Everything

Suzie Brown

February 27, 2020

1 Notation

2 Summary of KJJS proof

3 Relaxing assumptions of KJJS theorem

4 Extension to conditional SMC

4.1 The conditional SMC update

write something here...
referencing the CSMC algorithm...

4.2 Conditional SMC genealogies

Theorem 1. *Under the conditions of Koskela et al. (2018, Lemma 3), the genealogy of any n particles from a conditional SMC algorithm with multinomial resampling converges to Kingman's n-coalescent in the sense of finite-dimensional distributions, under the time-scaling defined in (**??**).*

Proof. In the derivation of (**??**) – (**??**) we will make extensive use of the formula for factorial moments of the multinomial distribution given in Mosimann (1962, p.67):

$$\mathbb{E}[(X_i)_a(X_j)_b] = (n)_{a+b}\, p_i^a p_j^b \tag{1}$$

where $(X_1, \dots X_k) \sim \text{Multinomial}(n, \mathbf{p})$. To apply this formula we need to write everything in terms of falling factorial powers. The required conversions are summarised in Table 1.

In standard SMC with multinomial resampling, the marginal offspring distributions, conditioned on the filtration $\mathcal{F}_{t-1}$ generated by the previous offspring counts, are

$$\nu_t^{(i)} \overset{d}{=} \text{Binomial}(N, w_t^{(i)}), \qquad i = 1, \dots, N$$

where $\nu_t^{(i)}$ is the number of offspring in generation $t+1$ of the ith particle in generation t, N is the number of particles and $w_t^{(i)}$ is the weight associated with the ith particle in generation t.

In conditional SMC we condition on the immortal trajectory surviving each resampling step. By exchangeability we can set without loss of generality that the immortal trajectory consists of particle 1 in each generation. At each resampling step, particle 1 must therefore choose particle 1 as its parent, while the remaining $N-1$ offspring are assigned multinomially to the N possible parents. The marginal offspring distributions are then

$$\tilde{\nu}_t^{(1)} \overset{d}{=} 1 + \text{Binomial}(N-1, w_t^{(1)})$$
$$\tilde{\nu}_t^{(i)} \overset{d}{=} \text{Binomial}(N-1, w_t^{(i)}), \qquad i = 2, \dots, N.$$

Algorithm 1 Conditional SMC with multinomial resampling

Require: $N, T, \mu, \{K_t\}, \{g_t\}, y_{0:T}, x^*_{0:T}$

for $i \in \{1, \dots, N\}$ **do**
    Sample $X_0^{(i)} \sim \mu(\cdot)$ ▷ initialise
end for
Sample $a_0^* \sim \text{Uniform}(\{1, \dots, N\})$
$X_0^{(a_0^*)} \leftarrow x_0^*$
for $i \in \{1, \dots, N\}$ **do**
    $w_0^{(i)} \leftarrow g_0(X_0^{(i)})$
end for
$W \leftarrow \sum_{j=1}^{N} w_0^{(j)}$
for $i \in \{1, \dots, N\}$ **do**
    $w_0^{(i)} \leftarrow \frac{1}{W} w_0^{(i)}$ ▷ normalise weights
end for
for $t \in \{0, \dots, T-1\}$ **do**
    Sample $a_t^{(1:N)} \sim \text{Categorical}(\{1, \dots, N\}, w_t^{(1:N)})$ ▷ resample particles
    Sample $a_{t+1}^* \sim \text{Uniform}(\{1, \dots, N\})$
    $a_t^{(a_{t+1}^*)} \leftarrow a_t^*$
    for $i \in \{1, \dots, N\}$ **do**
        Sample $X_{t+1}^{(i)} \sim K_{t+1}(X_t^{(a_t^{(i)})}, \cdot)$ ▷ propagate particles
    end for
    $X_{t+1}^{(a_{t+1}^*)} \leftarrow X_{t+1}^*$
    for $i \in \{1, \dots, N\}$ **do**
        $w_{t+1}^{(i)} \leftarrow g_{t+1}(X_t^{(a_t^{(i)})}, X_{t+1}^{(i)})$ ▷ calculate weights
    end for
    $W \leftarrow \sum_{j=1}^{N} w_{t+1}^{(j)}$
    for $i \in \{1, \dots, N\}$ **do**
        $w_{t+1}^{(i)} \leftarrow \frac{1}{W} w_{t+1}^{(i)}$ ▷ normalise weights
    end for
end for

x	$=$	$(x)_1$
x^2	$=$	$(x)_2 + (x)_1$
x^3	$=$	$(x)_3 + 3(x)_2 + (x)_1$
x^4	$=$	$(x)_4 + 6(x)_3 + 7(x)_2 + (x)_1$
xy	$=$	$(x)_1(y)_1$
x^2y	$=$	$(x)_2(y)_1 + (x)_1(y)_1$
xy^2	$=$	$(x)_1(y)_2 + (x)_1(y)_1$
x^2y^2	$=$	$(x)_2(y)_2 + (x)_2(y)_1 + (x)_1(y)_2 + (x)_1(y)_1$
$(x+1)_2$	$=$	$(x)_2 + 2(x)_1$
$(x+1)^2$	$=$	$(x)_2 + 3(x)_1 + 1$
$(x+1)_2(x+1)$	$=$	$(x)_3 + 5(x)_2 + 4(x)_1$
$(x+1)_2^2$	$=$	$(x)_4 + 8(x)_3 + 14(x)_2 + 4(x)_1$

Table 1: Conversion of ordinary powers into falling factorial powers

First let us consider the coalescence rate

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

For standard SMC the expected value is, using the tower rule,

$$\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}] = \frac{1}{(N)_2}\sum_{i=1}^{N}\mathbb{E}\left[\mathbb{E}[(\nu_t^{(i)})_2]|\mathcal{F}_{t-1}\right] = \frac{1}{(N)_2}\sum_{i=1}^{N}\mathbb{E}\left[(N)_2(w_t^{(i)})^2|\mathcal{F}_{t-1}\right] = \sum_{i=1}^{N}\mathbb{E}\left[(w_t^{(i)})^2|\mathcal{F}_{t-1}\right]$$

as stated in Koskela et al. (2018, Remark 3). In the case of conditional SMC we separate the first term (corresponding to the immortal particle) from the sum to get

$$\begin{aligned}
\mathbb{E}[\tilde{c}_N(t)|\mathcal{F}_{t-1}] &= \frac{1}{(N)_2}\sum_{i=1}^{N}\mathbb{E}\left[(\tilde{\nu}_t^{(i)})_2|\mathcal{F}_{t-1}\right] = \frac{1}{(N)_2}\mathbb{E}\left[(\tilde{\nu}_t^{(1)})_2|\mathcal{F}_{t-1}\right] + \frac{1}{(N)_2}\sum_{i=2}^{N}\mathbb{E}\left[(\tilde{\nu}_t^{(i)})_2|\mathcal{F}_{t-1}\right] \\
&= \frac{1}{(N)_2}\left\{(N-1)_2\mathbb{E}[(w_t^{(1)})^2|\mathcal{F}_{t-1}] + 2(N-1)\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}] + \sum_{i=2}^{N}(N-1)_2\mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]\right\} \\
&= \frac{(N-1)_2}{(N)_2}\sum_{i=1}^{N}\mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}] + \frac{2(N-1)}{(N)_2}\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}] \\
&= \frac{N-2}{N}\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}] + \frac{2}{N}\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}]
\end{aligned}$$

which gives us (**??**).

An upper bound on the rate of super-binary mergers is given by

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

In the standard case this quantity has expectation

$$\begin{aligned}
\mathbb{E}[D_N(t)|\mathcal{F}_{t-1}] = \frac{1}{N(N)_2}\sum_{i=1}^{N}&\left\{(N)_3\mathbb{E}[(w_t^{(i)})^3|\mathcal{F}_{t-1}] + 2(N)_2\mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]\right\} \\
&+ \frac{1}{N^2(N)_2}\sum_{i=1}^{N}\sum_{j\neq i}\left\{(N)_4\mathbb{E}[(w_t^{(i)})^2(w_t^{(j)})^2|\mathcal{F}_{t-1}] + (N)_3\mathbb{E}[(w_t^{(i)})^2 w_t^{(j)}|\mathcal{F}_{t-1}]\right\}
\end{aligned}$$

while in the conditional case, again separating the terms involving particle 1,

$$
\begin{aligned}
\tilde{D}_N(t) &= \frac{1}{N(N)_2}(\tilde{\nu}_t^{(1)})_2 \left(\tilde{\nu}_t^{(1)} + \frac{1}{N}\sum_{j\neq 1}(\tilde{\nu}_t^{(j)})^2 \right) + \frac{1}{N(N)_2}\sum_{i\neq 1}(\tilde{\nu}_t^{(i)})_2 \left(\tilde{\nu}_t^{(i)} + \frac{1}{N}(\tilde{\nu}_t^{(1)})^2 + \frac{1}{N}\sum_{1\neq j\neq i}(\tilde{\nu}_t^{(j)})^2 \right) \\
&= \frac{1}{N(N)_2}\left\{ (\tilde{\nu}_t^{(1)})_2 \nu_t^{(1)} + \frac{1}{N}\sum_{j\neq 1}(\tilde{\nu}_t^{(1)})_2(\tilde{\nu}_t^{(j)})^2 + \frac{1}{N}\sum_{i\neq 1}(\tilde{\nu}_t^{(i)})_2(\tilde{\nu}_t^{(1)})^2 \right\} \\
&\quad + \frac{1}{N(N)_2}\sum_{i\neq 1}\left\{ (\tilde{\nu}_t^{(i)})_2\tilde{\nu}_t^{(i)} + \frac{1}{N}\sum_{1\neq j\neq i}(\tilde{\nu}_t^{(i)})_2(\tilde{\nu}_t^{(j)})^2 \right\} \\
&= \frac{1}{N(N)_2}\left\{ (\tilde{\nu}_t^{(1)})_2\tilde{\nu}_t^{(1)} + \frac{1}{N}\sum_{i\neq 1}\Big((\tilde{\nu}_t^{(1)})_2(\tilde{\nu}_t^{(i)})^2 + (\tilde{\nu}_t^{(i)})_2(\tilde{\nu}_t^{(1)})^2 \Big) \right\} \\
&\quad + \frac{1}{N(N)_2}\sum_{i\neq 1}\left\{ (\tilde{\nu}_t^{(i)})_3 + 2(\tilde{\nu}_t^{(i)})_2 + \frac{1}{N}\sum_{1\neq j\neq i}\Big((\tilde{\nu}_t^{(i)})_2(\tilde{\nu}_t^{(j)})_2 + (\tilde{\nu}_t^{(i)})_2\tilde{\nu}_t^{(j)} \Big) \right\}
\end{aligned}
$$

and so by applying the moments from (1) and Table 1 we find the expectation

$$
\begin{aligned}
&\mathbb{E}[\tilde{D}_N(t)|\mathcal{F}_{t-1}] = \\
&= \frac{1}{N(N)_2}\Big\{(N-1)_3\mathbb{E}[(w_t^{(1)})^3|\mathcal{F}_{t-1}] + 5(N-1)_2\mathbb{E}[(w_t^{(1)})^2|\mathcal{F}_{t-1}] + 4(N-1)\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}]\Big\} \\
&\quad + \frac{1}{N^2(N)_2}\sum_{i=2}^N \Big\{2(N-1)_4\mathbb{E}[(w_t^{(1)})^2(w_t^{(i)})^2|\mathcal{F}_{t-1}] + (N-1)_3\mathbb{E}[(w_t^{(1)})^2 w_t^{(i)}|\mathcal{F}_{t-1}] \\
&\qquad + 5(N-1)_3\mathbb{E}[w_t^{(1)}(w_t^{(i)})^2|\mathcal{F}_{t-1}] + 2(N-1)_2\mathbb{E}[w_t^{(1)}w_t^{(i)}|\mathcal{F}_{t-1}] + (N-1)_2\mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]\Big\} \\
&\quad + \frac{1}{N(N)_2}\sum_{i=2}^N \Big\{(N-1)_3\mathbb{E}[(w_t^{(i)})^3|\mathcal{F}_{t-1}] + 2(N-1)_2\mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]\Big\} \\
&\quad + \frac{1}{N^2(N)_2}\sum_{i=2}^N \sum_{1\neq j\neq i} \Big\{(N-1)_4\mathbb{E}[(w_t^{(i)})^2(w_t^{(j)})^2|\mathcal{F}_{t-1}] + (N-1)_3\mathbb{E}[(w_t^{(i)})^2 w_t^{(j)}|\mathcal{F}_{t-1}]\Big\} \\
&= \frac{1}{N(N)_2}\sum_{i=1}^N \Big\{(N-1)_3\mathbb{E}[(w_t^{(i)})^3|\mathcal{F}_{t-1}] + 2(N-1)_2\mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]\Big\} \\
&\quad + \frac{1}{N^2(N)_2}\sum_{i=1}^N \sum_{j\neq i} \Big\{(N-1)_4\mathbb{E}[(w_t^{(i)})^2(w_t^{(j)})^2|\mathcal{F}_{t-1}] + (N-1)_3\mathbb{E}[(w_t^{(i)})^2 w_t^{(j)}|\mathcal{F}_{t-1}]\Big\} \\
&\quad + \frac{1}{N(N)_2}\Big\{3(N-1)_2\mathbb{E}[(w_t^{(1)})^2|\mathcal{F}_{t-1}] + 4(N-1)\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}]\Big\} \\
&\quad + \frac{1}{N^2(N)_2}\sum_{i=2}^N \Big\{4(N-1)_3\mathbb{E}[w_t^{(1)}(w_t^{(i)})^2|\mathcal{F}_{t-1}] + 2(N-1)_2\mathbb{E}[w_t^{(1)}w_t^{(i)}|\mathcal{F}_{t-1}] + (N-1)_2\mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]\Big\} \\
&\le \mathbb{E}[D_N(t)|\mathcal{F}_{t-1}] + \frac{1}{N(N)_2}\Big\{3(N-1)_2\mathbb{E}[(w_t^{(1)})^2|\mathcal{F}_{t-1}] + 4(N-1)\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}]\Big\} \\
&\quad + \frac{1}{N^2(N)_2}\sum_{i=2}^N \Big\{4(N-1)_3\mathbb{E}[w_t^{(1)}(w_t^{(i)})^2|\mathcal{F}_{t-1}] + 2(N-1)_2\mathbb{E}[w_t^{(1)}w_t^{(i)}|\mathcal{F}_{t-1}] + (N-1)_2\mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]\Big\} \\
&\le \mathbb{E}[D_N(t)|\mathcal{F}_{t-1}] + \frac{3}{N}\mathbb{E}[(w_t^{(1)})^2|\mathcal{F}_{t-1}] + \frac{4}{N^2}\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}] \\
&\quad + \frac{4}{N}\sum_{i=2}^N \mathbb{E}[w_t^{(1)}(w_t^{(i)})^2|\mathcal{F}_{t-1}] + \frac{2}{N^2}\sum_{i=2}^N \mathbb{E}[w_t^{(1)}w_t^{(i)}|\mathcal{F}_{t-1}] + \frac{1}{N^2}\sum_{i=2}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]
\end{aligned}
$$

The second line of the first equality relies on multiplying the relevant terms in Table 1. For the second equality we recombine the terms in particle 1 into the sum. The inequalities follow by bounding e.g. $N-1$ by N, and identifying the first two lines with $\mathbb{E}[D_N(t)|\mathcal{F}_{t-1}]$. This gives us the inequality (**??**).

We also need control of the squared coalescence rate:

$$
c_N(t)^2 = \frac{1}{(N)_2^2}\left(\sum_{i=1}^N (\nu_t^{(i)})_2\right)^2 = \frac{1}{(N)_2^2}\left\{\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_2^2] + \sum_{i=1}^N\sum_{j\neq i}\mathbb{E}[(\nu_t^{(i)})_2(\nu_t^{(j)})_2]\right\}
$$

A bound on its expected value is proved in Koskela et al. (2018), but here we will use a different, more explicit bound to allow direct comparison between the standard and conditional cases. For standard SMC we have:

$$
\begin{aligned}
\mathbb{E}[c_N(t)^2|\mathcal{F}_{t-1}] &= \frac{1}{(N)_2^2}\left\{(N)_4 \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^4|\mathcal{F}_{t-1}] + 4(N)_3 \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^3|\mathcal{F}_{t-1}] + 2(N)_2 \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]\right\} \\
&\quad + \frac{1}{(N)_2^2}(N)_4 \sum_{i=1}^N \sum_{j\neq i} \mathbb{E}[(w_t^{(i)})^2 (w_t^{(j)})^2|\mathcal{F}_{t-1}] \\
&= \frac{1}{(N)_2}\left\{(N-2)_2 \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^4|\mathcal{F}_{t-1}] + 4(N-2) \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^3|\mathcal{F}_{t-1}] + 2 \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]\right. \\
&\qquad \left. + (N-2)_2 \sum_{i=1}^N \sum_{j\neq i} \mathbb{E}[(w_t^{(i)})^2 (w_t^{(j)})^2|\mathcal{F}_{t-1}]\right\}
\end{aligned}
$$

For conditional SMC, we again separate the terms involving particle 1:

$$
\begin{aligned}
\tilde{c}_N(t)^2 &= \frac{1}{(N)_2^2}\left\{\sum_{i=1}^N \mathbb{E}[(\tilde{\nu}_t^{(i)})_2^2] + \sum_{i=1}^N \sum_{j\neq i} \mathbb{E}[(\tilde{\nu}_t^{(i)})_2 (\tilde{\nu}_t^{(j)})_2]\right\} \\
&= \frac{1}{(N)_2^2}\left\{\sum_{i=2}^N \mathbb{E}[(\tilde{\nu}_t^{(i)})_2^2] + \mathbb{E}[(\tilde{\nu}_t^{(1)})_2^2] + \sum_{i=2}^N \sum_{1\neq j\neq i} \mathbb{E}[(\tilde{\nu}_t^{(i)})_2 (\tilde{\nu}_t^{(j)})_2] + 2\sum_{i=2}^N \mathbb{E}[(\tilde{\nu}_t^{(1)})_2 (\tilde{\nu}_t^{(i)})_2]\right\}
\end{aligned}
$$

and then use the same techniques as for $\tilde{D}_N(t)$ to calculate the expectation:

$$
\begin{aligned}
&\mathbb{E}[\tilde{c}_N(t)^2|\mathcal{F}_{t-1}] = \\
&= \frac{1}{(N)_2^2}\left\{(N-1)_4\sum_{i=2}^{N}\mathbb{E}[(w_t^{(i)})^4|\mathcal{F}_{t-1}] + 4(N-1)_3\sum_{i=2}^{N}\mathbb{E}[(w_t^{(i)})^3|\mathcal{F}_{t-1}] + 2(N-1)_2\sum_{i=2}^{N}\mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]\right\} \\
&\quad + \frac{1}{(N)_2^2}\left\{(N-1)_4\mathbb{E}[(w_t^{(1)})^4|\mathcal{F}_{t-1}] + 8(N-1)_3\mathbb{E}[(w_t^{(1)})^3|\mathcal{F}_{t-1}] + 14(N-1)_2\mathbb{E}[(w_t^{(1)})^2|\mathcal{F}_{t-1}]\right\} \\
&\quad + \frac{1}{(N)_2^2}4(N-1)\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}] + \frac{1}{(N)_2^2}(N-1)_4\sum_{i=2}^{N}\sum_{1\neq j\neq i}\mathbb{E}[(w_t^{(i)})^2(w_t^{(j)})^2|\mathcal{F}_{t-1}] \\
&\quad + \frac{2}{(N)_2^2}\sum_{i=2}^{N}\Big((N-1)_4\mathbb{E}[(w_t^{(1)})^2(w_t^{(i)})^2|\mathcal{F}_{t-1}] + 2(N-1)_3\mathbb{E}[w_t^{(1)}(w_t^{(i)})^2|\mathcal{F}_{t-1}]\Big) \\
&= \frac{1}{(N)_2^2}\left\{(N-1)_4\sum_{i=1}^{N}\mathbb{E}[(w_t^{(i)})^4|\mathcal{F}_{t-1}] + 4(N-1)_3\sum_{i=1}^{N}\mathbb{E}[(w_t^{(i)})^3|\mathcal{F}_{t-1}] + 2(N-1)_2\sum_{i=1}^{N}\mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]\right\} \\
&\quad + \frac{(N-1)_4}{(N)_2^2}\sum_{i=1}^{N}\sum_{j\neq i}\mathbb{E}[(w_t^{(i)})^2(w_t^{(j)})^2|\mathcal{F}_{t-1}] \\
&\quad + \frac{1}{(N)_2^2}\left\{4(N-1)_3\mathbb{E}[(w_t^{(1)})^3|\mathcal{F}_{t-1}] + 12(N-1)_2\mathbb{E}[(w_t^{(1)})^2|\mathcal{F}_{t-1}] + 4(N-1)\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}]\right\} \\
&\quad + \frac{1}{(N)_2^2}\left\{4(N-1)_3\sum_{i=2}^{N}\mathbb{E}[w_t^{(1)}(w_t^{(i)})^2|\mathcal{F}_{t-1}]\right\} \\
&\leq \mathbb{E}[c_N(t)^2|\mathcal{F}_{t-1}] + \frac{1}{(N)_2^2}\left\{4(N-1)_3\mathbb{E}[(w_t^{(1)})^3|\mathcal{F}_{t-1}] + 12(N-1)_2\mathbb{E}[(w_t^{(1)})^2|\mathcal{F}_{t-1}]\right\} \\
&\quad + \frac{1}{(N)_2^2}\left\{4(N-1)\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}] + 4(N-1)_3\sum_{i=2}^{N}\mathbb{E}[w_t^{(1)}(w_t^{(i)})^2|\mathcal{F}_{t-1}]\right\} \\
&\leq \mathbb{E}[c_N(t)^2|\mathcal{F}_{t-1}] + \frac{4}{N}\mathbb{E}[(w_t^{(1)})^3|\mathcal{F}_{t-1}] + \frac{12}{N^2}\mathbb{E}[(w_t^{(1)})^2|\mathcal{F}_{t-1}] + \frac{4}{N(N)_2}\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}] \\
&\quad + \frac{4}{N}\sum_{i=2}^{N}\mathbb{E}[w_t^{(1)}(w_t^{(i)})^2|\mathcal{F}_{t-1}]
\end{aligned}
$$

The conditions (18) and (19) of Koskela et al. (2018, Lemma 3) give us control over the weights so that we have $w_t^{(i)} = O(1)$ for all i. Under these conditions, in the limit as $N \to \infty$, the three modified expectations derived above simplify to:

$$
\begin{aligned}
&\mathbb{E}[\tilde{c}_N(t)|\mathcal{F}_{t-1}] \leq \mathbb{E}[c_N(t)|\mathcal{F}_{t-1}] + O(N^{-2}) \\
&\mathbb{E}[\tilde{D}_N(t)|\mathcal{F}_{t-1}] \leq \mathbb{E}[D_N(t)|\mathcal{F}_{t-1}] + O(N^{-3}) \\
&\mathbb{E}[\tilde{c}_N(t)^2|\mathcal{F}_{t-1}] \leq \mathbb{E}[c_N(t)^2|\mathcal{F}_{t-1}] + O(N^{-3})
\end{aligned}
$$

This shows that each of these quantities for conditional SMC is bounded above by the corresponding standard SMC quantity, plus some vanishing error term. This will allow us to apply Koskela et al. (2018, Theorem 1), as we will show in the following.

Next we apply the result of Koskela et al. (2018, Lemma 3), so that for our modified quantity $\tilde{c}_N(t)$ we have the upper bound:

$$\begin{aligned}
\mathbb{E}[\tilde{c}_N(t)|\mathcal{F}_{t-1}] &= \frac{N-2}{N}\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}] + \frac{2}{N}\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}] \\
&\le \mathbb{E}[c_N(t)|\mathcal{F}_{t-1}] + \frac{2}{N}\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}] \\
&\le \frac{a^4}{N\varepsilon^4} + \frac{2}{N}\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}] \\
&= \frac{a^4}{N\varepsilon^4} + O(N^{-2}) \qquad (2)
\end{aligned}$$

and lower bound:

$$\begin{aligned}
\mathbb{E}[\tilde{c}_N(t)|\mathcal{F}_{t-1}] &= \frac{N-2}{N}\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}] + \frac{2}{N}\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}] \\
&\ge \frac{N-2}{N}\frac{\varepsilon^4}{Na^4} + O(N^{-2}) \\
&= \frac{\varepsilon^4}{Na^4} - \frac{2\varepsilon^4}{N^2a^4} + O(N^{-2}) \\
&= \frac{\varepsilon^4}{Na^4} + O(N^{-2}) \qquad (3)
\end{aligned}$$

corresponding to (22) in Koskela et al. (2018), except for the addition of a vanishing error term. Furthermore, we obtain for the other quantities (where the constant C may change from one line to the next):

$$\begin{aligned}
\mathbb{E}[\tilde{D}_N(t)|\mathcal{F}_{t-1}] &\le \mathbb{E}[D_N(t)|\mathcal{F}_{t-1}] + O(N^{-3}) \\
&\le \frac{C}{N}\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}] + O(N^{-3}) \\
&= \frac{C}{N}\mathbb{E}[\tilde{c}_N(t)|\mathcal{F}_{t-1}] + O(N^{-3}) \qquad (4)
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}[\tilde{c}_N(t)^2|\mathcal{F}_{t-1}] &\le \mathbb{E}[c_N(t)^2|\mathcal{F}_{t-1}] + O(N^{-3}) \\
&\le \frac{C}{N}\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}] + O(N^{-3}) \\
&= \frac{C}{N}\mathbb{E}[\tilde{c}_N(t)|\mathcal{F}_{t-1}] + O(N^{-3}) \qquad (5)
\end{aligned}$$

according to equations (20) and (21) in Koskela et al. (2018), again with additional error terms.

Now let us define the time-scaling:

$$\tilde{\tau}_N(t) := \min\left\{ s \ge 1 : \sum_{r=1}^{s} \tilde{c}_N(r) \ge t \right\}$$

which is a generalised inverse of $\tilde{c}_N(t)$ and thus satisfies the property:

$$t - s - 1 \le \sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \tilde{c}_N(r) \le t - s + 1. \qquad (6)$$

We are finally ready to verify the conditions of Koskela et al. (2018, Theorem 1). The conditions are the following.

(Standing Assumption) The conditional distribution of parental indices $a_t^{(1:N)}$ given offspring counts $v_t^{(1:N)}$ is uniform over all valid assignments.

(A) $\lim_{N\to\infty} \mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \tilde{D}_N(r)\right] = 0$

(B) $\lim_{N\to\infty} \mathbb{E}[\tilde{c}_N(t)] = 0$

(C) $\lim_{N\to\infty} \mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \tilde{c}_N(r)^2\right] = 0$

(D) $\mathbb{E}[\tilde{\tau}_N(t) - \tilde{\tau}_N(s)] \le C_{t,s} N$

These five conditions are verified below.

(Standing Assumption) This holds by the exchangeability of offspring assignments arising from Algorithm 1.

(B) Using (2) and applying the tower rule, we find

$$\mathbb{E}[\tilde{c}_N(t)] = \mathbb{E}[\mathbb{E}[\tilde{c}_N(t)|\mathcal{F}_{t-1}]] \le \frac{a^4}{N\varepsilon^4} + O(N^{-2}) \overset{N\to\infty}{\longrightarrow} 0$$

(C) Using Koskela et al. (2018, Lemma 2) along with (5) and the upper bound in (6),

$$\begin{aligned}
\mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \tilde{c}_N(r)^2\right] &= \mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \mathbb{E}[\tilde{c}_N(r)^2|\mathcal{F}_{t-1}]\right] \le \mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \left(\frac{C}{N}\mathbb{E}[\tilde{c}_N(t)|\mathcal{F}_{t-1}] + O(N^{-3})\right)\right] \\
&= \frac{C}{N}\mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \mathbb{E}[\tilde{c}_N(t)|\mathcal{F}_{t-1}]\right] + O(N^{-2}) = \frac{C}{N}\mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \tilde{c}_N(r)\right] + O(N^{-2}) \\
&\le \frac{C}{N}(t-s+1) + O(N^{-2}) \overset{N\to\infty}{\longrightarrow} 0
\end{aligned}$$

(A) The above calculation replacing (5) with (4) yields

$$\begin{aligned}
\mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \tilde{D}_N(r)\right] &= \mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \mathbb{E}[\tilde{D}_N(r)|\mathcal{F}_{t-1}]\right] \\
&\le \mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \left(\frac{C}{N}\mathbb{E}[\tilde{c}_N(t)|\mathcal{F}_{t-1}] + O(N^{-3})\right)\right] \overset{N\to\infty}{\longrightarrow} 0
\end{aligned}$$

(D) Using (3), the upper bound in (6) and Koskela et al. (2018, Lemma 2),

$$\begin{aligned}
\mathbb{E}[\tilde{\tau}_N(t) - \tilde{\tau}_N(s)] = \mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} 1\right] &= \mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \frac{\mathbb{E}[\tilde{c}_N(r)|\mathcal{F}_{t-1}]}{\mathbb{E}[\tilde{c}_N(r)|\mathcal{F}_{t-1}]}\right] \le \mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \frac{\mathbb{E}[\tilde{c}_N(r)|\mathcal{F}_{t-1}]}{\frac{\varepsilon^4}{Na^4} + O(N^{-2})}\right] \\
&= \frac{1}{\frac{\varepsilon^4}{Na^4} + O(N^{-2})}\mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \mathbb{E}[\tilde{c}_N(r)|\mathcal{F}_{t-1}]\right] = \frac{1}{\frac{\varepsilon^4}{Na^4} + O(N^{-2})}\mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \tilde{c}_N(r)\right] \\
&\le \frac{t-s+1}{\frac{\varepsilon^4}{Na^4} + O(N^{-2})} = \frac{(t-s+1)a^4 N}{\varepsilon^4 + O(N^{-1})} = (t-s+1)\frac{a^4}{\varepsilon^4}N + O(1)
\end{aligned}$$

where the last equality follows by a Taylor expansion of $(\frac{\varepsilon^4}{Na^4} + O(N^{-2}))^{-1}$.
Similarly we derive a lower bound using (2), the lower bound in (6) and Koskela et al. (2018, Lemma

2):

$$\mathbb{E}[\tilde{\tau}_N(t)-\tilde{\tau}_N(s)] = \mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} 1\right] = \mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \frac{\mathbb{E}[\tilde{c}_N(r)|\mathcal{F}_{t-1}]}{\mathbb{E}[\tilde{c}_N(r)|\mathcal{F}_{t-1}]}\right] \geq \mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \frac{\mathbb{E}[\tilde{c}_N(r)|\mathcal{F}_{t-1}]}{\frac{a^4}{N\varepsilon^4}+O(N^{-2})}\right]$$

$$= \frac{1}{\frac{a^4}{N\varepsilon^4}+O(N^{-2})}\mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \mathbb{E}[\tilde{c}_N(r)|\mathcal{F}_{t-1}]\right] = \frac{1}{\frac{a^4}{N\varepsilon^4}+O(N^{-2})}\mathbb{E}\left[\sum_{r=\tilde{\tau}_N(s)+1}^{\tilde{\tau}_N(t)} \tilde{c}_N(r)\right]$$

$$\geq \frac{t-s-1}{\frac{a^4}{N\varepsilon^4}+O(N^{-2})} = \frac{(t-s-1)\varepsilon^4 N}{a^4+O(N^{-1})} = (t-s-1)\frac{\varepsilon^4}{a^4}N + O(1)$$

Therefore we have as required

$$\mathbb{E}[\tilde{\tau}_N(t)-\tilde{\tau}_N(s)] \sim C_{t,s}N$$

as $N \to \infty$.

□

4.3 With the updated theorem assumptions...

Theorem 2. *Under the time scaling of Koskela et al. (2018, Theorem 1) and the conditions of Koskela et al. (2018, Lemma 3), genealogies of SMC algorithms with multinomial resampling converge to Kingman's n-coalescent in the sense of finite-dimensional distributions as $N \to \infty$.*

Proof. The standing assumption holds by exchangeability of the Multinomial distribution. We also need to show that there exists a deterministic sequence $(b_N)_{N\in\mathbb{N}}$ such that $\lim_{N\to\infty} b_N = 0$ and

$$\frac{\frac{1}{(N)_3}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_3|\mathcal{F}_{t-1}}{\frac{1}{(N)_2}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_2|\mathcal{F}_{t-1}]} \leq b_N \tag{7}$$

for all $N \in \mathbb{N}$. For the denominator, we apply Koskela et al. (2018, Lemma 3) directly to obtain, for some constants $a \geq \varepsilon \geq 0$,

$$\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_2|\mathcal{F}_{t-1}] \geq \frac{(N)_2\varepsilon^4}{Na^4}$$

For the numerator, we use that $\nu_t^{(i)} \longrightarrow (\nu_t^{(i)})_3$ is $\{i\}$-increasing, along with the argument from the proof of Koskela et al. (2018, Lemma 3), to obtain

$$\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_3|\mathcal{F}_{t-1}] \leq \sum_{i=1}^N \mathbb{E}[(\tilde{a}_t^{(i)})_3]$$

where

$$\tilde{a}_t^{(j)} \sim^{iid} \text{Categorical}\left(\left(\frac{a}{\varepsilon}\right)^{\mathbb{I}_{\{i=1\}}-\mathbb{I}_{\{i\neq 1\}}}, \left(\frac{a}{\varepsilon}\right)^{\mathbb{I}_{\{i=2\}}-\mathbb{I}_{\{i\neq 2\}}}, \ldots, \left(\frac{a}{\varepsilon}\right)^{\mathbb{I}_{\{i=N\}}-\mathbb{I}_{\{i\neq N\}}}\right) \tag{8}$$

We can calculate the expectation for $\tilde{a}_t^{(1:N)}$, so we obtain

$$\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_3|\mathcal{F}_{t-1}] \leq (N)_3\left(\frac{1}{(N-1)\varepsilon/a + a/\varepsilon}\right)^3 \left[(N-1)\left(\frac{\varepsilon}{a}\right)^3 + \left(\frac{a}{\varepsilon}\right)^3\right]$$

$$= \frac{(N)_3}{((N-1)\varepsilon/a + a/\varepsilon)^3}\left(a^6 + (N-1)\varepsilon^6\right)$$

$$\leq \frac{(N)_3}{N^3\varepsilon^6}(Na^6) = \frac{(N)_3}{N^2}\frac{a^6}{\varepsilon^6}$$

where the last inequality follows because $a \geq \varepsilon$. Putting these together, we can bound the ratio above by

$$b_N := \frac{(N)_3 \frac{(N)_3}{N^2} \frac{a^6}{\varepsilon^6}}{(N)_2 \frac{(N)_2 \varepsilon^4}{N a^4}} = \frac{1}{N} \frac{a^{10}}{\varepsilon^{10}} \quad \overset{N \to \infty}{\longrightarrow} 0 \tag{9}$$

We conclude the result by applying Theorem ? [the KJJS Thm1 with updated assns]. $\square$

Theorem 3. *Under the time scaling of Koskela et al. (2018, Theorem 1) and the conditions of Koskela et al. (2018, Lemma 3), genealogies of conditional SMC algorithms with multinomial resampling converge to Kingman's n-coalescent in the sense of finite-dimensional distributions as $N \to \infty$.*

Proof. Denote the vector of particle weights $w_t^{(1:N)} = (w_t^{(1)}, w_t^{(2)}, \dots, w_t^{(N)})$. Let $\tilde{\nu}_t^{(1:N)} = (\tilde{\nu}_t^{(1)}, \tilde{\nu}_t^{(2)}, \dots, \tilde{\nu}_t^{(N)})$ denote the associated offspring counts under conditional SMC. Assuming WLOG that the immortal particle is particle 1, the offspring counts are distributed according to

$$\tilde{\nu}_t^{(1:N)} \mid w_t^{(1:N)} \overset{d}{=} (1, 0, 0, \dots, 0) + \text{Multinomial}(N-1, w_t^{(1:N)})$$

The standing assumption holds by exchangeability of the offspring assignments in Algorithm 1. As before, we need to show that

$$\frac{\frac{1}{(N)_3} \sum_{i=1}^N \mathbb{E}[(\tilde{\nu}_t^{(i)})_3 | \mathcal{F}_{t-1}}{\frac{1}{(N)_2} \sum_{i=1}^N \mathbb{E}[(\tilde{\nu}_t^{(i)})_2 | \mathcal{F}_{t-1}]} \leq b_N \tag{10}$$

for some deterministic sequence $b_N \lim_{N\to\infty} 0$. For the denominator, we find

$$\begin{aligned}
\sum_{i=1}^N \mathbb{E}[(\tilde{\nu}_t^{(i)})_2 | \mathcal{F}_{t-1} &= \mathbb{E}\left[(\tilde{\nu}_t^{(1)})_2 | \mathcal{F}_{t-1}\right] + \sum_{i=2}^N \mathbb{E}\left[(\tilde{\nu}_t^{(i)})_2 | \mathcal{F}_{t-1}\right] \\
&= (N-1)_2 \mathbb{E}[(w_t^{(1)})^2 | \mathcal{F}_{t-1}] + 2(N-1)\mathbb{E}[w_t^{(1)} | \mathcal{F}_{t-1}] + \sum_{i=2}^N (N-1)_2 \mathbb{E}[(w_t^{(i)})^2 | \mathcal{F}_{t-1}] \\
&= (N-1)_2 \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2 | \mathcal{F}_{t-1}] + 2(N-1)\mathbb{E}[w_t^{(1)} | \mathcal{F}_{t-1}]
\end{aligned}$$

using that $(X+1)_2 = (X)_2 + 2(X)_1$ and the factorial moments of the Multinomial distribution (Mosimann, 1962). For the numerator, we have

$$\begin{aligned}
\sum_{i=1}^N \mathbb{E}[(\tilde{\nu}_t^{(i)})_3 | \mathcal{F}_{t-1} &= \mathbb{E}\left[(\tilde{\nu}_t^{(1)})_3 | \mathcal{F}_{t-1}\right] + \sum_{i=2}^N \mathbb{E}\left[(\tilde{\nu}_t^{(i)})_3 | \mathcal{F}_{t-1}\right] \\
&= (N-1)_3 \mathbb{E}[(w_t^{(1)})^3 | \mathcal{F}_{t-1}] + 3(N-1)_2 \mathbb{E}[(w_t^{(1)})^2 | \mathcal{F}_{t-1}] + \sum_{i=2}^N (N-1)_3 \mathbb{E}[(w_t^{(i)})^3 | \mathcal{F}_{t-1}] \\
&= (N-1)_3 \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^3 | \mathcal{F}_{t-1}] + 3(N-1)_2 \mathbb{E}[(w_t^{(1)})^2 | \mathcal{F}_{t-1}]
\end{aligned}$$

using similarly that $(X+1)_3 = (X)_3 + 3(X)_2$. Combining these expressions, the ratio in (10) becomes

$$\begin{aligned}
\frac{\frac{1}{(N)_3} \sum_{i=1}^N \mathbb{E}[(\tilde{\nu}_t^{(i)})_3 | \mathcal{F}_{t-1}}{\frac{1}{(N)_2} \sum_{i=1}^N \mathbb{E}[(\tilde{\nu}_t^{(i)})_2 | \mathcal{F}_{t-1}]} &= \frac{1}{N-2} \frac{(N-1)_3 \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^3 | \mathcal{F}_{t-1}] + 3(N-1)_2 \mathbb{E}[(w_t^{(1)})^2 | \mathcal{F}_{t-1}]}{(N-1)_2 \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2 | \mathcal{F}_{t-1}] + 2(N-1)\mathbb{E}[w_t^{(1)} | \mathcal{F}_{t-1}]} \\
&\leq \frac{1}{N-2} \frac{(N-1)_3 \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^3 | \mathcal{F}_{t-1}]}{(N-1)_2 \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2 | \mathcal{F}_{t-1}]} + \frac{1}{N-2} \frac{3(N-1)_2 \mathbb{E}[(w_t^{(1)})^2 | \mathcal{F}_{t-1}]}{(N-1)_2 \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2 | \mathcal{F}_{t-1}]} \\
&\leq \frac{1}{N-2} \frac{(N-1)_3 \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^3 | \mathcal{F}_{t-1}]}{(N-1)_2 \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2 | \mathcal{F}_{t-1}]} + \frac{1}{N-2} \frac{3(N-1)_2 \mathbb{E}[(w_t^{(1)})^2 | \mathcal{F}_{t-1}]}{(N-1)_2 \mathbb{E}[(w_t^{(1)})^2 | \mathcal{F}_{t-1}]} \\
&< \frac{1}{N-2} \frac{\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_3 | \mathcal{F}_{t-1}]}{\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_2 | \mathcal{F}_{t-1}]} + \frac{3}{N-2}
\end{aligned}$$

where the last inequality comes from $\frac{(N-1)_3}{(N-1)_2} = N-3 < N-2 = \frac{(N)_3}{(N)_2}$. Then, using Theorem 2, we can bound this by

$$b_N := \frac{1}{N}\frac{a^6}{\varepsilon^6} + \frac{3}{N-2} \quad \xrightarrow{N\to\infty} 0$$

as required, where a and ε are constants defined in the conditions of Koskela et al. (2018, Lemma 3). We conclude the result by applying Theorem ? [the KJJS Thm1 with updated assns]. □

5 Residual resampling

The offspring counts are sampled according to:

$$\begin{aligned}
&\nu_t^{(i)} = \lfloor N w_t^{(i)} \rfloor + X_i \\
&X_i \sim \text{Multinomial}(N-k, (\bar{w}_t^{(1)}, \dots, \bar{w}_t^{(N)}))
\end{aligned}$$

where $k := \sum_{i=1}^N \lfloor N w_t^{(i)} \rfloor$ is the number of offspring assigned deterministically, and $\bar{w}_t^{(i)} := \frac{N w_t^{(i)} - \lfloor N w_t^{(i)} \rfloor}{N-k}$ are the residual weights. Let us also define the residuals $r_i := N w_t^{(i)} - \lfloor N w_t^{(i)} \rfloor$. So $\sum_{i=1}^N r_i = N - k$.

5.1 Coalescence rate

The coalescence rate is defined as

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

We will use $c_N^m(t)$ and $c_N^r(t)$ to denote the coalescence rates with multinomial and residual resampling respectively. The expectation then comes out as

$$\begin{aligned}
\mathbb{E}[(\nu_t^{(i)})_2 | \mathcal{F}_{t-1}] &= \mathbb{E}[(\nu_t^{(i)})^2 | \mathcal{F}_{t-1}] - \mathbb{E}[\nu_t^{(i)} | \mathcal{F}_{t-1}] \\
&= \mathbb{E}[\lfloor N w_t^{(i)} \rfloor^2 | \mathcal{F}_{t-1}] + 2\mathbb{E}[\lfloor N w_t^{(i)} \rfloor r_i | \mathcal{F}_{t-1}] + \mathbb{E}\left[r_i \left(1 - \frac{r_i}{N-k} + r_i \right) | \mathcal{F}_{t-1} \right] - \mathbb{E}[N w_t^{(i)} | \mathcal{F}_{t-1}] \\
&= \mathbb{E}[\lfloor N w_t^{(i)} \rfloor^2 | \mathcal{F}_{t-1}] - \mathbb{E}[\lfloor N w_t^{(i)} \rfloor | \mathcal{F}_{t-1}] + 2\mathbb{E}[\lfloor N w_t^{(i)} \rfloor r_i | \mathcal{F}_{t-1}] + \mathbb{E}\left[r_i^2 \left(1 - \frac{1}{N-k} \right) | \mathcal{F}_{t-1} \right] \\
&= \mathbb{E}[(N w_t^{(i)})^2 | \mathcal{F}_{t-1}] - \mathbb{E}[\lfloor N w_t^{(i)} \rfloor | \mathcal{F}_{t-1}] - \mathbb{E}\left[\frac{r_i^2}{N-k} | \mathcal{F}_{t-1} \right]
\end{aligned}$$

so we get

$$\begin{aligned}
\mathbb{E}[c_N^r(t) | \mathcal{F}_{t-1}] &= \frac{1}{(N)_2} \sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_2 | \mathcal{F}_{t-1}] \\
&= \frac{N}{N-1} \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2 | \mathcal{F}_{t-1}] - \frac{1}{(N)_2} \sum_{i=1}^N \mathbb{E}\left[\frac{r_i^2}{N-k} | \mathcal{F}_{t-1} \right] - \frac{1}{(N)_2} \mathbb{E}[k | \mathcal{F}_{t-1}] \\
&= \mathbb{E}[c_N^m(t) | \mathcal{F}_{t-1}] \left(1 + \frac{1}{N-1} \right) - \frac{1}{(N)_2} \mathbb{E}\left[\frac{\sum_{i=1}^N (N w_t^{(i)} - \lfloor N w_t^{(i)} \rfloor)^2}{\sum_{j=1}^N (N w_t^{(j)} - \lfloor N w_t^{(j)} \rfloor)} \middle| \mathcal{F}_{t-1} \right] \\
&\qquad - \frac{1}{(N)_2} \mathbb{E}\left[\sum_{i=1}^N \lfloor N w_t^{(i)} \rfloor \middle| \mathcal{F}_{t-1} \right] \qquad (11)
\end{aligned}$$

Sanity check:
When the weights are all equal, $w_t^{(i)} \equiv 1/N$, we should have $\mathbb{E}[c_N^r(t) | \mathcal{F}_{t-1}] = 0$ since each particle will have

exactly one offspring so it is impossible for any lineages to coalesce. In this case we have $\mathbb{E}[c_N^m(t)|\mathcal{F}_{t-1}] = \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}] = 1/N$ for multinomial resampling. We also have that $Nw_t^{(i)} \equiv \lfloor Nw_t^{(i)} \rfloor \equiv 1$ and hence $r_i = 0$ and $k = N$. Thus the RHS comes out as

$$\frac{1}{N}\frac{N}{N-1} - 0 - \frac{1}{(N)_2}N = \frac{1}{N-1} - \frac{1}{N-1} = 0$$

as expected.

We can write it in a different form by combining the second and third terms of (11):

$$\begin{aligned}
&-\frac{1}{(N)_2}\mathbb{E}\left[\frac{\sum_{i=1}^N (Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor)^2}{\sum_{j=1}^N (Nw_t^{(j)} - \lfloor Nw_t^{(j)} \rfloor)}|\mathcal{F}_{t-1}\right] - \frac{1}{(N)_2}\mathbb{E}\left[\sum_{k=1}^N \lfloor Nw_t^{(k)} \rfloor \middle| \mathcal{F}_{t-1}\right] \\
&=: -\frac{1}{(N)_2}\mathbb{E}[A|\mathcal{F}_{t-1}]
\end{aligned}$$

Then

$$\begin{aligned}
A &= \frac{\sum_{i=1}^N (Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor)^2}{\sum_{j=1}^N (Nw_t^{(j)} - \lfloor Nw_t^{(j)} \rfloor)} + \sum_{k=1}^N \lfloor Nw_t^{(k)} \rfloor \\
&= \frac{\sum_{i=1}^N (Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor)^2 + \sum_{i=1}^N \sum_{k=1}^N \lfloor Nw_t^{(k)} \rfloor (Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor)}{\sum_{j=1}^N (Nw_t^{(j)} - \lfloor Nw_t^{(j)} \rfloor)} \\
&=: \frac{A'}{\sum_{j=1}^N (Nw_t^{(j)} - \lfloor Nw_t^{(j)} \rfloor)}
\end{aligned}$$

Then

$$\begin{aligned}
A' &= \sum_{i=1}^N (Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor)^2 + \sum_{i=1}^N \sum_{k=1}^N \lfloor Nw_t^{(k)} \rfloor (Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor) \\
&= \sum_{i=1}^N \left\{ \left(Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor\right)^2 + \lfloor Nw_t^{(i)} \rfloor \left(Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor\right) + \sum_{k \neq i} \lfloor Nw_t^{(k)} \rfloor \left(Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor\right) \right\} \\
&= \sum_{i=1}^N \left\{ (Nw_t^{(i)})^2 - Nw_t^{(i)} \lfloor Nw_t^{(i)} \rfloor + \sum_{k \neq i} \lfloor Nw_t^{(k)} \rfloor \left(Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor\right) \right\} \\
&= \sum_{i=1}^N \left\{ \left(Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor\right) \left(Nw_t^{(i)} + \sum_{k \neq i} \lfloor Nw_t^{(k)} \rfloor\right) \right\}
\end{aligned}$$

So we have

$$\begin{aligned}
\mathbb{E}[c_N^r(t)|\mathcal{F}_{t-1}] = \mathbb{E}[c_N^m(t)|\mathcal{F}_{t-1}] \left(1 + \frac{1}{N-1}\right) \\
- \frac{1}{(N)_2}\mathbb{E}\left[\frac{\sum_{i=1}^N \left(Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor\right)\left(Nw_t^{(i)} + \sum_{k \neq i} \lfloor Nw_t^{(k)} \rfloor\right)}{\sum_{j=1}^N \left(Nw_t^{(j)} - \lfloor Nw_t^{(j)} \rfloor\right)} \middle| \mathcal{F}_{t-1}\right]
\end{aligned}$$

5.2 Non-asymptotic coalescence rate vs. Multinomial resampling

5.2.1 Case $N = 2$

Lemma 1. *For all weight vectors $w_t^{(1:2)}$, $\mathbb{E}[c_2^m(t)|w_t^{(1:2)}] \geq \mathbb{E}[c_2^r(t)|w_t^{(1:2)}]$.*

Proof. With only $N = 2$ particles, the coalescence rate becomes

$$\mathbb{E}[c_N(t)|w_t^{(1:2)}] = \frac{1}{(N)_2}\sum_{i=1}^{N}\mathbb{E}\left[(\nu_t^{(i)})_2|w_t^{(1:2)}\right] = \mathbb{P}[\nu_t^{(1)} = 0] + \mathbb{P}[\nu_t^{(1)} = 2].$$

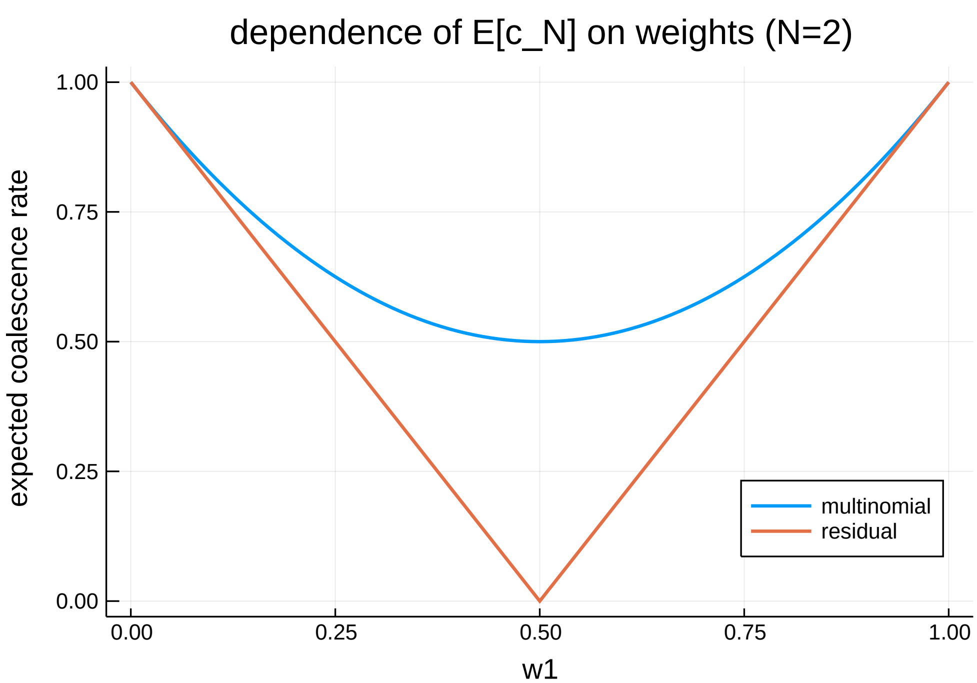
For residual resampling,

$$\mathbb{E}[c_2^r(t)|w_t^{(1:2)}] = \mathbb{I}_{\{w_t^{(1)}\geq 1/2\}}(2w_t^{(1)} - 1) + \mathbb{I}_{\{w_t^{(1)}<1/2\}}(2w_t^{(2)} - 1)$$

And for multinomial resampling,

$$\begin{aligned}\mathbb{E}[c_2^m(t)|w_t^{(1:2)}] &= (w_t^{(1)})^2 + (w_t^{(2)})^2\\ &= \mathbb{I}_{\{w_t^{(1)}\geq 1/2\}}((w_t^{(1)})^2 + (w_t^{(2)})^2) + \mathbb{I}_{\{w_t^{(1)}<1/2\}}((w_t^{(1)})^2 + (w_t^{(2)})^2)\\ &\geq \mathbb{I}_{\{w_t^{(1)}\geq 1/2\}}(w_t^{(1)})^2 + \mathbb{I}_{\{w_t^{(1)}<1/2\}}(w_t^{(2)})^2\end{aligned}$$

Then since $(w_t^{(i)} - 1)^2 = (w_t^{(i)})^2 - 2w_t^{(i)} + 1 \geq 0$, we have that $(w_t^{(i)})^2 \geq 2w_t^{(i)} - 1$ and hence we conclude the proof. $\square$



The red line on the plot is correct for stratified resampling and for stochastic rounding, as well as residual-multinomial, for $N = 2$. This is easy to show using the same kind of technique as above. (Pre-sorting the weights does not have any effect when $N = 2$, since they are already sorted in either ascending or descending order.)

We could add another line for CSMC-multinomial, which has expected coalescence rate w_1. It therefore dominates multinomial when $w_1 \leq 0.5$ but is dominated by multinomial when $w_1 \geq 0.5$. This asymmetry is due to the (wlog) assignment of the immortal particle to index 1.

5.2.2 Case $N = 3$

Given a weight vector $(w_t^{(1)}, w_t^{(2)}, w_t^{(3)})$, let $w_{(1)} \geq w_{(2)} \geq w_{(3)}$ denote the weights sorted from high to low. With $N = 3$ there are many more cases than with $N = 2$, and these are described below, using the sorted weights.

In each case for the conditions on the sorted weights, the possible offspring count vectors (sorted in the same order as the weights) are listed, along with the probability of each (conditional on the given case). Finally, using these outcomes and associated probabilities, the conditional expectation of interest is calculated.

Case	Weights	Offspring counts	Conditional probabilities	$\mathbb{E}[c_2^r(t)\|w_t^{(1:3)}]$
(A)	$w_{(1)} = 1$	$(3,0,0)$	1	1
(B)	$2/3 < w_{(1)} < 1$	$(3,0,0)$ $(2,1,0)$ $(2,0,1)$	$3w_{(1)} - 2$ $3w_{(2)}$ $3w_{(3)}$	$2w_{(1)} - 1$
(C)	$w_{(1)} = 2/3$	$(2,1,0)$ $(2,0,1)$	$3w_{(2)}$ $3w_{(3)}$	$1/3$
(D1)	$1/3 < w_{(1)} < 2/3$ and $1/3 \leq w_{(2)} < 2/3$	$(2,1,0)$ $(1,2,0)$ $(1,1,1)$	$3w_{(1)} - 1$ $3w_{(2)} - 1$ $3w_{(3)}$	$1/3 - w_{(3)}$
(D2)	$1/3 < w_{(1)} < 2/3$ and $w_{(2)} < 1/3$	$(3,0,0)$ $(2,1,0)$ $(2,0,1)$ $(1,2,0)$ $(1,0,2)$ $(1,1,1)$	$(3/2)^2(w_{(1)} - 1/3)^2$ $(3/2)^2 2(w_{(1)} - 1/3)w_{(2)}$ $(3/2)^2 2(w_{(1)} - 1/3)w_{(3)}$ $(3/2)^2 w_{(2)}^2$ $(3/2)^2 w_{(3)}^2$ $(3/2)^2 2w_{(2)}w_{(3)}$	$(1/4)(3w_{(1)} - 1)(w_{(1)} + 1)$
(E)	$w_{(1)} = 1/3$	$(1,1,1)$	1	0

Lemma 2. *For all weight vectors* $w_t^{(1:3)}$, $\mathbb{E}[c_3^r(t)|w_t^{(1:3)}] \leq \mathbb{E}[c_3^m(t)|w_t^{(1:3)}]$.

Proof. We have the following expression in the case of residual resampling:

$$\begin{aligned}\mathbb{E}[c_3^r(t)|w_t^{(1:3)}] = &(2w_{(1)} - 1)\mathbb{I}_{\{2/3 \leq w_{(1)} \leq 1\}} \\ &+ (1/3 - w_{(3)})\mathbb{I}_{\{1/3 \leq w_{(1)} < 2/3\}}\mathbb{I}_{\{1/3 \leq w_{(2)} < 2/3\}} \\ &+ (1/4)(3w_{(1)} - 1)(w_{(1)} + 1)\mathbb{I}_{\{1/3 \leq w_{(1)} < 2/3\}}\mathbb{I}_{\{w_{(2)} < 1/3\}}\end{aligned}$$

compared to the following in the case of multinomial resampling:

$$\begin{aligned}\mathbb{E}[c_3^m(t)|w_t^{(1:3)}] &= w_{(1)}^2 + w_{(2)}^2 + w_{(3)}^2 \\ &= (w_{(1)}^2 + w_{(2)}^2 + w_{(3)}^2)\mathbb{I}_{\{2/3 \leq w_{(1)} \leq 1\}} \\ &\quad + (w_{(1)}^2 + w_{(2)}^2 + w_{(3)}^2)\mathbb{I}_{\{1/3 \leq w_{(1)} < 2/3\}}\mathbb{I}_{\{1/3 \leq w_{(2)} < 2/3\}} \\ &\quad + (w_{(1)}^2 + w_{(2)}^2 + w_{(3)}^2)\mathbb{I}_{\{1/3 \leq w_{(1)} < 2/3\}}\mathbb{I}_{\{w_{(2)} < 1/3\}}.\end{aligned}$$

Hence it suffices to show the following:

(i) $(2w_{(1)} - 1) \leq (w_{(1)}^2 + w_{(2)}^2 + w_{(3)}^2)$

(ii) $(1/3 - w_{(3)}) \leq (w_{(1)}^2 + w_{(2)}^2 + w_{(3)}^2)$

(iii) $(1/4)(3w_{(1)} - 1)(w_{(1)} + 1) \leq (w_{(1)}^2 + w_{(2)}^2 + w_{(3)}^2)$

First consider (ii). Since $w_{(3)}$ is defined as the smallest of the three weights, we know that $w_{(3)} \in [0, 1/3]$. Meanwhile, the RHS is the sum of the squared weights, which is always between 1/3 and 1. Therefore (ii) is true.

Now let us consider (i). We have the identity $(w_{(1)} - 1)^2 = w_{(1)}^2 - 2w_{(1)} + 1$, which implies that $w_{(1)}^2 \geq 2_w(1) - 1$. Since $w_{(2)}^2 + w_{(3)}^2 \geq 0$ we can therefore conclude that (i) is true.

Finally consider (iii). In this case it is sufficient to show that $(1/4)(3w_{(1)} - 1)(w_{(1)} + 1) \leq w_{(1)}^2$. Note that

$$\begin{aligned}
&w_{(1)}^2 - \frac{1}{4}(3w_{(1)} - 1)(w_{(1)} + 1) \\
&= \frac{1}{4}w_{(1)}^2 - \frac{1}{2}w_{(1)} + \frac{1}{4} \\
&= \frac{1}{4}(w_{(1)} - 1)^2 \geq 0.
\end{aligned}$$

Therefore (iii) is also true, concluding the proof. $\square$

Remark 1. *Examining the table, we can see that the function $\mathbb{E}[c_3^r(t)|w_t^{(1:3)}]$ is continuous in $w_{(1)}$. However, the plot clearly shows discontinuities. These occur at the boundaries between different orderings when we sort the weights from high to low.*

5.2.3 General N

Lemma 3. *In the case $w_{(1)} \geq \frac{N-1}{N}$, we have*

$$\mathbb{E}[c_N^r(t)|w_t^{(1:N)}] = 2w_{(1)} - 1.$$

Proof. In this case, (N-1) offspring are deterministically assigned to particle 1. The one remaining offspring can either be assigned to particle 1 or to some other particle.

- The first option yields offspring vector $(N, 0, 0, \dots)$ and occurs with probability $(w_{(1)} - \frac{N-1}{N})/(1/N) = Nw_{(1)} - (N-1)$. The resulting value of c_N^r is $N(N-1)/(N)_2 = 1$.
- The second option yields an offspring vector that is some permutation of $(N-1, 1, 0, 0, \dots)$, and occurs with probability $1 - Nw_{(1)} - (N-1) = N - Nw_{(1)}$. The resulting value of c_N^r is $(N-1)(N-2)/(N)_2 = (N-2)/N$.

So, under this constraint on the weights, we have the following expectation:

$$\mathbb{E}[c_N^r(t)|w_t^{(1:N)}] = Nw_{(1)} - (N-1) + \frac{N-2}{N}(N - Nw_{(1)}) = 2w_{(1)} - 1.$$

$\square$

Lemma 4. *For $N \geq 3$, in the case $w_{(1)}, w_{(2)}, \dots, w_{(N-1)} \in \left[\frac{1}{N}, \frac{2}{N}\right]$, we have*

$$\mathbb{E}[c_N^r(t)|w_t^{(1:N)}] = \frac{2}{N-1}(1/N - w_{(N)}).$$

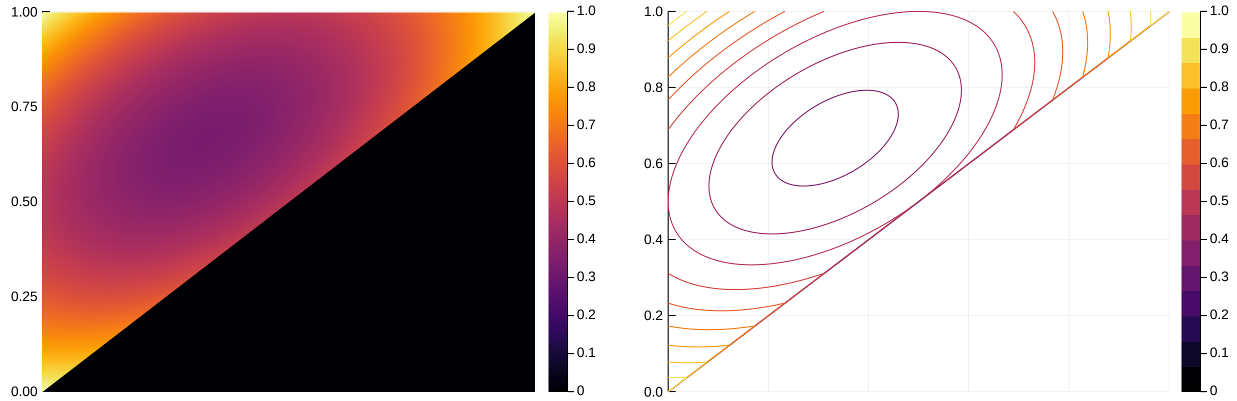
Proof. In this case, one offspring is deterministically assigned to each of the particles $1, 2, \dots, N-1$. The one remaining offspring can be assigned either to particle N or to one of the others.

- The first option yields offspring vector $(1, 1, \dots, 1)$ and occurs with probability $w_{(N)}/(1/N) = Nw_{(N)}$. The resulting value of c_N^r is 0.
- The second option yields an offspring vector that is some permutation of $(2, 0, 1, 1, \dots, 1)$ and occurs with probability $1 - Nw_{(N)}$. The resulting value of c_N^r is 2.

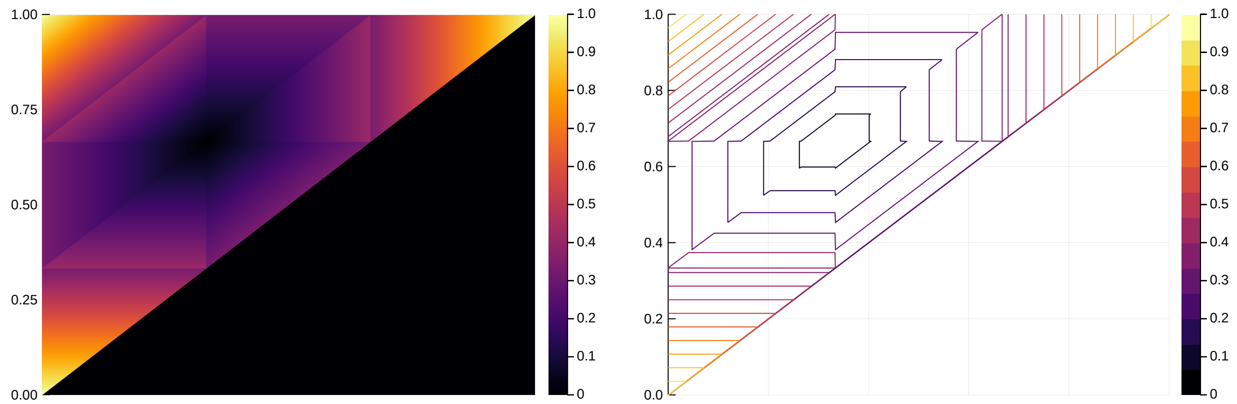
So, under this constraint on the weights, we have the following expectation:

$$\mathbb{E}[c_N^r(t)|w_t^{(1:N)}] = 2(1 - Nw_{(N)})/(N)_2 = \frac{2}{N-1}(1/N - w_{(N)}).$$

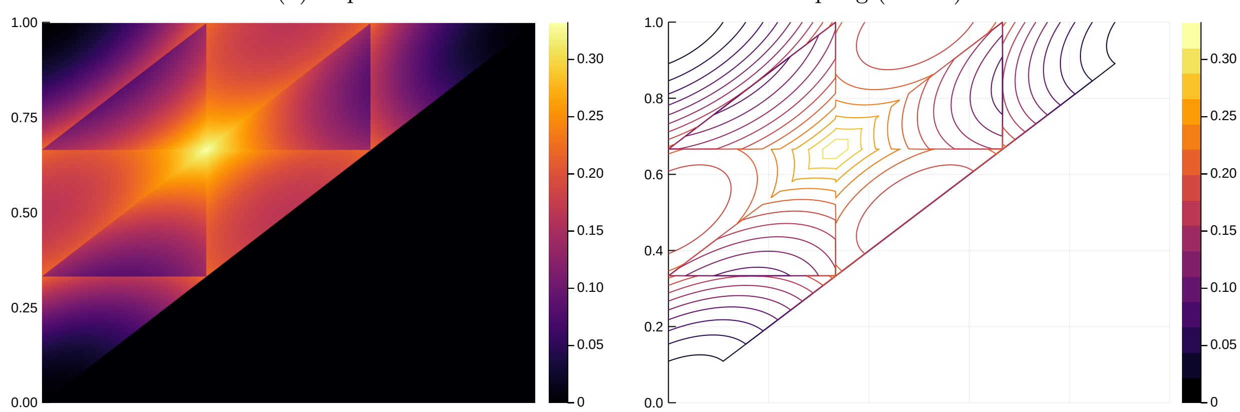
$\square$



(a) Expected coalescence rate with multinomial resampling ($N = 3$)



(b) Expected coalescence rate with residual resampling ($N = 3$)



(c) Difference between expected coalescence rate with multinomial or residual resampling ($N = 3$)

Conjecture 1. *The number of cases to consider for population size N (excluding zero-measure cases which can be included in other cases) is*

$$p(N-1) + \frac{(N-1)(N-2)}{2}$$

where $p(\cdot)$ denotes the partition function (number of integer partitions).

This expression was determined by pattern-spotting in low-N cases. I believe a combinatorial proof could be found but, since it is not a particularly useful result, I shan't bother.

Conjecture 2. *For all weight vectors $w_t^{(1:N)}$, $\mathbb{E}[c_2^m(t)|w_t^{(1:N)}] \geq \mathbb{E}[c_2^r(t)|w_t^{(1:N)}]$.*

5.3 Squared coalescence rate and other awful expressions

First let's write down various moments that will be needed (where $i \neq j$).

$$\begin{aligned}
\mathbb{E}[(\nu_t^{(i)})^2|\mathcal{F}_{t-1}] &= \lfloor Nw_i \rfloor^2 + 2\lfloor Nw_i \rfloor N w_t^{(i)} + (N)_2 (w_t^{(i)})^2 \\
\mathbb{E}[(\nu_t^{(i)})^3|\mathcal{F}_{t-1}] &= \lfloor Nw_i \rfloor^3 + 3\lfloor Nw_i \rfloor^2 N w_t^{(i)} + 3\lfloor Nw_i \rfloor (N)_2 (w_t^{(i)})^2 + (N)_3 (w_t^{(i)})^3 \\
\mathbb{E}[(\nu_t^{(i)})^4|\mathcal{F}_{t-1}] &= \lfloor Nw_i \rfloor^4 + 4\lfloor Nw_i \rfloor^3 N w_t^{(i)} + 6\lfloor Nw_i \rfloor^2 (N)_2 (w_t^{(i)})^2 + 4\lfloor Nw_i \rfloor (N)_3 (w_t^{(i)})^3 + (N)_4 (w_t^{(i)})^4 \\
\mathbb{E}[\nu_t^{(i)}\nu_t^{(j)}|\mathcal{F}_{t-1}] &= \lfloor Nw_i \rfloor \lfloor Nw_i \rfloor[j] + \lfloor Nw_i \rfloor N w_t^{(j)} + \lfloor Nw_i \rfloor[j] N w_t^{(i)} + (N)_2 w_t^{(i)} w_t^{(j)} \\
\mathbb{E}[(\nu_t^{(i)})^2\nu_t^{(j)}|\mathcal{F}_{t-1}] &= \lfloor Nw_i \rfloor^2 \lfloor Nw_i \rfloor[j] + \lfloor Nw_i \rfloor^2 N w_t^{(j)} + 2\lfloor Nw_i \rfloor \lfloor Nw_i \rfloor[j] N w_t^{(i)} + 2\lfloor Nw_i \rfloor (N)_2 w_t^{(i)} w_t^{(j)} \\
&\quad + \lfloor Nw_i \rfloor[j] (N)_2 (w_t^{(i)})^2 + (N)_3 (w_t^{(i)})^2 w_t^{(j)} \\
\mathbb{E}[(\nu_t^{(i)})^2(\nu_t^{(j)})^2|\mathcal{F}_{t-1}] &= \lfloor Nw_i \rfloor^2 \lfloor Nw_i \rfloor[j]^2 + 2\lfloor Nw_i \rfloor^2 \lfloor Nw_i \rfloor[j] N w_t^{(j)} + 2\lfloor Nw_i \rfloor \lfloor Nw_i \rfloor[j]^2 N w_t^{(i)} \\
&\quad + \lfloor Nw_i \rfloor^2 (N)_2 (w_t^{(j)})^2 + 4\lfloor Nw_i \rfloor \lfloor Nw_i \rfloor[j] (N)_2 w_t^{(i)} w_t^{(j)} + \lfloor Nw_i \rfloor[j]^2 (N)_2 (w_t^{(i)})^2 \\
&\quad + 2\lfloor Nw_i \rfloor (N)_3 w_t^{(i)} (w_t^{(j)})^2 + 2\lfloor Nw_i \rfloor[j] (N)_3 (w_t^{(i)})^2 w_t^{(j)} + (N)_4 (w_t^{(i)})^2 (w_t^{(j)})^2
\end{aligned}$$

For the squared coalescence rate, expanding the falling factorials appropriately, we get

$$\begin{aligned}
\mathbb{E}[(c_N^r(t))^2|\mathcal{F}_{t-1}] &= \frac{1}{(N)_2^2} \left(\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_2^2|\mathcal{F}_{t-1}] + \sum_{i=1}^N \sum_{j\neq i} \mathbb{E}[(\nu_t^{(i)})_2 (\nu_t^{(j)})_2|\mathcal{F}_{t-1}] \right) \\
&= \frac{1}{(N)_2^2} \sum_{i=1}^N \Big(\mathbb{E}[(\nu_t^{(i)})^4|\mathcal{F}_{t-1}] - 2\mathbb{E}[(\nu_t^{(i)})^3|\mathcal{F}_{t-1}] + \mathbb{E}[(\nu_t^{(i)})^2|\mathcal{F}_{t-1}] \Big) \\
&\quad + \frac{1}{(N)_2^2} \sum_{i=1}^N \sum_{j\neq i} \Big(\mathbb{E}[(\nu_t^{(i)})^2(\nu_t^{(j)})^2|\mathcal{F}_{t-1}] - \mathbb{E}[(\nu_t^{(i)})^2\nu_t^{(j)}|\mathcal{F}_{t-1}] - \mathbb{E}[\nu_t^{(i)}(\nu_t^{(j)})^2|\mathcal{F}_{t-1}] + \mathbb{E}[\nu_t^{(i)}\nu_t^{(j)}|\mathcal{F}_{t-1}
\end{aligned}$$

Then we can try plugging in the expressions derived above and find that none of the terms cancel. Maybe if we're clever we can factorise it or something.

5.4 Mega-merger rate

Now the rate of super-binary mergers...

$$\begin{aligned}
\mathbb{E}[D_N(t)|\mathcal{F}_{t-1}] &= \frac{1}{N(N)_2} \sum_{i=1}^N \Big(\mathbb{E}[(\nu_t^{(i)})^3|\mathcal{F}_{t-1}] - \mathbb{E}[(\nu_t^{(i)})^2|\mathcal{F}_{t-1}] \Big) \\
&\quad + \frac{1}{N(N)_2} \sum_{i=1}^N \sum_{j\neq i} \Big(\mathbb{E}[(\nu_t^{(i)})^2(\nu_t^{(j)})^2|\mathcal{F}_{t-1}] - \mathbb{E}[\nu_t^{(i)}(\nu_t^{(j)})^2|\mathcal{F}_{t-1}] \Big)
\end{aligned}$$

6 Randomised roundings

Definition 1. Let $X \geq 0$. A random variable $Y : \mathbb{R}_+ \to \mathbb{N}$ is a *randomised rounding* of X if Y takes the values

$$Y = \begin{cases} \lfloor X \rfloor & \text{with probability } 1 - X + \lfloor X \rfloor \\ \lfloor X \rfloor + 1 & \text{with probability } X - \lfloor X \rfloor \end{cases}$$

Note that we therefore have $\mathbb{E}[Y] = X$.

Definition 1 generalises easily to multivariate X. A randomised rounding of a vector $X \in \mathbb{R}_+^N$ produces a vector $Y \in \mathbb{N}^N$ such that $\mathbb{E}[Y] = X$. Y can thus be used to construct an unbiased resampling scheme.

Such popular resampling schemes as systematic resampling and residual resampling with systematic residuals are based on randomised roundings. In each scheme, the marginal distributions of the offspring counts are as in Definition 1, where $X = Nw_t^{(i)}$ and $Y = \nu_t^{(i)}$. The variety in rounding-based schemes arises from the interactions between these individual offspring counts.

In this document we will consider a general resampling scheme based on randomised rounding, and compare its genealogical properties to those of multinomial resampling. Let $(w_1, \dots, w_N)$ be a normalised weight vector (that is, $w_i \geq 0$ and $\sum w_i = 1$), let $N \geq 2$ be an integer denoting the number of particles, and let $(v_1, \dots v_N)$ denote the resampled offspring numbers. We will attach a superscript m or r to denote offspring numbers corresponding to multinomial and rounding-based resampling respectively.

Lemma 5. *For any $i \in \{1, \dots, N\}$,*

$$\mathbb{E}[(v_i)_2^{(m)} | w_i] \geq \mathbb{E}[(v_i)_2^{(r)} | w_i].$$

Proof. Using properties of the Multinomial distribution (Mosimann, 1962), we have

$$\mathbb{E}[(v_i)_2^{(m)} | w_i] = N(N-1)w_i^2$$

Directly from Definition 1, we calculate the corresponding quantity in the case of randomised rounding to be

$$\begin{aligned} \mathbb{E}[(v_i)_2^{(r)} | w_i] &= \lfloor Nw_i \rfloor (\lfloor Nw_i \rfloor - 1)(1 - Nw_i + \lfloor Nw_i \rfloor) + (\lfloor Nw_i \rfloor + 1)\lfloor Nw_i \rfloor (Nw_i - \lfloor Nw_i \rfloor) \\ &= \lfloor Nw_i \rfloor \left(2(Nw_i - \lfloor Nw_i \rfloor) + \lfloor Nw_i \rfloor - 1\right) \end{aligned}$$

We define the difference

$$\begin{aligned} \Delta_i &:= \mathbb{E}[(v_i)_2^{(m)} | w_i] - \mathbb{E}[(v_i)_2^{(r)} | w_i] \\ &= N^2 w_i^2 - N w_i^2 - 2\lfloor Nw_i \rfloor (Nw_i - \lfloor Nw_i \rfloor) - \lfloor Nw_i \rfloor^2 + \lfloor Nw_i \rfloor \\ &= N^2 w_i^2 + \lfloor Nw_i \rfloor^2 - 2Nw_i \lfloor Nw_i \rfloor - Nw_i^2 + \lfloor Nw_i \rfloor \\ &= (Nw_i - \lfloor Nw_i \rfloor)^2 - Nw_i^2 + \lfloor Nw_i \rfloor \end{aligned}$$

Then we have to show that $\Delta_i \geq 0$ for all $N \geq 2$ and $w_i \in [0, 1]$. Consider the following cases.

Case $w_i = k/N$ for some $k \in \{1, \dots, N-1\}$.
Then $\lfloor Nw_i \rfloor = Nw_i = k$ and we have
$\Delta_i = 0 - k^2/N + k = k(1 - k/N) \geq 0$.

Case $w_i \in (k/N, (k+1)/N)$ for some $k \in \{0, \dots, N-1\}$

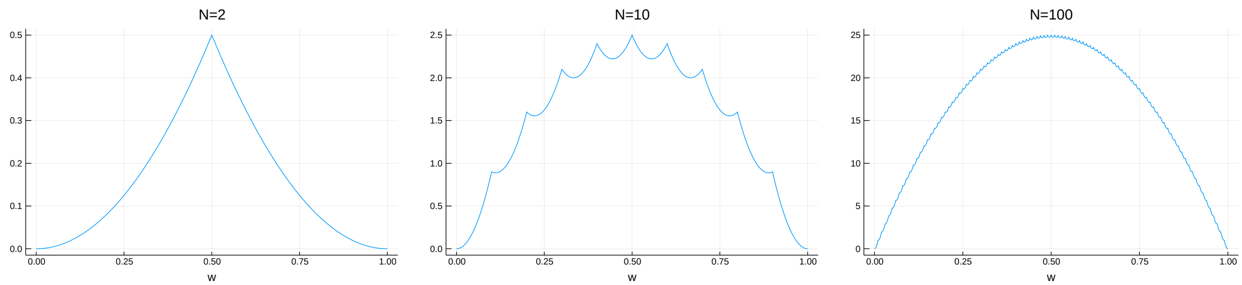


Figure 2: Δ_i as a function of w_i for various values of N. The function is piecewise quadratic between the points $w_i = k/N$. The limiting shape appears to be a parabola.

Then $\lfloor Nw_i \rfloor = k$, and

$$\begin{aligned}
\Delta_i &= (Nw_i - k)^2 - Nw_i^2 + k \\
&= N(N-1)w_i^2 - 2Nkw_i + k(k+1) \\
&= N(N-1)\left[\left(w - \frac{k}{N-1}\right)^2 - \frac{k^2}{(N-1)^2} + \frac{k(k+1)}{N(N-1)}\right] \\
&= N(N-1)\left(w - \frac{k}{N-1}\right)^2 - \frac{k^2 N}{N-1} + k^2 + k \\
&\geq -\frac{k^2 N}{N-1} + k^2 + k \\
&= k\left(1 - \frac{k}{N-1}\right) \\
&\geq 0
\end{aligned}$$

For each N, any $w_i \in [0,1]$ falls into one of the above cases, so we conclude that $\Delta_i \geq 0$ for all N, w_i. $\square$

6.1 Third moment & required ratio

$$\begin{aligned}
\mathbb{E}[(\nu_t^{(i)})_3 | w_t^{(1:N)}] &= (\lfloor Nw_i \rfloor)_3 (1 - Nw_t^{(i)} + \lfloor Nw_i \rfloor) + (\lfloor Nw_i \rfloor + 1)_3 (Nw_t^{(i)} - \lfloor Nw_i \rfloor) \\
&= (\lfloor Nw_i \rfloor)_2 \left\{ (\lfloor Nw_i \rfloor - 2)(1 - Nw_t^{(i)} + \lfloor Nw_i \rfloor) + (\lfloor Nw_i \rfloor + 1)(Nw_t^{(i)} - \lfloor Nw_i \rfloor) \right\} \\
&= (\lfloor Nw_i \rfloor)_2 \left(3Nw_t^{(i)} - 2\lfloor Nw_i \rfloor - 2 \right)
\end{aligned}$$

So for the ratio we have

$$
\begin{aligned}
\frac{\frac{1}{(N)_3}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_3|\mathcal{F}_{t-1}]}{\frac{1}{(N)_2}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_2|\mathcal{F}_{t-1}]} &= \frac{1}{N-2}\frac{\sum_{i=1}^N \mathbb{E}[(\lfloor Nw_i\rfloor)_2(3Nw_t^{(i)}-2\lfloor Nw_i\rfloor-2)|\mathcal{F}_{t-1}]}{\sum_{i=1}^N \mathbb{E}[\lfloor Nw_i\rfloor(2Nw_t^{(i)}-\lfloor Nw_i\rfloor-1)|\mathcal{F}_{t-1}]} \\
&\le \frac{1}{N-2}\frac{\sum_{i=1}^N \mathbb{E}[(\lfloor Nw_i\rfloor)_2 Nw_t^{(i)}|\mathcal{F}_{t-1}]}{\sum_{i=1}^N \mathbb{E}[\lfloor Nw_i\rfloor(Nw_t^{(i)}-1)|\mathcal{F}_{t-1}]} \\
&\le \frac{1}{N-2}\frac{\sum_{i=1}^N \mathbb{E}[(Nw_t^{(i)})_2 Nw_t^{(i)}|\mathcal{F}_{t-1}]}{\sum_{i=1}^N \mathbb{E}[(Nw_t^{(i)}-1)^2|\mathcal{F}_{t-1}]} \\
&= \frac{1}{N-2}\frac{\sum_{i=1}^N \mathbb{E}[(Nw_t^{(i)})^3|\mathcal{F}_{t-1}]-\sum_{i=1}^N \mathbb{E}[(Nw_t^{(i)})^2|\mathcal{F}_{t-1}]}{N^2\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]-2N\sum_{i=1}^N \mathbb{E}[w_t^{(i)}|\mathcal{F}_{t-1}]+\sum_{i=1}^N \mathbb{E}[1|\mathcal{F}_{t-1}]} \\
&= \frac{1}{N-2}\frac{N^3\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^3|\mathcal{F}_{t-1}]-N^2\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]}{N^2\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]-2N+N} \\
&= \frac{1}{N-2}\frac{N^2\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^3|\mathcal{F}_{t-1}]-N\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]}{N\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]-1} \\
&= \frac{1}{N-2}\left[\frac{N^2\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^3|\mathcal{F}_{t-1}]}{N\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]-1}-\frac{N\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]}{N\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]-1}\right] \\
&= \frac{1}{N-2}\left[-1+\frac{N^2\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^3|\mathcal{F}_{t-1}]-1}{N\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]-1}\right]
\end{aligned}
$$

Then, using that $w_t^{(i)} = \Theta(N^{-1})$, the sum in the denominator is $O(N^{-1})$ and the sum in the numerator is $O(N^{-2})$. Hence the whole expression is $O(N^{-1})$ as $N \to \infty$, so we can find a suitable sequence $b_N \overset{N\to\infty}{\longrightarrow} 0$ to satisfy the conditions of Theorem (?) [the one with new assns].

References

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* .

Mosimann, J. E. (1962), 'On the compound multinomial distribution, the multivariate β-distribution, and correlations among proportions', *Biometrika* **49**(1/2), 65–82.