Weak convergence proof (in progress)

Suzie Brown

September 17, 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*

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

for all N, uniformly in $t \geq 1$. Then the rescaled genealogical process $(G^{(n,N)}_{\tau_N(t)})_{t\geq 0}$ 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 $\{1, \dots, n\}$ 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 ($1 \leq k \leq n-1$), 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,

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

Discarding the zero summands,

$$p_{\xi'\xi'}(t) = \frac{1}{(N)_k(N-k)} \sum_{\substack{i_1,\dots,i_k \\ \text{all distinct:} \\ \nu_t^{(i_1)},\dots,\nu_t^{(i_k)}>0}} \left\{ \nu_t^{(i_1)} \cdots \nu_t^{(i_k)} \sum_{\substack{i_{k+1}=1 \\ \text{also distinct}}}^{N} \nu_t^{(i_{k+1})} \right\}. \tag{4}$$

The inner sum is

$$\sum_{\substack{i_{k+1}=1 \\ \text{also distinct}}}^{N} \nu_t^{(i_{k+1})} = \left\{ \sum_{i=1}^{N} \nu_t^{(i)} - \sum_{i \in \{i_1,\dots,i_k\}} \nu_t^{(i)} \right\} \leq N-k \tag{5}$$

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

$$p_{\xi'\xi'}(t) \leq \frac{N-k}{(N)_k(N-k)} \sum_{\substack{i_1,\dots,i_k \\ \text{all distinct:} \\ \nu_t^{(i_1)},\dots,\nu_t^{(i_k)}>0}} \nu_t^{(i_1)} \cdots \nu_t^{(i_k)} = p_{\xi\xi}(t). \tag{6}$$

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.

$$\lim_{N\to\infty} \mathbb{E}\left[\prod_{r=1}^{\tau_N(t)} (1-p_r) \right] = e^{-\alpha t} \tag{7}$$

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

Proof.

Lower Bound

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

$$1 - p_t = p_{\Delta\Delta}(t) \geq 1 - \alpha(1 + O(N^{-1})) \left[\frac{B_n}{\alpha} D_N(t) + c_N(t) \right] \tag{8}$$

where $B_n > 0$. Hence, by a multinomial expansion,

$$\begin{aligned}
\prod_{r=1}^{\tau_N(t)} (1-p_r) &\geq \prod_{r=1}^{\tau_N(t)} \left\{ 1 - \alpha(1 + O(N^{-1}) \left[\frac{B_n}{\alpha} D_N(r) + c_N(r) \right] \right\} \\
&= 1 + \sum_{k=1}^{\infty} \sum_{\substack{r_1 < \dots < r_k \\ =1}}^{\tau_N(t)} \prod_{j=1}^{k} \left\{ -\alpha(1 + O(N^{-1})) \left[\frac{B_n}{\alpha} D_N(r_j) + c_N(r_j) \right] \right\} \\
&= 1 + \sum_{k=1}^{\infty} \left\{ -\alpha(1 + O(N^{-1})) \right\}^k \sum_{\substack{r_1 < \dots < r_k \\ =1}}^{\tau_N(t)} \prod_{j=1}^{k} \left\{ \frac{B_n}{\alpha} D_N(r_j) + c_N(r_j) \right\}
\end{aligned}$$

where the empty sum is taken to be zero. Taking expectations,

$$\mathbb{E}\left[\prod_{r=1}^{\tau_N(t)} (1-p_r) \right] \geq 1 + \sum_{k=1}^{\infty} \left\{ -\alpha(1 + O(N^{-1})) \right\}^k \mathbb{E}\left[\sum_{\substack{r_1 < \dots < r_k \\ =1}}^{\tau_N(t)} \prod_{j=1}^{k} \left\{ \frac{B_n}{\alpha} D_N(r_j) + c_N(r_j) \right\} \right] \tag{9}$$

(the infinite sum has only finitely many non-zero summands, since the inner sum is empty for $k > \tau_N(t)$, which justifies swapping the sum and expectation.) 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. From Koskela et al. (2018, Equation (8)),

$$
\begin{aligned}
\sum_{\substack{r_1<\cdots<r_k\\=1}}^{\tau_N(t)} \prod_{j=1}^{k} \left\{ \frac{B_n}{\alpha} D_N(r_j) + c_N(r_j) \right\} &\geq \sum_{\substack{r_1<\cdots<r_k\\=1}}^{\tau_N(t)} \prod_{j=1}^{k} c_N(r_j) \\
&\geq \frac{1}{k!} \left(\sum_{s=1}^{\tau_N(t)} c_N(s) \right)^k - \frac{1}{k!} \binom{k}{2} \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \right) \left(\sum_{s=1}^{\tau_N(t)} c_N(s) \right)^{k-2} \\
&\geq \frac{1}{k!} t^k - \frac{1}{k!} \binom{k}{2} (t+1)^{k-2} \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \right).
\end{aligned}
$$

Then

$$
\mathbb{E} \left[\sum_{\substack{r_1<\cdots<r_k\\=1}}^{\tau_N(t)} \prod_{j=1}^{k} \left\{ \frac{B_n}{\alpha} D_N(r_j) + c_N(r_j) \right\} \right] \geq \frac{1}{k!} t^k - \frac{1}{k!} \binom{k}{2} (t+1)^{k-2} \mathbb{E} \left[\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \right] \longrightarrow \frac{1}{k!} t^k \tag{10}
$$

as $N \to \infty$ using Brown et al. (2020, Equation (5)), via lemmata 1 and 3 therein.

Now for the upper bound.

$$
\begin{aligned}
\sum_{\substack{r_1<\cdots<r_k\\=1}}^{\tau_N(t)} \prod_{j=1}^{k} \left\{ \frac{B_n}{\alpha} D_N(r_j) + c_N(r_j) \right\} &= \frac{1}{k!} \sum_{\substack{r_1\neq\cdots\neq r_k\\ \text{all distinct}}}^{\tau_N(t)} \prod_{j=1}^{k} \left\{ \frac{B_n}{\alpha} D_N(r_j) + c_N(r_j) \right\} \\
&= \frac{1}{k!} \sum_{\substack{r_1\neq\cdots\neq r_k\\ \text{all distinct}}}^{\tau_N(t)} \sum_{\mathcal{I}\subseteq\{1,\dots,k\}} \left(\frac{B_n}{\alpha} \right)^{k-|\mathcal{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_{\mathcal{I}\subseteq\{1,\dots,k\}} \left(\frac{B_n}{\alpha} \right)^{k-|\mathcal{I}|} \sum_{\substack{r_1\neq\cdots\neq r_k\\ \text{all distinct}}}^{\tau_N(t)} \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_{I=0}^{k} \binom{k}{I} \left(\frac{B_n}{\alpha} \right)^{k-I} \sum_{\substack{r_1\neq\cdots\neq r_k\\ \text{all distinct}}}^{\tau_N(t)} \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{r_1\neq\cdots\neq r_k\\ \text{all distinct}}}^{\tau_N(t)} \left\{ \prod_{i=1}^{k} c_N(r_i) \right\} \\
&\quad + \frac{1}{k!} \sum_{I=0}^{k-1} \binom{k}{I} \left(\frac{B_n}{\alpha} \right)^{k-I} \sum_{\substack{r_1\neq\cdots\neq r_k\\ \text{all distinct}}}^{\tau_N(t)} \left\{ \prod_{i=1}^{I} c_N(r_i) \right\} \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} \left(\frac{B_n}{\alpha} \right)^{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!} \{ t + c_N(\tau_N(t)) \}^k + \frac{1}{k!} \sum_{I=0}^{k-1} \binom{k}{I} \left(\frac{B_n}{\alpha} \right)^{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\} \left\{ \sum_{r_k=1}^{\tau_N(t)} D_N(r_k) \right\} \\
&\leq \frac{1}{k!} \{ t + c_N(\tau_N(t)) \}^k + \frac{1}{k!} \sum_{I=0}^{k-1} \binom{k}{I} \left(\frac{B_n}{\alpha} \right)^{k-I} \left(\sum_{r=1}^{\tau_N(t)} c_N(r) \right)^{k-1} \left(\sum_{r=1}^{\tau_N(t)} D_N(r) \right) \\
&\leq \frac{1}{k!} \{ t + c_N(\tau_N(t)) \}^k + \frac{1}{k!} \sum_{I=0}^{k-1} \binom{k}{I} \left(\frac{B_n}{\alpha} \right)^{k-I} (t+1)^{k-1} \left(\sum_{r=1}^{\tau_N(t)} D_N(r) \right).
\end{aligned}
$$

Taking expectations,

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

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

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

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, $\mathbb{E}[c_N(s)^k]\le\mathbb{E}[c_N(s)]$ for all $k\ge 1$, and the expansion

$$\mathbb{E}\left[\frac{1}{k!}\{t+c_N(\tau_N(t))\}^k\right]=\mathbb{E}\left[\frac{1}{k!}\sum_{i=0}^{k}\binom{k}{i}t^i c_N(\tau_N(t))^{k-i}\right]=\frac{1}{k!}\left\{t^k+kt^{k-1}\mathbb{E}[c_N(\tau_N(t))]+\dots\right\}\longrightarrow\frac{1}{k!}t^k. \quad (11)$$

Combining these upper and lower limits, we conclude that

$$1+\sum_{k=1}^{\infty}\left\{-\alpha(1+O(N^{-1}))\right\}^k\mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k\\=1}}^{\tau_N(t)}\prod_{j=1}^{k}\left\{\frac{B_n}{\alpha}D_N(r_j)+c_N(r_j)\right\}\right]\longrightarrow 1+\sum_{k=1}^{\infty}(-\alpha)^k\frac{1}{k!}t^k=e^{-\alpha t} \quad (12)$$

as $N\to\infty$.

Upper Bound

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

$$1-p_t=p_{\Delta\Delta}(t)\le 1-\alpha(1+O(N^{-1}))\left[c_N(t)-\binom{n-1}{2}D_N(t)\right]. \quad (13)$$

A multinomial expansion as before yields

$$\prod_{r=1}^{\tau_N(t)}(1-p_r)\le 1+\sum_{k=1}^{\infty}\left\{-\alpha(1+O(N^{-1}))\right\}^k\sum_{\substack{r_1<\cdots<r_k\\=1}}^{\tau_N(t)}\prod_{j=1}^{k}\left\{c_N(r_j)-\binom{n-1}{2}D_N(r_j)\right\}. \quad (14)$$

Similarly to (10), an upper bound for the inner sum is

$$\sum_{\substack{r_1<\cdots<r_k\\=1}}^{\tau_N(t)}\prod_{j=1}^{k}\left\{c_N(r_j)-\binom{n-1}{2}D_N(r_j)\right\}\le\sum_{\substack{r_1<\cdots<r_k\\=1}}^{\tau_N(t)}\prod_{j=1}^{k}c_N(r_j)\le\frac{1}{k!}\left(\sum_{r=1}^{\tau_N(t)}c_N(r)\right)^k\le\frac{1}{k!}\{t+c_N(\tau_N(t))\}^k \quad (15)$$

with $\mathbb{E}[\{t+c_N(\tau_N(t))\}^k/k!]\longrightarrow t^k/k!$.

For the lower bound, stealing some results from the mega-align earlier,

$$\sum_{\substack{r_1<\cdots<r_k\\=1}}^{\tau_N(t)}\prod_{j=1}^{k}\left\{c_N(r_j)-\binom{n-1}{2}D_N(r_j)\right\}=\frac{1}{k!}\sum_{\substack{r_1\neq\cdots\neq r_k\\\text{all distinct}}}^{\tau_N(t)}\left\{\prod_{i=1}^{k}c_N(r_i)\right\}$$

$$+\frac{1}{k!}\sum_{I=0}^{k-1}\binom{k}{I}\left(-\binom{n-1}{2}\right)^{k-I}\sum_{\substack{r_1\neq\cdots\neq r_k\\\text{all distinct}}}\left\{\prod_{i=1}^{I}c_N(r_i)\right\}\left\{\prod_{j=I+1}^{k}D_N(r_j)\right\}.$$

First let us treat the first term:

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

$$\ge\frac{1}{k!}t^k-\frac{1}{k!}\binom{k}{2}(t+1)^{k-2}\left(\sum_{s=1}^{\tau_N(t)}c_N(s)^2\right)$$

and as before we have

$$\lim_{N\to\infty} \mathbb{E}\left[\frac{1}{k!}t^k - \frac{1}{k!}\binom{k}{2}(t+1)^{k-2}\left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right)\right] = \frac{1}{k!}t^k. \tag{16}$$

It remains to show that the expectation of the second term converges to zero.

$$\begin{aligned}
&\frac{1}{k!}\sum_{I=0}^{k-1}\binom{k}{I}\left(-\binom{n-1}{2}\right)^{k-I} \sum_{\substack{r_1\neq\cdots\neq r_k \\ \text{all distinct}}} \left\{\prod_{i=1}^{I} c_N(r_i)\right\}\left\{\prod_{j=I+1}^{k} D_N(r_j)\right\} \\
&\qquad = \frac{1}{k!}\sum_{\substack{I=0 \\ (k-I)\text{even}}}^{k-1}\binom{k}{I}\binom{n-1}{2}^{k-I} \sum_{\substack{r_1\neq\cdots\neq r_k \\ \text{all distinct}}} \left\{\prod_{i=1}^{I} c_N(r_i)\right\}\left\{\prod_{j=I+1}^{k} D_N(r_j)\right\} \\
&\qquad\quad - \frac{1}{k!}\sum_{\substack{I=0 \\ (k-I)\text{odd}}}^{k-1}\binom{k}{I}\binom{n-1}{2}^{k-I} \sum_{\substack{r_1\neq\cdots\neq r_k \\ \text{all distinct}}} \left\{\prod_{i=1}^{I} c_N(r_i)\right\}\left\{\prod_{j=I+1}^{k} D_N(r_j)\right\} \\
&\qquad \geq 0 - \frac{1}{k!}\sum_{\substack{I=0 \\ (k-I)\text{odd}}}^{k-1}\binom{k}{I}\binom{n-1}{2}^{k-I}(t+1)^{k-1}\left\{\sum_{s=1}^{\tau_N(t)} D_N(s)\right\}
\end{aligned}$$

using that $c_N(r), D_N(r) \geq 0$ for all r to bound the even terms below, and arguments from the mega-align earlier to bound the odd terms above. Taking the expectation and limit yields the desired result:

$$\begin{aligned}
&\lim_{N\to\infty} \mathbb{E}\left[\frac{1}{k!}\sum_{I=0}^{k-1}\binom{k}{I}\left(-\binom{n-1}{2}\right)^{k-I} \sum_{\substack{r_1\neq\cdots\neq r_k \\ \text{all distinct}}} \left\{\prod_{i=1}^{I} c_N(r_i)\right\}\left\{\prod_{j=I+1}^{k} D_N(r_j)\right\}\right] \\
&\qquad \geq -\frac{1}{k!}\sum_{\substack{I=0 \\ (k-I)\text{odd}}}^{k-1}\binom{k}{I}\binom{n-1}{2}^{k-I}(t+1)^{k-1}\lim_{N\to\infty}\mathbb{E}\left[\sum_{s=1}^{\tau_N(t)} D_N(s)\right] = 0.
\end{aligned}$$

Combining these upper and lower limits, we conclude that

$$1 + \sum_{k=1}^{\infty}\left\{-\alpha(1+O(N^{-1}))\right\}^k \mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k \\ =1}}^{\tau_N(t)} \prod_{j=1}^{k}\left\{c_N(r_j) - \binom{n-1}{2}D_N(r_j)\right\}\right] \longrightarrow 1 + \sum_{k=1}^{\infty}(-\alpha)^k\frac{1}{k!}t^k = e^{-\alpha t} \tag{17}$$

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 $e^{-\alpha t}$, so we're done. □

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.