\epsilon > 0$, $t > 0$, there exist $m \in \mathbb{N}$, $\delta > 0$ such that

$$\liminf_{N\to\infty} \mathbb{P}[T_m^{(N)} \ge t, \varpi_i^{(N)} > \delta\, \forall i \in \{1, \ldots, m\}] \ge 1 - \epsilon. \tag{4.3}$$

Since $T_m^{(N)} = \varpi_1^{(N)} + \cdots + \varpi_m^{(N)}$, there is a positive correlation between $T_m^{(N)}$ and each of the $\varpi_i^{(N)}$, and the $\varpi_i^{(N)}$'s are independent conditionally... should all these probabilities be conditioned on $\mathcal{F}_\infty$?, so

$$\begin{aligned}\mathbb{P}[T_m^{(N)} \geq& t, \varpi_i^{(N)} > \delta\, \forall i \in \{1,\dots,m\}] \\ &= \mathbb{P}[T_m^{(N)} \geq t \mid \varpi_i^{(N)} > \delta\, \forall i \in \{1,\dots,m\}]\, \mathbb{P}[\varpi_i^{(N)} > \delta\, \forall i \in \{1,\dots,m\}] \\ &\geq \mathbb{P}[T_m^{(N)} \geq t]\, \mathbb{P}[\varpi_i^{(N)} > \delta\, \forall i \in \{1,\dots,m\}].\end{aligned}$$

Due to Lemma 4.3, the limiting distributions of $\varpi_i^{(N)}$ are i.i.d. $\mathrm{Exp}(\alpha_n)$, so

$$\liminf_{N\to\infty} \mathbb{P}[\varpi_i^{(N)} > \delta\, \forall i \in \{1,\dots,m\}] = (e^{-\alpha_n \delta})^m$$

and

$$\liminf_{N\to\infty} \mathbb{P}[T_m^{(N)} \geq t] = \liminf_{N\to\infty} \mathbb{P}[\varpi_1^{(N)} + \dots + \varpi_m^{(N)} \geq t] = e^{-\alpha_n \delta} \sum_{i=0}^{m-1} \frac{(\alpha_n t)^i}{i!}.$$

using the series expansion for the Erlang CDF (see for example Forbes et al. 2011, Chapter 15). Hence

$$\liminf_{N\to\infty} \mathbb{P}[T_m^{(N)} \geq t, \varpi_i^{(N)} > \delta\, \forall i \in \{1,\dots,m\}] \geq (e^{-\alpha_n \delta})^{m+1} \sum_{i=0}^{m-1} \frac{(\alpha_n t)^i}{i!},$$

which can be made $\geq 1-\epsilon$ by taking m sufficiently large and δ sufficiently small. Since this argument applies for any ϵ and t, (4.3) and hence Condition 2 is satisfied, and the proof is complete. ■

Lemma 4.2. $\max_{\xi\in\mathcal{P}_n}(1 - p_{\xi\xi}(t)) = 1 - p_{\Delta\Delta}(t)$.

Proof. Consider any $\xi \in E$ consisting of k blocks ($1 \leq k \leq n-1$), and any $\xi' \in E$ consisting of $k+1$ blocks. Setting $\eta = \xi$ in (3.4),

$$p_{\xi\xi}(t) = \frac{1}{(N)_k} \sum_{\substack{i_1,\dots,i_k \\ \text{all distinct}}} \nu_t^{(i_1)} \cdots \nu_t^{(i_k)}.$$

Similarly,

$$\begin{aligned}p_{\xi'\xi'}(t) &= \frac{1}{(N)_{k+1}} \sum_{\substack{i_1,\dots,i_k,i_{k+1} \\ \text{all distinct}}} \nu_t^{(i_1)} \cdots \nu_t^{(i_k)} \nu_t^{(i_{k+1})} \\ &= \frac{1}{(N)_k (N-k)} \sum_{\substack{i_1,\dots,i_k \\ \text{all distinct}}} \left\{ \nu_t^{(i_1)} \cdots \nu_t^{(i_k)} \sum_{\substack{i_{k+1}=1 \\ \text{also distinct}}}^{N} \nu_t^{(i_{k+1})} \right\}.\end{aligned}$$

Discarding the zero summands,

$$p_{\xi'\xi'}(t) = \frac{1}{(N)_k(N-k)} \sum_{\substack{i_1,\ldots,i_k \\ \text{all distinct:} \\ \nu_t^{(i_1)},\ldots,\nu_t^{(i_k)}>0}} \left\{ \nu_t^{(i_1)} \cdots \nu_t^{(i_k)} \sum_{\substack{i_{k+1}=1 \\ \text{also distinct}}}^{N} \nu_t^{(i_{k+1})} \right\}.$$

The inner sum is

$$\sum_{\substack{i_{k+1}=1 \\ \text{also distinct}}}^{N} \nu_t^{(i_{k+1})} = \left\{ \sum_{i=1}^{N} \nu_t^{(i)} - \sum_{i \in \{i_1,\ldots,i_k\}} \nu_t^{(i)} \right\} \leq N-k$$

since $\nu_t^{(i_1)}, \ldots, \nu_t^{(i_k)}$ are all at least 1. Hence

$$p_{\xi'\xi'}(t) \leq \frac{N-k}{(N)_k(N-k)} \sum_{\substack{i_1,\ldots,i_k \\ \text{all distinct:} \\ \nu_t^{(i_1)},\ldots,\nu_t^{(i_k)}>0}} \nu_t^{(i_1)} \cdots \nu_t^{(i_k)} = p_{\xi\xi}(t).$$

Thus $p_{\xi\xi}(t)$ is decreasing in the number of blocks of ξ, and is therefore minimised by taking $\xi = \Delta$, which achieves the maximum n blocks. This choice in turn maximises $1 - p_{\xi\xi}(t)$, as required. ∎

Lemma 4.3. *The finite-dimensional distributions of $\varpi_1^{(N)}, \varpi_2^{(N)}, \ldots$ converge as $N \to \infty$ to those of $\varpi_1, \varpi_2, \ldots$, where the ϖ_i are independent* $\mathrm{Exp}(\alpha_n)$*-distributed random variables.*

Proof. There is a continuous bijection between the jump times $T_1^{(N)}, T_2^{(N)}, \ldots$ and the holding times $\varpi_1^{(N)}, \varpi_2^{(N)}, \ldots$, so convergence of the holding times to $\varpi_1, \varpi_2, \ldots$ is equivalent to convergence of the jump times to $T_1, T_2, \ldots$, where $T_i := \varpi_1 + \cdots + \varpi_i$. We will work with the jump times, following the structure of Möhle (1999, Lemma 3.2).

The idea is to prove by induction that, for any $k \in \mathbb{N}$ and $t_1, \ldots, t_k > 0$,

$$\lim_{N\to\infty} \mathbb{P}[T_1^{(N)} \leq t_1, \ldots, T_k^{(N)} \leq t_k] = \mathbb{P}[T_1 \leq t_1, \ldots, T_k \leq t_k]. \tag{4.4}$$

Take the basis case $k = 1$. Then

$$\mathbb{P}[T_1 \leq t] = \mathbb{P}[\varpi_1 \leq t] = 1 - e^{-\alpha_n t}$$

and $T_1^{(N)} > t$ if and only if Z has no jumps up to time t: Expectation appears by tower property to remove (implicit) conditioning in transition probabilities

$$\mathbb{P}[T_1^{(N)} > t] = \mathbb{E}\left[\prod_{r=1}^{\tau_N(t)} (1 - p_r) \right].$$

Lemma 4.7 shows that this probability converges to $e^{-\alpha_n t}$ as required.

For the induction step, assume that (4.4) holds for some k. We have the following decomposition:

$$\begin{aligned}\mathbb{P}[T_1^{(N)} \le t_1, \dots, T_{k+1}^{(N)} \le t_{k+1}] = \mathbb{P}[T_1^{(N)} \le t_1, \dots, T_k^{(N)} \le t_k] \\ - \mathbb{P}[T_1^{(N)} \le t_1, \dots, T_k^{(N)} \le t_k, T_{k+1}^{(N)} > t_{k+1}].\end{aligned}$$

The first term on the RHS converges to $\mathbb{P}[T_1 \le t_1, \dots, T_k \le t_k]$ by the induction hypothesis, and it remains to show that

$$\lim_{N\to\infty} \mathbb{P}[T_1^{(N)} \le t_1, \dots, T_k^{(N)} \le t_k, T_{k+1}^{(N)} > t_{k+1}] = \mathbb{P}[T_1 \le t_1, \dots, T_k \le t_k, T_{k+1} > t_{k+1}].$$

As shown in Möhle (1999), the RHS

$$\mathbb{P}[T_1 \le t_1, \dots, T_k \le t_k, T_{k+1} > t_{k+1}] = \alpha_n^k e^{-\alpha_n t} \sum_{\substack{i_1 \le \cdots \le i_{k-1} \\ \in \{0,\dots,k\}: \\ i_j \ge j \forall j}} \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!}.$$

The event on the LHS can be written (Möhle 1999)

$$\mathbb{P}[T_1^{(N)} \le t_1, \dots, T_k^{(N)} \le t_k, T_{k+1}^{(N)} > t_{k+1}] = \mathbb{E}\left[\sum_{\substack{r_1 < \cdots < r_k: \\ r_i \le \tau_N(t_i) \forall i}} \left(\prod_{i=1}^{k} p_{r_i} \right) \left(\prod_{\substack{r=1 \\ \notin \{r_1,\dots,r_k\}}}^{\tau_N(t)} (1 - p_r) \right) \right],$$

that is, there are jumps at some times $r_1, \dots, r_k$ and identity transitions at all other times. Lemmata 4.8 and 4.9 show that this probability converges to the correct limit. This completes the induction. ■

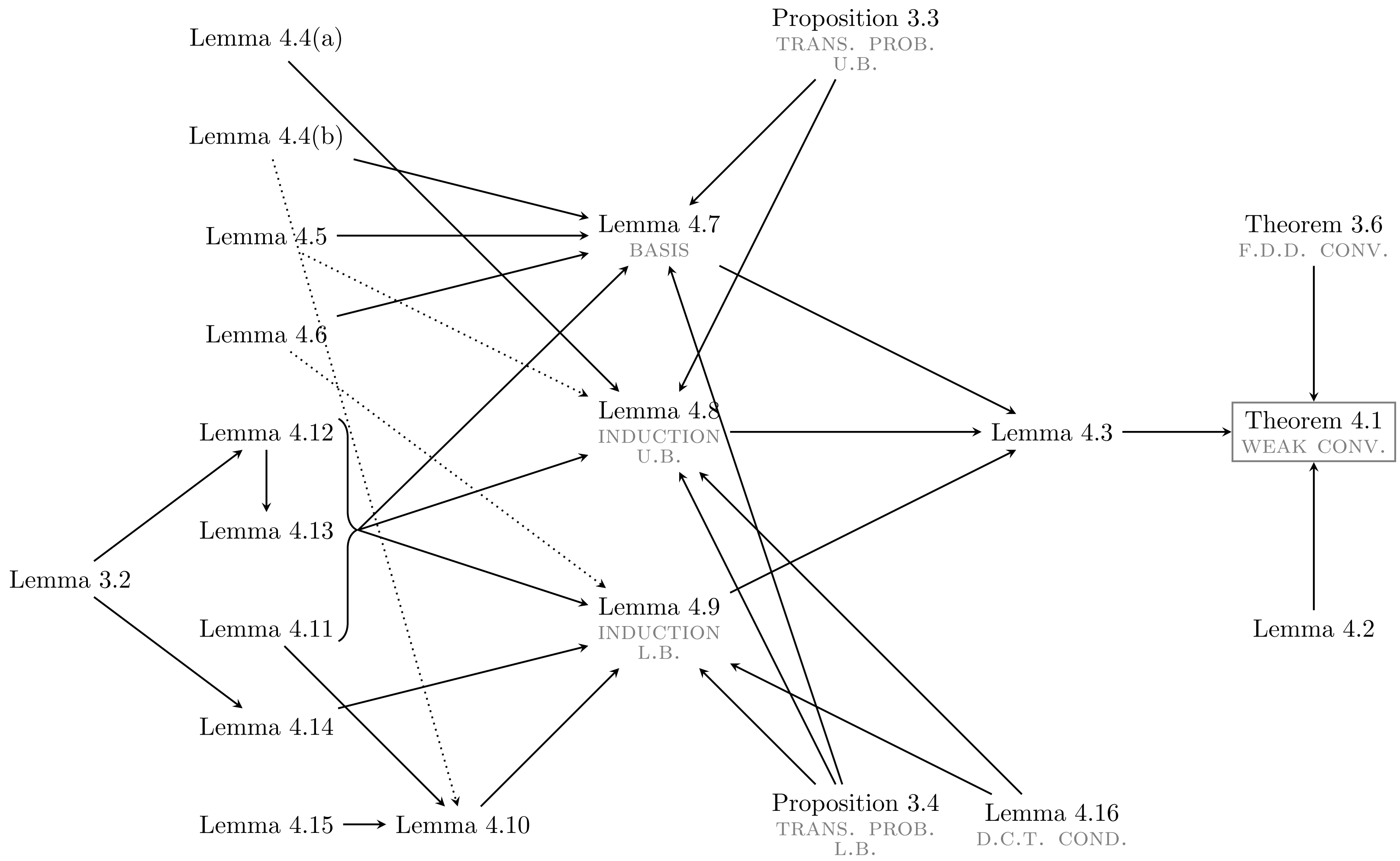


Figure 4.1: Graph showing dependencies between the lemmata used to prove weak convergence. Dotted arrows indicate dependence via a slight modification of the preceding lemma. Dependencies preceding Theorem 3.6 are not shown; these are shown in Figure ?? draw a corresponding figure for fdd proof, or delete this sentence?.

4.1 Bounds on sum-products

Lemma 4.4. *Fix $t > 0$, $l \in \mathbb{N}$.*

(a) $$\sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) \leq (t+1)^l$$

(b) $$t^l - \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \right) \binom{l}{2} (t+1)^{l-2} \leq \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) \leq t^l + c_N(\tau_N(t))(t+1)^l$$

Proof. **(a)** It is a true fact that

$$\sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) \leq \left(\sum_{s=0}^{\tau_N(t)} c_N(s) \right)^l ,$$

as can be seen by considering the multinomial expansion of the RHS. Applying (4.16),

$$\sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) \leq (t+1)^l. \tag{4.5}$$

(b) As pointed out in Koskela et al. (2018, Equation (8)),

$$\sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) \geq \left(\sum_{s=0}^{\tau_N(t)} c_N(s) \right)^l - \binom{l}{2} \left(\sum_{s=0}^{\tau_N(t)} c_N(s)^2 \right) \left(\sum_{s=0}^{\tau_N(t)} c_N(s) \right)^{l-2} . \tag{4.6}$$

Applying (4.16) on the RHS of (4.6) yields the lower bound.

For the upper bound we have

$$\sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) \leq \left(\sum_{s=0}^{\tau_N(t)} c_N(s) \right)^l \leq \left(\sum_{s=0}^{\tau_N(t)-1} c_N(s) + c_N(\tau_N(t)) \right)^l \leq \left[t + c_N(\tau_N(t)) \right]^l ,$$

using the definition of τ_N. A binomial expansion yields

$$[t + c_N(\tau_N(t))]^l = t^l + \sum_{i=0}^{l-1} \binom{l}{i} t^i c_N(\tau_N(t))^{l-i} = t^l + c_N(\tau_N(t)) \sum_{i=0}^{l-1} \binom{l}{i} t^i c_N(\tau_N(t))^{l-1-i},$$

then by (4.14),

$$\sum_{i=0}^{l-1} \binom{l}{i} t^i c_N(\tau_N(t))^{l-1-i} \leq \sum_{i=0}^{l-1} \binom{l}{i} t^i \leq (t+1)^l.$$

Putting this together yields the upper bound. ∎

Lemma 4.5. *Fix $t > 0$, $l \in \mathbb{N}$. Let B be a positive constant which may depend on n.*

$$\sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} [c_N(s_j) + BD_N(s_j)] \leq \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) + \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right) (t+1)^{l-1} (1+B)^l.$$

Proof. We start with a binomial expansion:

$$\begin{aligned} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} [c_N(s_j) + BD_N(s_j)] &= \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \sum_{\mathcal{I} \subseteq [l]} B^{l-|\mathcal{I}|} \left(\prod_{i \in \mathcal{I}} c_N(s_i) \right) \left(\prod_{j \notin \mathcal{I}} D_N(s_j) \right) \\ &= \sum_{\mathcal{I} \subseteq [l]} B^{l-|\mathcal{I}|} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \left(\prod_{i \in \mathcal{I}} c_N(s_i) \right) \left(\prod_{j \notin \mathcal{I}} D_N(s_j) \right) \end{aligned} \tag{4.7}$$

where $[l] := \{1, \ldots, l\}$. Since the sum is over all permutations of $s_1, \ldots, s_l$, we may arbitrarily choose an ordering for $\{1, \ldots, l\}$ such that $\mathcal{I} = \{1, \ldots, |\mathcal{I}|\}$:

$$\begin{aligned} &\sum_{\mathcal{I} \subseteq [l]} B^{l-|\mathcal{I}|} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \left(\prod_{i \in \mathcal{I}} c_N(s_i) \right) \left(\prod_{j \notin \mathcal{I}} D_N(s_j) \right) \\ &\qquad = \sum_{I=0}^{l} \binom{l}{I} B^{l-I} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \left(\prod_{i=1}^{I} c_N(s_i) \right) \left(\prod_{j=I+1}^{l} D_N(s_j) \right). \end{aligned}$$

Separating the term $I = l$,

$$\begin{aligned} &\sum_{I=0}^{l} \binom{l}{I} B^{l-I} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \left(\prod_{i=1}^{I} c_N(s_i) \right) \left(\prod_{j=I+1}^{l} D_N(s_j) \right) \\ &\qquad = \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) + \sum_{I=0}^{l-1} \binom{l}{I} B^{l-I} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \left(\prod_{i=1}^{I} c_N(s_i) \right) \left(\prod_{j=I+1}^{l} D_N(s_j) \right). \end{aligned} \tag{4.8}$$

In the second term on the RHS, there is always at least one D_N term, so using (4.15) we can write

$$\begin{aligned}
&\sum_{I=0}^{l-1}\binom{l}{I}B^{l-I}\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\left(\prod_{i=1}^{I}c_N(s_i)\right)\left(\prod_{j=I+1}^{l}D_N(s_j)\right)\\
&\qquad\leq\sum_{I=0}^{l-1}\binom{l}{I}B^{l-I}\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\left(\prod_{i=1}^{l-1}c_N(s_i)\right)D_N(s_l)\\
&\qquad\leq\sum_{I=0}^{l-1}\binom{l}{I}B^{l-I}\left(\sum_{s_1\neq\cdots\neq s_{l-1}}^{\tau_N(t)}\prod_{i=1}^{l-1}c_N(s_i)\right)\sum_{s_l=1}^{\tau_N(t)}D_N(s_l)\\
&\qquad\leq\sum_{I=0}^{l-1}\binom{l}{I}B^{l-I}(t+1)^{l-1}\sum_{s=1}^{\tau_N(t)}D_N(s) \qquad (4.9)
\end{aligned}$$

using (4.5). Finally, by the Binomial Theorem,

$$\sum_{I=0}^{l-1}\binom{l}{I}B^{l-I}(t+1)^{l-1}\sum_{s=1}^{\tau_N(t)}D_N(s)\leq\left(\sum_{s=1}^{\tau_N(t)}D_N(s)\right)(t+1)^{l-1}(1+B)^{l}, \qquad (4.10)$$

which, together with (4.8), concludes the proof. ■

Lemma 4.6. *Fix $t>0$, $l\in\mathbb{N}$. Let B be a positive constant which may depend on n.*

$$\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\prod_{j=1}^{l}[c_N(s_j)-BD_N(s_j)]\geq\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\prod_{j=1}^{l}c_N(s_j)-\left(\sum_{s=1}^{\tau_N(t)}D_N(s)\right)(t+1)^{l-1}(1+B)^{l}.$$

Proof. A binomial expansion and subsequent manipulation as in (4.7)–(4.8) gives

$$\begin{aligned}
&\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\prod_{j=1}^{l}[c_N(s_j)-BD_N(s_j)]\\
&\qquad=\sum_{\mathcal{I}\subseteq[l]}(-B)^{l-|\mathcal{I}|}\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\left(\prod_{i\in\mathcal{I}}c_N(s_i)\right)\left(\prod_{j\notin\mathcal{I}}D_N(s_j)\right)\\
&\qquad=\sum_{I=0}^{l}\binom{l}{I}(-B)^{l-I}\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\left(\prod_{i=1}^{I}c_N(s_i)\right)\left(\prod_{j=I+1}^{l}D_N(s_j)\right)\\
&\qquad=\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\prod_{j=1}^{l}c_N(s_j)+\sum_{I=0}^{l-1}\binom{l}{I}(-B)^{l-I}\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\left(\prod_{i=1}^{I}c_N(s_i)\right)\left(\prod_{j=I+1}^{l}D_N(s_j)\right)\\
&\qquad\geq\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\prod_{j=1}^{l}c_N(s_j)-\sum_{I=0}^{l-1}\binom{l}{I}B^{l-I}\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\left(\prod_{i=1}^{I}c_N(s_i)\right)\left(\prod_{j=I+1}^{l}D_N(s_j)\right)
\end{aligned}$$

where the last inequality just multiplies some positive terms by -1. Then (4.9)–(4.10) can be applied directly (noting that an upper bound on negative terms gives a lower bound overall):

$$-\sum_{I=0}^{l-1}\binom{l}{I}B^{l-I}\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\left(\prod_{i=1}^{I}c_N(s_i)\right)\left(\prod_{j=I+1}^{l}D_N(s_j)\right)\geq-\left(\sum_{s=1}^{\tau_N(t)}D_N(s)\right)(t{+}1)^{l-1}(1{+}B)^l$$

which concludes the proof. ■

4.2 Main components of induction argument

Recall that the following conditions are all consequences of (4.1): for all $t > s > 0$,

$$\mathbb{E}\left[c_N(\tau_N(t))\right]\to 0 \tag{4.11}$$

$$\mathbb{E}\left[\sum_{r=\tau_N(s)+1}^{\tau_N(t)}c_N(r)^2\right]\to 0 \tag{4.12}$$

$$\mathbb{E}\left[\sum_{r=\tau_N(s)+1}^{\tau_N(t)}D_N(r)\right]\to 0 \tag{4.13}$$

as $N\to\infty$. Also recall the following properties from Proposition 3.1:

$$c_N(t), D_N(t)\in[0,1] \tag{4.14}$$

$$D_N(t)\leq c_N(t) \tag{4.15}$$

$$t'\leq\sum_{r=1}^{\tau_N(t')}c_N(r)\leq t'+1. \tag{4.16}$$

Lemma 4.7 (Basis step)**.** *Assume* (4.1) *holds. For any* $0<t<\infty$*,*

$$\lim_{N\to\infty}\mathbb{E}\left[\prod_{r=1}^{\tau_N(t)}(1-p_r)\right]=e^{-\alpha_n t}$$

where $\alpha_n := n(n-1)/2$.

Proof. We start by showing that $\lim_{N\to\infty}\mathbb{E}\left[\prod_{r=1}^{\tau_N(t)}(1-p_r)\right]\leq e^{-\alpha_n t}$. Setting $\xi=\Delta$ in Proposition 3.4, we have for each r

$$1-p_r=p_{\Delta\Delta}(r)\leq 1-\alpha_n 1_N\left[c_N(r)-B_n'D_N(r)\right]. \tag{4.17}$$

When $N\geq 3$, a sufficient condition to ensure the bound in (4.17) is non-negative is that the event

$$E_N^1(r):=\left\{c_N(r)<\alpha_n^{-1}A_N\right\} \tag{4.18}$$

occurs, where $A_N = 1_N$ as $N \to \infty$ and is independent of r but will not be specified explicitly. We will also need to control the sign of $c_N(r) - B_n' D_N(r)$, for which we define the event

$$E_N^2(r) := \left\{ c_N(r) \geq B_n' D_N(r) \right\}, \tag{4.19}$$

and we define $E_N^1 := \bigcap_{r=1}^{\tau_N(t)} E_N^1(r)$ and $E_N^2 := \bigcap_{r=1}^{\tau_N(t)} E_N^2(r)$. Then

$$1 - p_r = p_{\Delta\Delta}(r) \leq 1 - \alpha_n 1_N \left[c_N(r) - B_n' D_N(r) \right] \mathbb{1}_{E_N^1 \cap E_N^2}.$$

Applying a multinomial expansion and then separating the positive and negative terms,

$$\begin{aligned}
\prod_{r=1}^{\tau_N(t)} (1 - p_r) &\leq 1 + \sum_{l=1}^{\tau_N(t)} (-\alpha_n)^l 1_N \frac{1}{l!} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} \left[c_N(s_j) - B_n' D_N(s_j) \right] \mathbb{1}_{E_N^1 \cap E_N^2} \\
&= 1 + \sum_{\substack{l=2 \\ \text{even}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} \left[c_N(s_j) - B_n' D_N(s_j) \right] \mathbb{1}_{E_N^1 \cap E_N^2} \\
&\quad - \sum_{\substack{l=1 \\ \text{odd}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} \left[c_N(s_j) - B_n' D_N(s_j) \right] \mathbb{1}_{E_N^1 \cap E_N^2}.
\end{aligned} \tag{4.20}$$

This is further bounded by applying Lemma 4.6 and then both bounds of Lemma 4.4(b):

$$\begin{aligned}
\prod_{r=1}^{\tau_N(t)} (1 - p_r) \leq 1 + \mathbb{1}_{E_N^1 \cap E_N^2} &\Bigg\{ \sum_{\substack{l=2 \\ \text{even}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) \\
&- \sum_{\substack{l=1 \\ \text{odd}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!} \left[\sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) - \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right) (t+1)^{l-1} (1 + B_n')^l \right] \Bigg\} \\
\leq 1 + \Bigg\{ &\sum_{\substack{l=2 \\ \text{even}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!} \left\{ t^l + c_N(\tau_N(t)) (t+1)^l \right\} \\
&- \sum_{\substack{l=1 \\ \text{odd}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!} \left[t^l - \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \right) \binom{l}{2} (t+1)^{l-2} \right] \\
&- \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right) (t+1)^{l-1} (1 + B_n')^l \Bigg\} \mathbb{1}_{E_N^1 \cap E_N^2}.
\end{aligned}$$

Collecting some terms,

$$\begin{aligned}\prod_{r=1}^{\tau_N(t)}(1-p_r) &\le 1+\sum_{l=1}^{\tau_N(t)}(-\alpha_n)^l 1_N \frac{1}{l!}t^l \mathbb{1}_{E_N^1\cap E_N^2} + c_N(\tau_N(t))\sum_{\substack{l=2\\ \text{even}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!}(t+1)^l \\ &\quad + \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right)\sum_{\substack{l=1\\ \text{odd}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!}\binom{l}{2}(t+1)^{l-2} \\ &\quad + \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right)\sum_{\substack{l=1\\ \text{odd}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!}(t+1)^{l-1}(1+B_n')^l \\ &\le 1+\sum_{l=1}^{\infty}(-\alpha_n)^l 1_N \frac{1}{l!}t^l \mathbb{1}_{\{\tau_N(t)\ge l\}}\mathbb{1}_{E_N^1\cap E_N^2} + c_N(\tau_N(t))\exp[\alpha_n 1_N(t+1)] \\ &\quad + \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right)\frac{1}{2}\alpha_n^2\exp[\alpha_n 1_N(t+1)] \\ &\quad + \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right)\exp[\alpha_n 1_N(t+1)(1+B_n')]. \end{aligned} \tag{4.21}$$

Now, taking the expectation and limit, then applying (4.11)–(4.13), and Lemmata 4.12, 4.13 and 4.14 to deal with the indicators,

$$\begin{aligned}\lim_{N\to\infty}\mathbb{E}\left[\prod_{r=1}^{\tau_N(t)}(1-p_r)\right] &\le 1+\sum_{l=1}^{\infty}(-\alpha_n)^l\frac{1}{l!}t^l\lim_{N\to\infty}\mathbb{P}\left[\{\tau_N(t)\ge l\}\cap E_N^1\cap E_N^2\right] \\ &\quad + \lim_{N\to\infty}\mathbb{E}\left[c_N(\tau_N(t))\right]\exp[\alpha_n(t+1)] \\ &\quad + \lim_{N\to\infty}\mathbb{E}\left[\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right]\frac{1}{2}\alpha_n^2\exp[\alpha_n(t+1)] \\ &\quad + \lim_{N\to\infty}\mathbb{E}\left[\sum_{s=1}^{\tau_N(t)} D_N(s)\right]\exp[\alpha_n(t+1)(1+B_n')] \\ &= 1+\sum_{l=1}^{\infty}(-\alpha_n)^l\frac{1}{l!}t^l = e^{-\alpha_n t}. \end{aligned} \tag{4.22}$$

Passing the limit and expectation inside the infinite sum is justified by dominated convergence and Fubini.

It remains to show the corresponding lower bound

$$\lim_{N\to\infty}\mathbb{E}\left[\prod_{r=1}^{\tau_N(t)}(1-p_r)\right] \ge e^{-\alpha_n t}.$$

Setting $\xi = \Delta$ in Proposition 3.3, we have

$$1 - p_t = p_{\Delta\Delta}(t) \geq 1 - \frac{N^{n-2}}{(N-2)_{n-2}} \alpha_n [c_N(t) + B_n D_N(t)] \tag{4.23}$$

where $B_n > 0$. Due to (4.15), a sufficient condition for this bound to be non-negative is

$$E_N^3(r) := \left\{ c_N(r) \leq \frac{(N-2)_{n-2}}{N^{n-2}} \alpha_n^{-1} (1 + B_n)^{-1} \right\}, \tag{4.24}$$

and we again define $E_N^3 := \bigcap_{r=1}^{\tau_N(t)} E_N^3(r)$. We now apply a multinomial expansion to the product, and split into positive and negative terms:

$$\begin{aligned}
\prod_{r=1}^{\tau_N(t)} (1 - p_r) &\geq \left\{ 1 + \sum_{l=1}^{\tau_N(t)} (-\alpha_n)^l 1_N \frac{1}{l!} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} [c_N(s_j) + B_n D_N(s_j)] \right\} \mathbb{1}_{E_N^3} \\
&= \left\{ 1 + \sum_{\substack{l=2 \\ \text{even}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} [c_N(s_j) + B_n D_N(s_j)] \right. \\
&\qquad \left. - \sum_{\substack{l=1 \\ \text{odd}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} [c_N(s_j) + B_n D_N(s_j)] \right\} \mathbb{1}_{E_N^3}
\end{aligned}$$

This is further bounded by applying Lemma 4.5 and both bounds in Lemma 4.4(b):

$$\begin{aligned}
\prod_{r=1}^{\tau_N(t)} (1 - p_r) \geq \mathbb{1}_{E_N^3} &\left\{ 1 + \sum_{\substack{l=2 \\ \text{even}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) \right. \\
&\left. - \sum_{\substack{l=1 \\ \text{odd}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!} \left[\sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) + \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right) (t+1)^{l-1} (1 + B_n)^l \right] \right\} \\
\geq \mathbb{1}_{E_N^3} &\left\{ 1 + \sum_{\substack{l=2 \\ \text{even}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!} \left[t^l - \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \right) \binom{l}{2} (t+1)^{l-2} \right] \right. \\
&\left. - \sum_{\substack{l=1 \\ \text{odd}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!} \left[t^l + c_N(\tau_N(t))(t+1)^l + \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right) (t+1)^{l-1} (1 + B_n)^l \right] \right\}.
\end{aligned}$$

Collecting terms,

$$\begin{aligned}
\prod_{r=1}^{\tau_N(t)}(1-p_r) \geq \sum_{l=0}^{\tau_N(t)}(-\alpha_n)^l 1_N \frac{1}{l!}t^l \mathbb{1}_{E_N^3} &- \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right) \sum_{\substack{l=2\\ \text{even}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!}\binom{l}{2}(t+1)^{l-2} \\
&- c_N(\tau_N(t)) \sum_{\substack{l=1\\ \text{odd}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!}(t+1)^l \\
&- \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right) \sum_{\substack{l=1\\ \text{odd}}}^{\tau_N(t)} \alpha_n^l 1_N \frac{1}{l!}(t+1)^{l-1}(1+B_n)^l \\
\geq \sum_{l=0}^{\infty}(-\alpha_n)^l 1_N \frac{1}{l!}t^l \mathbb{1}_{E_N^3}\mathbb{1}_{\{\tau_N(t)\geq l\}} &- \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right)\frac{1}{2}\alpha_n^2 \exp[\alpha_n 1_N(t+1)] \\
&- c_N(\tau_N(t))\exp[\alpha_n 1_N(t+1)] \\
&- \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right)\exp[\alpha_n 1_N(t+1)(1+B_n)]. \qquad (4.25)
\end{aligned}$$

Now, taking the expectation and limit, and applying (4.11)–(4.13) to show that all but the first sum vanish, and Lemmata 4.13 and 4.12 to show that $\lim_{N\to\infty}\mathbb{P}[\{\tau_N(t)\geq l\}\cap E_N^3] = 1$,

$$\begin{aligned}
\lim_{N\to\infty}\mathbb{E}\left[\prod_{r=1}^{\tau_N(t)}(1-p_r)\right] &\geq \sum_{l=0}^{\infty}(-\alpha_n)^l 1_N \frac{1}{l!}t^l \lim_{N\to\infty}\mathbb{P}\left[\{\tau_N(t)\geq l\}\cap E_N^3\right] \\
&\quad - \lim_{N\to\infty}\mathbb{E}\left[\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right]\frac{1}{2}\alpha_n^2\exp[\alpha_n(t+1)] \\
&\quad - \lim_{N\to\infty}\mathbb{E}\left[c_N(\tau_N(t))\right]\exp[\alpha_n(t+1)] \\
&\quad - \lim_{N\to\infty}\mathbb{E}\left[\sum_{s=1}^{\tau_N(t)} D_N(s)\right]\exp[\alpha_n(t+1)(1+B_n)] \\
&= \sum_{l=0}^{\infty}(-\alpha_n)^l\frac{1}{l!}t^l = e^{-\alpha_n t}. \qquad (4.26)
\end{aligned}$$

Again, passing the limit and expectation inside the infinite sum is justified by dominated convergence and Fubini. Combining the upper and lower bounds in (4.22) and (4.26) respectively concludes the proof. ■

Lemma 4.8 (Induction step upper bound)**.** *Assume* (4.1) *holds. Fix* $k \in \mathbb{N}$, $i_0 := 0$, $i_k := k$. *For any sequence of times* $0 = t_0 \le t_1 \le \cdots \le t_k \le t$,

$$\lim_{N\to\infty} \mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i \le \tau_N(t_i) \forall i}} \left(\prod_{i=1}^{k} p_{r_i} \right) \left(\prod_{\substack{r=1\\ \notin\{r_1,\ldots,r_k\}}}^{\tau_N(t)} (1-p_r) \right) \right] \le \alpha_n^k e^{-\alpha_n t} \sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j \ge j \forall j}} \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!}.$$

Proof. We use the bound on $(1-p_r)$ from (4.17) and apply a multinomial expansion, defining as in (4.18) and (4.19) respectively the sequences of events E_N^1 and E_N^2 which ensure the bounds are non-negative:

$$\begin{aligned}
\prod_{\substack{r=1\\ \notin\{r_1,\ldots,r_k\}}}^{\tau_N(t)} (1-p_r) &\le \prod_{\substack{r=1\\ \notin\{r_1,\ldots,r_k\}}}^{\tau_N(t)} \left\{ 1 - \alpha_n 1_N [c_N(r) - B_n' D_N(r)] \mathbb{1}_{E_N^1 \cap E_N^2} \right\} \\
&= 1 + \sum_{l=1}^{\tau_N(t)-k} (-\alpha_n)^l 1_N \frac{1}{l!} \sum_{\substack{s_1\ne\cdots\ne s_l\\ \notin\{r_1,\ldots,r_k\}}}^{\tau_N(t)} \prod_{j=1}^{l} [c_N(s_j) - B_n' D_N(s_j)] \mathbb{1}_{E_N^1 \cap E_N^2} \\
&= 1 + \sum_{l=1}^{\tau_N(t)-k} (-\alpha_n)^l 1_N \frac{1}{l!} \sum_{s_1\ne\cdots\ne s_l}^{\tau_N(t)} \prod_{j=1}^{l} [c_N(s_j) - B_n' D_N(s_j)] \mathbb{1}_{E_N^1 \cap E_N^2} \\
&\quad - \sum_{l=1}^{\tau_N(t)-k} (-\alpha_n)^l 1_N \frac{1}{l!} \sum_{\substack{s_1\ne\cdots\ne s_l\\ \exists i,i': s_i = r_{i'}}} \prod_{j=1}^{l} [c_N(s_j) - B_n' D_N(s_j)] \mathbb{1}_{E_N^1 \cap E_N^2}.
\end{aligned} \tag{4.27}$$

The penultimate line above is exactly the expansion we had in the basis step (4.20), except for the limit on l, and as such following the same arguments gives a bound analogous to that in (4.21):

$$\begin{aligned}
1 &+ \sum_{l=1}^{\tau_N(t)-k} (-\alpha_n)^l 1_N \frac{1}{l!} \sum_{s_1\ne\cdots\ne s_l}^{\tau_N(t)} \prod_{j=1}^{l} [c_N(s_j) - B_n' D_N(s_j)] \mathbb{1}_{E_N^1 \cap E_N^2} \\
&\le 1 + \sum_{l=1}^{\tau_N(t)-k} (-\alpha_n)^l 1_N \frac{1}{l!} t^l \mathbb{1}_{E_N^1 \cap E_N^2} + c_N(\tau_N(t)) \exp[\alpha_n 1_N (t+1)] \\
&\quad + \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \right) \frac{1}{2} \alpha_n^2 \exp[\alpha_n 1_N (t+1)] \\
&\quad + \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right) \exp[\alpha_n 1_N (t+1)(1 + B_n')].
\end{aligned}$$

For the last line of (4.27),

$$
\begin{aligned}
-\sum_{l=1}^{\tau_N(t)-k} (-\alpha_n)^l 1_N \frac{1}{l!} \sum_{\substack{s_1\neq\cdots\neq s_l\\ \exists i,i':s_i=r_{i'}}} \prod_{j=1}^{l} &\{c_N(s_j) - B_n' D_N(s_j)\} \mathbb{1}_{E_N^1\cap E_N^2} \\
&\leq \sum_{l=1}^{\tau_N(t)-k} \alpha_n^l 1_N \frac{1}{l!} \sum_{\substack{s_1\neq\cdots\neq s_l\\ \exists i,i':s_i=r_{i'}}} \prod_{j=1}^{l} \{c_N(s_j) + B_n' D_N(s_j)\} \\
&\leq \sum_{l=1}^{\tau_N(t)-k} \alpha_n^l 1_N \frac{1}{l!} \sum_{\substack{s_1\neq\cdots\neq s_l\\ \exists i,i':s_i=r_{i'}}} (1+B_n')^l \prod_{j=1}^{l} c_N(s_j) \\
&\leq \sum_{l=1}^{\tau_N(t)-k} \alpha_n^l 1_N \frac{1}{(l-1)!} \sum_{s_1\in\{r_1,\ldots,r_k\}} \sum_{s_2\neq\cdots\neq s_l}^{\tau_N(t)} (1+B_n')^l \prod_{j=1}^{l} c_N(s_j) \\
&= \sum_{s\in\{r_1,\ldots,r_k\}} c_N(s) \sum_{l=1}^{\tau_N(t)-k} \alpha_n^l 1_N \frac{1}{(l-1)!} (1+B_n')^l \sum_{s_1\neq\cdots\neq s_{l-1}}^{\tau_N(t)} \prod_{j=1}^{l-1} c_N(s_j) \\
&\leq \sum_{j=1}^{k} c_N(r_j) \sum_{l=1}^{\tau_N(t)-k} \alpha_n^l 1_N \frac{1}{(l-1)!} (1+B_n')^l (t+1)^{l-1} \\
&\leq \left(\sum_{j=1}^{k} c_N(r_j)\right) \alpha_n (1+B_n') \exp[\alpha_n 1_N (1+B_n')(t+1)],
\end{aligned}
$$

where the penultimate inequality uses Lemma 4.4(a). Putting these together, we have

$$
\begin{aligned}
\prod_{\substack{r=1\\ \notin\{r_1,\ldots,r_k\}}}^{\tau_N(t)} (1-p_r) \leq 1 + \sum_{l=1}^{\tau_N(t)-k} &(-\alpha_n)^l 1_N \frac{1}{l!} t^l \mathbb{1}_{E_N^1\cap E_N^2} + c_N(\tau_N(t)) \exp[\alpha_n 1_N (t+1)] \\
&+ \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right) \frac{1}{2} \alpha_n^2 \exp[\alpha_n 1_N (t+1)] \\
&+ \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right) \exp[\alpha_n 1_N (t+1)(1+B_n')] \\
&+ \left(\sum_{j=1}^{k} c_N(r_j)\right) \alpha_n (1+B_n') \exp[\alpha_n 1_N (1+B_n')(t+1)]. \qquad (4.28)
\end{aligned}
$$

Meanwhile, using the bound on p_r from (4.23) then applying a modification of Lemma 4.5

where the sum is over ordered indices rather than distinct indices,

$$
\begin{aligned}
\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^k p_{r_i} &\le \alpha_n^k 1_N \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^k [c_N(r_i)+B_nD_N(r_i)] \\
&\le \alpha_n^k 1_N \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^k c_N(r_i) + \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right) \alpha_n^k 1_N (t+1)^{k-1}(1+B_n)^k.
\end{aligned}
\tag{4.29}
$$

A more liberal (but simpler) bound can be arrived at thus:

$$
\begin{aligned}
\prod_{i=1}^k p_{r_i} &\le \alpha_n^k 1_N \prod_{i=1}^k [c_N(r_i)+B_nD_N(r_i)] \\
&\le \alpha_n^k 1_N \prod_{i=1}^k c_N(r_i)(1+B_n) \\
&\le \alpha_n^k 1_N (1+B_n)^k \prod_{i=1}^k c_N(r_i)
\end{aligned}
$$

which, using Lemma 4.4(a), also leads to the deterministic bound

$$
\begin{aligned}
\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^k p_{r_i} &\le \alpha_n^k 1_N (1+B_n)^k \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^k c_N(r_i) \\
&\le \alpha_n^k 1_N (1+B_n)^k \frac{1}{k!} \sum_{r_1\ne\cdots\ne r_k}^{\tau_N(t)} \prod_{i=1}^k c_N(r_i) \\
&\le \alpha_n^k 1_N (1+B_n)^k \frac{1}{k!}(t+1)^k.
\end{aligned}
\tag{4.30}
$$

Combining (4.28) with the other product, the expression inside the expectation in Lemma 4.8

is bounded above by

$$
\begin{aligned}
&\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\dots,r_k\}}}^{\tau_N(t)} (1-p_r)\right)\\
&\qquad \le \left\{1+\sum_{l=1}^{\tau_N(t)-k}(-\alpha_n)^l 1_N \frac{1}{l!}t^l \mathbb{1}_{E_N^1\cap E_N^2}\right\} \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^{k} p_{r_i}\\
&\qquad\quad + \left\{c_N(\tau_N(t))\exp[\alpha_n 1_N(t+1)] + \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right)\frac{1}{2}\alpha_n^2 \exp[\alpha_n 1_N(t+1)]\right.\\
&\qquad\qquad \left. + \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right)\exp[\alpha_n 1_N(t+1)(1+B_n')]\right\} \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^{k} p_{r_i}\\
&\qquad\quad + \exp[\alpha_n 1_N(1+B_n')(t+1)]\alpha_n(1+B_n') \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \sum_{j=1}^{k} c_N(r_j)\prod_{i=1}^{k} p_{r_i}.
\end{aligned}
$$

Applying the various bounds (4.29)–(4.30), we have

$$
\begin{aligned}
&\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\dots,r_k\}}}^{\tau_N(t)} (1-p_r)\right)\\
&\qquad \le \alpha_n^k 1_N \left\{1+\sum_{l=1}^{\tau_N(t)-k}(-\alpha_n)^l 1_N \frac{1}{l!}t^l \mathbb{1}_{E_N^1\cap E_N^2}\right\} \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\\
&\qquad\quad + \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right)\alpha_n^k 1_N (t+1)^{k-1}(1+B_n)^k \sum_{l=0}^{\tau_N(t)} (\alpha_n)^l 1_N \frac{1}{l!}t^l\\
&\qquad\quad + \left\{c_N(\tau_N(t))\exp[\alpha_n 1_N(t+1)] + \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right)\frac{1}{2}\alpha_n^2 \exp[\alpha_n 1_N(t+1)]\right.\\
&\qquad\qquad \left. + \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right)\exp[\alpha_n 1_N(t+1)(1+B_n')]\right\}\alpha_n^k 1_N(1+B_n)^k\frac{1}{k!}(t+1)^k\\
&\qquad\quad + \exp[\alpha_n(1+B_n')(t+1)]\alpha_n(1+B_n')\alpha_n^k 1_N(1+B_n)^k\\
&\qquad\qquad\qquad \times \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \sum_{j=1}^{k} c_N(r_j)\prod_{i=1}^{k} c_N(r_i).
\end{aligned}
$$

Upon taking the expectation and limit, we have

$$\begin{aligned}
&\lim_{N\to\infty} \mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\dots,r_k\}}}^{\tau_N(t)} (1-p_r)\right)\right] \\
&\quad\le \alpha_n^k \lim_{N\to\infty} \mathbb{E}\left[\left(1+\sum_{l=1}^{\tau_N(t)-k} (-\alpha_n)^l \frac{1}{l!} t^l \mathbb{1}_{E_N^1\cap E_N^2}\right) \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\right] \\
&\qquad + \lim_{N\to\infty} \mathbb{E}\left[\sum_{s=1}^{\tau_N(t)} D_N(s)\right] \alpha_n^k (t+1)^{k-1}(1+B_n)^k \exp[\alpha_n t] \\
&\qquad + \Bigg\{ \lim_{N\to\infty} \mathbb{E}\left[c_N(\tau_N(t))\right] \exp[\alpha_n(t+1)] + \lim_{N\to\infty} \mathbb{E}\left[\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right] \frac{1}{2}\alpha_n^2 \exp[\alpha_n(t+1)] \\
&\qquad\qquad + \lim_{N\to\infty} \mathbb{E}\left[\sum_{s=1}^{\tau_N(t)} D_N(s)\right] \exp[\alpha_n(t+1)(1+B_n')]\Bigg\} \alpha_n^k (1+B_n)^k \frac{1}{k!}(t+1)^k \\
&\qquad + \exp[\alpha_n(1+B_n')(t+1)]\alpha_n^{k+1}(1+B_n')(1+B_n)^k \\
&\qquad\qquad \times \lim_{N\to\infty} \mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \sum_{j=1}^{k} c_N(r_j) \prod_{i=1}^{k} c_N(r_i)\right]. \qquad (4.31)
\end{aligned}$$

The middle terms vanish due to (4.11)–(4.13) and the expression becomes

$$\begin{aligned}
&\lim_{N\to\infty} \mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\dots,r_k\}}}^{\tau_N(t)} (1-p_r)\right)\right] \le \alpha_n^k \lim_{N\to\infty} \mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\right] \\
&\qquad + \alpha_n^k \sum_{l=1}^{\infty} (-\alpha_n)^l \frac{1}{l!} t^l \lim_{N\to\infty} \mathbb{E}\left[\mathbb{1}_{\{\tau_N(t)\ge k+l\}} \mathbb{1}_{E_N^1\cap E_N^2} \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\right] \\
&\qquad + \exp[\alpha_n(1+B_n')(t+1)]\alpha_n^{k+1}(1+B_n')(1+B_n)^k \\
&\qquad\qquad \times \lim_{N\to\infty} \mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \sum_{j=1}^{k} c_N(r_j) \prod_{i=1}^{k} c_N(r_i)\right], \qquad (4.32)
\end{aligned}$$

where passing the limit and expectation inside the infinite sum is justified by dominated convergence and Fubini; see Lemma 4.16. To simplify the last line,

$$\begin{aligned}
\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \sum_{j=1}^{k} c_N(r_j)\prod_{i=1}^{k} c_N(r_i) &\le \frac{1}{k!}\sum_{r_1\ne\cdots\ne r_k}^{\tau_N(t)}\sum_{j=1}^{k} c_N(r_j)\prod_{i=1}^{k} c_N(r_i)\\
&= \frac{1}{k!}\sum_{r_1\ne\cdots\ne r_k}^{\tau_N(t)}\sum_{j=1}^{k} c_N(r_j)^2\prod_{i\ne j} c_N(r_i)\\
&\le \frac{1}{k!}\sum_{j=1}^{k}\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \sum_{r_1\ne\cdots\ne r_{k-1}}^{\tau_N(t)}\prod_{i=1}^{k-1} c_N(r_i)\\
&\le \frac{1}{(k-1)!}\sum_{s=1}^{\tau_N(t)} c_N(s)^2(t+1)^{k-1},
\end{aligned}$$

using Lemma 4.4(a) for the final inequality. Hence

$$\lim_{N\to\infty}\mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}}\sum_{s\in\{r_1,\ldots,r_k\}} c_N(s)\prod_{i=1}^{k} c_N(r_i)\right] \le \frac{1}{(k-1)!}(t+1)^{k-1}\lim_{N\to\infty}\mathbb{E}\left[\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right] = 0$$

by (4.12). By Lemmata 4.13, 4.12 and 4.14, $\lim_{N\to\infty}\mathbb{P}[\{\tau_N(t)\ge k+l\}\cap E_N^1\cap E_N^2] = 1$, so we can apply Lemma 4.10 to the remaining expectations in (4.32), yielding

$$\begin{aligned}
&\lim_{N\to\infty}\mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}}\left(\prod_{i=1}^{k} p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\ldots,r_k\}}}^{\tau_N(t)}(1-p_r)\right)\right]\\
&\qquad\le \alpha_n^k\sum_{l=0}^{\infty}(-\alpha_n)^l\frac{1}{l!}t^l\sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\ge j\forall j}}\prod_{j=1}^{k}\frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}\\
&\qquad= \alpha_n^k e^{-\alpha_n t}\sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\ge j\forall j}}\prod_{j=1}^{k}\frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}
\end{aligned}$$

as required. ∎

Lemma 4.9 (Induction step lower bound)**.** *Assume* (4.1) *holds. Fix* $k \in \mathbb{N}$, $i_0 := 0$, $i_k := k$. *For any sequence of times* $0 = t_0 \le t_1 \le \cdots \le t_k \le t$,

$$\lim_{N\to\infty} \mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\dots,r_k\}}}^{\tau_N(t)} (1-p_r)\right)\right] \ge \alpha_n^k e^{-\alpha_n t} \sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\dots,k\}:\\ i_j\ge j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}.$$

Proof. Firstly,

$$\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\dots,r_k\}}}^{\tau_N(t)} (1-p_r)\right) \ge \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right)\left(\prod_{r=1}^{\tau_N(t)} (1-p_r)\right). \quad (4.33)$$

Now the second product does not depend on $r_1,\dots,r_k$, and we can use the lower bound from (4.25):

$$\begin{aligned}\prod_{r=1}^{\tau_N(t)} (1-p_r) \ge \sum_{l=0}^{\tau_N(t)} (-\alpha_n)^l 1_N \frac{1}{l!} t^l \mathbb{1}_{E_N^3} - \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right) \frac{1}{2}\alpha_n^2 \exp[\alpha_n 1_N (t+1)] \\ - c_N(\tau_N(t)) \exp[\alpha_n 1_N(t+1)] \\ - \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right) \exp[\alpha_n 1_N (t+1)(1+B_n)] \qquad (4.34)\end{aligned}$$

where E_N^3 is defined as in (4.24). We will also need an upper bound on this product, which is formed from (4.21) with a further deterministic bound:

$$\begin{aligned}\prod_{r=1}^{\tau_N(t)} (1-p_r) &\le \sum_{l=0}^{\tau_N(t)} (-\alpha_n)^l 1_N \frac{1}{l!} t^l \mathbb{1}_{\{\tau_N(t)\ge l\}} \mathbb{1}_{E_N^1\cap E_N^2} + c_N(\tau_N(t)) \exp[\alpha_n 1_N (t+1)] \\ &\quad + \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right) \frac{1}{2}\alpha_n^2 \exp[\alpha_n 1_N(t+1)] \\ &\quad + \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right) \exp[\alpha_n 1_N(t+1)(1+B_n')] \\ &\le \exp[\alpha_n 1_N t] + \exp[\alpha_n 1_N (t+1)] \\ &\quad + \frac{1}{2}\alpha_n^2 (t+1)\exp[\alpha_n 1_N(t+1)] + (t+1)\exp[\alpha_n 1_N (t+1)(1+B_n')] \\ &\le \left(2 + \frac{\alpha_n^2(t+1)}{2}\right) \exp[\alpha_n 1_N (t+1)] + (t+1) \exp[\alpha_n 1_N(t+1)(1+B_n')].\end{aligned} \quad (4.35)$$

Now let us consider the remaining sum-product on the RHS of (4.33). We use the same bound on p_r as in (4.17):

$$p_r = 1 - p_{\Delta\Delta}(r) \geq \alpha_n 1_N \left[c_N(r) - B'_n D_N(r)\right] \tag{4.36}$$

where the $O(N^{-1})$ term does not depend on r. When N is large enough for the factor of 1_N to be non-negative, the condition that the bound in (4.36) is non-negative holds on the event E_N^2 that was defined in (4.19). Then

$$\prod_{i=1}^{k} p_{r_i} \geq \alpha_n^k 1_N \prod_{i=1}^{k} \left[c_N(r_i) - B'_n D_N(r_i)\right] \mathbb{1}_{E_N^2}.$$

Applying a modification of Lemma 4.6 where the sum is over ordered indices rather than distinct indices,

$$\begin{aligned}\sum_{\substack{r_1<\cdots<r_k:\\ r_i\leq\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right) &\geq \alpha_n^k 1_N \sum_{\substack{r_1<\cdots<r_k:\\ r_i\leq\tau_N(t_i)\forall i}} \prod_{i=1}^{k} \left[c_N(r_i) - B'_n D_N(r_i)\right] \mathbb{1}_{E_N^2} \\ &\geq \alpha_n^k 1_N \left\{ \sum_{\substack{r_1<\cdots<r_k:\\ r_i\leq\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i) \mathbb{1}_{E_N^2} - \frac{1}{k!}\left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right)(t+1)^{k-1}(1+B'_n)^k \right\}.\end{aligned}$$

The above expression is already split into positive and negative terms; a lower bound on (4.33) can be formed by multiplying the positive terms by the lower bound (4.34) and the negative terms by the upper bound (4.35). Thus

$$
\begin{aligned}
&\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\ldots,r_k\}}}^{\tau_N(t)} (1-p_r)\right) \\
&\qquad \ge \alpha_n^k \mathbb{1}_N \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\mathbb{1}_{E_N^2}\Bigg\{\sum_{l=0}^{\tau_N(t)}(-\alpha_n)^l \mathbb{1}_N \frac{1}{l!}t^l \mathbb{1}_{E_N^3} \\
&\qquad\qquad - \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right)\frac{1}{2}\alpha_n^2 \exp[\alpha_n \mathbb{1}_N(t+1)] \\
&\qquad\qquad - c_N(\tau_N(t))\exp[\alpha_n \mathbb{1}_N(t+1)] \\
&\qquad\qquad - \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right)\exp[\alpha_n \mathbb{1}_N(t+1)(1+B_n)]\Bigg\} \\
&\qquad - \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right)\alpha_n^k \mathbb{1}_N \frac{1}{k!}(t+1)^{k-1}(1+B_n')^k\Bigg\{ \\
&\qquad\qquad \left(2+\frac{\alpha_n^2(t+1)}{2}\right)\exp[\alpha_n \mathbb{1}_N(t+1)] \\
&\qquad\qquad + (t+1)\exp[\alpha_n \mathbb{1}_N(t+1)(1+B_n')]\Bigg\}.
\end{aligned}
$$

Due to (4.11)–(4.13), all but the first line on the RHS of the above have vanishing expectation, leaving

$$
\begin{aligned}
&\lim_{N\to\infty}\mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\ldots,r_k\}}}^{\tau_N(t)} (1-p_r)\right)\right] \\
&\qquad \ge \lim_{N\to\infty}\mathbb{E}\left[\alpha_n^k \mathbb{1}_N \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\mathbb{1}_{E_N^2}\sum_{l=0}^{\tau_N(t)}(-\alpha_n)^l \mathbb{1}_N \frac{1}{l!}t^l \mathbb{1}_{E_N^3}\right] \\
&\qquad = \alpha_n^k \sum_{l=0}^{\infty}(-\alpha_n)^l\frac{1}{l!}t^l \lim_{N\to\infty}\mathbb{E}\left[\mathbb{1}_{\{\tau_N(t)\ge l\}}\mathbb{1}_{E_N^2\cap E_N^3}\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\right].
\end{aligned} \tag{4.37}
$$

Passing the limit and expectation inside the infinite sum is justified by dominated convergence and Fubini; see Lemma 4.16. Lemmata 4.12 and 4.14 establish that $\lim_{N\to\infty}\mathbb{P}[E_N^2\cap E_N^3] = 1$ and Lemma 4.13 deals with the other indicator. We can therefore apply

Lemma 4.10 to conclude that

$$\lim_{N\to\infty} \mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\dots,r_k\}}}^{\tau_N(t)} (1-p_r)\right)\right]$$

$$\ge \alpha_n^k \sum_{l=0}^{\infty}(-\alpha_n)^l \frac{1}{l!}t^l \sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\dots,k\}:\\ i_j\ge j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}$$

$$= \alpha_n^k e^{-\alpha_n t} \sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\dots,k\}:\\ i_j\ge j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}$$

as required. ■

Lemma 4.10. *Assume (4.1) holds. Fix $k\in\mathbb{N}$, $i_0 := 0$, $i_k := k$. Let E_N be a sequence of events such that $\lim_{N\to\infty}\mathbb{P}[E_N]=1$. Then for any sequence of times $0=t_0\le t_1\le\cdots\le t_k\le t$,*

$$\lim_{N\to\infty} \mathbb{E}\left[\mathbb{1}_{E_N} \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\right] = \sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\dots,k\}:\\ i_j\ge j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}.$$

Proof. As pointed out by Möhle (1999, p. 460), the sum-product on the left hand side can be expanded as

$$\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i) = \sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\dots,k\}:\\ i_j\ge j\forall j}} \prod_{j=1}^{k} \frac{1}{(i_j-i_{j-1})!} \sum_{\substack{r_{i_{j-1}+1}\ne\cdots\ne r_{i_j}\\ =\tau_N(t_{j-1})+1}}^{\tau_N(t_j)} \prod_{i=i_{j-1}+1}^{i_j} c_N(r_i).$$

By a modification of the upper bound in Lemma 4.4(b) where the lower limit of the sum is a general time rather than 1,

$$\sum_{\substack{r_{i_{j-1}+1}\ne\cdots\ne r_{i_j}\\ =\tau_N(t_{j-1})+1}}^{\tau_N(t_j)} \prod_{i=i_{j-1}+1}^{i_j} c_N(r_i) \le (t_j-t_{j-1})^{i_j-i_{j-1}} + c_N(\tau_N(t_j))(t_j-t_{j-1}+1)^{i_j-i_{j-1}}$$

Now, taking the product on the outside,

$$\prod_{j=1}^{k} \frac{1}{(i_j - i_{j-1})!} \sum_{\substack{r_{i_{j-1}+1} \neq \cdots \neq r_{i_j} \\ = \tau_N(t_{j-1})+1}}^{\tau_N(t_j)} \prod_{i=i_{j-1}+1}^{i_j} c_N(r_i)$$

$$\begin{aligned}
&\leq \prod_{j=1}^{k} \left\{ \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} + c_N(\tau_N(t_j)) \frac{(t_j - t_{j-1} + 1)^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \right\} \\
&\leq \prod_{j=1}^{k} \left\{ \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} + c_N(\tau_N(t_j))(t_j - t_{j-1} + 1)^{i_j - i_{j-1}} \right\} \\
&= \sum_{\mathcal{I} \subseteq [k]} \left(\prod_{j \in \mathcal{I}} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \right) \left(\prod_{j \notin \mathcal{I}} c_N(\tau_N(t_j))(t_j - t_{j-1} + 1)^{i_j - i_{j-1}} \right) \\
&= \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \\
&\qquad + \sum_{\mathcal{I} \subset [k]} \left(\prod_{j \in \mathcal{I}} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \right) \left(\prod_{j \notin \mathcal{I}} c_N(\tau_N(t_j))(t_j - t_{j-1} + 1)^{i_j - i_{j-1}} \right) \\
&\leq \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \\
&\qquad + \sum_{\mathcal{I} \subset [k]} \left(\prod_{j \in \mathcal{I}} t^{i_j - i_{j-1}} \right) \left(\prod_{j \notin \mathcal{I}} c_N(\tau_N(t_j))(t+1)^{i_j - i_{j-1}} \right) \\
&\leq \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} + \sum_{\mathcal{I} \subset [k]} c_N(\tau_N(t_{j^\star(\mathcal{I})})) \prod_{j=1}^{k} (t+1)^{i_j - i_{j-1}} \\
&= \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} + \sum_{\mathcal{I} \subset [k]} c_N(\tau_N(t_{j^\star(\mathcal{I})}))(t+1)^{k}
\end{aligned}$$

where, say, $j^\star(\mathcal{I}) := \min\{j \notin \mathcal{I}\}$. Now we are in a position to evaluate the limit in

Lemma 4.10:

$$\begin{aligned}
\lim_{N\to\infty} \mathbb{E}\left[\mathbb{1}_{E_N} \sum_{\substack{r_1<\cdots<r_k:\\ r_i\leq\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\right] &\leq \lim_{N\to\infty} \mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\leq\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\right] \\
&\leq \sum_{\substack{i_1\leq\cdots\leq i_{k-1}\\ \in\{0,\dots,k\}:\\ i_j\geq j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!} + \sum_{\substack{i_1\leq\cdots\leq i_{k-1}\\ \in\{0,\dots,k\}:\\ i_j\geq j\forall j}} \sum_{\mathcal{I}\subset[k]} \lim_{N\to\infty} \mathbb{E}\left[c_N(\tau_N(t_{j^\star(\mathcal{I})}))\right](t+1)^k \\
&= \sum_{\substack{i_1\leq\cdots\leq i_{k-1}\\ \in\{0,\dots,k\}:\\ i_j\geq j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}
\end{aligned}$$

using (4.11)

For the corresponding lower bound, by a modification of the lower bound in Lemma 4.4(b) where the lower limit of the sum is a general time rather than 1,

$$\begin{aligned}
&\sum_{\substack{r_{i_{j-1}+1}\neq\cdots\neq r_{i_j}\\ =\tau_N(t_{j-1})+1}}^{\tau_N(t_j)} \prod_{i=i_{j-1}+1}^{i_j} c_N(r_i) \\
&\quad\geq (t_j-t_{j-1})^{i_j-i_{j-1}} - \binom{i_j-i_{j-1}}{2}\left(\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2\right)(t_j-t_{j-1}+1)^{i_j-i_{j-1}-2} \\
&\quad\geq (t_j-t_{j-1})^{i_j-i_{j-1}} - (i_j-i_{j-1})!\left(\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2\right)(t_j-t_{j-1}+1)^{i_j-i_{j-1}-2}.
\end{aligned}$$

Define the events

$$E_N^4(j) = \left\{\left(\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2\right) \leq \frac{1}{(i_j-i_{j-1})!}\left(\frac{t_j-t_{j-1}}{t_j-t_{j-1}+1}\right)^{i_j-i_{j-1}}\right\},$$

which is sufficient to ensure the j^{th} term in the following product is non-negative, and define $E_N^4 := \bigcap_{j=1}^{k} E_N^4(j)$. (If $t_j = t_{j-1}$ then $E_N^4(j)$ has probability one automatically; otherwise the constant on the right is strictly positive and so satisfies the conditions of

Lemma 4.15.) Now, taking a product over j,

$$
\begin{aligned}
&\prod_{j=1}^{k} \frac{1}{(i_j - i_{j-1})!} \sum_{\substack{r_{i_{j-1}+1} \neq \cdots \neq r_{i_j} \\ = \tau_N(t_{j-1})+1}}^{\tau_N(t_j)} \prod_{i=i_{j-1}+1}^{i_j} c_N(r_i) \\
&\quad \geq \prod_{j=1}^{k} \left\{ \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} - \left(\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2 \right) (t_j - t_{j-1} + 1)^{i_j - i_{j-1} - 2} \right\} \mathbb{1}_{E_N^4} \\
&\quad = \sum_{\mathcal{I} \subseteq [k]} (-1)^{k - |\mathcal{I}|} \left(\prod_{j \in \mathcal{I}} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \right) \\
&\qquad \times \left(\prod_{j \notin \mathcal{I}} \left(\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2 \right) (t_j - t_{j-1} + 1)^{i_j - i_{j-1} - 2} \right) \mathbb{1}_{E_N^4} \\
&\quad = \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \mathbb{1}_{E_N^4} \\
&\qquad + \sum_{\mathcal{I} \subset [k]} (-1)^{k - |\mathcal{I}|} \left(\prod_{j \in \mathcal{I}} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \right) \\
&\qquad \times \left(\prod_{j \notin \mathcal{I}} \left(\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2 \right) (t_j - t_{j-1} + 1)^{i_j - i_{j-1} - 2} \right) \mathbb{1}_{E_N^4} \\
&\quad \geq \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \mathbb{1}_{E_N^4} \\
&\qquad - \sum_{\mathcal{I} \subset [k]} \left(\prod_{j \in \mathcal{I}} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \right) \left(\prod_{j \notin \mathcal{I}} \left(\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2 \right) (t_j - t_{j-1} + 1)^{i_j - i_{j-1} - 2} \right) \\
&\quad \geq \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \mathbb{1}_{E_N^4} \\
&\qquad - \sum_{\mathcal{I} \subset [k]} \left(\prod_{j \in \mathcal{I}} t^{i_j - i_{j-1}} \right) \left(\prod_{j \notin \mathcal{I}} \left(\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2 \right) (t+1)^{i_j - i_{j-1} - 2} \right)
\end{aligned}
$$

$$
\begin{aligned}
&\geq \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \mathbb{1}_{E_N^4} \\
&\qquad - \sum_{\mathcal{I} \subset [k]} \left(\sum_{s=\tau_N(t_{j^\star(\mathcal{I})-1})+1}^{\tau_N(t_{j^\star(\mathcal{I})})} c_N(s)^2 \right) \left(\prod_{j \in \mathcal{I}} t^{i_j - i_{j-1}} \right) \left(\prod_{j \notin \mathcal{I}} (t+1)^{i_j - i_{j-1} - 1} \right) \\
&\geq \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \mathbb{1}_{E_N^4} - \sum_{\mathcal{I} \subset [k]} \left(\sum_{s=\tau_N(t_{j^\star(\mathcal{I})-1})+1}^{\tau_N(t_{j^\star(\mathcal{I})})} c_N(s)^2 \right) \prod_{j=1}^{k} (t+1)^{i_j - i_{j-1}} \\
&= \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \mathbb{1}_{E_N^4} - \sum_{\mathcal{I} \subset [k]} \left(\sum_{s=\tau_N(t_{j^\star(\mathcal{I})-1})+1}^{\tau_N(t_{j^\star(\mathcal{I})})} c_N(s)^2 \right) (t+1)^{k},
\end{aligned}
$$

where again we have arbitrarily set $j^\star(\mathcal{I}) := \min\{j \notin \mathcal{I}\}$. We can now evaluate the limit:

$$
\begin{aligned}
&\lim_{N\to\infty} \mathbb{E}\left[\mathbb{1}_{E_N} \sum_{\substack{r_1<\cdots<r_k:\\ r_i\leq\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\right] \\
&\quad\geq \lim_{N\to\infty} \mathbb{E}\left[\mathbb{1}_{E_N\cap E_N^4} \sum_{\substack{i_1\leq\cdots\leq i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\geq j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}\right] \\
&\qquad - \lim_{N\to\infty} \mathbb{E}\left[\mathbb{1}_{E_N} \sum_{\substack{i_1\leq\cdots\leq i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\geq j\forall j}} \sum_{\mathcal{I}\subset[k]} \left(\sum_{s=\tau_N(t_{j^\star(\mathcal{I})-1})+1}^{\tau_N(t_{j^\star(\mathcal{I})})} c_N(s)^2\right)(t+1)^k\right] \\
&\quad\geq \sum_{\substack{i_1\leq\cdots\leq i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\geq j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!} \lim_{N\to\infty} \mathbb{E}[\mathbb{1}_{E_N\cap E_N^4}] \\
&\qquad - \lim_{N\to\infty} \mathbb{E}\left[\sum_{\substack{i_1\leq\cdots\leq i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\geq j\forall j}} \sum_{\mathcal{I}\subset[k]} \left(\sum_{s=\tau_N(t_{j^\star(\mathcal{I})-1})+1}^{\tau_N(t_{j^\star(\mathcal{I})})} c_N(s)^2\right)(t+1)^k\right] \\
&\quad= \sum_{\substack{i_1\leq\cdots\leq i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\geq j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!} \lim_{N\to\infty} \mathbb{P}[E_N\cap E_N^4] \\
&\qquad - \sum_{\substack{i_1\leq\cdots\leq i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\geq j\forall j}} \sum_{\mathcal{I}\subset[k]} \lim_{N\to\infty} \mathbb{E}\left[\sum_{s=\tau_N(t_{j^\star(\mathcal{I})-1})+1}^{\tau_N(t_{j^\star(\mathcal{I})})} c_N(s)^2\right](t+1)^k \\
&\quad= \sum_{\substack{i_1\leq\cdots\leq i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\geq j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}
\end{aligned}
$$

where for the last equality we use (4.12) to show that the second sum vanishes and Lemma 4.15 to show that $\lim_{N\to\infty} \mathbb{P}[E_N \cap E_N^4] = 1$. We have shown that the upper and lower bounds coincide, so the result follows. ■

4.3 Indicators

Lemma 4.11. *Let* $(A_N), (B_N)$ *be sequences of events. If* $\lim_{N\to\infty} \mathbb{P}[A_N] = 1$ *and* $\lim_{N\to\infty} \mathbb{P}[B_N] = 1$ *then* $\lim_{N\to\infty} \mathbb{P}[A_N \cap B_N] = 1$.

The above might be so obvious as to go unstated, but it is very important because it means we don't have to deal with intersections of dependent events! Here is a little proof just to be sure:

Proof.

$$\begin{aligned}
&\lim_{N\to\infty} \mathbb{P}[A_N] = 1 \textbf{ and } \lim_{N\to\infty} \mathbb{P}[B_N] = 1 \\
\Leftrightarrow &\lim_{N\to\infty} \mathbb{P}[A_N^c] = 0 \textbf{ and } \lim_{N\to\infty} \mathbb{P}[B_N^c] = 0 \\
\Rightarrow &\lim_{N\to\infty} \{\mathbb{P}[A_N^c] + \mathbb{P}[B_N^c]\} = 0 \\
\Rightarrow &\lim_{N\to\infty} \mathbb{P}[A_N^c \cup B_N^c] = 0 \\
\Leftrightarrow &\lim_{N\to\infty} \mathbb{P}[A_N \cap B_N] = 1.
\end{aligned}$$

The only part of this argument that I find potentially controversial is going from the third to the fourth line, which is an application of the sandwich theorem (since $0 \leq \mathbb{P}[A_N^c \cup B_N^c] \leq \mathbb{P}[A_N^c] + \mathbb{P}[B_N^c]$). ■

Lemma 4.12. *Assume* (4.12) *holds. Let* $K > 0$ *be a constant which may depend on* n, N *but not on* r*, such that* $K^{-2} = O(1)$ *as* $N \to \infty$*. Define the events* $E_N(r) := \{c_N(r) < K\}$ *and denote* $E_N := \bigcap_{r=1}^{\tau_N(t)} E_N(r)$*. Then* $\lim_{N\to\infty} \mathbb{P}[E_N] = 1$.

Proof.

$$\begin{aligned}
\mathbb{P}[E_N] = 1 - \mathbb{P}[E_N^c] = 1 - \mathbb{P}\left[\bigcup_{r=1}^{\tau_N(t)} E_N^c(r)\right] &= 1 - \mathbb{E}\left[\mathbb{1}_{\bigcup E_N^c(r)}\right] \geq 1 - \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{1}_{E_N^c(r)}\right] \\
= 1 - \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{E}\left[\mathbb{1}_{E_N^c(r)} \mid \mathcal{F}_{r-1}\right]\right] &= 1 - \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{P}\left[E_N^c(r) \mid \mathcal{F}_{r-1}\right]\right]
\end{aligned} \tag{4.38}$$

where for the second line we apply Lemma 3.2 with $f(r) = \mathbb{1}_{E_N^c(r)}$. By the generalised Markov inequality,

$$\mathbb{P}[E_N^c(r) \mid \mathcal{F}_{r-1}] = \mathbb{P}[c_N(r) \geq K \mid \mathcal{F}_{r-1}] \leq K^{-2}\mathbb{E}[c_N(r)^2 \mid \mathcal{F}_{r-1}].$$

Substituting this into (4.38) and applying Lemma 3.2 again, this time with $f(r) = c_N(r)^2$,

$$\mathbb{P}[E_N] \geq 1 - K^{-2}\mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{E}[c_N(r)^2 \mid \mathcal{F}_{r-1}]\right] = 1 - K^{-2}\mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} c_N(r)^2\right].$$

Applying (4.12), the limit is

$$\lim_{N\to\infty} \mathbb{P}[E_N] = 1 - O(1) \times 0 = 1$$

as required. ■

Lemma 4.13. *Fix $t > 0$. For any $l \in \mathbb{N}$, $\lim_{N\to\infty} \mathbb{P}[\tau_N(t) \geq l] = 1$.*

Proof. We can replace the event $\{\tau_N(t) \geq l\}$ with an event of the form of E_N in Lemma 4.12:

$$\begin{aligned}\{\tau_N(t) \geq l\} &= \left\{\min\left\{s \geq 1 : \sum_{r=1}^{s} c_N(r) \geq t\right\} \geq l\right\} = \left\{\sum_{r=1}^{l-1} c_N(r) < t\right\} \\ &\supseteq \bigcap_{r=1}^{l-1} \left\{c_N(r) < \frac{t}{l}\right\} \supseteq \bigcap_{r=1}^{\tau_N(t)} \left\{c_N(r) < \frac{t}{l}\right\}.\end{aligned}$$

Hence

$$\lim_{N\to\infty} \mathbb{P}[\tau_N(t) \geq l] \geq \lim_{N\to\infty} \mathbb{P}\left[\bigcap_{r=1}^{\tau_N(t)} \left\{c_N(r) < \frac{t}{l}\right\}\right] = 1$$

by applying Lemma 4.12 with $K = t/l$. ■

Lemma 4.14. *Assume* (4.13) *holds. Fix $t > 0$. Let K be a constant not depending on N, r, but which may depend on n.*

$$\lim_{N\to\infty} \mathbb{P}\left[\bigcap_{r=1}^{\tau_N(t)} \{c_N(r) \geq K D_N(r)\}\right] = 1.$$

Proof.

$$\begin{aligned}\mathbb{P}\left[\bigcap_{r=1}^{\tau_N(t)} \{c_N(r) \geq K D_N(r)\}\right] &\geq \mathbb{P}\left[\bigcap_{r=1}^{\tau_N(t)} \{c_N(r) > K D_N(r)\}\right] \\ &= 1 - \mathbb{P}\left[\bigcup_{r=1}^{\tau_N(t)} \{c_N(r) \leq K D_N(r)\}\right] \\ &= 1 - \mathbb{E}\left[\mathbb{1}_{\bigcup\{c_N(r) \leq K D_N(r)\}}\right] \\ &\geq 1 - \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{1}_{\{c_N(r) \leq K D_N(r)\}}\right] \\ &= 1 - \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{P}[c_N(r) \leq K D_N(r) \mid \mathcal{F}_{r-1}]\right] \qquad (4.39)\end{aligned}$$

where the final inequality is an application of Lemma 3.2 with $f(r) = \mathbb{1}_{\{c_N(r)\le KD_N(r)\}}$.

Fix $0<\varepsilon<K^{-1}/2$ and assume $N>\max\{\varepsilon^{-1},(K^{-1}-2\varepsilon)^{-1}\}$. For each r,i define the event $A_i(r) := \{\nu_r^{(i)} \le N\varepsilon\}$. Conditional on $\mathcal{F}_{r-1}$, we have

$$D_N(r) = \frac{1}{N(N)_2}\sum_{i=1}^N (\nu_r^{(j)})_2 \left[\nu_r^{(i)} + \frac{1}{N}\sum_{j\neq i}(\nu_r^{(i)})^2\right]\mathbb{1}_{A_i^c(r)}$$
$$+ \frac{1}{N(N)_2}\sum_{i=1}^N (\nu_r^{(i)})_2 \left[\nu_r^{(i)} + \frac{1}{N}\sum_{j\neq i}(\nu_r^{(j)})^2\right]\mathbb{1}_{A_i(r)}.$$

For the first term,

$$\frac{1}{N(N)_2}\sum_{i=1}^N (\nu_r^{(i)})_2 \left[\nu_r^{(i)} + \frac{1}{N}\sum_{j\neq i}(\nu_r^{(j)})^2\right]\mathbb{1}_{A_i^c(r)} \le \sum_{i=1}^N \mathbb{1}_{A_i^c(r)}.$$

For the second term,

$$\begin{aligned}
&\frac{1}{N(N)_2}\sum_{i=1}^N (\nu_r^{(i)})_2 \left[\nu_r^{(i)} + \frac{1}{N}\sum_{j\neq i}(\nu_r^{(j)})^2\right]\mathbb{1}_{A_i(r)}\\
&\qquad \le \frac{1}{N(N)_2}\sum_{i=1}^N (\nu_r^{(i)})_2\nu_r^{(i)}\mathbb{1}_{A_i(r)} + \frac{1}{N^2(N)_2}\sum_{i=1}^N(\nu_r^{(i)})_2\sum_{j=1}^N(\nu_r^{(j)})^2\mathbb{1}_{A_i(r)}\\
&\qquad \le \frac{1}{N}c_N(r)N\varepsilon + \frac{1}{N^2(N)_2}\sum_{i=1}^N(\nu_r^{(i)})_2\sum_{j=1}^N(\nu_r^{(j)})_2\mathbb{1}_{A_i(r)}\\
&\qquad\quad + \frac{1}{N^2(N)_2}\sum_{i=1}^N(\nu_r^{(i)})_2\sum_{j=1}^N(\nu_r^{(j)})\mathbb{1}_{A_i(r)}\\
&\qquad \le \varepsilon c_N(r) + \frac{1}{N^2}\sum_{i=1}^N \nu_r^{(i)} N\varepsilon c_N(r) + \frac{1}{N^2}c_N(r)N\\
&\qquad = c_N(r)\left(2\varepsilon + \frac{1}{N}\right).
\end{aligned}$$

Altogether we have

$$D_N(r) \le c_N(r)\left(2\varepsilon+\frac{1}{N}\right) + \sum_{i=1}^N \mathbb{1}_{A_i^c(r)}.$$

Hence, still conditional on $\mathcal{F}_{r-1}$,

$$\begin{aligned}
\{c_N(r)\le KD_N(r)\} &\subseteq \left\{c_N(r)\le Kc_N(r)(2\varepsilon+N^{-1}) + K\sum_{i=1}^N \mathbb{1}_{A_i^c(r)}\right\}\\
&= \left\{K^{-1} - 2\varepsilon - \frac{1}{N} \le \sum_{i=1}^N \frac{\mathbb{1}_{A_i^c(r)}}{c_N(r)}\right\}
\end{aligned}$$

where the ratio $\mathbb{1}_{A_i^c(r)}/c_N(r)$ is well-defined because

$$A_i^c(r) \Rightarrow c_N(r) := \frac{1}{(N)_2}\sum_{j=1}^{N}(\nu_r^{(j)})_2 \geq \frac{1}{(N)_2}(\nu_r^{(i)})_2 \geq \frac{\varepsilon(N\varepsilon-1)}{N-1} \geq \varepsilon\left(\varepsilon-\frac{1}{N}\right) > 0.$$

Hence by Markov's inequality (the conditions on ε, N ensuring the constant is always strictly positive),

$$\begin{aligned}
\mathbb{P}\left[c_N(r) \leq KD_N(r) \mid \mathcal{F}_{r-1}\right] &\leq \mathbb{P}\left[\sum_{i=1}^{N}\mathbb{1}_{A_i^c(r)} \geq \left(K^{-1}-2\varepsilon-\frac{1}{N}\right)\varepsilon\left(\varepsilon-\frac{1}{N}\right)\middle|\mathcal{F}_{r-1}\right] \\
&\leq \frac{1}{\left(K^{-1}-2\varepsilon-\frac{1}{N}\right)\varepsilon\left(\varepsilon-\frac{1}{N}\right)}\mathbb{E}\left[\sum_{i=1}^{N}\mathbb{1}_{A_i^c(r)}\middle|\mathcal{F}_{r-1}\right] \\
&\leq \frac{1}{\left(K^{-1}-2\varepsilon-\frac{1}{N}\right)\varepsilon\left(\varepsilon-\frac{1}{N}\right)}\mathbb{E}\left[\sum_{i=1}^{N}\frac{(\nu_r^{(i)})_3}{(N\varepsilon)_3}\middle|\mathcal{F}_{r-1}\right] \\
&\leq \frac{1}{\left(K^{-1}-2\varepsilon-\frac{1}{N}\right)\varepsilon\left(\varepsilon-\frac{1}{N}\right)}\mathbb{E}\left[\frac{N(N)_2}{(N\varepsilon)_3}D_N(r)\middle|\mathcal{F}_{r-1}\right].
\end{aligned}$$

Applying Lemma 3.2 once more, with $f(r) = D_N(r)$,

$$\begin{aligned}
&\mathbb{E}\left[\sum_{r=1}^{\tau_N(t)}\mathbb{P}[c_N(r) \leq KD_N(r) \mid \mathcal{F}_{r-1}]\right] \\
&\qquad\leq \frac{1}{\left(K^{-1}-2\varepsilon-\frac{1}{N}\right)\varepsilon\left(\varepsilon-\frac{1}{N}\right)}\frac{N(N)_2}{(N\varepsilon)_3}\mathbb{E}\left[\sum_{r=1}^{\tau_N(t)}\mathbb{E}[D_N(r) \mid \mathcal{F}_{r-1}]\right] \\
&\qquad= \frac{1}{\left(K^{-1}-2\varepsilon-\frac{1}{N}\right)\varepsilon\left(\varepsilon-\frac{1}{N}\right)}\frac{N(N)_2}{(N\varepsilon)_3}\mathbb{E}\left[\sum_{r=1}^{\tau_N(t)}D_N(r)\right] \\
&\qquad\xrightarrow[N\to\infty]{} \frac{1}{(K^{-1}-2\varepsilon)\varepsilon^5} \times 0 = 0
\end{aligned}$$

due to (4.13). Substituting this back into (4.39) concludes the proof. ∎

Lemma 4.15. *Assume* (4.12) *holds. Fix* $k \in \mathbb{N}$, *a sequence of times* $0 = t_0 \leq t_1 \leq \cdots \leq t_k \leq t$, *and let* $K_1, \ldots, K_k$ *be strictly positive constants. Define the events*

$$E_N := \bigcap_{j=1}^{k}\left\{\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2 \leq K_j\right\}.$$

Then $\lim_{N\to\infty}\mathbb{P}[E_N] = 1$.

Proof.

$$\mathbb{P}[E_N] = 1 - \mathbb{P}[E_N^c] = 1 - \mathbb{P}\left[\bigcup_{j=1}^{k}\left\{\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2 > K_j\right\}\right]$$
$$\geq 1 - \sum_{j=1}^{k}\mathbb{P}\left[\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2 \geq K_j\right].$$

Applying Markov's inequality,

$$\mathbb{P}[E_N] \geq 1 - \sum_{j=1}^{k} K_j^{-1}\mathbb{E}\left[\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2\right] \underset{N\to\infty}{\longrightarrow} 1 - \sum_{j=1}^{k} O(1) \times 0 = 1$$

by (4.12). ∎

4.4 Fubini & dominated convergence conditions

There are a few instances where Fubini's Theorem and the Dominated Convergence Theorem are needed in order to pass a limit and expectation through an infinite sum. Now we verify that the conditions of these theorems indeed hold. This result, analogous to that in Koskela et al. (2018, Appendix), is used once in Lemma 4.8 at (4.31) and once in Lemma 4.9 at (4.37).

Lemma 4.16. *For any fixed $t > 0$,*

$$\mathbb{E}\left[\sum_{l=0}^{\infty}\left|(-\alpha_n)^l 1_N \frac{1}{l!}t^l \sum_{\substack{r_1<\cdots<r_k:\\ r_i\leq\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\right|\right] < \infty.$$

Proof.

$$\mathbb{E}\left[\sum_{l=0}^{\infty}\left|(-\alpha_n)^l 1_N \frac{1}{l!}t^l \sum_{\substack{r_1<\cdots<r_k:\\ r_i\leq\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\right|\right] \leq \mathbb{E}\left[\sum_{l=0}^{\infty}\alpha_n^l 1_N \frac{1}{l!}t^l (t+1)^k\right]$$
$$= \mathbb{E}[\exp\{\alpha_n t 1_N\}(t+1)^k] = \exp\{\alpha_n t 1_N\}(t+1)^k < \infty.$$

∎

5 Applications

Theorem 4.1 gives verifiable conditions under which interacting particle systems with dynamics in the form of Algorithm 1 have asymptotically Kingman genealogies. The work was motivated by SMC algorithms, which have the required form. However, certain choices of state space and dynamics within the context of Algorithm 1 yield systems that are not very SMC-like but may have applications in other fields such as population genetics. For instance, we have generally imagined that the resampling scheme is unbiased, but this is by no means necessary for Theorem 4.1 (or indeed Theorem 3.6); it is just that biased resampling schemes are of little use in SMC.

The applications presented in this chapter are all motivated by SMC, but an interesting area of future research would be to explore the implications of Theorem 4.1 in other contexts. From the population genetics point-of-view, Theorem 4.1 may be seen as a complement to the convergence criteria for neutral models (e.g. Möhle 1999) discussed in Section ?? add the section reference once that part of Chapter 2 is written, so it would be interesting to construct some corollaries for classical non-neutral population models.

5.1 Multinomial resampling

Multinomial resampling is often preferred in theoretical studies of SMC, because it renders the parental indices conditionally i.i.d. given the weights, making it relatively simple to analyse the resulting algorithm. The convergence of finite-dimensional distributions for multinomial resampling was proved in Koskela et al. (2018, Corollary 1), but we are now able to prove an analogous weak convergence result. The following proof also demonstrates the relative ease with which we can verify Theorem 3.6 as opposed to Koskela et al. (2018, Theorem 1).

Corollary 5.1. *Consider an SMC algorithm using multinomial resampling, such that (A1) is satisfied. Assume there exist constants $\varepsilon \in (0,1], a \in [1,\infty)$ and probability density h such that for all x, x', t,*

$$\frac{1}{a} \leq g_t(x,x') \leq a, \quad \varepsilon h(x') \leq q_t(x,x') \leq \frac{1}{\varepsilon} h(x'). \tag{5.1}$$

Let $(G_t^{(n,N)})_{t\geq 0}$ denote the genealogy of a random sample of n terminal particles from the output of the algorithm among a total of N particles. Then, for any fixed n, the time-scaled genealogy $(G_{\tau_N(t)}^{(n,N)})_{t\geq 0}$ converges weakly to Kingman's n-coalescent as $N \to \infty$.

The bounds on g_t and q_t in (5.1) are rather strong; they can only reasonably be expected to hold if the state space is compact. However, they are widespread in the literature, where they are known as the *strong mixing conditions* (Del Moral 2004, Section 3.5.2), because they greatly facilitate the theoretical analysis of SMC algorithms. It is often possible to relax these conditions at the expense of considerable technical complication. The conditions on g_t in (5.1) ensure that the weights are all $O(N^{-1})$, none of them being too close to zero or one. Together with the bounds on q_t, this is enough to control the relative rate of multiple mergers, as seen in the following proof.

Proof. Recall that the sequence of σ-algebras

$$\mathcal{H}_t := \sigma(X_{t-1}^{(1:N)}, X_t^{(1:N)}, w_{t-1}^{(1:N)}, w_t^{(1:N)}) \tag{5.2}$$

are such that $\nu_t^{(1:N)}$ is conditionally independent of the filtration $\mathcal{F}_{t-1}$ given $\mathcal{H}_t$. Conditional on $\mathcal{H}_t$ the parental indices are independent, with conditional law

$$\mathbb{P}\left[a_t^{(i)} = a_i \mid \mathcal{H}_t\right] \propto g_t(X_{t+1}^{a_{t+1}^{(a_i)}}, X_t^{(a_i)}) q_{t-1}(X_t^{(a_i)}, X_{t-1}^{(i)}) \tag{5.3}$$

for each i, so the joint law is

$$\mathbb{P}\left[a_t^{(1:N)} = a_{1:N} \mid \mathcal{H}_t\right] \propto \prod_{i=1}^{N} g_t(X_{t+1}^{a_{t+1}^{(a_i)}}, X_t^{(a_i)}) q_{t-1}(X_t^{(a_i)}, X_{t-1}^{(i)}).$$

Using the bounds (5.1) and the balls-in-bins coupling of Koskela et al. (2018, Proof of Lemma 3), we can obtain bounds on expectations of functions of $a_t^{(1:N)}$. For any $k \in \mathbb{N}$ the function $a_t^{(1:N)} \to (\nu_t^{(i)})_k$ is $\{i\}$-increasing in the sense of Koskela et al. (2018), so we may apply the bounds

$$\mathbb{E}[(V_1^{(i)})_k] \leq \mathbb{E}[(\nu_t^{(i)})_k \mid \mathcal{H}_t] \leq \mathbb{E}[(V_2^{(i)})_k],$$

where

$$V_1^{(i)} \sim \text{Binomial}\left(N, \frac{\varepsilon/a}{(\varepsilon/a)+(N-1)(a/\varepsilon)}\right),$$
$$V_2^{(i)} \sim \text{Binomial}\left(N, \frac{a/\varepsilon}{(a/\varepsilon)+(N-1)(\varepsilon/a)}\right).$$

independently for each i and independently of $\mathcal{F}_\infty$. Furthermore, using the moments of the Binomial distribution (see for example Mosimann 1962, p. 67)

$$\mathbb{E}[(V_1^{(i)})_k] = (N)_k \left(\frac{\varepsilon/a}{(\varepsilon/a)+(N-1)(a/\varepsilon)}\right)^k \geq (N)_k \left(\frac{\varepsilon/a}{N(a/\varepsilon)}\right)^k = \frac{(N)_k}{N^k}\frac{\varepsilon^{2k}}{a^{2k}}.$$

Similarly,

$$\mathbb{E}[(V_2^{(i)})_k] \leq \frac{(N)_k}{N^k}\frac{a^{2k}}{\varepsilon^{2k}}.$$

We therefore have the bounds

$$\frac{(N)_k}{N^k}\frac{\varepsilon^{2k}}{a^{2k}} \leq \mathbb{E}[(\nu_t^{(i)})_k \mid \mathcal{H}_t] \leq \frac{(N)_k}{N^k}\frac{a^{2k}}{\varepsilon^{2k}}.$$

for each k. Consequently,

$$\frac{1}{(N)_2}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_2 \mid \mathcal{H}_t] \geq \frac{\varepsilon^4}{Na^4} \tag{5.4}$$

and

$$\frac{1}{(N)_3}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_3 \mid \mathcal{H}_t] \leq \frac{a^6}{N^2\varepsilon^6}. \tag{5.5}$$

The definition of $\mathcal{H}_t$ is such that, for any suitable function f, by the tower property and conditional independence we have

$$\mathbb{E}_t[f(\nu_t^{(1:N)})] = \mathbb{E}_t\left[\mathbb{E}[f(\nu_t^{(1:N)}) \mid \mathcal{H}_t, \mathcal{F}_{t-1}]\right] = \mathbb{E}_t\left[\mathbb{E}[f(\nu_t^{(1:N)}) \mid \mathcal{H}_t]\right]. \tag{5.6}$$

Applying this identity to (5.4) and (5.5) we find

$$\frac{\frac{1}{(N)_3}\sum_{i=1}^N \mathbb{E}_t[(\nu_t^{(i)})_3]}{\frac{1}{(N)_2}\sum_{i=1}^N \mathbb{E}_t[(\nu_t^{(i)})_2]} \leq \frac{a^6/(N^2\varepsilon^6)}{\varepsilon^4/(Na^4)} = \frac{a^{10}}{N\varepsilon^{10}} =: b_N \underset{N\to\infty}{\longrightarrow} 0.$$

Thus (3.11) is satisfied. It remains to show that, for N sufficiently large, $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all finite t, a technicality which is proved in Lemma 5.2. Applying Theorem 4.1 then yields the result. ■

Lemma 5.2. *Consider an SMC algorithm using multinomial resampling, satisfying (A1) and* (5.1)*. Then, for all* $N > 2$, $\mathbb{P}[\tau_N(t) = \infty] = 0$ *for all finite* t.

Proof. Since $c_N(t) \in [0,1]$ almost surely and has strictly positive expectation, for any fixed N the distribution of $c_N(t)$ with given expectation that maximises $\mathbb{P}[c_N(t) = 0 \mid \mathcal{F}_{t-1}]$ is two atoms, at 0 and 1 respectively. To ensure the correct expectation, the atom at 1 should have mass $\mathbb{P}[c_N(t) = 1 \mid \mathcal{F}_{t-1}] = \mathbb{E}_t[c_N(t)]$, which is bounded below by (5.4). If $c_N(t) > 0$ then $c_N(t) \geq 2/(N)_2 > 2/N^2$. Hence, in general $\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{F}_{t-1}] \geq \mathbb{E}_t[c_N(t)]$ the above explanation could be a bit more verbose/explicit. Applying (5.4) along with (5.6), we have for any finite N

$$\sum_{t=0}^{\infty} \mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{F}_{t-1}] \geq \sum_{t=0}^{\infty} \mathbb{E}_t[c_N(t)] \geq \sum_{t=0}^{\infty} \frac{\varepsilon^4}{Na^4} = \infty.$$

By a filtered version of the second Borel–Cantelli lemma (see for example Durrett 2019, Theorem 4.3.4), this implies that $c_N(t) > 2/N^2$ for infinitely many t, almost surely. This ensures, for all $t < \infty$, that $\mathbb{P}\left[\exists s < \infty : \sum_{r=1}^{s} c_N(r) \geq t\right] = 1$, which by definition of $\tau_N(t)$ is equivalent to $\mathbb{P}[\tau_N(t) = \infty] = 0$. ■

5.2 Stratified resampling

Corollary 5.3. *Consider an SMC algorithm using stratified resampling, such that (A1) is satisfied. Assume that there exists a constant $a \in [1, \infty)$ such that for all x, x', t,*

$$\frac{1}{a} \leq g_t(x, x') \leq a. \tag{5.7}$$

Assume that $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all finite t. Let $(G_t^{(n,N)})_{t \geq 0}$ denote the genealogy of a random sample of n terminal particles from the output of the algorithm among a total of N particles. Then, for any fixed n, the time-scaled genealogy $(G_{\tau_N(t)}^{(n,N)})_{t \geq 0}$ converges weakly to Kingman's n-coalescent as $N \to \infty$.

Stratified resampling is, by design, much more restrictive than multinomial resampling. Once the weights are known there is little freedom in the offspring counts, so it is not surprising that control over the weights such as (5.7) provides is sufficient without any additional control over the transition densities q_t. This is in contrast to multinomial resampling (Corollary 5.1), where g_t and q_t are more or less on an equal footing, and we require both to be bounded.

It is not immediately clear that the finite time scale condition $\mathbb{P}[\tau_N(t) = \infty] = 0$ holds under conditions (5.7), so it is included in the statement of the corollary. Proposition 5.6 presents some sufficient conditions for the finite time scale, but these are by no means necessary.

By the way, does the lack of conditions of q_t here imply that we do not even need the transition kernels to admit densities?

Proof. Define the σ-algebras $\mathcal{H}_t$ as in (5.2). With stratified resampling, conditional on the weights each offspring count almost surely takes one of four values: $\nu_t^{(i)} \in \{\lfloor Nw_t^{(i)} \rfloor - 1, \lfloor Nw_t^{(i)} \rfloor, \lfloor Nw_t^{(i)} \rfloor + 1, \lfloor Nw_t^{(i)} \rfloor + 2\}$. Define for each $k \in \mathbb{Z}$

$$p_k^{(i)} := \mathbb{P}\left[\nu_t^{(i)} = \lfloor Nw_t^{(i)} \rfloor + k \,\middle|\, \mathcal{H}_t\right]. \tag{5.8}$$

Then $p_k^{(i)} \equiv 0$ for $k \notin \{-1, 0, 1, 2\}$. Now

$$\begin{aligned}\mathbb{E}[(\nu_t^{(i)})_2 \mid \mathcal{H}_t] = p_{-1}^{(i)}(\lfloor Nw_t^{(i)} \rfloor - 1)_2 + p_0^{(i)}(\lfloor Nw_t^{(i)} \rfloor)_2 + p_1^{(i)}(\lfloor Nw_t^{(i)} \rfloor + 1)_2 \\ + p_2^{(i)}(\lfloor Nw_t^{(i)} \rfloor + 2)_2\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}[(\nu_t^{(i)})_3 \mid \mathcal{H}_t] &= p_{-1}^{(i)}(\lfloor Nw_t^{(i)} \rfloor - 1)_3 + p_0^{(i)}(\lfloor Nw_t^{(i)} \rfloor)_3 + p_1^{(i)}(\lfloor Nw_t^{(i)} \rfloor + 1)_3 \\ &\qquad + p_2^{(i)}(\lfloor Nw_t^{(i)} \rfloor + 2)_3 \\ &= p_{-1}^{(i)}(\lfloor Nw_t^{(i)} \rfloor - 3)(\lfloor Nw_t^{(i)} \rfloor - 1)_2 + p_0^{(i)}(\lfloor Nw_t^{(i)} \rfloor - 2)(\lfloor Nw_t^{(i)} \rfloor)_2 \\ &\qquad + p_1^{(i)}(\lfloor Nw_t^{(i)} \rfloor - 1)(\lfloor Nw_t^{(i)} \rfloor + 1)_2 + p_2^{(i)}\lfloor Nw_t^{(i)} \rfloor(\lfloor Nw_t^{(i)} \rfloor + 2)_2 \\ &\leq \lfloor Nw_t^{(i)} \rfloor \Big\{ p_{-1}^{(i)}(\lfloor Nw_t^{(i)} \rfloor - 1)_2 + p_0^{(i)}(\lfloor Nw_t^{(i)} \rfloor)_2 + p_1^{(i)}(\lfloor Nw_t^{(i)} \rfloor + 1)_2 \\ &\qquad + p_2^{(i)}(\lfloor Nw_t^{(i)} \rfloor + 2)_2 \Big\} \\ &= \lfloor Nw_t^{(i)} \rfloor \mathbb{E}[(\nu_t^{(i)})_2 \mid \mathcal{H}_t] \\ &\leq a^2 \mathbb{E}[(\nu_t^{(i)})_2 \mid \mathcal{H}_t]. \end{aligned} \tag{5.9}$$

The last line uses the almost sure bound $w_t^{(i)} \leq a^2/N$ which follows from (5.7) along with the form of the weights in Algorithm 1. Note that some terms in the above expressions may be equal to zero when $w_t^{(i)}$ is small enough, but the bound still holds in these cases. Since (5.9) holds for all i, applying the tower rule we have

$$\frac{1}{(N)_3} \sum_{i=1}^{N} \mathbb{E}_t[(\nu_t^{(i)})_3] \leq \frac{a^2}{N-2} \frac{1}{(N)_2} \sum_{i=1}^{N} \mathbb{E}_t[(\nu_t^{(i)})_2],$$

satisfying (3.11) with $b_N := a^2/(N-2) \to 0$. The result follows by applying Theorem 4.1. ■

Proposition 5.4. *Consider an SMC algorithm using stratified resampling. Suppose that for each t? — actually, isn't it sufficient that these bounds exist for infinitely many t? there exists a constant $\varepsilon \in (0, 1]$ and a probability density h such that*

$$\varepsilon h(x') \leq q_t(x, x') \leq \varepsilon^{-1} h(x')$$

uniformly in x, and that there exist $\zeta > 0$ and $\delta > 0$ such that

$$\mathbb{P}[\max_i w_t^{(i)} - \min_i w_t^{(i)} \geq 2\delta/N \mid \mathcal{F}_{t-1}] \geq \zeta \tag{5.10}$$

for infinitely many t. Then, for all $N > 1$, $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all finite t.

We now assume q_t is bounded above and away from zero, as in (5.1). We saw that such a condition was not necessary for Corollary 5.3, and we do not believe it to be necessary here either; it is merely a convenient way to control the contributions from the transition density. The bounds established in the following proof are rather crude, particularly the terms in ε; it may well be possible to achieve similar bounds under less restrictive conditions.

The second condition (5.10) is required to ensure that, at least infinitely often, the weights are not equal to $(1, \ldots, 1)/N$, since stratified resampling is degenerate under equal weights, which could cause the time scale to explode. It is hardly conceivable that any real SMC algorithm would fail to satisfy this very mild condition, which effectively ensures that the weights cannot be "too well-behaved".

Proof. As argued in Lemma 5.2, it is sufficient to prove that under the stated conditions

$$\sum_{r=0}^{\infty} \mathbb{P}[c_N(r) > 2/N^2 \mid \mathcal{F}_{r-1}] = \infty.$$

Firstly,

$$\begin{aligned}\mathbb{P}[c_N(t) \leq 2/N^2 \mid \mathcal{H}_t] = \mathbb{P}[c_N(t) = 0 \mid \mathcal{H}_t] &= \mathbb{P}[\nu_t^{(i)} = 1\, \forall i \in \{1, \ldots, N\} \mid \mathcal{H}_t] \\ &\leq \mathbb{P}[\nu_t^{(i^\star)} = 1 \mid \mathcal{H}_t], \end{aligned} \tag{5.11}$$

where $i^\star := \operatorname{argmax}_i\{w_t^{(i)}\}$ (but note that the inequality holds when $i^\star$ is taken to be any particular index). Define $p_k^{(i)}$ as in (5.8) and recall that, under stratified resampling, $p_k^{(i)} \equiv 0$ for $k \notin \{-1, 0, 1, 2\}$ and

$$\sum_{k=-1}^{2} p_k^{(i)} = \sum_{k=-1}^{2} \mathbb{P}\left[\nu_t^{(i)} = \lfloor N w_t^{(i)} \rfloor + k \,\middle|\, w_t^{(1:N)}\right] = 1.$$

Up to a proportionality constant C,

$$p_k^{(i)} = C\,\mathbb{P}\left[\nu_t^{(i)} = \lfloor N w_t^{(i)} \rfloor + k \,\middle|\, w_t^{(1:N)}\right]$$
$$\times \sum_{\substack{a_{1:N} \in \{1,\dots,N\}^N: \\ |\{j: a_j = i\}| = \lfloor N w_t^{(i)} \rfloor + k}} \mathbb{P}\left[a_t^{(1:N)} = a_{1:N} \,\middle|\, \nu_t^{(i)}, w_t^{(1:N)}\right] \prod_{j=1}^N q_{t-1}(X_t^{(a_j)}, X_{t-1}^{(j)})$$

for each $k \in \{-1, 0, 1, 2\}$. We can bound each probability above and below using the almost sure bounds on q_{t-1} from the statement of the Proposition (once the bounds on q_{t-1} are brought outside, the remaining sum of probabilities is equal to one):

$$p_k^{(i)} \geq C\,\mathbb{P}\left[\nu_t^{(i)} = \lfloor N w_t^{(i)} \rfloor + k \,\middle|\, w_t^{(1:N)}\right] \varepsilon^N \prod_{j=1}^N h(X_{t-1}^{(j)}),$$
$$p_k^{(i)} \leq C\,\mathbb{P}\left[\nu_t^{(i)} = \lfloor N w_t^{(i)} \rfloor + k \,\middle|\, w_t^{(1:N)}\right] \varepsilon^{-N} \prod_{j=1}^N h(X_{t-1}^{(j)}).$$

We then eliminate the proportionality constant C by normalising, to obtain lower bounds

$$\begin{aligned} p_k^{(i)} &\geq \frac{C\,\mathbb{P}[\nu_t^{(i)} = \lfloor N w_t^{(i)} \rfloor + k \mid w_t^{(1:N)}] \varepsilon^N \prod_{j=1}^N h(X_{t-1}^{(j)})}{\sum_{j=-1}^2 C\,\mathbb{P}[\nu_t^{(i)} = \lfloor N w_t^{(i)} \rfloor + j \mid w_t^{(1:N)}] \varepsilon^{-N} \prod_{j=1}^N h(X_{t-1}^{(j)})} \\ &= \mathbb{P}[\nu_t^{(i)} = \lfloor N w_t^{(i)} \rfloor + k \mid w_t^{(1:N)}] \varepsilon^{2N} \end{aligned} \tag{5.12}$$

for each k, which also imply

$$1 - p_k^{(i)} \geq \left(1 - \mathbb{P}[\nu_t^{(i)} = \lfloor N w_t^{(i)} \rfloor + k \mid w_t^{(1:N)}]\right) \varepsilon^{2N}. \tag{5.13}$$

Suppose that $\max_i w_t^{(i)} - \min_i w_t^{(i)} \geq 2\delta/N$. Then that at least one of $\{\max_i w_t^{(i)} \geq (1+\delta)/N\}$ and $\{\min_i w_t^{(i)} \leq (1-\delta)/N\}$ occurs. We will now examine each of these possibilities.

We can always write the maximum weight as $w_t^{(i^\star)} = \frac{1+\gamma}{N}$ for some $\gamma \geq 0$. Then, using (5.11),

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \geq 1 - \mathbb{P}[\nu_t^{(i^\star)} = 1 \mid \mathcal{H}_t] = \begin{cases} 0 & \text{if } \gamma = 0 \\ 1 - p_0^{(i^\star)} & \text{if } \gamma \in (0,1) \\ 1 - p_{-1}^{(i^\star)} & \text{if } \gamma \in [1,2) \\ 1 & \text{if } \gamma \geq 2. \end{cases}$$

If $\gamma \in (0,1)$ then the "overhang" in the sense of Figure 2.7 is γ, and

$$1 - p_0^{(i^\star)} \geq \frac{3\gamma}{4} \varepsilon^{2N}$$

using Table 2.2 (upper bound on p_0) and (5.13). Similarly, if $\gamma \in [1,2)$ then the overhang is $\gamma - 1$ and by Table 2.2 (upper bound on p_{-1}),

$$1 - p_{-1}^{(i^\star)} \geq \left(1 - \frac{1}{4}\right) \varepsilon^{2N} \geq \frac{3}{4}\varepsilon^{2N}.$$

Overall, under the constraint $\max_i w_t^{(i)} \geq (1+\delta)/N$, we have

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \geq \min_{\gamma \geq \delta} \left\{ \frac{3\gamma}{4}\varepsilon^{2N} \mathbb{1}_{\{\gamma \in [0,1)\}} + \frac{3}{4}\varepsilon^{2N} \mathbb{1}_{\{\gamma \in [1,2)\}} + \mathbb{1}_{\{\gamma \geq 2\}} \right\} = \frac{3}{4}\delta\varepsilon^{2N}.$$

We now construct a similar argument for the minimum weight. Let $j^\star := \operatorname{argmin}_i\{w_t^{(i)}\}$ and write $w_t^{(j^\star)} = \frac{1-\gamma}{N}$, for some $\gamma \in [0,1]$. Then by (5.11) we have

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \geq 1 - \mathbb{P}[\nu_t^{(j^\star)} = 1 \mid \mathcal{H}_t] = \begin{cases} 1 - p_1^{(j^\star)} & \text{if } \gamma \in (0,1] \\ 0 & \text{if } \gamma = 0. \end{cases}$$

If $\gamma \in (0,1]$ then the "overhang" in the sense of Figure 2.7 is $1-\gamma$, and

$$1 - p_1^{(j^\star)} \geq \left(1 - \frac{1 + (1-\gamma)}{2}\right) \varepsilon^{2N} = \frac{\gamma}{2}\varepsilon^{2N},$$

using Table 2.2 (upper bound on p_1). Therefore, under the constraint $\min_i w_t^{(i)} \leq (1-\delta)/N$, we have

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \geq \min_{\gamma \geq \delta} \left\{ \frac{\gamma}{2}\varepsilon^{2N} \right\} = \frac{1}{2}\delta\varepsilon^{2N}.$$

Combining both cases, we find for arbitrary r

$$\mathbb{P}[c_N(r) > 2/N^2 \mid \mathcal{H}_r] \geq \frac{1}{2}\delta\varepsilon^{2N} \mathbb{1}_{\{\max_i w_r^{(i)} - \min_i w_r^{(i)} \geq 2\delta/N\}}$$

so, by the tower rule and conditional independence,

$$\begin{aligned} \mathbb{P}[c_N(r) > 2/N^2 \mid \mathcal{F}_{r-1}] &= \mathbb{E}_r \left[\mathbb{P}[c_N(r) > 2/N^2 \mid \mathcal{H}_r] \right] \\ &\geq \frac{1}{2}\delta\varepsilon^{2N} \mathbb{P}[\max_i w_r^{(i)} - \min_i w_r^{(i)} \geq 2\delta/N \mid \mathcal{F}_{r-1}] \\ &\geq \frac{1}{2}\delta\varepsilon^{2N}\zeta > 0 \end{aligned}$$

for infinitely many r. Hence

$$\sum_{r=0}^{\infty} \mathbb{P}[c_N(r) > 2/N^2 \mid \mathcal{F}_{r-1}] = \infty$$

as required. ∎

5.3 Stochastic rounding

Corollary 5.5. *Consider an SMC algorithm using any stochastic rounding as its resampling scheme, such that (A1) is satisfied. Assume that there exists a constant* $a \in [1, \infty)$ *such that for all* x, x', t,

$$\frac{1}{a} \leq g_t(x, x') \leq a.$$

Assume that $\mathbb{P}[\tau_N(t) = \infty] = 0$ *for all finite* t. *Let* $(G_t^{(n,N)})_{t \geq 0}$ *denote the genealogy of a random sample of* n *terminal particles from the output of the algorithm among a total of* N *particles. Then, for any fixed* n, *the time-scaled genealogy* $(G_{\tau_N(t)}^{(n,N)})_{t \geq 0}$ *converges weakly to Kingman's* n*-coalescent as* $N \to \infty$.

Proof. We can apply exactly the proof of Corollary 5.3, except that stochastic rounding is more restrictive than stratified resampling, so that conditional on $w_t^{(1:N)}$ the only possible offspring counts (almost surely) are $\nu_t^{(i)} \in \{\lfloor N w_t^{(i)} \rfloor, \lfloor N w_t^{(i)} \rfloor + 1\}$. We simply set $p_{-1}^{(i)} = p_2^{(i)} = 0$ in the proof of Corollary 5.3 to see that

$$\frac{1}{(N)_3} \sum_{i=1}^{N} \mathbb{E}_t[(\nu_t^{(i)})_3] \leq \frac{a^2}{N-2} \frac{1}{(N)_2} \sum_{i=1}^{N} \mathbb{E}_t[(\nu_t^{(i)})_2]$$

as required. The result then follows by applying Theorem 4.1. ∎

We can also show, under additional conditions, that the assumption $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all finite t holds.

Proposition 5.6. *Consider an SMC algorithm using any stochastic rounding as its resampling scheme. Suppose that there exists a constant* $\varepsilon \in (0, 1]$ *and a probability density* h *such that*

$$\varepsilon h(x') \leq q_t(x, x') \leq \varepsilon^{-1} h(x')$$

uniformly in x, *and that there exist* $\zeta > 0$ *and* $\delta > 0$ *such that*

$$\mathbb{P}[\max_i w_t^{(i)} - \min_i w_t^{(i)} \geq 2\delta / N \mid \mathcal{F}_{t-1}] \geq \zeta$$

for infinitely many t. *Then, for all* $N > 1$, $\mathbb{P}[\tau_N(t) = \infty] = 0$ *for all finite* t.

This result was published in Brown et al. (2021, Lemma B.1) with the slightly stronger conditions where the bounds on q_t are also uniform in x'. It has since been noted that the conditions given here are sufficient; the h terms can be cancelled as in (5.12).

Proof. Define $p_k^{(i)}$ for $k \in \mathbb{Z}$ as in (5.8). In the case of stochastic rounding, $p_k^{(i)} \equiv 0$ for all

$k \notin \{0,1\}$, and we also have

$$\mathbb{P}[\nu_t^{(i)} = \lfloor N w_t^{(i)} \rfloor \mid w_t^{(1:N)}] = 1 - N w_t^{(i)} + \lfloor N w_t^{(i)} \rfloor,$$
$$\mathbb{P}[\nu_t^{(i)} = \lfloor N w_t^{(i)} \rfloor + 1 \mid w_t^{(1:N)}] = N w_t^{(i)} - \lfloor N w_t^{(i)} \rfloor.$$

Combining this with (5.12),

$$\begin{aligned} p_0^{(i)} &\geq (1 - N w_t^{(i)} + \lfloor N w_t^{(i)} \rfloor)\varepsilon^{2N}, \\ p_1^{(i)} &\geq (N w_t^{(i)} - \lfloor N w_t^{(i)} \rfloor)\varepsilon^{2N}. \end{aligned} \tag{5.14}$$

Define $i^\star := \operatorname{argmax}_i\{w_t^{(i)}\}$ and write $w_t^{(i^\star)} = \frac{1+\gamma}{N}$, for some $\gamma \geq 0$. Then, using (5.11),

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \geq 1 - \mathbb{P}[\nu_t^{(i^\star)} = 1 \mid \mathcal{H}_t] = \begin{cases} 1 - p_0^{(i^\star)} & \text{if } \gamma \in [0,1) \\ 1 & \text{if } \gamma \geq 1. \end{cases}$$

In the case $\gamma \in [0,1)$ we have $N w_t^{(i^\star)} - \lfloor N w_t^{(i^\star)} \rfloor = \gamma$, so

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \geq 1 - p_0^{(i^\star)} = p_1^{(i^\star)} \geq \gamma \varepsilon^{2N},$$

due to (5.14). Therefore, subject to $\max_i w_t^{(i)} \geq (1+\delta)/N$,

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \geq \min_{\gamma \geq \delta} \left\{\gamma \varepsilon^{2N}\right\} = \delta \varepsilon^{2N}.$$

Similarly, write $j^\star := \operatorname{argmin}_i\{w_t^{(i)}\}$ and $w_t^{(j^\star)} = \frac{1-\gamma}{N}$, for some $\gamma \in [0,1]$. Then, again using (5.11),

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \geq 1 - \mathbb{P}[\nu_t^{(j^\star)} = 1 \mid \mathcal{H}_t] = \begin{cases} 0 & \text{if } \gamma = 0 \\ 1 - p_1^{(j^\star)} & \text{if } \gamma \in (0,1) \\ 1 & \text{if } \gamma = 1. \end{cases}$$

If $\gamma \in (0,1)$ then $N w_t^{(i^\star)} - \lfloor N w_t^{(i^\star)} \rfloor = 1 - \gamma$, so

$$1 - p_1^{(j^\star)} = p_0^{(j^\star)} \geq (1 - (1-\gamma))\varepsilon^{2N} = \gamma \varepsilon^{2N}.$$

Therefore, subject to $\min_i w_t^{(i)} \leq (1-\delta)/N$,

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \geq \min_{\gamma \geq \delta} \left\{\gamma \varepsilon^{2N}\right\} = \delta \varepsilon^{2N}.$$

Combining the cases for the maximum and minimum weight we have that

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \geq \delta \varepsilon^{2N} \mathbb{1}_{\{\max_i w_t^{(i)} - \min_i w_t^{(i)} \geq 2\delta/N\}}$$

and we conclude as in Proposition 5.4. ■

5.4 Residual resampling with stratified residuals

Corollary 5.7. *Consider an SMC algorithm using residual resampling with stratified residuals, such that (A1) is satisfied. Assume that there exists a constant $a \in [1, \infty)$ such that for all x, x', t,*

$$\frac{1}{a} \leq g_t(x, x') \leq a.$$

Assume that $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all finite t. Let $(G_t^{(n,N)})_{t \geq 0}$ denote the genealogy of a random sample of n terminal particles from the output of the algorithm among a total of N particles. Then, for any fixed n, the time-scaled genealogy $(G_{\tau_N(t)}^{(n,N)})_{t \geq 0}$ converges weakly to Kingman's n-coalescent as $N \to \infty$.

Proof. We can apply exactly the proof of Corollary 5.3, except that residual-stratified resampling is more restrictive than stratified resampling, so that conditional on $w_t^{(1:N)}$ the only possible offspring counts (almost surely) are $\nu_t^{(i)} \in \{\lfloor N w_t^{(i)} \rfloor, \lfloor N w_t^{(i)} \rfloor + 1, \lfloor N w_t^{(i)} \rfloor + 2\}$. We simply set $p_{-1}^{(i)} = 0$ in the proof of Corollary 5.3 to see that

$$\frac{1}{(N)_3} \sum_{i=1}^{N} \mathbb{E}_t[(\nu_t^{(i)})_3] \leq \frac{a^2}{N-2} \frac{1}{(N)_2} \sum_{i=1}^{N} \mathbb{E}_t[(\nu_t^{(i)})_2]$$

as required. The result then follows by applying Theorem 4.1. ■

We can also show, under additional conditions, that the assumption $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all finite t holds.

Proposition 5.8. *Consider an SMC algorithm using residual resampling with stratified residuals. Suppose that there exists a constant $\varepsilon \in (0, 1]$ and a probability density h such that*

$$\varepsilon h(x') \leq q_t(x, x') \leq \varepsilon^{-1} h(x')$$

uniformly in x, and that there exist $\zeta > 0$ and $\delta > 0$ such that

$$\mathbb{P}[\max_i w_t^{(i)} - \min_i w_t^{(i)} \geq 2\delta/N \mid \mathcal{F}_{t-1}] \geq \zeta$$

for infinitely many t. Then, for all $N > 1$, $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all finite t.

Proof. Define $p_k^{(i)}$ for $k \in \mathbb{Z}$ as in (5.8). In the case of residual resampling with stratified residuals, $p_k^{(i)} \equiv 0$ for all $k \notin \{0, 1, 2\}$. Define $i^\star := \operatorname{argmax}_i\{w_t^{(i)}\}$ and write $w_t^{(i^\star)} = \frac{1+\gamma}{N}$, for some $\gamma \geq 0$. Then, using (5.11),

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \geq 1 - \mathbb{P}[\nu_t^{(i^\star)} = 1 \mid \mathcal{H}_t] = \begin{cases} 0 & \text{if } \gamma = 0 \\ 1 - p_0^{(i^\star)} & \text{if } \gamma \in (0,1) \\ 1 & \text{if } \gamma \geq 1. \end{cases}$$

In the case $\gamma \in (0,1)$ we have

$$1 - p_0^{(i^\star)} = p_1^{(i^\star)} + p_2^{(i^\star)} \geq p_1^{(i^\star)} \geq \mathbb{P}[\nu_t^{(i^\star)} = \lfloor N w_t^{(i^\star)} \rfloor + 1 \mid w_t^{(1:N)}]\varepsilon^{2N}$$

by (5.12). Also, the residual weight in this case is $r_{i^\star} = \gamma/R$, for some $R \in \{1, \dots, N-1\}$ (since $\gamma > 0$, $R \neq 0$). Therefore $\mathbb{P}[\nu_t^{(i^\star)} = \lfloor N w_t^{(i^\star)} \rfloor + 1 \mid w_t^{(1:N)}]$ is the probability that stratified resampling with R individuals assigns exactly 1 offspring to a parent with weight γ/R. According to Table 2.2 (lower bound on p_1), this probability is at least $\gamma/2$. Hence

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \geq \frac{\gamma}{2}\varepsilon^{2N}.$$

This means that, subject to $\max_i w_t^{(i)} \geq (1+\delta)/N$,

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \geq \min_{\gamma \geq \delta} \left\{ \frac{\gamma}{2}\varepsilon^{2N} \right\} = \frac{1}{2}\delta\varepsilon^{2N}.$$

Now a similar calculation for the minimum weight: let $j^\star := \operatorname{argmin}_i\{w_t^{(i)}\}$ and write $w_t^{(j^\star)} = \frac{1-\gamma}{N}$, for some $\gamma \in [0,1]$. Using (5.11),

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \geq 1 - \mathbb{P}[\nu_t^{(j^\star)} = 1 \mid \mathcal{H}_t] = \begin{cases} 0 & \text{if } \gamma = 0 \\ 1 - p_1^{(j^\star)} & \text{if } \gamma \in (0,1) \\ 1 & \text{if } \gamma = 1. \end{cases}$$

If $\gamma \in (0,1)$ then $r_{j^\star} = (1-\gamma)/R$, for some $R \in \{1, \dots, N-1\}$, and

$$1 - p_1^{(j^\star)} = p_0^{(j^\star)} + p_2^{(j^\star)} \geq p_0^{(j^\star)} \geq \mathbb{P}[\nu_t^{(j^\star)} = \lfloor N w_t^{(j^\star)} \rfloor \mid w_t^{(1:N)}]\varepsilon^{2N}$$

by (5.12). Now $\mathbb{P}[\nu_t^{(j^\star)} = \lfloor N w_t^{(j^\star)} \rfloor \mid w_t^{(1:N)}]$ is the probability that stratified resampling with R individuals assigns exactly 0 offspring to a parent with weight $(1-\gamma)/R$. According to Table 2.2 (lower bound on p_0), this probability is at least $\gamma/2$. Hence

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \geq \frac{\gamma}{2}\varepsilon^{2N}.$$

Therefore, using (5.12), we have that subject to $\min_i w_t^{(i)} \leq (1-\delta)/N$

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \geq \min_{\gamma \geq \delta} \left\{ \frac{\gamma}{2}\varepsilon^{2N} \right\} = \frac{1}{2}\delta\varepsilon^{2N}.$$

Combining the cases for the maximum and minimum weight we have

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \geq \frac{1}{2}\delta\varepsilon^{2N}\mathbb{1}_{\{\max_i w_t^{(i)} - \min_i w_t^{(i)} \geq 2\delta/N\}}$$

and we conclude as in Proposition 5.4. ■

5.5 Residual resampling with multinomial residuals

We believe that an analogous result holds when the resampling scheme used is residual resampling with multinomial residuals. Considering the ordering by variance presented in Proposition 2.3, the residual-multinomial scheme sits between the multinomial scheme and the residual-stratified scheme, both of which admit the desired convergence result (Corollaries 5.1 and 5.7).

However, we have so far been unable to prove a similar corollary for the residual-multinomial scheme. The techniques used for other residual schemes (see Section 5.4) fail here because the number of offspring assigned to each individual is not upper bounded by $\lfloor N w_t^{(i)} \rfloor$ plus a constant; as many as $R = O(N)$ residual offspring may be assigned to a single individual. The technique used for multinomial resampling (Section 5.1) also fails here: although we have a closed-form expression for the joint distribution of parental indices, it is not a straightforward product form because of the additional dependence between offspring counts induced by the deterministic assignments, so it is unclear how to recover the marginal distributions.

If I manage to prove this corollary, it would make this chapter satisfyingly complete :-) Res-star might prove an easier starting point.

5.6 Star resampling

One might ask the question: is it possible to construct an SMC algorithm whose genealogies converge to some non-trivial limit other than the n-coalescent? The answer is yes, as we now illustrate.

Recall that star resampling assigns all of the offspring to a single parent which is sampled from the Categorical distribution parametrised by $w_t^{(1:N)}$. It is easy enough to show that such a resampling scheme does not satisfy (3.11). The vector of offspring counts is at every generation some permutation of $(N, 0, \dots, 0)$, and hence we calculate

$$\frac{1}{(N)_2}\sum_{i=1}^{N}\mathbb{E}[(\nu_t^{(i)})_2 \mid \mathcal{H}_t] = \frac{1}{(N)_2}(N)_2 = 1,$$

$$\frac{1}{(N)_3}\sum_{i=1}^{N}\mathbb{E}[(\nu_t^{(i)})_3 \mid \mathcal{H}_t] = \frac{1}{(N)_3}(N)_3 = 1,$$

so no suitable sequence b_N can be found. Now we know that Theorem 3.6 does not apply, but this is not enough because condition (3.11) was not proved to be necessary. But in fact we know exactly what the genealogy of n particles from this SMC algorithm looks like (Figure 5.1). Whatever time scale is used, we cannot get away from the fact that this

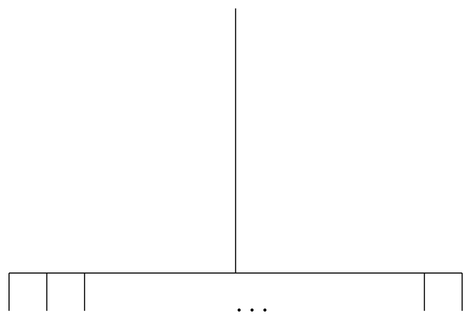


Figure 5.1: Sample genealogy induced by star resampling

genealogy involves multiple mergers; it cannot converge to the n-coalescent.

The limiting genealogy is more like a *star coalescent* (Pitman 1999; Griffiths and Mano 2016). This is the coalescent process comprising an Exp(1)-distributed event time at which all of the lineages merge into one.

In the case of star resampling we have $c_N(t) \equiv 1$, so the time-scaling function $\tau_N(t)$ defined in (3.2) converges pointwise to the identity function $\tau(t) \equiv t$ as $N \to \infty$, and this does not yield a continuous-time limit. Under any time scale that results in a continuous-time limiting process, the coalescent event time converges to 0, rather than the usual Exp(1)-distributed random variable. The resulting genealogy is a variant star coalescent where the distribution of the event time is a point mass at 0. A fun consequence of this is that this coalescent comes down from infinity, while the star coalescent does not if I decide not to write about cdf∞ in ch2 then remove this last sentence.

5.7 Conditional SMC

In conditional SMC, one "immortal" particle is treated differently to the others when it comes to assigning offspring to parents. The immortal particle is guaranteed at least one offspring, and has on average one more offspring than each of the other parents in each generation. This results in genealogies that are qualitatively different to those of a corresponding standard SMC algorithm. For one thing, the population MRCA is *guaranteed* to be an immortal particle; there is a sense in which the immortal lineage *attracts* coalescence events.

Given this, we should not have been surprised if conditional SMC genealogies converged to a quite different coalescent process, perhaps a *structured coalescent* (Notohara 1990). As it turns out, we still recover Kingman's n-coalescent in the large population limit (Corollary 5.9). A possible explanation for this is that, as $N \to \infty$, the probability of a given sample of size n interacting with the immortal lineage (before its within-sample MRCA) vanishes, leaving a process that looks very much like the one induced by the corresponding standard SMC algorithm.

Corollary 5.9. *Consider a conditional SMC algorithm using multinomial resampling, such that (A1) is satisfied. Assume there exist constants $\varepsilon \in (0,1]$ and $a \in [1,\infty)$ and probability density h such that for all x, x', t,*

$$\frac{1}{a} \le g_t(x,x') \le a, \quad \varepsilon h(x') \le q_t(x,x') \le \frac{1}{\varepsilon} h(x'). \tag{5.15}$$

Let $(G_t^{(n,N)})_{t\ge 0}$ denote the genealogy of a random sample of n terminal particles from the output of the algorithm among a total of N particles. Then, for any fixed n, the time-scaled genealogy $(G_{\tau_N(t)}^{(n,N)})_{t\ge 0}$ converges weakly to Kingman's n-coalescent as $N \to \infty$.

We restrict here to the case of multinomial resampling, which seems to be the most commonly-used resampling scheme within conditional SMC. Implementing other resampling schemes while maintaining the immortal lineage is more involved, though by no means impossible (Lee, Murray, and Johansen 2019). We conjecture that similar results hold for conditional SMC with other resampling schemes, as in the preceding corollaries.

The conditions (5.15) are, as one might expect, identical to those assumed in the case of standard SMC with multinomial resampling (Corollary 5.1).

Proof. Assume, without loss of generality, that the immortal particle takes index 1 in each generation. This assumption is valid due to (A1), and significantly lightens the notation, but the same argument holds if the immortal indices are taken to be $a^\star_{(0:T)}$ rather than $(1,\dots,1)$.

Define $\mathcal{H}_t$ as in (5.2). The parental indices are conditionally independent given $\mathcal{H}_t$, as in standard SMC with multinomial resampling, but we have to treat $i=1$ as a special case. The conditional law on the i^{th} parental index is

$$\mathbb{P}\left[a_t^{(i)} = a_i \mid \mathcal{H}_t\right] \propto \begin{cases} \mathbb{1}_{a_i=1} & i=1 \\ g_t(X_{t+1}^{a_t^{(a_i)}}, X_t^{(a_i)}) q_{t-1}(X_t^{(a_i)}, X_{t-1}^{(i)}) & i=2,\dots,N, \end{cases}$$

resulting in the joint law

$$\mathbb{P}\left[a_t^{(1:N)} = a_{1:N} \mid \mathcal{H}_t\right] \propto \mathbb{1}_{a_1=1} \prod_{i=2}^N g_t(X_{t+1}^{a_t^{(a_i)}}, X_t^{(a_i)}) q_{t-1}(X_t^{(a_i)}, X_{t-1}^{(i)}).$$

As in Corollary 5.1, under (5.15) we have bounds

$$\mathbb{E}[(V_1^{(i)})_k] \le \mathbb{E}[(\nu_t^{(i)})_k \mid \mathcal{H}_t] \le \mathbb{E}[(V_2^{(i)})_k],$$

where now

$$V_1^{(i)} \overset{d}{=} \mathbb{1}_{i=1} + \text{Binomial}\left(N-1, \frac{\varepsilon/a}{(\varepsilon/a)+(N-1)(a/\varepsilon)}\right),$$
$$V_2^{(i)} \overset{d}{=} \mathbb{1}_{i=1} + \text{Binomial}\left(N-1, \frac{a/\varepsilon}{(a/\varepsilon)+(N-1)(\varepsilon/a)}\right).$$

independently for each i and independently of $\mathcal{F}_\infty$. Furthermore, using the Binomial moments and the identity $(X+1)_2 \equiv 2(X)_1 + (X)_2$, one can show that

$$\mathbb{E}[(V_1^{(i)})_2] \geq \begin{cases} \frac{(N-1)_2}{N^2}\frac{\varepsilon^4}{a^4} + \frac{2(N-1)}{N}\frac{\varepsilon^2}{a^2} & \text{if } i=1 \\ \frac{(N-1)_2}{N^2}\frac{\varepsilon^4}{a^4} & \text{if } i \neq 1. \end{cases}$$

Using the identity $(X+1)_3 \equiv 3(X)_2 + (X)_3$, we also have

$$\mathbb{E}[(V_2^{(i)})_3] \leq \begin{cases} \frac{(N-1)_3}{N^3}\frac{a^6}{\varepsilon^6} + \frac{3(N-1)_2}{N^2}\frac{a^4}{\varepsilon^4} & \text{if } i=1 \\ \frac{(N-1)_3}{N^3}\frac{a^6}{\varepsilon^6} & \text{if } i \neq 1. \end{cases}$$

We therefore have

$$\begin{aligned}\frac{1}{(N)_2}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_2 \mid \mathcal{H}_t] &\geq \frac{1}{(N)_2}\sum_{i=1}^N \mathbb{E}[(V_1^{(i)})_2] \geq \frac{1}{(N)_2}\left[\frac{2(N-1)}{N}\frac{\varepsilon^2}{a^2} + \sum_{i=1}^N \frac{(N-1)_2}{N^2}\frac{\varepsilon^4}{a^4}\right] \\ &= \frac{1}{N^2}\left[2\frac{\varepsilon^2}{a^2} + (N-2)\frac{\varepsilon^4}{a^4}\right] \geq \frac{\varepsilon^4}{Na^4} \end{aligned} \tag{5.16}$$

and

$$\begin{aligned}\frac{1}{(N)_3}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_3 \mid \mathcal{H}_t] &\leq \frac{1}{(N)_3}\sum_{i=1}^N \mathbb{E}[(V_2^{(i)})_3] \leq \frac{1}{(N)_3}\left[\frac{3(N-1)_2}{N^2}\frac{a^4}{\varepsilon^4} + \sum_{i=1}^N \frac{(N-1)_3}{N^3}\frac{a^6}{\varepsilon^6}\right] \\ &= \frac{1}{N^3}\left[3\frac{a^4}{\varepsilon^4} + (N-3)\frac{a^6}{\varepsilon^6}\right] \leq \frac{a^6}{N^2\varepsilon^6}. \end{aligned}$$

Hence, applying (5.6), we can upper bound the ratio

$$\frac{\frac{1}{(N)_3}\sum_{i=1}^N \mathbb{E}_t[(\nu_t^{(i)})_3]}{\frac{1}{(N)_2}\sum_{i=1}^N \mathbb{E}_t[(\nu_t^{(i)})_2]} \leq \frac{a^{10}}{N\varepsilon^{10}} =: b_N \underset{N\to\infty}{\longrightarrow} 0$$

so (3.11) is satisfied. Proof that the time scale is finite is relegated to Lemma 5.10, whence we conclude by applying Theorem 4.1.

∎

Lemma 5.10. *Consider a conditional SMC algorithm using multinomial resampling, satisfying (A1) and* (5.15)*. Then, for all* $N > 2$, $\mathbb{P}[\tau_N(t) = \infty] = 0$ *for all finite* t.

Proof. The proof is identical to that of Lemma 5.2, since (5.16) gives us exactly the same lower bound on $\mathbb{E}_t[c_N(t)]$ that we had in standard SMC with multinomial resampling. ■

5.7.1 Effect of ancestor sampling

Ancestor sampling breaks up the immortal lineage into sections, so it is not really a lineage anymore, and we do not really have a pure coalescent process backwards in time. Regardless, we shall throw caution to the wind and examine the resulting "genealogies".

Using the parent sampling probabilities specified in (2.10), now with time reversed and the notation made to fit preferably the presentation in Chapter 2 should use the notation we want to use here with this study of genealogies, we obtain add one more step of working below to make it less "magic"?

$$\mathbb{P}[a_t^{(i)} = a_i \mid \mathcal{H}_t] \propto \begin{cases} w_t^{(a_i)} q_{t-1}(X_t^{(a_i)}, X_{t-1}^{(i)}) & i \in \text{non-immortal particles} \\ w_t^{(a_i)} q_{t-1}(X_t^{(a_i)}, X_{t-1}^{*}) & i = \text{immortal particle.} \end{cases}$$

But when i is the index of the immortal particle, $X_{t-1}^{(i)} = X_{t-1}^{*}$, so the above simplifies to

$$\mathbb{P}[a_t^{(i)} = a_i \mid \mathcal{H}_t] \propto w_t^{(a_i)} q_{t-1}(X_t^{(a_i)}, X_{t-1}^{(i)})$$

for each i, which is exactly (5.3), the law on parental indices under standard SMC with multinomial resampling. In other words, when parental indices are chosen, the immortal particle is treated exactly like all of the other particles; it has completely lost its "reproductive advantage". This means it is no more likely for lineages to coalesce onto the "immortal" lineage than onto any other lineage, so we do not see the behaviour of Figure 2.8 which caused the particle Gibbs chain to mix slowly over the sequential component. This supports the claim of Section 2.5.3: particle Gibbs with ancestor sampling still experiences ancestral degeneracy, but this no longer causes the sequential component to get stuck.

Bibliography

[1] Suzie Brown et al. "Simple Conditions for Convergence of Sequential Monte Carlo Genealogies with Applications". In: *Electronic Journal of Probability* 26.1 (2021), pp. 1–22. ISSN: 1083-6489. DOI: 10.1214/20-EJP561.

[2] Pierre Del Moral. *Feynman–Kac Formulae: Genealogical and Interacting Particle Systems with Applications*. Springer, 2004.

[3] Richard Durrett. *Probability: Theory and Examples*. 5th ed. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2019. DOI: 10.1017/9781108591034.

[4] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence*. John Wiley & Sons, 2009.

[5] Catherine Forbes et al. *Statistical Distributions*. John Wiley & Sons, 2011.

[6] Y.-X. Fu. "Exact Coalescent for the Wright–Fisher Model". In: *Theoretical Population Biology* 69 (4 2006), pp. 385–394.

[7] Robert Griffiths and Shuhei Mano. "The Star-shaped Λ-coalescent and Fleming–Viot Process". In: *Stochastic Models* 32.4 (2016), pp. 606–631.

[8] Godfrey Harold Hardy and Srinivasa Aaiyangar Ramanujan. "Asymptotic Formulae in Combinatory Analysis". In: *Proceedings of the London Mathematical Society* s2-17.1 (1918), pp. 75–115.

[9] Jere Koskela et al. *Asymptotic genealogies of interacting particle systems with an application to sequential Monte Carlo*. Mathematics e-print 1804.01811. ArXiv, 2018.

[10] Anthony Lee, Lawrence Murray, and Adam M. Johansen. "Resampling in Conditional SMC Algorithms". Unpublished. 2019.

[11] Martin Möhle. "Robustness Results for the Coalescent". In: *Journal of Applied Probability* 35.2 (1998), pp. 438–447.

[12] Martin Möhle. "Weak Convergence to the Coalescent in Neutral Population Models". In: *Journal of Applied Probability* 36.2 (1999), pp. 446–460.

[13] Martin Möhle and Serik Sagitov. "Coalescent Patterns in Exchangeable Diploid Population Models". In: *Journal of Mathematical Biology* 47 (2003), pp. 337–352.

[14] James E. Mosimann. "On the Compound Multinomial Distribution, the Multivariate β-Distribution, and Correlations among Proportions". In: *Biometrika* 49.1/2 (1962), pp. 65–82.

[15] M. Notohara. “The Coalescent and the Genealogical Process in Geographically Structured Population”. In: *Journal of Mathematical Biology* 29.1 (1990), pp. 59–75.

[16] Jim Pitman. “Coalescents with Multiple Collisions”. In: *Annals of Probability* (1999), pp. 1870–1902.