Weak convergence proof (in progress)

Suzie Brown

September 28, 2020

Theorem 1. Let $\nu_t^{(1:N)}$ denote the offspring numbers in an interacting particle system satisfying the standing assumption and 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

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

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

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

$$\mathbb{P}(Z_{t} = j, S_{t} = \eta \mid Z_{t-1} = i, S_{t-1} = \xi) = \begin{cases}
1 - p_{t} & \text{if } j = i \text{ and } \xi = \eta \\
p_{\xi\xi}(t) + p_{t} - 1 & \text{if } j = i + 1 \text{ and } \xi = \eta \\
p_{\xi\eta}(t) & \text{if } j = i + 1 \text{ and } \xi \neq \eta \\
0 & \text{otherwise.}
\end{cases} \tag{2}$$

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 is non-negative, attaining the value zero when $\xi = \Delta$.)

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

Lemma 1. $\max_{\xi \in E} (1 - p_{\xi\xi}(t)) = 1 - p_{\Delta\Delta}(t)$.

Proof. Consider any $\xi \in E$ consisting of k blocks $(1 \le k \le n-1)$, and any $\xi' \in E$ consisting of k+1 blocks. From the definition of $p_{\xi\eta}(t)$ (Koskela et al., 2018, Equation (1)),

$$p_{\xi\xi}(t) = \frac{1}{(N)_k} \sum_{\substack{i_1,\dots,i_k \text{all distinct}}} \nu_t^{(i_1)} \cdots \nu_t^{(i_k)}. \tag{3}$$

Similarly,

$$p_{\xi'\xi'}(t) = \frac{1}{(N)_{k+1}} \sum_{\substack{i_1, \dots, i_k, i_{k+1} \\ \text{all distinct}}} \nu_t^{(i_1)} \cdots \nu_t^{(i_k)} \nu_t^{(i_{k+1})}$$

$$= \frac{1}{(N)_k (N-k)} \sum_{\substack{i_1, \dots, i_k \\ \text{all distinct}}} \left\{ \nu_t^{(i_1)} \cdots \nu_t^{(i_k)} \sum_{\substack{i_{k+1} = 1 \\ \text{also distinct}}}^N \nu_t^{(i_{k+1})} \right\}. \tag{4}$$

Discarding the zero summands,

$$p_{\xi'\xi'}(t) = \frac{1}{(N)_k(N-k)} \sum_{\substack{i_1,\dots,i_k \\ \text{all distinct:} \\ \nu_t^{(i_1)},\dots,\nu_t^{(i_k)} > 0}} \left\{ \nu_t^{(i_1)} \cdots \nu_t^{(i_k)} \sum_{\substack{i_{k+1}=1 \\ \text{also distinct}}}^N \nu_t^{(i_{k+1})} \right\}.$$
 (5)

The inner sum is

$$\sum_{\substack{i_{k+1}=1\\\text{also distinct}}}^{N} \nu_t^{(i_{k+1})} = \left\{ \sum_{i=1}^{N} \nu_t^{(i)} - \sum_{i \in \{i_1, \dots, i_k\}} \nu_t^{(i)} \right\} \le N - k$$
 (6)

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

$$p_{\xi'\xi'}(t) \le \frac{N-k}{(N)_k(N-k)} \sum_{\substack{i_1,\dots,i_k \\ \text{all distinct:} \\ \nu_t^{(i_1)},\dots,\nu_t^{(i_k)} > 0}} \nu_t^{(i_1)} \cdots \nu_t^{(i_k)} = p_{\xi\xi}(t).$$
 (7)

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 2. For any t > 0,

$$\lim_{N \to \infty} \mathbb{E} \left[\prod_{r=1}^{\tau_N(t)} (1 - p_r) \right] = e^{-\alpha t}$$
 (8)

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

Proof. The strategy is to find upper and lower bounds on $\mathbb{E}\left[\prod_{r=1}^{\tau_N(t)}(1-p_r)\right]$, both of which converge to $e^{-\alpha t}$.

Lower Bound

From Brown et al. (2020, Equation (14)), taking $\xi = \Delta$, we have

$$1 - p_t = p_{\Delta\Delta}(t) \ge 1 - \alpha(1 + O(N^{-1})) \left[c_N(t) + B_n D_N(t) \right]$$
(9)

where $B_n > 0$ and the $O(N^{-1})$ term does not depend on t. In particular,

$$1 - p_t = p_{\Delta\Delta}(t) \ge 1 - \frac{N^{n-2}}{(N-2)_{n-2}} \alpha c_N(t) - \frac{N^{n-3}}{(N-3)_{n-3}} B_n D_N(t)$$
(10)

Hence, by a multinomial expansion,

$$\prod_{r=1}^{\tau_{N}(t)} (1 - p_{r}) \ge \prod_{r=1}^{\tau_{N}(t)} \left\{ 1 - \alpha (1 + O(N^{-1})) \left[c_{N}(r) + B_{n} D_{N}(r) \right] \right\}$$

$$= 1 + \sum_{k=1}^{\tau_{N}(t)} \sum_{\substack{r_{1} < \dots < r_{k} \\ = 1}}^{\tau_{N}(t)} \prod_{j=1}^{k} \left\{ -\alpha (1 + O(N^{-1})) \left[c_{N}(r_{j}) + B_{n} D_{N}(r_{j}) \right] \right\}$$

$$= 1 + \sum_{k=1}^{\tau_{N}(t)} \left\{ -\alpha (1 + O(N^{-1})) \right\}^{k} \sum_{\substack{r_{1} < \dots < r_{k} \\ -1}}^{\tau_{N}(t)} \prod_{j=1}^{k} \left\{ c_{N}(r_{j}) + B_{n} D_{N}(r_{j}) \right\}. \tag{11}$$

Taking expectations,

$$\mathbb{E}\left[\prod_{r=1}^{\tau_N(t)} (1 - p_r)\right] \ge 1 + \sum_{k=1}^{\tau_N(t)} \left\{-\alpha(1 + O(N^{-1}))\right\}^k \mathbb{E}\left[\sum_{\substack{r_1 < \dots < r_k \\ =1}}^{\tau_N(t)} \prod_{j=1}^k \left\{c_N(r_j) + B_n D_N(r_j)\right\}\right]. \tag{12}$$

We want to show that the expectation on the right converges to $t^k/k!$, for reasons that will become clear. The strategy is to upper and lower bound this expectation by quantities that converge to $t^k/k!$.

First the lower bound. Assume that $k \leq \tau_N(t)$, i.e. the sum is non-empty. From Koskela et al. (2018, Equation (8)),

$$\sum_{\substack{r_1 < \dots < r_k \\ = 1}}^{\tau_N(t)} \prod_{j=1}^k \left\{ c_N(r_j) + B_n D_N(r_j) \right\} \ge \sum_{\substack{r_1 < \dots < r_k \\ = 1}}^{\tau_N(t)} \prod_{j=1}^k c_N(r_j) \\
\ge \frac{1}{k!} \left(\sum_{s=1}^{\tau_N(t)} c_N(s) \right)^k - \frac{1}{k!} \binom{k}{2} \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \right) \left(\sum_{s=1}^{\tau_N(t)} c_N(s) \right)^{k-2} \\
\ge \frac{1}{k!} t^k - \frac{1}{k!} \binom{k}{2} (t+1)^{k-2} \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \right) \tag{13}$$

by the definition of τ_N . Then

$$\mathbb{E}\left[\sum_{\substack{r_1 < \dots < r_k \\ = 1}}^{\tau_N(t)} \prod_{j=1}^k \left\{ c_N(r_j) + B_n D_N(r_j) \right\} \right] \ge \frac{1}{k!} t^k - \frac{1}{k!} \binom{k}{2} (t+1)^{k-2} \mathbb{E}\left[\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right] \longrightarrow \frac{1}{k!} t^k \tag{14}$$

as $N \to \infty$ using Brown et al. (2020, Equation (5)), which is shown to hold via lemmata 1 and 3 therein. Now for the upper bound. We start with a multinomial expansion and some manipulations of the sums:

$$\sum_{\substack{r_1 < \dots < r_k \\ = 1}}^{r_N(t)} \prod_{j=1}^k \left\{ c_N(r_j) + B_n D_N(r_j) \right\} = \frac{1}{k!} \sum_{\substack{r_1 \neq \dots \neq r_k \\ \text{all distinct}}}^{r_N(t)} \prod_{j=1}^k \left\{ c_N(r_j) + B_n D_N(r_j) \right\}$$

$$= \frac{1}{k!} \sum_{\substack{r_1 \neq \dots \neq r_k \\ \text{all distinct}}}^{r_N(t)} \sum_{\substack{I \subseteq \{1,\dots,k\}}} (B_n)^{k-|I|} \left\{ \prod_{i \in \mathcal{I}} c_N(r_i) \right\} \left\{ \prod_{j \notin \mathcal{I}} D_N(r_j) \right\}$$

$$= \frac{1}{k!} \sum_{\substack{I \subseteq \{1,\dots,k\}}} (B_n)^{k-|I|} \sum_{\substack{r_1 \neq \dots \neq r_k \\ \text{all distinct}}}^{r_N(t)} \left\{ \prod_{i \in \mathcal{I}} c_N(r_i) \right\} \left\{ \prod_{j \notin I} D_N(r_j) \right\}$$

$$= \frac{1}{k!} \sum_{\substack{I = 0 \\ I \text{ lidistinct}}}^{r_N(t)} \left\{ \prod_{i=1}^k c_N(r_i) \right\} \left\{ \prod_{j=I+1}^k D_N(r_j) \right\}$$

$$= \frac{1}{k!} \sum_{\substack{r_1 \neq \dots \neq r_k \\ \text{all distinct}}}^{r_N(t)} \left\{ \prod_{i=1}^k c_N(r_i) \right\} \left\{ \prod_{j=I+1}^k D_N(r_j) \right\}. \quad (15)$$

Then, using that $D_N(s) \leq c_N(s)$ for all s (Koskela et al., 2018, p.9), along with the definition of τ_N ,

$$\frac{1}{k!} \sum_{\substack{r_1 \neq \cdots \neq r_k \\ \text{all distinct}}}^{\tau_N(t)} \left\{ \prod_{i=1}^k c_N(r_i) \right\} + \frac{1}{k!} \sum_{I=0}^{k-1} \binom{k}{I} (B_n)^{k-I} \sum_{\substack{r_1 \neq \cdots \neq r_k \\ \text{all distinct}}}^{\tau_N(t)} \left\{ \prod_{j=I+1}^{k} D_N(r_j) \right\} \\
\leq \frac{1}{k!} \left(\sum_{r=1}^{\tau_N(t)} c_N(r) \right)^k + \frac{1}{k!} \sum_{I=0}^{k-1} \binom{k}{I} (B_n)^{k-I} \sum_{\substack{r_1 \neq \cdots \neq r_k \\ \text{all distinct}}}^{\tau_N(t)} \left\{ \prod_{i=1}^{k-1} c_N(r_i) \right\} D_N(r_k) \\
\leq \frac{1}{k!} \left\{ t + c_N(\tau_N(t)) \right\}^k + \frac{1}{k!} \sum_{I=0}^{k-1} \binom{k}{I} (B_n)^{k-I} \left\{ \sum_{\substack{r_1 \neq \cdots \neq r_{k-1} \\ \text{all distinct}}}^{\tau_N(t)} \prod_{i=1}^{k-1} c_N(r_i) \right\} \binom{\tau_N(t)}{r_k} \\
\leq \frac{1}{k!} \left\{ t + c_N(\tau_N(t)) \right\}^k + \frac{1}{k!} \sum_{I=0}^{k-1} \binom{k}{I} (B_n)^{k-I} \left(\sum_{r=1}^{\tau_N(t)} c_N(r) \right)^{k-1} \binom{\tau_N(t)}{r_{k-1}} D_N(r) \\
\leq \frac{1}{k!} \left\{ t + c_N(\tau_N(t)) \right\}^k + \frac{1}{k!} \sum_{I=0}^{k-1} \binom{k}{I} (B_n)^{k-I} (t+1)^{k-1} \binom{\tau_N(t)}{r_{k-1}} D_N(r) \right). \tag{16}$$

Taking expectations,

$$\lim_{N \to \infty} \mathbb{E} \left[\sum_{\substack{r_1 < \dots < r_k \\ = 1}}^{\tau_N(t)} \prod_{j=1}^k \left\{ c_N(r_j) + B_n D_N(r_j) \right\} \right] \le \frac{1}{k!} \lim_{N \to \infty} \mathbb{E} \left[\left\{ t + c_N(\tau_N(t)) \right\}^k \right]$$

$$+ \frac{1}{k!} \sum_{I=0}^{k-1} \binom{k}{I} (B_n)^{k-I} (t+1)^{k-1} \lim_{N \to \infty} \mathbb{E} \left[\left(\sum_{r=1}^{\tau_N(t)} D_N(r) \right) \right]$$

$$= \frac{1}{k!} t^k.$$

$$(17)$$

The limit follows from Brown et al. (2020, Equations (3),(4)) along with the fact that, since $c_N(s) \in [0,1]$ for all s, $\mathbb{E}[c_N(s)^k] \leq \mathbb{E}[c_N(s)]$ for all $k \geq 1$, and the expansion

$$\mathbb{E}\left[\frac{1}{k!}\left\{t + c_N(\tau_N(t))\right\}^k\right] = \frac{1}{k!}\sum_{i=0}^k \binom{k}{i}t^i \mathbb{E}\left[c_N(\tau_N(t))^{k-i}\right] \longrightarrow \frac{1}{k!}t^k.$$
(18)

Combining these upper and lower limits, we conclude that

$$\lim_{N \to \infty} \mathbb{E} \left[\sum_{\substack{r_1 < \dots < r_k \\ -1}}^{\tau_N(t)} \prod_{j=1}^k \left\{ c_N(r_j) + B_n D_N(r_j) \right\} \right] = \frac{1}{k!} t^k$$
 (19)

and thus

$$\lim_{N \to \infty} \mathbb{E} \left[1 + \sum_{k=1}^{\tau_N(t)} \left\{ -\alpha (1 + O(N^{-1})) \right\}^k \sum_{r_1 < \dots < r_k}^{\tau_N(t)} \prod_{j=1}^k \left\{ c_N(r_j) + B_n D_N(r_j) \right\} \right] \\
= 1 + \lim_{N \to \infty} \sum_{k=1}^{\infty} \left\{ -\alpha (1 + O(N^{-1})) \right\}^k \mathbb{E} \left[\mathbb{I}_{\{k \le \tau_N(t)\}} \sum_{r_1 < \dots < r_k}^{\tau_N(t)} \prod_{j=1}^k \left\{ c_N(r_j) + B_n D_N(r_j) \right\} \right] \\
= 1 + \sum_{k=1}^{\infty} (-\alpha)^k \lim_{N \to \infty} \mathbb{E} \left[\sum_{r_1 < \dots < r_k}^{\tau_N(t)} \prod_{j=1}^k \left\{ c_N(r_j) + B_n D_N(r_j) \right\} \right] k \le \tau_N(t) \\
= 1 + \sum_{k=1}^{\infty} (-\alpha)^k \frac{t^k}{k!} \times 1 = e^{-\alpha t} \tag{20}$$

as $N \to \infty$, where the last line follows from (19) and Lemma [??].

Upper Bound

From Koskela et al. (2018, Lemma 1 Case 1), taking $\xi = \Delta$, we have

$$1 - p_t = p_{\Delta\Delta}(t) \le 1 - \alpha(1 + O(N^{-1})) \left[c_N(t) - B_n' D_N(t) \right]. \tag{21}$$

A multinomial expansion as before yields

$$\prod_{r=1}^{\tau_N(t)} (1 - p_r) \le 1 + \sum_{k=1}^{\tau_N(t)} \left\{ -\alpha (1 + O(N^{-1})) \right\}^k \sum_{r_1 < \dots < r_k}^{\tau_N(t)} \prod_{j=1}^k \left\{ c_N(r_j) - B'_n D_N(r_j) \right\}. \tag{22}$$

Analogously to (15), assuming $k \leq \tau_N(t)$ we can write

$$\sum_{\substack{r_1 < \dots < r_k \\ = 1}}^{\tau_N(t)} \prod_{j=1}^k \left\{ c_N(r_j) - B'_n D_N(r_j) \right\} = \frac{1}{k!} \sum_{\substack{r_1 \neq \dots \neq r_k \\ \text{all distinct}}}^{\tau_N(t)} \left\{ \prod_{i=1}^k c_N(r_i) \right\} \\
+ \frac{1}{k!} \sum_{I=0}^{k-1} \binom{k}{I} \left(-B'_n \right)^{k-I} \sum_{\substack{r_1 \neq \dots \neq r_k \\ \text{all distinct}}} \left\{ \prod_{i=1}^I c_N(r_i) \right\} \left\{ \prod_{j=I+1}^k D_N(r_j) \right\}.$$
(23)

We start by dealing with the second term:

$$\frac{1}{k!} \sum_{I=0}^{k-1} {k \choose I} (-B'_n)^{k-I} \sum_{\substack{r_1 \neq \cdots \neq r_k \\ \text{all distinct}}} \left\{ \prod_{i=1}^{I} c_N(r_i) \right\} \left\{ \prod_{j=I+1}^{k} D_N(r_j) \right\}$$

$$= \frac{1}{k!} \sum_{\substack{I=0: \\ k-I \text{ even}}}^{k-1} {k \choose I} (B'_n)^{k-I} \sum_{\substack{r_1 \neq \cdots \neq r_k \\ \text{all distinct}}} \left\{ \prod_{i=1}^{I} c_N(r_i) \right\} \left\{ \prod_{j=I+1}^{k} D_N(r_j) \right\}$$

$$- \frac{1}{k!} \sum_{\substack{I=0: \\ k-I \text{ odd}}}^{k-1} {k \choose I} (B'_n)^{k-I} \sum_{\substack{r_1 \neq \cdots \neq r_k \\ \text{all distinct}}} \left\{ \prod_{i=1}^{I} c_N(r_i) \right\} \left\{ \prod_{j=I+1}^{k} D_N(r_j) \right\}. \tag{24}$$

This is lower bounded by

$$0 - \frac{1}{k!} \sum_{\substack{I=0\\k-I \text{ odd}}}^{k-1} {k \choose I} (B'_n)^{k-I} (t+1)^{k-1} \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right)$$
 (25)

using that $c_N(r), D_N(r) \ge 0$ for all r to bound the even terms below, and arguments from (16) to bound the odd terms above. The same arguments lead to the upper bound

$$\frac{1}{k!} \sum_{\substack{I=0\\k-I \text{ even}}}^{k-1} {k \choose I} (B'_n)^{k-I} (t+1)^{k-1} \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right) - 0.$$
 (26)

By Brown et al. (2020, Equation (4)), both bounds have vanishing expectation as $N \to \infty$. We are left with the first term in (23), which is upper bounded by

$$\frac{1}{k!} \sum_{\substack{r_1 \neq \dots \neq r_k \text{all distinct}}}^{\tau_N(t)} \left\{ \prod_{i=1}^k c_N(r_i) \right\} \leq \frac{1}{k!} \left(\sum_{s=1}^{\tau_N(t)} c_N(s) \right)^k \leq \frac{1}{k!} \{ t + c_N(\tau_N(t)) \}^k \tag{27}$$

the expectation of which converges to $t^k/k!$ as in (18). We use Koskela et al. (2018, Equation (8)) to construct a

lower bound:

$$\frac{1}{k!} \sum_{\substack{r_1 \neq \dots \neq r_k \text{all distinct}}}^{\tau_N(t)} \left\{ \prod_{i=1}^k c_N(r_i) \right\} \ge \frac{1}{k!} \left(\sum_{s=1}^{\tau_N(t)} c_N(s) \right)^k - \frac{1}{k!} \binom{k}{2} \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \right) \left(\sum_{s=1}^{\tau_N(t)} c_N(s) \right)^{k-2} \\
\ge \frac{1}{k!} t^k - \frac{1}{k!} \binom{k}{2} (t+1)^{k-2} \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \right) \tag{28}$$

The expectation of this bound also converges to $t^k/k!$, using Brown et al. (2020, Equation (5)). We can now conclude that

$$\lim_{N \to \infty} \mathbb{E} \left[\sum_{\substack{r_1 < \dots < r_k \\ 1}}^{\tau_N(t)} \prod_{j=1}^k \left\{ c_N(r_j) - B'_n D_N(r_j) \right\} \right] = \frac{1}{k!} t^k$$
 (29)

and thus, by calculations analogous to (20),

$$\lim_{N \to \infty} \mathbb{E} \left[1 + \sum_{k=1}^{\tau_N(t)} \left\{ -\alpha (1 + O(N^{-1})) \right\}^k \sum_{\substack{r_1 < \dots < r_k \\ =1}}^{\tau_N(t)} \prod_{j=1}^k \left\{ c_N(r_j) - B'_n D_N(r_j) \right\} \right] = 1 + \sum_{k=1}^{\infty} (-\alpha)^k \frac{1}{k!} t^k = e^{-\alpha t}$$
 (30)

as $N \to \infty$.

We now have upper and lower bounds on $\lim_{N\to\infty} \mathbb{E}\left[\prod_{r=1}^{\tau_N(t)}(1-p_r)\right]$, both of which are equal to $e^{-\alpha t}$, and the result follows.

References

Brown, S., Jenkins, P. A., Johansen, A. M. and Koskela, J. (2020), 'Simple conditions for convergence of sequential Monte Carlo genealogies with applications', arXiv preprint arXiv:2007.00096.

Koskela, J., Jenkins, P. A., Johansen, A. M. and Spanò, D. (2018), 'Asymptotic genealogies of interacting particle systems with an application to sequential Monte Carlo', arXiv preprint arXiv:1804.01811.

Möhle, M. (1999), 'Weak convergence to the coalescent in neutral population models', *Journal of Applied Probability* **36**(2), 446–460.