Non-triviality condition (fuller details of paper appendix)

Suzie Brown

July 7, 2020

The following theorem will be used in each section. It is a filtered version of the second Borel–Cantelli lemma, which can be found for instance in Durrett (2019, Theorem 4.3.4).

Lemma 1. *Let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration with $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Let $(B_t)_{t\geq 0}$ be a sequence of events such that $B_t \in \mathcal{F}_t$ for all t. Then the events $\{B_t \text{ i.o.}\}$ and $\{\sum_{t=1}^{\infty} \mathbb{P}[B_t \mid \mathcal{F}_{t-1}] = \infty\}$ are almost surely equal.*

We will also use the following equivalence in each section.

Lemma 2. *Let τ_N denote the generalised inverse of c_N, i.e.*

$$\tau_N(t) = \min \left\{ s \geq 1 : \sum_{r=1}^{s} c_N(r) \geq t \right\}.$$

Suppose that there exists $N_0 \in \mathbb{N}$ such that almost surely for all $N > N_0$, $c_N(t)$ is bounded away from zero for infinitely many t. Then for all $N > N_0$, for all finite t, $\mathbb{P}[\tau_N(t) = \infty] = 0$.

Proof. Applying the definition of $\tau_N(t)$,

$$\begin{aligned}
\mathbb{P}[\tau_N(t) = \infty] = 0 &\Leftrightarrow \mathbb{P}[\tau_N(t) < \infty] = 1 \\
&\Leftrightarrow \mathbb{P}\left[\min \left\{ s \geq 1 : \sum_{r=1}^{s} c_N(r) \geq t \right\} < \infty \right] = 1 \\
&\Leftrightarrow \mathbb{P}\left[\exists s < \infty : \sum_{r=1}^{s} c_N(r) \geq t \right] = 1
\end{aligned}$$

A sufficient condition for the last line is that, almost surely for all $N > N_0$, $c_N(r)$ is bounded away from zero infinitely often in r. $\square$

Multinomial resampling

Lemma 3. *For all $N \geq 2$, for all t,*

$$\mathbb{P}\left[c_N(t) > \frac{2}{N^2} \middle| \mathbf{w} = \left(\frac{1}{N}, \dots, \frac{1}{N} \right) \right] = 1 - \frac{N!}{N^N}.$$

Proof. Fix arbitrary t and $N \geq 2$. Since $2/(N)_2 > 2/(N^2)$ is the smallest possible non-zero value for $c_N(t)$,

$$\begin{aligned}
\mathbb{P}\left[c_N(t) > \frac{2}{N^2} \mid \mathbf{w} = (1/N, \dots, 1/N) \right] &= 1 - \mathbb{P}[c_N(t) = 0 \mid \mathbf{w} = (1/N, \dots, 1/N)] \\
&= 1 - \mathbb{P}[\nu_t^{(1:N)} = (1, \dots, 1) \mid \mathbf{w} = (1/N, \dots, 1/N)].
\end{aligned}$$

Conditional on the weights, $\nu_t^{(1:N)} \sim \text{Multinomial}(N, (1/N, \dots, 1/N))$, so the probability of interest is

$$\mathbb{P}[\nu_t^{(1:N)} = (1, \dots, 1) \mid \mathbf{w} = (1/N, \dots, 1/N)] = N! \prod_{i=1}^{N} \frac{1}{N} = \frac{N!}{N^N}.$$

$\square$

Lemma 4. *In the neutral case (i.e. when all weights are equal at every time step) with multinomial resampling, there exists N_0 such that for all $N > N_0$, for all finite t, $\mathbb{P}[\tau_N(t) = \infty] = 0$.*

Proof. Let us rewrite the event of interest in a different way.

$$\begin{aligned}
\mathbb{P}[\tau_N(t) = \infty] = 0 \Leftrightarrow \mathbb{P}[\tau_N(t) < \infty] = 1 \\
&\Leftrightarrow \mathbb{P}\left[\min\left\{s \geq 1 : \sum_{r=1}^{s} c_N(r) \geq t\right\} < \infty\right] = 1 \\
&\Leftrightarrow \mathbb{P}\left[\exists s < \infty : \sum_{r=1}^{s} c_N(r) \geq t\right] = 1
\end{aligned}$$

It is sufficient to show that, almost surely for all $N > N_0$, $c_N(r)$ is bounded away from zero infinitely often in r. We consider the sequence of events $E_r := \{c_N(r) > 2/N^2\}$ for $r \in \mathbb{N}$. In the neutral case, the resampled family sizes at each generation are independent, hence the events E_r are independent. By the second Borel-Cantelli lemma, E_r occurs infinitely often if $\sum_{r=1}^{\infty} \mathbb{P}(E_r) = \infty$. An expression for $\mathbb{P}(E_r)$ is given in Lemma 3. For any fixed $N \geq 2$, the probability is strictly positive and constant in r, so the Borel-Cantelli condition is satisfied, thus we conclude that E_r occurs infinitely often. Hence, taking $N_0 = 1$, we have that $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all $N > N_0$ and all finite t, as required. □

Lemma 5. *For all $N \geq 2$, for all t, for any weight vector $(w_1, \dots, w_N) \in \mathcal{S}_{N-1}$,*

$$\mathbb{P}\left[c_N(t) > \frac{2}{N^2} \middle| \mathbf{w} = (w_1, \dots, w_N)\right] \geq \mathbb{P}\left[c_N(t) > \frac{2}{N^2} \middle| \mathbf{w} = \left(\frac{1}{N}, \dots, \frac{1}{N}\right)\right].$$

That is, the probability of having at least one merger is minimised by the vector of equal weights.

Proof. Fix arbitrary t and $N \geq 2$. Recall that

$$1 - \mathbb{P}\left[c_N(t) > \frac{2}{N^2} \mid \mathbf{w} = (w_1, \dots, w_N)\right] = N! \prod_{i=1}^{N} w_i. \tag{1}$$

We will show that the global maximum of this function on the simplex $\mathcal{S}_{N-1}$ is attained at $\mathbf{w} = (1/N, \dots, 1/N)$. This weight vector will therefore minimise the probability of the complementary event, implying the desired result.

First, since we are working on the simplex, we encode the constraint $\sum w_j = 1$ by rewriting the function to optimise as

$$f(\mathbf{w}) := \prod_{i=1}^{N} w_i = \left(1 - \sum_{j=1}^{N-1} w_j\right) \prod_{i=1}^{N-1} w_i$$

where we have also dropped the constant positive factor $N!$. Note that this objective function is non-negative and obtains its minimal value zero whenever one or more of the weights is equal to zero; since we are looking for a maximum we can assume that $w_i > 0$ for all i. Now, for every $k \in \{1, \dots, N-1\}$, we solve

$$\frac{\partial f(\mathbf{w})}{\partial w_k} = \left(1 - w_k - \sum_{j=1}^{N-1} w_j\right) \prod_{i \neq k}^{N-1} w_i = 0.$$

The product over $i \neq k$ is constant and positive for each k, so this amounts to solving

$$w_k = 1 - \sum_{j=1}^{N-1} w_j = w_N$$

for all k. The unique solution is $w_1 = w_2 = \cdots = w_N = 1/N$.

To verify that the critical point is a maximum, we evaluate the Hessian H:

$$H_{kl}(\mathbf{w}) = \begin{cases} -2 \prod_{i \neq k}^{N-1} w_i & k = l \\ \left(1 - w_k - w_l - \sum_{j=1}^{N-1} w_j\right) \prod_{i \neq k,l}^{N-1} w_i & k \neq l \end{cases}$$

$$H_{kl}(1/N, \dots, 1/N) = \begin{cases} -2 \left(\frac{1}{N}\right)^{N-2} & k = l \\ -\left(\frac{1}{N}\right)^{N-2} & k \neq l \end{cases}$$

and show that H is negative definite at $(1/N,\dots,1/N)$: for any $\mathbf{x}\in\mathbb{R}^{N-1}\setminus\{\mathbf{0}\}$,

$$\mathbf{x}^T H\left(\frac{1}{N},\dots,\frac{1}{N}\right)\mathbf{x} = \sum_{k=1}^{N-1}\left[-2\left(\frac{1}{N}\right)^{N-2} x_k^2 - \sum_{l\neq k}^{N-1}\left(\frac{1}{N}\right)^{N-2} x_k x_l\right] = \left(\frac{1}{N}\right)^{N-2}\left[-\sum_{k=1}^{N-1} 2x_k^2 - \sum_{k=1}^{N-1}\sum_{l\neq k}^{N-1} x_k x_l\right]$$
$$= \left(\frac{1}{N}\right)^{N-2}\left[-\sum_{k=1}^{N-1} x_k^2 - \sum_{k=1}^{N-1}\sum_{l=1}^{N-1} x_k x_l\right] = \left(\frac{1}{N}\right)^{N-2}\left[-\sum_{k=1}^{N-1} x_k^2 - \left(\sum_{k=1}^{N-1} x_k\right)^2\right] < 0.$$

□

Theorem 1. *With multinomial resampling, conditional on any sequence of weight vectors* $\mathbf{w}_r^{(1:N)}\in\mathcal{S}_{N-1}; r\in\mathbb{N}$, *there exists* N_0 *such that for all* $N>N_0$, *for all finite* t, $\mathbb{P}[\tau_N(t)=\infty]=0$.

Proof. As in Lemma 4, denote the sequence of events $E_r := \{c_N(r) > 2/N^2\}$ for $r\in\mathbb{N}$. We know from Lemma 4 that, in the neutral case, E_r occurs infinitely often. Lemma 5 tells us that $\mathbb{P}[E_r \mid \mathbf{w}=(w_1,\dots,w_N)] \geq \mathbb{P}[E_r \mid \mathbf{w} = (1/N,\dots,1/N)]$ for all r. Therefore, by a coupling argument, we conclude that E_r occurs infinitely often in the non-neutral case as well. □

Conditional SMC with multinomial resampling

Define $\mathbf{w}^* := \frac{1}{N-1}\left[(1,\dots,1)-\mathbf{e}_{i^*}\right]$, where i^* is the immortal index at generation t, and $\mathbf{e}_i$ denotes the i^{th} canonical basis vector.

Lemma 6. *For all* $N\geq 2$, *for all* t,

$$\mathbb{P}\left[c_N(t) > \frac{2}{N^2}\middle|\mathcal{H}_t, \mathbf{w}_t = \mathbf{w}^*\right] = 1 - (\varepsilon(N-1))^{1-N}\prod_{i\neq i^*}^{N} h(X_{t-1}^{(i)}).$$

Proof. Under $\mathbf{w}^*$, the immortal parent has zero weight and is therefore assigned exactly one offspring (the immortal particle). The remaining $N-1$ offspring are assigned to the remaining $N-1$ parents according to a Multinomial distribution with equal weights. We therefore have

$$\begin{aligned}
\mathbb{P}\left[c_N(t)\leq \frac{2}{N^2}\mid \mathcal{H}_t\right] &= \frac{1}{(N-1)!}\sum_{\mathbf{a}_t:\nu_t^{(i)}=1\forall i}\mathbb{P}[\mathbf{a}_t\mid\mathcal{H}_t]\\
&= \frac{1}{(N-1)!}\sum_{\mathbf{a}_t:\nu_t^{(i)}=1\forall i}\prod_{i\neq i^*}^{N} w_t^{(i)} q_{t-1}(X_t^{(a_t^{(i)})}, X_{t-1}^{(i)})\\
&\leq \frac{1}{(N-1)!}\sum_{\mathbf{a}_t:\nu_t^{(i)}=1\forall i}\prod_{i\neq i^*}^{N} w_t^{(i)}\varepsilon^{-1} h(X_{t-1}^{(i)})\\
&= \prod_{i\neq i^*}^{N} w_t^{(i)}\varepsilon^{-1} h(X_{t-1}^{(i)})\\
&= \varepsilon^{1-N}\left(\prod_{j\neq i^*}^{N} w_t^{(j)}\right)\left(\prod_{i\neq i^*}^{N} h(X_{t-1}^{(i)})\right)\\
&\propto \prod_{j\neq i^*}^{N} w_t^{(j)} =: f(\mathbf{w}_t).
\end{aligned}$$

Evaluated at $\mathbf{w}^*$, we have

$$\mathbb{P}\left[c_N(t) > \frac{2}{N^2}\middle|\mathcal{H}_t, \mathbf{w}_t = \mathbf{w}^*\right] = 1-\varepsilon^{1-N}\left(\prod_{j\neq i^*}^{N}\frac{1}{N-1}\right)\left(\prod_{i\neq i^*}^{N} h(X_{t-1}^{(i)})\right) = 1-(\varepsilon(N-1))^{1-N}\prod_{i\neq i^*}^{N} h(X_{t-1}^{(i)})$$

as required. □

Lemma 7. *In conditional SMC with multinomial resampling, in the optimal case where the weight vector is equal to $\mathbf{w}^*$ at every time step, there exists N_0 such that for all $N > N_0$, for all finite t, $\mathbb{P}[\tau_N(t) = \infty] = 0$.*

Proof. The proof is exactly the same as for Lemma 4; Lemma 6 provides the expression for $P(E_r)$ which is strictly positive and constant in r. □

Lemma 8. *For all $N \geq 2$, for all t, for any weight vector $(w_1, \dots, w_N) \in \mathcal{S}_{N-1}$,*

$$\mathbb{P}\left[c_N(t) > \frac{2}{N^2} \middle| \mathcal{H}_t\right] \geq \mathbb{P}\left[c_N(t) > \frac{2}{N^2} \middle| \mathcal{H}_t, \mathbf{w}_t = \mathbf{w}^*\right].$$

Proof. In the proof of Lemma [4] we found an expression for the probability of interest, $\mathbb{P}[c_N(t) \leq 2/N^2 \mid \mathcal{H}_t]$. To see that $\mathbf{w}^*$ is the "worst case" weight vector (i.e maximising that probability), consider the optimisation of

$$f(\mathbf{w}) := \prod_{\substack{i=1\\i\neq i^*}}^{N} w_i = \left(1 - \sum_{j=1}^{N-1} w_j\right) \prod_{\substack{i=1\\i\neq i^*}}^{N-1} w_i \propto \mathbb{P}\left[c_N(t) \leq \frac{2}{N^2} \mid \mathcal{H}_t, \mathbf{w}_t = \mathbf{w}\right]$$

over $\mathbf{w} \in \mathcal{S}_{N-1}$. This objective function is non-negative and obtains its minimal value zero whenever one or more of the non-immortal weights is equal to zero; since we are looking for a maximum we can assume that $w_i > 0$ for all $i \neq i^*$. Now, for every $k \in \{1, \dots, N-1\} \setminus \{i^*\}$, we solve

$$\frac{\partial f(\mathbf{w})}{\partial w_k} = \left(1 - w_k - \sum_{j=1}^{N-1} w_j\right) \prod_{\substack{i=1\\i\neq k,i^*}}^{N-1} w_i = 0.$$

The product is constant and positive for each k, so this amounts to solving

$$w_k = 1 - \sum_{\substack{j=1\\j\neq i^*}}^{N-1} w_j = w_N$$

simultaneously for all $k \in \{1, \dots, N-1\} \setminus \{i^*\}$. The locus of solutions is the ridge $\mathbf{w}_a = \{(1, \dots, 1) + a\mathbf{e}_{i^*}\}/(N+a)$ for some constant $a \in [-1, \infty)$. (See Figure 1 for an illustration in the case N=3.) It can be shown that the Hessian in indices $\{1, \dots, N\} \setminus \{i^*\}$ is negative definite. The Hessian in the non-immortal indices comes out exactly as in the standard multinomial case; a proof analogous to that one could easily be included. On this ridge the objective function takes values $f(\mathbf{w}_a) = (N+a)^{1-N}$. Further optimising over a, the unique maximum is at $a = -1$, thus $\mathbf{w}^* = \{(1, \dots, 1) - \mathbf{e}_{i^*}\}/(N-1)$. This weight vector thus minimises the probability of the complementary event, and we conclude the result. □

Theorem 2. *In conditional SMC with multinomial resampling, there exists N_0 such that for all $N > N_0$, for all finite t, $\mathbb{P}[\tau_N(t) = \infty] = 0$.*

Proof. Combining Lemmata [4] and [6], we see that $\mathbb{P}\left[c_N(t) > \frac{2}{N^2} \middle| \mathcal{H}_t\right] \geq 1 - (\varepsilon(N-1))^{1-N} \prod_{i\neq i^*}^{N} h(X_{t-1}^{(i)})$ for all t. For sufficiently large N, say $N > N_0$, this probability is bounded away from zero. [A D-separation + Borel–Cantelli argument analogous to the stochastic rounding case will lead to the result.] □

There's an easier way...

Theorem 3. *In conditional SMC with multinomial resampling, there exists N_0 such that for all $N > N_0$, for all finite t, $\mathbb{P}[\tau_N(t) = \infty] = 0$.*

Proof. We have, from the proof of Corollary 2 in the draft paper,

$$\mathbb{E}[c_N(t) \mid \mathcal{F}_{t-1}] \geq \frac{1}{(N)_2} \left\{ \frac{2\varepsilon^2}{a^2} + \frac{(N)_3 \varepsilon^4}{(N-1)^2 a^4} \right\}.$$

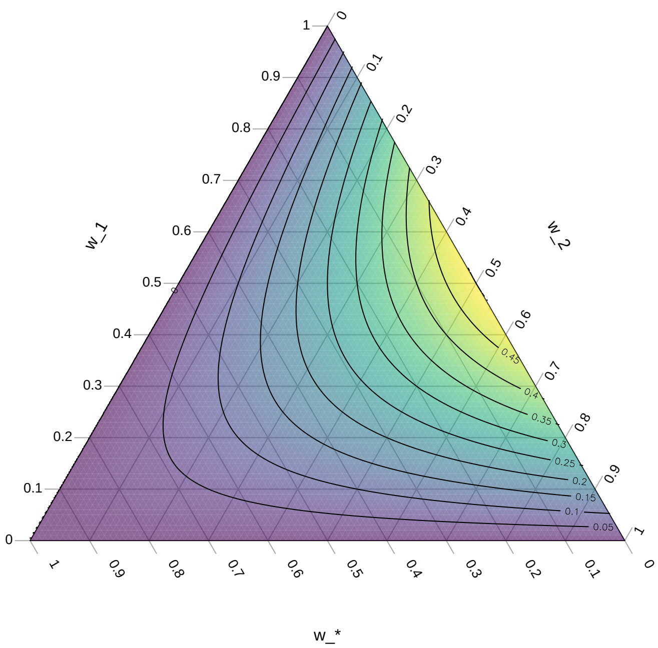

Figure 1: Plot of objective function $f(\mathbf{w})$ in the case $N = 3$, where the immortal index is $i^* = 3$.

Since $c_N(t) \in [0,1]$ almost surely, for any fixed N the "worst-case" distribution of $c_N(t)$ (i.e. maximising $\mathbb{P}[c_N(t) = 0 \mid \mathcal{F}_{t-1}]$) is two atoms, at 0 and 1. To ensure the correct expectation, the atom at 1 must have weight $\mathbb{E}[c_N(t) \mid \mathcal{F}_{t-1}]$, which is bounded below by the above inequality. Hence for any finite N,

$$\sum_{t=0}^{\infty} \mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{F}_{t-1}] \geq \sum_{t=0}^{\infty} \frac{1}{(N)_2} \left\{ \frac{2\varepsilon^2}{a^2} + \frac{(N)_3 \varepsilon^4}{(N-1)^2 a^4} \right\} = \infty.$$

By Borel–Cantelli, we therefore have almost surely for all $N > 2$ that $c_N(t) > 2/N^2$ for infinitely many t, which as argued earlier implies $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all finite t. □

Stochastic rounding

Lemma 9. *Let $\mathbf{w} = (w_1, \dots, w_N) \in \mathcal{S}_{N-1}$ and resample by stochastic rounding.*

(i) If $w_i \geq 2/N$ for some i, then $\mathbb{P}[c_N(t) = 0 | \mathbf{w}] = 0$.

(ii) If $w_i = 0$ for some i, then $\mathbb{P}[c_N(t) = 0 | \mathbf{w}] = 0$.

Proof. In case (i) particle i is assigned at least two offspring, so $c_N(t)$ cannot be equal to zero. In case (ii) particle i is assigned zero offspring, so at least one other particle must be assigned more than one offspring, thus $c_N(t)$ cannot be equal to zero. □

The upshot of Lemma 9 is that in these cases of "extreme weights" we have $c_N(t) > 2/N^2$ almost surely, so we can exclude these cases while we go about bounding $\mathbb{P}[c_N(t) > 2/N^2 | \mathbf{w}]$ away from zero.

Lemma 10. *Define $\mathbf{w}^\delta := \frac{1}{N}\{(1, \dots, 1) + \delta \mathbf{e}_i - \delta \mathbf{e}_j\}$ for any $i \neq j$ and $0 < \delta < 1$. Then $\mathbb{P}[c_N(t) > 2/N^2 | \mathcal{H}_t, \mathbf{w}_t = \mathbf{w}_\delta] \geq \delta \varepsilon^3$.*

Proof. We use a bound on $\mathbb{P}[\nu_t^{(i)} = \lfloor N w_t^{(i)} \rfloor]$ from the proof of Corollary 1 in the draft paper:

$$\mathbb{P}[\nu_t^{(i)} = \lfloor N w_t^{(i)} \rfloor \mid \mathcal{H}_t] =: p_0 = 1 - p_1 \leq 1 - (N w_t^{(i)} - \lfloor N w_t^{(i)} \rfloor) \varepsilon^{(2\lfloor N w_t^{(i)} \rfloor + 1)}.$$

Then

$$\begin{aligned}\mathbb{P}[c_N(t) \le 2/N^2 | \mathcal{H}_t, \mathbf{w}_t = \mathbf{w}_\delta] &= \mathbb{P}[\nu_t^{(1:N)} = (1, \dots, 1) \mid \mathcal{H}_t, \mathbf{w}_t = \mathbf{w}_\delta] \\ &= \mathbb{P}[\nu_t^{(i)} = 1, \nu_t^{(j)} = 1 \mid \mathcal{H}_t, \mathbf{w}_t = \mathbf{w}_\delta] \\ &= \mathbb{P}[\nu_t^{(i)} = 1 \mid \mathcal{H}_t, \mathbf{w}_t = \mathbf{w}_\delta] \\ &\le 1 - (N w_\delta^{(i)} - \lfloor N w_\delta^{(i)} \rfloor) \varepsilon^{(2\lfloor N w_\delta^{(i)} \rfloor + 1)} \\ &= 1 - \{N(1+\delta)/N - 1\}\varepsilon^3 \\ &= 1 - \delta\varepsilon^3,\end{aligned}$$

since the offspring counts are deterministically equal to one apart from particles i and j, and it remains that $\nu_t^{(i)} = 1$ if and only if $\nu_t^{(j)} = 1$. □

Lemma 11. *For any $\delta \in (0,1)$, denote $\mathcal{S}_{N-1}^\delta := \{\mathbf{w} \in \mathcal{S}_{N-1} : \forall i,\, 0 < w_i < \frac{2}{N};\, \max_i w_i \ge \frac{1+\delta}{N}\}$. Then for all $\mathbf{w} \in \mathcal{S}_{N-1}^\delta$, $\mathbb{P}[c_N(t) > 2/N^2 | \mathbf{w}] \ge \mathbb{P}[c_N(t) > 2/N^2 | \mathbf{w}_\delta]$.*

Proof. Fix arbitrary $\mathbf{w} \in \mathcal{S}_{N-1}^\delta$. Let i^* be then index of the particle with the largest weight. Denote $\mathcal{I} := \{i \in \{1, \dots, N\} : w_i > 1/N\}$. Notice that

$$\mathbb{P}[c_N(t) \le 2/N^2 | \mathbf{w}] = \mathbb{P}[\nu_t^{(i)} = 1 \,\forall i \in \{1, \dots, N\} | \mathbf{w}] = \mathbb{P}[\nu_t^{(i)} = 1 \,\forall i \in \mathcal{I} | \mathbf{w}].$$

This is true because all weights are in $(0, 2/N)$, so for $i \in \mathcal{I}, \nu_t^{(i)} \in \{1, 2\}$, and for $i \notin \mathcal{I}, \nu_t^{(i)} \in \{0, 1\}$; and the offspring counts must sum to N (a generalisation of the argument used in Lemma 10).

We can then decompose this probability into a product of conditional probabilities:

$$\begin{aligned}\mathbb{P}[\nu_t^{(i)} = 1 \,\forall i \in \mathcal{I} | \mathbf{w}] &= \prod_{i \in \mathcal{I}} \mathbb{P}[\nu_t^{(i)} = 1 | \nu_t^{(j)} = 1 \,\forall j < i \in \mathcal{I}; \mathbf{w}] \\ &= \mathbb{P}[\nu_t^{(i^*)} = 1 | \mathbf{w}] \prod_{i \ne i^* \in \mathcal{I}} \mathbb{P}[\nu_t^{(i)} = 1 | \nu_t^{(i^*)} = 1; \nu_t^{(j)} = 1 \,\forall j < i \in \mathcal{I}; \mathbf{w}] \\ &\le \mathbb{P}[\nu_t^{(i^*)} = 1 | \mathbf{w}].\end{aligned}$$

The last line is equal to the probability $\mathbb{P}[c_N(t) \le 2/N^2 | \mathbf{w}]$ in the case where $|\mathcal{I}| = 1$, i.e. the only weight larger than $1/N$ is w_{i^*}.

In other words, $\mathbb{P}[c_N(t) > 2/N^2 | \mathbf{w}]$ is minimised on $\mathcal{S}_{N-1}^\delta$ by having only one weight larger than $1/N$, in which case the values of the other weights do not affect this probability.

We therefore find that a minimum of $\mathbb{P}[c_N(t) > 2/N^2 | \mathbf{w}]$ on $\mathcal{S}_{N-1}^\delta$ is given by $\mathbf{w}_{\delta'}$, for some $\delta' \ge \delta$. It only remains to show that taking $\delta' > \delta$ does not decrease the probability. This is a consequence of Lemma 10, where we see that $\mathbb{P}[c_N(t) > 2/N^2 | \mathbf{w}_{\delta'}]$ is monotonically increasing in δ'. Thus the minimum of $\mathbb{P}[c_N(t) > 2/N^2 | \mathbf{w}]$ is attained at $\mathbf{w} = \mathbf{w}_\delta$, as required. (Although this minimum is not unique, we have shown explicitly that it is a global minimum on $\mathcal{S}_{N-1}^\delta$.) □

Theorem 4. *Consider a sequential Monte Carlo algorithm using any stochastic rounding as its resampling scheme. If there exists $\mu > 0$ such that $\mathbb{P}\{\max_i w_t^{(i)} \ge (1+\delta)/N \mid \mathcal{H}_t\} \ge \mu$ for infinitely many t then $\mathbb{P}\{\tau_N(t) = \infty\} = 0$ for all $N > 1$ and for all finite t.*

Proof. Combining Lemmata [7–9] we see that, for any $\mathbf{w} \in \mathcal{S}_{N-1}$ such that $\max_i w_i \ge \frac{1+\delta}{N}$, we have the bound $\mathbb{P}[c_N(t) > 2/N^2 | \mathbf{w}] \ge \delta\varepsilon^3$. By the law of total probability,

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \ge \mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t, \max w_i \ge (1+\delta)/N]\, \mathbb{P}[\max w_i \ge (1+\delta)/N \mid \mathcal{H}_t] \ge \mu\delta\varepsilon^3$$

for those infinitely many t where $\mathbb{P}\{\max_i w_t^{(i)} \ge (1+\delta)/N \mid \mathcal{H}_t\} \ge \mu$. Using the D-separation established in [draft paper, Cor 1 proof], we can write

$$\begin{aligned}\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{F}_{t-1}] &= \mathbb{E}[\mathbb{I}_{\{c_N(t) > 2/N^2\}} \mid \mathcal{F}_{t-1}] \\ &= \mathbb{E}[\mathbb{E}[\mathbb{I}_{\{c_N(t) > 2/N^2\}} \mid \mathcal{H}_t] \mid \mathcal{F}_{t-1}] \\ &= \mathbb{E}[\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] \mid \mathcal{F}_{t-1}].\end{aligned}$$

Hence this probability is bounded below by $\mu\delta\varepsilon^3$ for infinitely many t. We therefore have

$$\sum_{t=0}^{\infty} \mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{F}_{t-1}] \geq \sum_{j=0}^{\infty} \mu\delta\varepsilon^3 = \infty, \tag{2}$$

and applying Theorem [1 - that BC2 statement], almost surely $c_N(t) > 2/N^2$ for infinitely many t. As argued in Lemma [2], this is sufficient for the result. $\square$

The lemma below is here to clear up any uncertainty about the tower property / D-separation argument, as used in this proof in the paper.

Lemma 12. *Let A, B be events such that A is measurable with respect to $\mathcal{F}_t$, and B is measurable with respect to $\mathcal{H}_t$ (but not vice versa), and neither event is measurable with respect to $\mathcal{F}_{t-1}$. (In the real proof we have $A := \{c_N(t) > 2/N^2\}$ and $B := \{\max_i w_t^{(i)} - \min_i w_t^{(i)} \geq 2\delta/N\}$). Then*

$$\mathbb{P}[A \mid \mathcal{F}_{t-1}, B] = \mathbb{E}[\mathbb{P}[A \mid \mathcal{H}_t] \mid \mathcal{F}_{t-1}, B] \tag{3}$$

and

$$\mathbb{P}[A \mid \mathcal{H}_t] \geq \mathbb{P}[A \mid \mathcal{H}_t, B]\, \mathbb{I}_B. \tag{4}$$

Proof. For the first point,

$$\begin{aligned}
\mathbb{P}[A \mid \mathcal{F}_{t-1}, B] = \mathbb{E}[\mathbb{I}_A \mid \mathcal{F}_{t-1}, B] &= \frac{\mathbb{E}[\mathbb{I}_A \mathbb{I}_B \mid \mathcal{F}_{t-1}]}{\mathbb{P}[B \mid \mathcal{F}_{t-1}]} = \frac{\mathbb{E}[\mathbb{E}[\mathbb{I}_A \mathbb{I}_B \mid \mathcal{H}_t, \mathcal{F}_{t-1}] \mid \mathcal{F}_{t-1}]}{\mathbb{P}[B \mid \mathcal{F}_{t-1}]} = \frac{\mathbb{E}[\mathbb{E}[\mathbb{I}_A \mathbb{I}_B \mid \mathcal{H}_t] \mid \mathcal{F}_{t-1}]}{\mathbb{P}[B \mid \mathcal{F}_{t-1}]} \\
&= \frac{\mathbb{E}[\mathbb{E}[\mathbb{I}_A \mid \mathcal{H}_t] \mathbb{I}_B \mid \mathcal{F}_{t-1}]}{\mathbb{P}[B \mid \mathcal{F}_{t-1}]} = \frac{\mathbb{E}[\mathbb{E}[\mathbb{I}_A \mid \mathcal{H}_t] \mid \mathcal{F}_{t-1}, B] \mathbb{P}[B \mid \mathcal{F}_{t-1}]}{\mathbb{P}[B \mid \mathcal{F}_{t-1}]} \\
&= \mathbb{E}[\mathbb{E}[\mathbb{I}_A \mid \mathcal{H}_t] \mid \mathcal{F}_{t-1}, B] = \mathbb{E}[\mathbb{P}[A \mid \mathcal{H}_t] \mid \mathcal{F}_{t-1}, B].
\end{aligned}$$

For the second point,

$$\mathbb{P}[A \mid \mathcal{H}_t] = \mathbb{P}[A \mid \mathcal{H}_t, B]\mathbb{P}[B \mid \mathcal{H}_t] + \mathbb{P}[A \mid \mathcal{H}_t, B^c]\mathbb{P}[B^c \mid \mathcal{H}_t] \geq \mathbb{P}[A \mid \mathcal{H}_t, B]\mathbb{P}[B \mid \mathcal{H}_t] = \mathbb{P}[A \mid \mathcal{H}_t, B]\mathbb{I}_B$$

since B is $\mathcal{H}_t$-measurable. $\square$

The next Lemma shows how these two results are helpful in our scenario of Corollary 1.

Lemma 13. *Let $A := \{c_N(t) > 2/N^2\}$. Let $B := \{\max_i w_t^{(i)} - \min_i w_t^{(i)} \geq 2\delta/N\}$. (Notice that these events satisfy the measurability properties in the previous Lemma.) As an assumption in Corollary 1 we have that $\mathbb{P}[B \mid \mathcal{F}_{t-1}] \geq \zeta > 0$ for infinitely many t. We showed in the proof of Corollary 1 that $\mathbb{P}[A \mid \mathcal{H}_t, B] \geq \delta\varepsilon^3$. Then, under this set-up, we have $\mathbb{P}[A \mid \mathcal{F}_{t-1}] \geq \zeta\delta\varepsilon^3$ for infinitely many t.*

Proof.

$$\begin{aligned}
\mathbb{P}[A \mid \mathcal{F}_{t-1}] = \mathbb{E}[\mathbb{I}_A \mid \mathcal{F}_{t-1}] = \mathbb{E}[\mathbb{E}[\mathbb{I}_A \mid \mathcal{F}_{t-1}, B] \mid \mathcal{F}_{t-1}] &= \mathbb{E}[\mathbb{P}[A \mid \mathcal{F}_{t-1}, B] \mid \mathcal{F}_{t-1}] = \mathbb{E}[\mathbb{E}[\mathbb{P}[A \mid \mathcal{H}_t] \mid \mathcal{F}_{t-1}, B] \mid \mathcal{F}_{t-1}] \\
&\geq \mathbb{E}[\mathbb{E}[\mathbb{P}[A \mid \mathcal{H}_t, B]\mathbb{I}_B \mid \mathcal{F}_{t-1}, B] \mid \mathcal{F}_{t-1}] = \mathbb{E}[\mathbb{E}[\mathbb{P}[A \mid \mathcal{H}_t, B] \mid \mathcal{F}_{t-1}, B]\mathbb{I}_B \mid \mathcal{F}_{t-1}] \\
&\geq \mathbb{E}[\mathbb{E}[\delta\varepsilon^3 \mid \mathcal{F}_{t-1}, B]\mathbb{I}_B \mid \mathcal{F}_{t-1}] = \mathbb{E}[\delta\varepsilon^3 \mathbb{I}_B \mid \mathcal{F}_{t-1}] = \delta\varepsilon^3 \mathbb{E}[\mathbb{I}_B \mid \mathcal{F}_{t-1}] = \delta\varepsilon^3 \mathbb{P}[B \mid \mathcal{F}_{t-1}] \\
&\geq \zeta\delta\varepsilon^3 \text{ for infinitely many } t.
\end{aligned}$$

$\square$

References

Durrett, R. (2019), *Probability: Theory and Examples*, Cambridge Series in Statistical and Probabilistic Mathematics, 5 edn, Cambridge University Press.