Non-triviality condition (fuller details of paper appendix)

Suzie Brown

July 14, 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 for infinitely many r. □

Combining Lemmata 1 and 2, we see that a sufficient condition for non-triviality is

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

Multinomial resampling

Proposition 1. *In standard 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 the following lower bound:

$$\mathbb{E}[c_N(t) \mid \mathcal{F}_{t-1}] \geq \frac{\varepsilon^4}{Na^4}.$$

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,

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

Applying Lemmata 1 and 2 yields the result. □

Conditional SMC with multinomial resampling

Proposition 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. We have the following lower bound:

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

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,

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

Applying Lemmata 1 and 2 yields the result. □

Stochastic rounding

The stochastic rounding case is more involved than the others, because there is not a positive lower bound on $\mathbb{E}_t[c_N(t)]$ that holds for all weight vectors. (In particular, $\mathbb{E}_t[c_N(t)] = 0$ when the weights are all exactly equal.) First notice the following.

Lemma 3. *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 $\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] = 1$.

(ii) If $w_i = 0$ for some i, then $\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] = 1$.

Proof. For any N it is easy to see, by considering the possible values of $c_N(t)$, that $c_N(t) > 2/N^2$ if and only if $c_N(t) \neq 0$. The only way to attain $c_N(t) = 0$ is to assign exactly one offspring to each particle. 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 $c_N(t)$ cannot be equal to zero. □

The upshot of Lemma 3 is that we need only prove (1) in the case where all weights are in $(0, 2/N)$, since it holds automatically otherwise.

Lemma 4. *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

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

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

Proof. Fix arbitrary $\mathbf{w} \in \mathcal{S}^{\delta}_{N-1}$. 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

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

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 4).

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

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

The last line is equal to the probability $\mathbb{P}[c_N(t) \leq 2/N^2 | \mathbf{w}]$ 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 $\mathcal{S}^{\delta}_{N-1}$ 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 $\mathcal{S}^{\delta}_{N-1}$ is given by $\mathbf{w}_{\delta'}$, for some $\delta' \geq \delta$. It only remains to show that taking $\delta' > \delta$ does not decrease the probability. This is a consequence of Lemma 4, 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 $\mathcal{S}^{\delta}_{N-1}$.) $\square$

Theorem 1. *Consider a sequential Monte Carlo algorithm using any stochastic rounding as its resampling scheme. If there exists $\mu > 0$ such that $\mathbb{P}\{\max_i w_t^{(i)} \geq (1+\delta)/N \mid \mathcal{H}_t\} \geq \mu$ 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 $\max_i w_i \geq \frac{1+\delta}{N}$, we have the bound $\mathbb{P}[c_N(t) > 2/N^2 | \mathbf{w}] \geq \delta\varepsilon^3$. By the law of total probability,

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

for those infinitely many t where $\mathbb{P}\{\max_i w_t^{(i)} \geq (1+\delta)/N \mid \mathcal{H}_t\} \geq \mu$. 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 6. *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. □

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

Lemma 7. *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.

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

□

DUMPED FROM PAPER:

Proof. Let $\mathcal{H}_t$ be defined as in (**??**). The first step is to show that whenever $\max_i w_t^{(i)} \geq (1+\delta)/N$, $\mathbb{P}\{c_N(t) > 2/N^2 | \mathcal{H}_t\} = \mathbb{P}\{c_N(t) \neq 0 | \mathcal{H}_t\} \geq \delta\varepsilon^3$. For this purpose we need consider only weight vectors such that $w_t^{(i)} \in (0, 2/N)$ for all i; otherwise $\mathbb{P}\{c_N(t) \neq 0 | \mathcal{H}_t\} = 1$ by the definition of stochastic rounding.

Denote $\mathcal{S}_{N-1}^{\delta} = \{w^{(1:N)} \in \mathcal{S}_{N-1} : \forall i,\, 0 < w^{(i)} < 2/N;\, \max_i w^{(i)} \geq (1+\delta)/N\}$ for any $\delta \in (0,1)$, where $\mathcal{S}_k$ denotes the k-dimensional simplex. Fix arbitrary $w_t^{(1:N)} \in \mathcal{S}_{N-1}^{\delta}$. Set $i^\star = \arg\max_i w_t^{(i)}$ and denote $\mathcal{I} = \{i \in \{1, \dots, N\} : w^{(i)} > 1/N\}$. Since all weights are in $(0, 2/N)$, for $i \in \mathcal{I}, \nu_t^{(i)} \in \{1, 2\}$ and for $i \notin \mathcal{I}, \nu_t^{(i)} \in \{0, 1\}$; and since the offspring counts must sum to N, we can write

$$\begin{aligned}\mathbb{P}\{c_N(t) \leq 2/N^2 | \mathcal{H}_t\} &= \mathbb{P}(\nu_t^{(i)} = 1 \, \forall i \in \{1, \dots, N\} | \mathcal{H}_t) \\ &= \mathbb{P}(\nu_t^{(i)} = 1 \, \forall i \in \mathcal{I} | \mathcal{H}_t) \\ &= \prod_{i \in \mathcal{I}} \mathbb{P}(\nu_t^{(i)} = 1 | \nu_t^{(j)} = 1 \, \forall j \in \mathcal{I} : j < i; \mathcal{H}_t) \\ &= \mathbb{P}(\nu_t^{(i^\star)} = 1 | \mathcal{H}_t) \prod_{\substack{i \in \mathcal{I} \\ i \neq i^\star}} \mathbb{P}(\nu_t^{(i)} = 1 | \nu_t^{(i^\star)} = 1; \nu_t^{(j)} = 1 \, \forall j \in \mathcal{I} : j < i; \mathcal{H}_t) \\ &\leq \mathbb{P}(\nu_t^{(i^\star)} = 1 | \mathcal{H}_t).\end{aligned}$$

The final inequality holds with equality when $|\mathcal{I}| = 1$, i.e. the only weight larger than $1/N$ is $w_t^{(i^\star)}$. Thus $\mathbb{P}\{c_N(t) > 2/N^2 | \mathcal{H}_t\}$ is minimised on $\mathcal{S}_{N-1}^{\delta}$ when only one weight is larger than $1/N$, in which case the values of the other weights do not affect this probability.

Define $w_{\delta'} = \{(1, \ldots, 1) + \delta' e_{i^\star} - \delta' e_{j^\star}\}/N$ for fixed $i^\star \neq j^\star$ and $\delta' \in (0, 1)$, where e_i denotes the ith canonical basis vector in $\mathbb{R}^N$. Using the lower bound (**??**) on $p_1(w_t^{(i)}) = \mathbb{P}(\nu_t^{(i)} = \lfloor N w_t^{(i)} \rfloor + 1 \mid \mathcal{H}_t)$,

$$\begin{aligned} \mathbb{P}\{c_N(t) \leq 2/N^2 \mid \mathcal{H}_t, w_t^{(1:N)} = w_{\delta'}\} &= \mathbb{P}(\nu_t^{(i^\star)} = 1 \mid \mathcal{H}_t, w_t^{(1:N)} = w_{\delta'}) \\ &= 1 - p_1\{(1 + \delta')/N\} \leq 1 - \delta' \varepsilon^3. \end{aligned}$$

We conclude that $\mathbb{P}\{c_N(t) > 2/N^2 | \mathcal{H}_t, \max_i w_t^{(i)} \geq (1 + \delta)/N\} \geq \min_{\delta' \geq \delta} \delta' \varepsilon^3 = \delta \varepsilon^3$.

A slight modification of this argument yields $\mathbb{P}\{c_N(t) > 2/N^2 | \mathcal{H}_t, \min_i w_t^{(i)} \leq (1 - \delta)/N\} \geq \delta \varepsilon^3$. Whenever $\max_i w_t^{(i)} - \min_i w_t^{(i)} \geq 2\delta/N$, either $\max_i w_t^{(i)} \geq (1 + \delta)/N$ or $\min_i w_t^{(i)} \leq (1 - \delta)/N$, so we have $\mathbb{P}\{c_N(t) > 2/N^2 | \mathcal{H}_t, \max_i w_t^{(i)} - \min_i w_t^{(i)} \geq 2\delta/N\} \geq \delta \varepsilon^3$. Thus $\mathbb{P}\{c_N(t) > 2/N^2 \mid \mathcal{H}_t\} \geq \delta \varepsilon^3 \mathbb{I}_{\max_i w_t^{(i)} - \min_i w_t^{(i)} \geq 2\delta/N}$. Using the D-separation established in Appendix **??** combined with the tower property, we have

$$\mathbb{P}\{c_N(t) > 2/N^2 \mid \mathcal{F}_{t-1}\} = \mathbb{E}_t[\mathbb{P}\{c_N(t) > 2/N^2 \mid \mathcal{H}_t\}] \geq \delta \varepsilon^3 \mathbb{P}(\max_i w_t^{(i)} - \min_i w_t^{(i)} \geq 2\delta/N \mid \mathcal{F}_{t-1}),$$

which is bounded below by $\zeta \delta \varepsilon^3$ for infinitely many t. Hence,

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

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

References

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