Recall the quantities

$$
\begin{aligned}
c_N(t) &:= \frac{1}{(N)_2} \sum_{i=1}^{N} (\nu_t^{(i)})_2, \\
D_N(t) &:= \frac{1}{N(N)_2} \sum_{i=1}^{N} (\nu_t^{(i)})_2 \left(\nu_t^{(i)} + \frac{1}{N} \sum_{j \neq i}^{N} (\nu_t^{(j)})^2 \right);
\end{aligned}
$$

the assumptions

$$
\mathbb{E}[c_N(t)] \to 0, \tag{1}
$$

$$
\mathbb{E}\left[\sum_{r=\tau_N(s)+1}^{\tau_N(t)} D_N(r) \right] \to 0, \tag{2}
$$

$$
\mathbb{E}\left[\sum_{r=\tau_N(s)+1}^{\tau_N(t)} c_N(r)^2 \right] \to 0; \tag{3}
$$

and the reverse-time filtration $\mathcal{F}_t := \sigma(\boldsymbol{\nu}_s; 1 \leq s \leq t)$.

Lemma 1. *We have*

$$
c_N(t)^2 \leq \frac{N}{N-1} D_N(t),
$$

so that (2) $\Rightarrow$ (3).

Proof.

$$
\begin{aligned}
c_N(t)^2 &= \frac{1}{N(N-1)(N)_2} \sum_{i=1}^{N} (\nu_t^{(i)})_2 \left(\nu_t^{(i)}(\nu_t^{(i)} - 1) + \sum_{j \neq i}^{N} (\nu_t^{(j)})_2 \right) \\
&= \frac{1}{N(N)_2} \sum_{i=1}^{N} (\nu_t^{(i)})_2 \left(\frac{\nu_t^{(i)}(\nu_t^{(i)} - 1)}{N-1} + \frac{1}{N-1} \sum_{j \neq i}^{N} (\nu_t^{(j)})_2 \right) \\
&\leq \frac{1}{N(N)_2} \sum_{i=1}^{N} (\nu_t^{(i)})_2 \left(\nu_t^{(i)} + \frac{1}{N-1} \sum_{j \neq i}^{N} (\nu_t^{(j)})_2 \right) \\
&\leq \frac{1}{N(N)_2} \sum_{i=1}^{N} (\nu_t^{(i)})_2 \left(\nu_t^{(i)} + \frac{N/(N-1)}{N} \sum_{j \neq i}^{N} (\nu_t^{(j)})^2 \right) \\
&\leq \frac{N/(N-1)}{N(N)_2} \sum_{i=1}^{N} (\nu_t^{(i)})_2 \left(\nu_t^{(i)} + \frac{1}{N} \sum_{j \neq i}^{N} (\nu_t^{(j)})^2 \right) = \frac{N}{N-1} D_N(t).
\end{aligned}
$$

□

We now introduce a new assumption: for some deterministic sequence $b_N \to 0$, we have

$$
\frac{1}{(N)_3} \sum_{i=1}^{N} \mathbb{E}[(\nu_t^{(i)})_3 | \mathcal{F}_{t-1}] \leq b_N \mathbb{E}[c_N(t) | \mathcal{F}_{t-1}] \tag{4}
$$

uniformly in $t \geq 1$.

Lemma 2. (4) $\Rightarrow$ (1) *and* (4) $\Rightarrow$ (2).

Proof. We prove the two implications separately, starting with the former. Following the proof of [Möhle and Sagitov, 2003, Lemma 5.5], we fix $\varepsilon > 0$ and define the event $A_i := \{\nu_t^{(i)} \leq N\varepsilon\}$. Now

$$\begin{aligned}
\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}] &= \frac{1}{(N)_2}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_2|\mathcal{F}_{t-1}] \\
&= \frac{1}{(N)_2}\sum_{i=1}^N \left\{\mathbb{E}[(\nu_t^{(i)})_2\mathbb{1}_{A_i}|\mathcal{F}_{t-1}] + \mathbb{E}[(\nu_t^{(i)})_2\mathbb{1}_{A_i^c}|\mathcal{F}_{t-1}]\right\} \\
&\leq \frac{\varepsilon}{N-1}\sum_{i=1}^N \mathbb{E}[\nu_t^{(i)}\mathbb{1}_{A_i}|\mathcal{F}_{t-1}] + \sum_{i=1}^N \mathbb{E}[\mathbb{1}_{A_i^c}|\mathcal{F}_{t-1}] \\
&\leq \{1 + O(N^{-1})\}\varepsilon + \sum_{i=1}^N \mathbb{P}(\nu_t^{(i)} > N\varepsilon|\mathcal{F}_{t-1}). \qquad (5)
\end{aligned}$$

For $N \geq 3/\varepsilon$, Markov's inequality yields

$$\begin{aligned}
\sum_{i=1}^N \mathbb{P}(\nu_t^{(i)} > N\varepsilon|\mathcal{F}_{t-1}) &\leq \frac{1}{(N\varepsilon)_3}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_3|\mathcal{F}_{t-1}] = \frac{\{1 + O(N^{-1})\}}{\varepsilon^3(N)_3}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_3|\mathcal{F}_{t-1}] \\
&\leq \{1 + O(N^{-1})\}\frac{b_N}{\varepsilon^3}\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}]. \qquad (6)
\end{aligned}$$

Substituting (6) into (5) and using $c_N(t) \leq 1$ results in

$$\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}] \leq \{1 + O(N^{-1})\}\left(\varepsilon + \frac{b_N}{\varepsilon^3}\right) \to \varepsilon$$

because $b_N \to 0$. Since $\varepsilon > 0$ was arbitrary, we have

$$\mathbb{E}[c_N(t)] = \mathbb{E}[\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}]] \to 0$$

as $N \to \infty$.

We will show (4) $\Rightarrow$ (2) in two parts, the first of which is

$$\begin{aligned}
\frac{1}{N(N)_2}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_2\nu_t^{(i)}|\mathcal{F}_{t-1}] &= \frac{1}{N(N)_2}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_3 + 2(\nu_t^{(i)})_2|\mathcal{F}_{t-1}] \\
&\leq \frac{1}{(N)_3}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_3|\mathcal{F}_{t-1}] + \frac{2}{N}\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}] \\
&\leq \left(b_N + \frac{2}{N}\right)\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}]. \qquad (7)
\end{aligned}$$

For the second, note

$$\begin{aligned}
\frac{1}{N^2(N)_2}\sum_{i\neq j}^N \mathbb{E}[(\nu_t^{(i)})_2(\nu_t^{(j)})^2|\mathcal{F}_{t-1}] &= \frac{1}{N^2(N)_2}\sum_{i\neq j=1}^N \mathbb{E}[(\nu_t^{(i)})_2(\nu_t^{(j)})_2 + (\nu_t^{(i)})_2\nu_t^{(j)}|\mathcal{F}_{t-1}] \\
&\leq \frac{1}{N^2(N)_2}\sum_{i\neq j}^N \mathbb{E}[(\nu_t^{(i)})_2(\nu_t^{(j)})_2|\mathcal{F}_{t-1}] + \frac{\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}]}{N}. \qquad (8)
\end{aligned}$$

Now

$$\begin{aligned}
\sum_{i\neq j}^{N} \mathbb{E}[(\nu_t^{(i)})_2(\nu_t^{(j)})_2|\mathcal{F}_{t-1}] &= \sum_{i\neq j}^{N} \left\{\mathbb{E}[(\nu_t^{(i)})_2(\nu_t^{(j)})_2 \mathbb{1}_{A_i}|\mathcal{F}_{t-1}] + \mathbb{E}[(\nu_t^{(i)})_2(\nu_t^{(j)})_2 \mathbb{1}_{A_i^c}|\mathcal{F}_{t-1}]\right\} \\
&\leq N\varepsilon \sum_{i\neq j}^{N} \mathbb{E}[\nu_t^{(i)}(\nu_t^{(j)})_2 \mathbb{1}_{A_i}|\mathcal{F}_{t-1}] + N^3 \sum_{i\neq j}^{N} \mathbb{E}[\nu_t^{(j)} \mathbb{1}_{A_i^c}|\mathcal{F}_{t-1}] \\
&\leq N^2(N)_2\varepsilon\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}] + N^4 \sum_{i=1}^{N} \mathbb{P}(\nu_t^{(i)} > N\varepsilon|\mathcal{F}_{t-1}). \qquad (9)
\end{aligned}$$

Substituting (6) into (9) yields

$$\sum_{i\neq j}^{N} \mathbb{E}[(\nu_t^{(i)})_2(\nu_t^{(j)})_2|\mathcal{F}_{t-1}] \leq N^4\{1 + O(N^{-1})\}\left(\varepsilon + \frac{b_N}{\varepsilon^3}\right)\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}], \qquad (10)$$

and substituting (10) into (8) gives

$$\frac{1}{N^2(N)_2} \sum_{i\neq j}^{N} \mathbb{E}[(\nu_t^{(i)})_2(\nu_t^{(j)})^2|\mathcal{F}_{t-1}] \leq \left(\{1 + O(N^{-1})\}\left[\varepsilon + \frac{b_N}{\varepsilon^3}\right] + \frac{1}{N}\right)\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}]. \qquad (11)$$

Finally, invoking Lemma 2 from our paper twice, with (7) and (11) in between, gives

$$\begin{aligned}
\mathbb{E}\left[\sum_{r=\tau_N(s)+1}^{\tau_N(t)} D_N(r)\right] &= \mathbb{E}\left[\sum_{r=\tau_N(s)+1}^{\tau_N(t)} \mathbb{E}[D_N(r)|\mathcal{F}_{t-1}]\right] \\
&\leq \left(\{1 + O(N^{-1})\}\left[\varepsilon + \frac{b_N}{\varepsilon^3}\right] + \frac{3}{N} + b_N\right)\mathbb{E}\left[\sum_{r=\tau_N(s)+1}^{\tau_N(t)} c_N(r)\right] \\
&\leq \left(\{1 + O(N^{-1})\}\left[\varepsilon + \frac{b_N}{\varepsilon^3}\right] + \frac{3}{N} + b_N\right)(t - s + 1) \to \varepsilon(t - s + 1),
\end{aligned}$$

and recalling that $\varepsilon > 0$ was arbitrary concludes the proof. □

References

M. Möhle and S. Sagitov. Coalescent patterns in exchangeable diploid population models. *J. Math. Biol.*, 47:337–352, 2003.