Resampling and genealogies in sequential Monte Carlo algorithms

Susanna Elizabeth Brown

A thesis submitted for the degree of
Doctor of Philosophy in Statistics

University of Warwick, Department of Statistics

January 2021

Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank...

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

The work presented (including data generated and data analysis) was carried out by the author except in the cases outlined below:

Parts of this thesis have been published by the author:

Notation and conventions

$\mathbb{N}$	the natural numbers starting from one, $\{1, 2, \dots\}$
$\mathbb{N}_0$	the natural numbers starting from zero, $\{0, 1, 2, \dots\}$
$[a]$	the set $\{1, 2, \dots, a\}$ where $a \in \mathbb{N}$
$(a)_b$	the falling factorial $a(a-1)\cdots(a-b+1)$ where $a, b \in \mathbb{N}$
$\prod_\emptyset$	the empty product is taken to be 1
$\sum_\emptyset$	the empty sum is taken to be 0, while the sum over an index vector of length zero is the identity operator ?

1 Introduction

2 Background

Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin.

John von Neumann

2.1 Sequential Monte Carlo

2.1.1 Motivation

Being Bayesian. SSMs/HMMs. Example(s) of SSM (1D train?).

2.1.2 Inference in SSMs

What quantities do we want to infer? Why is this generally difficult? Filtering, prediction, smoothing, likelihood/normalising constant.

2.1.3 Exact solutions

Which SSMs are tractable? Kalman filter, extended KF, unscented KF, other conjugate models.

2.1.4 Feynman-Kac models

Define a generic FK model. Show that this class includes all SSMs. Example of non-SSM that is FK?

2.1.5 Sequential Monte Carlo for Feynman-Kac models

Present generic algorithm. State the SMC estimators of the quantities of interest. Include the dependence diagram and note that the offspring counts are not independent at each time, but can be made so by conditioning on the separatrix $\mathcal{H}$.

2.1.6 Theoretical justification

How come SMC works? Convergence results (briefly!) e.g. CLT.

2.2 Coalescent theory

2.2.1 Kingman's coalescent

Define the n-coalescent, and Kingman's coalescent as extension of it. (Do I need to introduce random partitions first?)

2.2.2 Properties

Properties of Kingman's coalescent / n-coalescent. Distributions of branch length, waiting times, time to MRCA. Coming down from infinity.

2.2.3 Models in population genetics

Talk about KC's "domain of attraction". Introduce Wright-Fisher model, Moran model, Cannings models.

2.2.4 Particle populations

Particles = individuals, iterations = generations. In what ways is SMC like a population model? (constant population size, non-overlapping generations, discrete time). In what ways is SMC not like a population model? (non-neutral, non-Markov?)

2.3 Sequential Monte Carlo genealogies

2.3.1 From particles to genealogies

How does the SMC algorithm induce a genealogy? (resampling = parent-child relationship).

2.3.2 Performance

How do genealogies affect performance? Variance (and variance estimation?), storage cost. Ancestral degeneracy.

2.3.3 Mitigating ancestral degeneracy

Low-variance resampling (save details for next section). Adaptive resampling: idea of balancing weight/ancestral degeneracy; rule of thumb for implementing it; when is it effective or not?; necessary changes to our generic SMC algorithm (calculation of weights in particular). Backward sampling: when is it possible to do this?

2.3.4 Asymptotics

Why are large population asymptotics useful? Existing results (path storage, KJJS).

2.4 Resampling

2.4.1 Definition

The job of resampling (map weights to counts). Define "valid" resampling schemes (the three rules). Counter-examples where these rules are violated (the examples I've mentioned in previous writings, plus optimal transport resampling and that one FC told me about recently).

2.4.2 What makes a good resampling scheme?

Low-variance: variance of what? Different criteria/ definitions of optimality. Negative association. Link back to adaptive resampling: interaction between adaptive and low-variance resampling.

2.4.3 Examples

Tour of the key resampling schemes (multinomial, residual-*, stratified, systematic, and the worst possible scheme). Comparison of properties of these, existing results comparing schemes. Implementation considerations. Theoretical justification (or lack of).

2.4.4 Stochastic rounding

Define stochastic rounding. Resampling schemes contained by this class. General properties for this class (marginal distributions, negative association, minimum-variance).

2.5 Conditional SMC

2.5.1 Particle MCMC

Motivate particle MCMC methods.

2.5.2 Particle Gibbs algorithm

Present particle Gibbs algorithm. Explain why CSMC is required within particle Gibbs.

2.5.3 Ancestor sampling

Algorithm (or required changes to generic algorithm). Relation to backward sampling. When can it be implemented? Effect on performance (when is it effective?).

3 Limits

3.1 Encoding genealogies

3.1.1 The genealogical process

Encoding as process on space of partitions $\mathcal{P}_n$. Argue that this encodes everything we need. Initial and absorbing states. Intuit with diagram(s), explain relationship between partition blocks and genealogical tree.

3.1.2 Time scale

Introduce c_N, τ_N, D_N. Contrast to pop gen literature, e.g. our c_N/time scale is random. Properties of these quantities: $c_N, D_N \in [0,1]$, and $D_N \leq c_N$ and $\sum_{r=1}^{\tau_N(t)} c_N \in [t, t+1]$ (or rather the version of that with general start time).

3.1.3 Transition probabilities

Introduce $p_{\xi\eta}$. Present expression for that (or at least for $p_{\xi\xi}$), and hence the bounds on it that will be used later (keeping big-O terms explicit where possible).

3.2 An existing limit theorem

State KJJS theorem. Discuss the conditions in detail. Give outline of proof.

3.3 A new limit theorem

State our limit theorem. Give intuition for the new condition. Compare to KJJS: why our conditions might be considered "weaker" (Moran model example, and whatever else we said to our referee/in the BJJK article); our condition is easier to check (as demonstrated in later corollaries).

3.3.1 Proof of theorem

Proof that KJJS conditions are implied by ours. Modification of KJJS proof (or even write out a complete proof?) using weaker bound on $p_{\xi\xi}$ (that bound should have been stated and proved already in transition probabilities section).

4 Applications

4.1 Multinomial resampling

This is the easy-to-analyse scheme, because conditionally i.i.d., and was presented in KJJS already. Now (with our simpler conditions) it is easier to show.

4.1.1 Proof of main condition

4.1.2 Proof of finite time scale condition

4.2 Stochastic rounding

4.2.1 Proof of main condition

Corollary 4.1. *Consider an SMC algorithm using any stochastic rounding as its resampling scheme, such that the standing assumption is satisfied. Assume that there exists a constant $a \in [1, \infty)$ such that for all x, x', t,*

$$\frac{1}{a} \leq g_t(x, x') \leq a. \tag{4.1}$$

Assume that $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all finite t. Let $(G_t^{(n,N)})_{t\geq 0}$ denote the genealogy of a random sample of n terminal particles from the output of the algorithm when the total number of particles used is N. Then, for any fixed n, the time-scaled genealogy $(G_{\tau_N(t)}^{(n,N)})_{t\geq 0}$ converges to Kingman's n-coalescent as $N \to \infty$, in the sense of finite-dimensional distributions.

Proof. Using the forward-time Markov property of SMC, and the associated conditional dependence graph, for each N we establish a sequence of σ-algebras

$$\mathcal{H}_t := \sigma(X_{t-1}^{(1:N)}, X_t^{(1:N)}, w_{t-1}^{(1:N)}, w_t^{(1:N)}) \tag{4.2}$$

such that $\nu_t^{(1:N)}$ is conditionally independent of the filtration $\mathcal{F}_{t-1}$ given $\mathcal{H}_t$. The full D-separation argument is presented in Appendix **??**.

Defining the family sizes $\nu_t^{(i)} = |\{j : a_t^{(j)} = i\}|$ as functions of $a_t^{(1:N)}$, we have the almost sure constraint $\nu_t^{(i)} \in \{\lfloor Nw_t^{(i)} \rfloor, \lfloor Nw_t^{(i)} \rfloor + 1\}$. Denote $p_0^{(i)} := \mathbb{P}[\nu_t^{(i)} = \lfloor Nw_t^{(i)} \rfloor \mid \mathcal{H}_t]$ and $p_1^{(i)} := \mathbb{P}[\nu_t^{(i)} = \lfloor Nw_t^{(i)} \rfloor + 1 \mid \mathcal{H}_t] = 1 - p_0^{(i)}$.

We obtain the following upper bounds, using the almost sure bounds $w_t^{(i)} \leq a^2/N$ which follow from (4.1) along with the form of the weights in Algorithm **??**:

$$\begin{aligned}
\mathbb{E}[(\nu_t^{(i)})_3 \mid \mathcal{H}_t] &= p_0^{(i)}(\lfloor N w_t^{(i)} \rfloor)_3 + p_1^{(i)}(\lfloor N w_t^{(i)} \rfloor + 1)_3 \\
&= \lfloor N w_t^{(i)} \rfloor(\lfloor N w_t^{(i)} \rfloor - 1)\{p_0^{(i)}(\lfloor N w_t^{(i)} \rfloor - 2) + p_1^{(i)}(\lfloor N w_t^{(i)} \rfloor + 1)\} \\
&= \lfloor N w_t^{(i)} \rfloor(\lfloor N w_t^{(i)} \rfloor - 1)\{\lfloor N w_t^{(i)} \rfloor(p_0^{(i)} + p_1^{(i)}) - 2p_0^{(i)} + p_1^{(i)}\} \\
&= \lfloor N w_t^{(i)} \rfloor(\lfloor N w_t^{(i)} \rfloor - 1)\{\lfloor N w_t^{(i)} \rfloor - 2p_0^{(i)} + p_1^{(i)}\} \\
&\leq a^2(a^2 - 1)(a^2 - 0 + 1)\mathbb{1}_{\lfloor N w_t^{(i)} \rfloor \geq 2} \\
&\leq (a^2 + 1)^3 \mathbb{1}_{\lfloor N w_t^{(i)} \rfloor \geq 2}.
\end{aligned}$$

We also have the lower bounds

$$\begin{aligned}
\mathbb{E}[(\nu_t^{(i)})_2 \mid \mathcal{H}_t] &= p_0^{(i)}(\lfloor N w_t^{(i)} \rfloor)_2 + p_1^{(i)}(\lfloor N w_t^{(i)} \rfloor + 1)_2 \\
&= \lfloor N w_t^{(i)} \rfloor\{p_0^{(i)}(\lfloor N w_t^{(i)} \rfloor - 1) + p_1^{(i)}(\lfloor N w_t^{(i)} \rfloor + 1)\} \\
&= \lfloor N w_t^{(i)} \rfloor\{\lfloor N w_t^{(i)} \rfloor(p_0^{(i)} + p_1^{(i)}) - p_0^{(i)} + p_1^{(i)}\} \\
&= \lfloor N w_t^{(i)} \rfloor\{\lfloor N w_t^{(i)} \rfloor - p_0^{(i)} + p_1^{(i)}\} \\
&\geq 2(2 - 1 + 0)\mathbb{1}_{\lfloor N w_t^{(i)} \rfloor \geq 2} = 2\mathbb{1}_{\lfloor N w_t^{(i)} \rfloor \geq 2}.
\end{aligned}$$

Applying the tower property and conditional independence,

$$\begin{aligned}
\frac{1}{(N)_2}\sum_{i=1}^N \mathbb{E}_t[(\nu_t^{(i)})_2] &= \frac{1}{(N)_2}\mathbb{E}_t\left[\sum_{i=1}^N \mathbb{E}\left[(\nu_t^{(i)})_2 \mid \mathcal{H}_t, \mathcal{F}_{t-1}\right]\right] \\
&= \frac{1}{(N)_2}\mathbb{E}_t\left[\sum_{i=1}^N \mathbb{E}\left[(\nu_t^{(i)})_2 \mid \mathcal{H}_t\right]\right] \geq \frac{1}{(N)_2}2\mathbb{E}_t\left[|\{i : \lfloor N w_t^{(i)} \rfloor \geq 2\}|\right]
\end{aligned}$$

and similarly

$$\begin{aligned}
\frac{1}{(N)_3}\sum_{i=1}^N \mathbb{E}_t[(\nu_t^{(i)})_3] &\leq \frac{1}{(N)_3}(a^2+1)^3\mathbb{E}_t\left[|\{i : \lfloor N w_t^{(i)} \rfloor \geq 2\}|\right] \\
&\leq b_N \frac{1}{(N)_2}\sum_{i=1}^N \mathbb{E}_t[(\nu_t^{(i)})_2]
\end{aligned}$$

where

$$b_N := \frac{1}{N-2}\frac{(a^2+1)^3}{2} \underset{N\to\infty}{\longrightarrow} 0$$

is independent of $\mathcal{F}_\infty$, satisfying (**??**). The result follows by applying Theorem **??**. ■

4.2.2 Finite time scale

Lemma 4.1. *Consider an SMC algorithm using any stochastic rounding as its resampling scheme. Suppose that* $\varepsilon \leq q_t(x, x') \leq \varepsilon^{-1}$ *uniformly for some* $\varepsilon \in (0, 1]$*, and that there exist* $\zeta > 0$ *and* $\delta \in (0, 1)$ *such that* $\mathbb{P}[\max_i w_t^{(i)} - \min_i w_t^{(i)} \geq 2\delta/N \mid \mathcal{F}_{t-1}] \geq \zeta$ *for infinitely many* t*. Then, for all* $N > 1$*,* $\mathbb{P}[\tau_N(t) = \infty] = 0$ *for all finite* t*.*

Proof. Let $\mathcal{H}_t$ be defined as in (4.2). The first step is to show that whenever $\max_i w_t^{(i)} \geq (1+\delta)/N$, $\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t] = \mathbb{P}[c_N(t) \neq 0 \mid \mathcal{H}_t]$ is bounded below uniformly in t. For this purpose we need consider only weight vectors such that $w_t^{(i)} \in (0, 2/N)$ for all i; otherwise $\mathbb{P}[c_N(t) \neq 0 \mid \mathcal{H}_t] = 1$ 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 probability 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

$$\begin{aligned}
\mathbb{P}[c_N(t) \leq 2/N^2 \mid \mathcal{H}_t] &= \mathbb{P}[\nu_t^{(i)} = 1 \, \forall i \in \{1, \dots, N\} \mid \mathcal{H}_t] \\
&= \mathbb{P}[\nu_t^{(i)} = 1 \, \forall i \in \mathcal{I} \mid \mathcal{H}_t] \\
&= \prod_{i \in \mathcal{I}} \mathbb{P}[\nu_t^{(i)} = 1 \mid \nu_t^{(j)} = 1 \, \forall j \in \mathcal{I} : j < i; \mathcal{H}_t] \\
&= \mathbb{P}[\nu_t^{(i^\star)} = 1 \mid \mathcal{H}_t] \prod_{\substack{i \in \mathcal{I} \\ i \neq i^\star}} \mathbb{P}[\nu_t^{(i)} = 1 \mid \nu_t^{(i^\star)} = 1; \nu_t^{(j)} = 1 \, \forall j \in \mathcal{I} : j < i; \mathcal{H}_t] \\
&\leq \mathbb{P}[\nu_t^{(i^\star)} = 1 \mid \mathcal{H}_t].
\end{aligned} \tag{4.3}$$

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 $\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t]$ 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 $w_{\delta'} = \{(1, \dots, 1) + \delta' e_{i^\star} - \delta' e_{j^\star}\}/N$ for fixed $i^\star \neq j^\star$ and $\delta' \in (0, 1)$, where e_i denotes the ith canonical basis vector in $\mathbb{R}^N$. As in the proof of Corollary 4.1, define $p_0^{(i)} = \mathbb{P}[\nu_t^{(i)} = \lfloor N w_t^{(i)} \rfloor \mid \mathcal{H}_t]$ and $p_1^{(i)} = \mathbb{P}[\nu_t^{(i)} = \lfloor N w_t^{(i)} \rfloor + 1 \mid \mathcal{H}_t]$. Then from (4.3) we have

$$\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{H}_t, w_t^{(1:N)} = w_{\delta'}] = 1 - \mathbb{P}[\nu_t^{(i^\star)} = 1 \mid \mathcal{H}_t, w_t^{(1:N)} = w_{\delta'}] = p_1^{(i^\star)},$$

evaluated on $w_{\delta'}$. We will need a lower bound on $p_1^{(i^\star)}$ when $w_t^{(1:N)} = w_{\delta'}$. We first derive expressions for $p_0^{(i)}$ and $p_1^{(i)}$ up to a constant, then use $p_0^{(i)} + p_1^{(i)} = 1$ to get a normalised bound. We have

$$\begin{aligned}
p_0^{(i)} &= C(1 - Nw_t^{(i)} + \lfloor Nw_t^{(i)} \rfloor) \\
&\quad\times \sum_{\substack{a_{1:N} \in \{1,\dots,N\}^N: \\ |\{j: a_j = i\}| = \lfloor Nw_t^{(i)} \rfloor}} \mathbb{P}\left[a_t^{(1:N)} = a_{1:N} \mid \nu_t^{(i)}, w_t^{(1:N)}\right] \prod_{k=1}^N q_{t-1}(X_t^{(a_k)}, X_{t-1}^{(k)}), \\
p_1^{(i)} &= C(Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor) \\
&\quad\times \sum_{\substack{a_{1:N} \in \{1,\dots,N\}^N: \\ |\{j: a_j = i\}| = \lfloor Nw_t^{(i)} \rfloor + 1}} \mathbb{P}\left[a_t^{(1:N)} = a_{1:N} \mid \nu_t^{(i)}, w_t^{(1:N)}\right] \prod_{k=1}^N q_{t-1}(X_t^{(a_k)}, X_{t-1}^{(k)}).
\end{aligned}$$

Applying the bounds on q_t, we have

$$\begin{gathered}
C(1 - Nw_t^{(i)} + \lfloor Nw_t^{(i)} \rfloor)\varepsilon^N \leq p_0^{(i)} \leq C(1 - Nw_t^{(i)} + \lfloor Nw_t^{(i)} \rfloor)\varepsilon^{-N}, \\
C(Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor)\varepsilon^N \leq p_1^{(i)} \leq C(Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor)\varepsilon^{-N}
\end{gathered}$$

from which we construct the normalised bound

$$p_1^{(i)} \geq \frac{(Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor)\varepsilon^N}{(Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor)\varepsilon^{-N} + (1 - Nw_t^{(i)} + \lfloor Nw_t^{(i)} \rfloor)\varepsilon^{-N}} = (Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor)\varepsilon^{2N}.$$

When $w_t^{(1:N)} = w_{\delta'}$, we have $w_t^{(i^\star)} = (1+\delta')/N$, so $p_1^{(i^\star)} \geq \delta'\varepsilon^{2N}$, which is increasing in δ'. We conclude that $\mathbb{P}[c_N(t) > 2/N^2 | \mathcal{H}_t, \max_i w_t^{(i)} \geq (1+\delta)/N] \geq \min_{\delta' \geq \delta} \delta'\varepsilon^{2N} = \delta\varepsilon^{2N}$.

A slight modification of this argument yields $\mathbb{P}[c_N(t) > 2/N^2 | \mathcal{H}_t, \min_i w_t^{(i)} \leq (1-\delta)/N] \geq \delta\varepsilon^{2N}$. 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 $\mathbb{P}[c_N(t) > 2/N^2 | \mathcal{H}_t, \max_i w_t^{(i)} - \min_i w_t^{(i)} \geq 2\delta/N] \geq \delta\varepsilon^{2N}$. Thus

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

Using the D-separation established in Appendix **??** combined with the tower property, we have

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

which is bounded below by $\zeta\delta\varepsilon^{2N}$ 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 $\mathbb{P}[\tau_N(t) = \infty] = 0$. ■

4.3 The worst possible resampling scheme

Remark that this one doesn't converge to KC, but rather to a star-shaped coalescent.

4.4 Conditional SMC

Why CSMC is qualitatively different to, say, standard SMC with multinomial resampling (immortal particle etc.). Reasons for restriction to multinomial resampling, conjecture that limit theorem holds for other schemes in CSMC.

4.4.1 Proof of main condition

Corollary 4.2. *Consider a conditional SMC algorithm using multinomial resampling, such that the standing assumption is satisfied. Assume there exist constants $\varepsilon \in (0, 1], a \in [1, \infty)$ and probability density h such that for all x, x', t,*

$$\frac{1}{a} \leq g_t(x, x') \leq a, \quad \varepsilon h(x') \leq q_t(x, x') \leq \frac{1}{\varepsilon} h(x'). \tag{4.4}$$

Let $(G_t^{(n,N)})_{t \geq 0}$ denote the genealogy of a random sample of n terminal particles from the output of the algorithm when the total number of particles used is N. Then, for any fixed n, the time-scaled genealogy $(G_{\tau_N(t)}^{(n,N)})_{t \geq 0}$ converges to Kingman's n-coalescent as $N \to \infty$, in the sense of finite-dimensional distributions.

Proof. Define the conditioning σ-algebra $\mathcal{H}_t$ as in (4.2). We assume without loss of generality that the immortal particle takes index 1 in each generation. This significantly simplifies the notation, but the same argument holds if the immortal indices are taken to be $a^\star_{(0:T)}$ rather than $(1, \ldots, 1)$.

The parental indices are conditionally independent, as in standard SMC with multinomial resampling, but we have to treat $i = 1$ as a special case. We have the following conditional law on parental indices

$$\mathbb{P}\left[a_t^{(i)} = a_i \mid \mathcal{H}_t\right] \propto \begin{cases} \mathbb{1}_{a_i = 1} & i = 1 \\ w_t^{(a_i)} q_{t-1}(X_t^{(a_i)}, X_{t-1}^{(i)}) & i = 2, \ldots, N. \end{cases}$$

The joint conditional law is therefore

$$\mathbb{P}\left[a_t^{(1:N)} = a_{1:N} \mid \mathcal{H}_t\right] \propto \mathbb{1}_{a_1 = 1} \prod_{i=2}^{N} w_t^{(a_i)} q_{t-1}(X_t^{(a_i)}, X_{t-1}^{(i)}).$$

First we make the following observation, which follows from a balls-in-bins coupling. Assume (4.4). Then for any function $f : \{1, \ldots, N\}^N \to \mathbb{R}$ such that (for a fixed i) $f(a_t'^{(1:N)}) \geq f(a_t^{(1:N)})$ whenever $|\{j : a_t'^{(j)} = i\}| \geq |\{j : a_t^{(j)} = i\}|$,

$$\mathbb{E}[f(A_{1,i}^{(1:N)})] \leq \mathbb{E}[f(a_t^{(1:N)}) \mid \mathcal{H}_t] \leq \mathbb{E}[f(A_{2,i}^{(1:N)})] \tag{4.5}$$

where the elements of $A_{1,i}^{(1:N)}, A_{2,i}^{(1:N)}$ are all mutually independent and independent of $\mathcal{F}_\infty$, and distributed according to

$$A_{1,i}^{(j)} \sim \begin{cases} \delta_1 & j = 1 \\ \text{Categorical}\left((\varepsilon/a)^{\mathbb{1}_{i=1}-\mathbb{1}_{i\neq 1}}, \ldots, (\varepsilon/a)^{\mathbb{1}_{i=N}-\mathbb{1}_{i\neq N}}\right) & j \neq 1 \end{cases}$$
$$A_{2,i}^{(j)} \sim \begin{cases} \delta_1 & j = 1 \\ \text{Categorical}\left((a/\varepsilon)^{\mathbb{1}_{i=1}-\mathbb{1}_{i\neq 1}}, \ldots, (a/\varepsilon)^{\mathbb{1}_{i=N}-\mathbb{1}_{i\neq N}}\right) & j \neq 1 \end{cases}$$

where the vector of probabilities is given up to a constant in the argument of Categorical distributions. We use these random vectors to construct bounds that are independent of $\mathcal{F}_\infty$. Also define the corresponding offspring counts $V_1^{(i)} = |\{j : A_{1,i}^{(j)} = i\}|$, $V_2^{(i)} = |\{j : A_{2,i}^{(j)} = i\}|$, for $i = 1, \ldots, N$, which have marginal distributions

$$V_1^{(i)} \overset{d}{=} \mathbb{1}_{i=1} + \text{Binomial}\left(N-1, \frac{\varepsilon/a}{(\varepsilon/a) + (N-1)(a/\varepsilon)}\right),$$
$$V_2^{(i)} \overset{d}{=} \mathbb{1}_{i=1} + \text{Binomial}\left(N-1, \frac{a/\varepsilon}{(a/\varepsilon) + (N-1)(\varepsilon/a)}\right).$$

Now consider the function $f_i(a_t^{(1:N)}) := (\nu_t^{(i)})_2$. We can apply (4.5) to obtain the lower bound

$$\begin{aligned}
\frac{1}{(N)_2}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_2 \mid \mathcal{H}_t] &\geq \frac{1}{(N)_2}\sum_{i=1}^N \mathbb{E}[(V_1^{(i)})_2] = \frac{1}{(N)_2}\left[\mathbb{E}[(V_1^{(1)})_2] + \sum_{i=2}^N \mathbb{E}[(V_1^{(i)})_2]\right] \\
&= \frac{1}{(N)_2}\left[\frac{(N-1)_2(\varepsilon/a)^2}{\{(\varepsilon/a) + (N-1)(a/\varepsilon)\}^2} + \frac{2(N-1)(\varepsilon/a)}{(\varepsilon/a) + (N-1)(a/\varepsilon)}\right. \\
&\qquad \left. + \sum_{i=2}^N \frac{(N-1)_2(\varepsilon/a)^2}{\{(\varepsilon/a) + (N-1)(a/\varepsilon)\}^2}\right] \\
&= \frac{1}{(N)_2}\left[\frac{2(N-1)(\varepsilon/a)}{(\varepsilon/a) + (N-1)(a/\varepsilon)} + \sum_{i=1}^N \frac{(N-1)_2(\varepsilon/a)^2}{\{(\varepsilon/a) + (N-1)(a/\varepsilon)\}^2}\right]
\end{aligned}$$

using the moments of the Binomial distribution (see Mosimann 1962 for example) along with the identity $(X+1)_2 \equiv 2(X)_1 + (X)_2$. This is further bounded by

$$\begin{aligned}\frac{1}{(N)_2}\sum_{i=1}^{N}\mathbb{E}[(\nu_t^{(i)})_2 \mid \mathcal{H}_t] &\geq \frac{1}{(N)_2}\left\{\frac{2(N-1)(\varepsilon/a)}{N(a/\varepsilon)} + \frac{(N)_3(\varepsilon/a)^2}{N^2(a/\varepsilon)^2}\right\} \\ &= \frac{1}{N^2}\left\{\frac{2\varepsilon^2}{a^2} + \frac{(N-2)\varepsilon^4}{a^4}\right\}. \end{aligned} \tag{4.6}$$

Similarly, we derive an upper bound on $f_i(a_t^{(1:N)}) := (\nu_t^{(i)})_3$, this time using the identity $(X+1)_3 \equiv 3(X)_2 + (X)_3$:

$$\begin{aligned}\frac{1}{(N)_3}\sum_{i=1}^{N}\mathbb{E}[(\nu_t^{(i)})_3 \mid \mathcal{H}_t] &\leq \frac{1}{(N)_3}\left[\mathbb{E}[(V_2^{(1)})_3] + \sum_{i=2}^{N}\mathbb{E}[(V_2^{(i)})_3]\right] \\ &\leq \frac{1}{(N)_3}\left[\frac{3(N-1)_2(a/\varepsilon)^2}{\{(a/\varepsilon)+(N-1)(\varepsilon/a)\}^2} + \sum_{i=1}^{N}\frac{(N-1)_3(a/\varepsilon)^3}{\{(a/\varepsilon)+(N-1)(\varepsilon/a)\}^3}\right] \\ &\leq \frac{1}{(N)_3}\left\{\frac{3(N-1)_2(a/\varepsilon)^2}{N^2(\varepsilon/a)^2} + \frac{(N)_4(a/\varepsilon)^3}{N^3(\varepsilon/a)^3}\right\} \\ &= \frac{1}{(N)_3}\left\{\frac{3(N-1)_2}{N^2}\frac{a^4}{\varepsilon^4} + \frac{(N)_4}{N^3}\frac{a^6}{\varepsilon^6}\right\} \\ &= \frac{1}{N^3}\left\{\frac{3a^4}{\varepsilon^4} + \frac{(N-3)a^6}{\varepsilon^6}\right\}.\end{aligned}$$

We apply the tower property and conditional independence as in Corollary 4.1, upper bounding the ratio by

$$\begin{aligned}\frac{\frac{1}{(N)_3}\sum_{i=1}^{N}\mathbb{E}_t[(\nu_t^{(i)})_3]}{\frac{1}{(N)_2}\sum_{i=1}^{N}\mathbb{E}_t[(\nu_t^{(i)})_2]} &\leq \frac{N^2}{N^3}\frac{\frac{3a^4}{\varepsilon^4} + \frac{(N-3)a^6}{\varepsilon^6}}{\frac{2\varepsilon^2}{a^2} + \frac{(N-2)\varepsilon^4}{a^4}} \leq \frac{1}{N}\frac{a^6}{\varepsilon^6}\frac{3+(N-3)a^2/\varepsilon^2}{2+(N-2)\varepsilon^2/a^2} \\ &\leq \frac{1}{N}\frac{a^6}{\varepsilon^6}\left\{\frac{3}{2} + \frac{N-3}{N-2}\frac{a^4}{\varepsilon^4}\right\} \leq \frac{1}{N}\left\{\frac{3a^6}{2\varepsilon^6} + \frac{a^{10}}{\varepsilon^{10}}\right\} =: b_N \underset{N\to\infty}{\longrightarrow} 0.\end{aligned}$$

Thus (**??**) is satisfied. It remains to show that, for N sufficiently large, $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all finite t, a technicality which is proved in Lemma 4.2 in Appendix **??**. Applying Theorem **??** gives the result. ■

4.4.2 Finite time scale

Lemma 4.2. *Consider a conditional SMC algorithm using multinomial resampling, satisfying the standing assumption and* (4.4)*. Then, for all $N > 2$, $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all finite t.*

Proof. Since $c_N(t) \in [0,1]$ almost surely and has strictly positive expectation, for any fixed N the distribution of $c_N(t)$ with given expectation that maximises $\mathbb{P}[c_N(t) = 0 \mid \mathcal{F}_{t-1}]$ is two atoms, at 0 and 1 respectively. To ensure the correct expectation, the atom at 1 should have mass $\mathbb{P}[c_N(t) = 1 \mid \mathcal{F}_{t-1}] = \mathbb{E}_t[c_N(t)]$, which is bounded below by (4.6). If $c_N(t) > 0$

then $c_N(t) \geq 2/(N)_2 > 2/N^2$. Hence, in general $\mathbb{P}[c_N(t) > 2/N^2 \mid \mathcal{F}_{t-1}] \geq \mathbb{E}_t[c_N(t)]$. Applying (4.6), we have for any finite N,

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

By an argument analogous to the conclusion of Lemma 4.1, $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all $t < \infty$. ■

4.4.3 Effect of ancestor sampling

Argue that ancestor sampling removes bias towards assigning offspring to immortal line, and leaves exactly the same genealogy as standard SMC with multinomial resampling.

5 Weak Convergence

At the age of twenty-one he wrote a treatise upon the Binomial Theorem, which has had a European vogue. On the strength of it he won the Mathematical Chair at one of our smaller universities, and had, to all appearances, a most brilliant career before him.

Sherlock Holmes

We start by defining a suitable metric space. Let $\mathcal{P}_n$ be the space of partitions of $\{1, \ldots, n\}$. Denote by $\mathcal{X}$ the set of all functions mapping $[0, \infty)$ to $\mathcal{P}_n$ that are right-continuous with left limits. (Our rescaled genealogical process $(\mathcal{G}^{(n,N)}_{\tau_N(t)})_{t \geq 0}$ and our encoding of the n-coalescent are piecewise-constant functions mapping time $t \in [0, \infty)$ to partitions, and thus live in the space $\mathcal{X}$.) Finally, equip the space $\mathcal{P}_n$ with the zero-one metric,

$$\rho(\xi, \eta) = 1 - \delta_{\xi\eta} := \begin{cases} 0 & \text{if } \xi = \eta \\ 1 & \text{otherwise} \end{cases} \tag{5.1}$$

for any $\xi, \eta \in \mathcal{P}_n$.

Theorem 5.1. *Let $\nu_t^{(1:N)}$ denote the offspring numbers in an interacting particle system satisfying the standing assumption and such that, for any N sufficiently large, for all finite t, $\mathbb{P}\{\tau_N(t) = \infty\} = 0$. Suppose that there exists a deterministic sequence $(b_N)_{N \geq 1}$ such that $\lim_{N\to\infty} b_N = 0$ and*

$$\frac{1}{(N)_3} \sum_{i=1}^{N} \mathbb{E}_t\{(\nu_t^{(i)})_3\} \leq b_N \frac{1}{(N)_2} \sum_{i=1}^{N} \mathbb{E}_t\{(\nu_t^{(i)})_2\} \tag{5.2}$$

for all N, uniformly in $t \geq 1$. Then the rescaled genealogical process $(G^{(n,N)}_{\tau_N(t)})_{t \geq 0}$ converges weakly in $(\mathcal{X}, \rho)$ to Kingman's n-coalescent as $N \to \infty$.

Proof. The structure of the proof follows Möhle (1999), albeit with considerable technical complication due to the lack of independence between generations (non-neutrality) in our model. is this the main/only source of complication? Since we already have convergence of the finite-dimensional distributions (Theorem ?? refers to a previous chapter), strengthening this to weak convergence requires relative compactness of the sequence of processes $\{(G^{(n,N)}_{\tau_N(t)})_{t\geq 0}\}_{N\in\mathbb{N}}$.

We can apply Ethier and Kurtz (2009, Chapter 3, Corollary 7.4): $\mathcal{P}_n$ is finite and therefore complete and separable, and the sample paths of $(G^{(n,N)}_{\tau_N(t)})_{t\geq 0}$ live in $\mathcal{X}$, so the conditions of the corollary are satisfied. The corollary states that the sequence of processes $\{(G^{(n,N)}_{\tau_N(t)})_{t\geq 0}\}_{N\in\mathbb{N}}$ is relatively compact if and only if the following two conditions hold:

1. For every $\varepsilon > 0$, $t \geq 0$ there exists a compact set $\Gamma \subseteq \mathcal{P}_n$ such that

$$\liminf_{N\to\infty} \mathbb{P}[G^{(n,N)}_{\tau_N(t)} \in \Gamma] \geq 1 - \varepsilon \tag{5.3}$$

2. For every $\varepsilon > 0$, $t > 0$ there exists $\delta > 0$ such that

$$\liminf_{N\to\infty} \mathbb{P}[\omega(G^{(n,N)}_{\tau_N(\cdot)}, \delta, t) < \varepsilon] \geq 1 - \varepsilon \tag{5.4}$$

   where ω is the modulus of continuity:

$$\omega(G^{(n,N)}_{\tau_N(t)}, \delta, t) := \inf \max_{i\in[K]} \sup_{u,v\in[T_{i-1},T_i)} \rho(G^{(n,N)}_{\tau_N(u)}, G^{(n,N)}_{\tau_N(v)}) \tag{5.5}$$

   with the infimum taken over all partitions of the form $0 = T_0 < T_1 < \cdots < T_{K-1} < t \leq T_K$ such that $\min_{i\in[K]}(T_i - T_{i-1}) > \delta$. use a different letter not K?

In our case, Condition 1 is satisfied automatically with $\Gamma = \mathcal{P}_n$, since $\mathcal{P}_n$ is finite and hence compact. The intuition behind condition 2 is that the jumps of the process must be "well-separated". In our case where ρ is the zero-one metric, we see that $\rho(G^{(n,N)}_{\tau_N(u)}, G^{(n,N)}_{\tau_N(v)})$ is equal to 1 if there is a jump between times u and v, and 0 otherwise. Taking the supremum and maximum then indicates whether there is a jump within any of the intervals of the given partition; this can only be equal to zero if all of the jumps occur exactly at the times $T_0, \ldots, T_k$. The infimum over all allowed partitions, then, can only be equal to zero if no two jumps occur less than δ (unscaled) time apart, because of the restriction we placed on these partitions.

The proof is concentrated on proving Condition 2. To do this, we use a coupling with another process that contains all of the jumps of the genealogical process, with the addition of some extra jumps. This process is constructed in such a way that it can be shown to satisfy Condition 2, and hence so does the genealogical process.

Define $p_t := \max_{\xi\in E}\{1 - p_{\xi\xi}(t)\} = 1 - p_{\Delta\Delta}(t)$, where Δ denotes the trivial partition of $\{1, \ldots, n\}$ into singletons. For a proof that the maximum is attained at $\xi = \Delta$, see Lemma 5.12. Following Möhle (1999), we now construct the two-dimensional conditionally on $\mathcal{F}$

? Markov process $(Z_t, S_t)_{t\in\mathbb{N}_0}$ with transition probabilities

$$\mathbb{P}(Z_t = j, S_t = \eta \mid Z_{t-1} = i, S_{t-1} = \xi) = \begin{cases} 1 - p_t & \text{if } j = i \text{ and } \xi = \eta \\ p_{\xi\xi}(t) + p_t - 1 & \text{if } j = i+1 \text{ and } \xi = \eta \\ p_{\xi\eta}(t) & \text{if } j = i+1 \text{ and } \xi \neq \eta \\ 0 & \text{otherwise} \end{cases} \tag{5.6}$$

and initial state $Z_0 = 0$, $S_0 = \Delta$. The construction is such that the marginal (S_t) has the same distribution as the genealogical process of interest, and (Z_t) has jumps at all the times (S_t) does plus some extra jumps. (The definition of p_t ensures that the probability in the second case is non-negative, attaining the value zero when $\xi = \Delta$.)

Denote by $0 = T_0^{(N)} < T_1^{(N)} < \dots$ the jump times of the rescaled process $(Z_{\tau_N(t)})_{t\geq 0}$, and by $\varpi_i^{(N)} := T_i^{(N)} - T_{i-1}^{(N)}$ the corresponding holding times.

...

■

Lemma 5.1. $\max_{\xi\in E}(1 - p_{\xi\xi}(t)) = 1 - p_{\Delta\Delta}(t)$. *I need to fix notation; perhaps the set E could become $\mathcal{P}$ (for partitions) or something?*

Proof. Consider any $\xi \in E$ consisting of k blocks $(1 \leq k \leq n-1)$, and any $\xi' \in E$ consisting of $k+1$ blocks. From the definition of $p_{\xi\eta}(t)$ (Koskela et al. 2018, Equation (1)),

$$p_{\xi\xi}(t) = \frac{1}{(N)_k} \sum_{\substack{i_1,\dots,i_k \\ \text{all distinct}}} \nu_t^{(i_1)} \cdots \nu_t^{(i_k)}. \tag{5.7}$$

Similarly,

$$\begin{aligned} p_{\xi'\xi'}(t) &= \frac{1}{(N)_{k+1}} \sum_{\substack{i_1,\dots,i_k,i_{k+1} \\ \text{all distinct}}} \nu_t^{(i_1)} \cdots \nu_t^{(i_k)} \nu_t^{(i_{k+1})} \\ &= \frac{1}{(N)_k (N-k)} \sum_{\substack{i_1,\dots,i_k \\ \text{all distinct}}} \left\{ \nu_t^{(i_1)} \cdots \nu_t^{(i_k)} \sum_{\substack{i_{k+1}=1 \\ \text{also distinct}}}^{N} \nu_t^{(i_{k+1})} \right\}. \end{aligned} \tag{5.8}$$

Discarding the zero summands,

$$p_{\xi'\xi'}(t) = \frac{1}{(N)_k (N-k)} \sum_{\substack{i_1,\dots,i_k \\ \text{all distinct:} \\ \nu_t^{(i_1)},\dots,\nu_t^{(i_k)}>0}} \left\{ \nu_t^{(i_1)} \cdots \nu_t^{(i_k)} \sum_{\substack{i_{k+1}=1 \\ \text{also distinct}}}^{N} \nu_t^{(i_{k+1})} \right\}. \tag{5.9}$$

The inner sum is

$$\sum_{\substack{i_{k+1}=1\\ \text{also distinct}}}^{N} \nu_t^{(i_{k+1})} = \left\{ \sum_{i=1}^{N} \nu_t^{(i)} - \sum_{i \in \{i_1, \ldots, i_k\}} \nu_t^{(i)} \right\} \leq N - k \tag{5.10}$$

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

$$p_{\xi'\xi'}(t) \leq \frac{N-k}{(N)_k (N-k)} \sum_{\substack{i_1, \ldots, i_k \\ \text{all distinct:} \\ \nu_t^{(i_1)}, \ldots, \nu_t^{(i_k)} > 0}} \nu_t^{(i_1)} \cdots \nu_t^{(i_k)} = p_{\xi\xi}(t). \tag{5.11}$$

Thus $p_{\xi\xi}(t)$ is decreasing in the number of blocks of ξ, and is therefore minimised by taking $\xi = \Delta$, which achieves the maximum n blocks. This choice in turn maximises $1 - p_{\xi\xi}(t)$, as required. ■

5.1 Bounds on sum-products

Lemma 5.2. *Fix $t > 0$, $l \in \mathbb{N}$.*

$$t^l - \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \right) \binom{l}{2} (t+1)^{l-2} \leq \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) \leq t^l + c_N(\tau_N(t))(t+1)^l. \tag{5.12}$$

Proof. As pointed out in Koskela et al. (2018, Equation (8)),

$$\sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) \geq \left(\sum_{s=0}^{\tau_N(t)} c_N(s) \right)^l - \binom{l}{2} \left(\sum_{s=0}^{\tau_N(t)} c_N(s)^2 \right) \left(\sum_{s=0}^{\tau_N(t)} c_N(s) \right)^{l-2}. \tag{5.13}$$

By definition of τ_N,

$$t \leq \sum_{s=0}^{\tau_N(t)} c_N(s) \leq t + 1. \tag{5.14}$$

Substituting these bounds into the RHS of (5.8) yields the lower bound.

It is a true fact that

$$\sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) \leq \left(\sum_{s=0}^{\tau_N(t)} c_N(s) \right)^l, \tag{5.15}$$

as can be seen by considering the multinomial expansion of the RHS. This is further bounded by

$$\left(\sum_{s=0}^{\tau_N(t)} c_N(s) \right)^l \leq \left(\sum_{s=0}^{\tau_N(t)-1} c_N(s) + c_N(\tau_N(t)) \right)^l \leq \left[t + c_N(\tau_N(t)) \right]^l, \tag{5.16}$$

again using the definition of τ_N. A binomial expansion yields

$$[t + c_N(\tau_N(t))]^l = t^l + \sum_{i=0}^{l-1} \binom{l}{i} t^i c_N(\tau_N(t))^{l-i} = t^l + c_N(\tau_N(t)) \sum_{i=0}^{l-1} \binom{l}{i} t^i c_N(\tau_N(t))^{l-1-i}, \tag{5.17}$$

then since $c_N(s) \leq 1$ for all s,

$$\sum_{i=0}^{l-1} \binom{l}{i} t^i c_N(\tau_N(t))^{l-1-i} \leq \sum_{i=0}^{l-1} \binom{l}{i} t^i \leq (t+1)^l. \tag{5.18}$$

Putting this together yields the upper bound. ■

Lemma 5.3. *Fix $t > 0$, $l \in \mathbb{N}$. Let B be a positive constant which may depend on n.*

$$\sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} [c_N(s_j) + BD_N(s_j)] \leq \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) + \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right) (t+1)^{l-1} (1+B)^l. \tag{5.19}$$

Proof. We start with a binomial expansion:

$$\begin{aligned} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} [c_N(s_j) + BD_N(s_j)] &= \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \sum_{\mathcal{I} \subseteq [l]} B^{l-|\mathcal{I}|} \left(\prod_{i \in \mathcal{I}} c_N(s_i) \right) \left(\prod_{j \notin \mathcal{I}} D_N(s_j) \right) \\ &= \sum_{\mathcal{I} \subseteq [l]} B^{l-|\mathcal{I}|} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \left(\prod_{i \in \mathcal{I}} c_N(s_i) \right) \left(\prod_{j \notin \mathcal{I}} D_N(s_j) \right) \end{aligned} \tag{5.20}$$

where $[l] := \{1, \ldots, l\}$. Since the sum is over all permutations of $r_1, \ldots, r_l$, we may arbitrarily choose an ordering for $\{1, \ldots, l\}$ such that $\mathcal{I} = \{1, \ldots, |\mathcal{I}|\}$:

$$\sum_{\mathcal{I} \subseteq [l]} B^{l-|\mathcal{I}|} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \left(\prod_{i \in \mathcal{I}} c_N(s_i) \right) \left(\prod_{j \notin \mathcal{I}} D_N(s_j) \right) = \sum_{I=0}^{l} \binom{l}{I} B^{l-I} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \left(\prod_{i=1}^{I} c_N(s_i) \right) \left(\prod_{j=I+1}^{l} D_N(s_j) \right) \tag{5.21}$$

Separating the term $I = l$,

$$\begin{aligned} \sum_{I=0}^{l} \binom{l}{I} B^{l-I} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \left(\prod_{i=1}^{I} c_N(s_i) \right) \left(\prod_{j=I+1}^{l} D_N(s_j) \right) &= \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) \\ &\quad + \sum_{I=0}^{l-1} \binom{l}{I} B^{l-I} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \left(\prod_{i=1}^{I} c_N(s_i) \right) \left(\prod_{j=I+1}^{l} D_N(s_j) \right). \end{aligned} \tag{5.22}$$

In the second line, there is always at least one D_N term, and $c_N(s) \geq D_N(s)$ for all s

(Koskela et al. 2018, p.9), so we can write

$$\begin{aligned}
\sum_{I=0}^{l-1} \binom{l}{I} B^{l-I} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \left(\prod_{i=1}^{I} c_N(s_i) \right) \left(\prod_{j=I+1}^{l} D_N(s_j) \right) &\leq \sum_{I=0}^{l-1} \binom{l}{I} B^{l-I} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \left(\prod_{i=1}^{l-1} c_N(s_i) \right) D_N(s_l) \\
&\leq \sum_{I=0}^{l-1} \binom{l}{I} B^{l-I} \left(\sum_{s_1 \neq \cdots \neq s_{l-1}}^{\tau_N(t)} \prod_{i=1}^{l-1} c_N(s_i) \right) \sum_{s_l=1}^{\tau_N(t)} D_N(\\
&\leq \sum_{I=0}^{l-1} \binom{l}{I} B^{l-I} (t+1)^{l-1} \sum_{s=1}^{\tau_N(t)} D_N(s)
\end{aligned} \tag{5.23}$$

using (5.10) and (5.9). Finally, by the Binomial Theorem,

$$\sum_{I=0}^{l-1} \binom{l}{I} B^{l-I} (t+1)^{l-1} \sum_{s=1}^{\tau_N(t)} D_N(s) \leq \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right) (t+1)^{l-1} (1+B)^l, \tag{5.24}$$

which, together with (5.17), concludes the proof. ∎

Lemma 5.4. *Fix $t > 0$, $l \in \mathbb{N}$. Let B be a positive constant which may depend on n.*

$$\sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} [c_N(s_j) - B D_N(s_j)] \geq \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) - \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right) (t+1)^{l-1} (1+B)^l. \tag{5.25}$$

Proof. A binomial expansion and subsequent manipulation as in (5.15)–(5.17) gives

$$\begin{aligned}
\sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} [c_N(s_j) - B D_N(s_j)] &= \sum_{\mathcal{I} \subseteq [l]} (-B)^{l-|\mathcal{I}|} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \left(\prod_{i \in \mathcal{I}} c_N(s_i) \right) \left(\prod_{j \notin \mathcal{I}} D_N(s_j) \right) \\
&= \sum_{I=0}^{l} \binom{l}{I} (-B)^{l-I} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \left(\prod_{i=1}^{I} c_N(s_i) \right) \left(\prod_{j=I+1}^{l} D_N(s_j) \right) \\
&= \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) + \sum_{I=0}^{l-1} \binom{l}{I} (-B)^{l-I} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \left(\prod_{i=1}^{I} c_N(s_i) \right) \left(\prod_{j=I+1}^{l} \right. \\
&\geq \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) - \sum_{I=0}^{l-1} \binom{l}{I} B^{l-I} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \left(\prod_{i=1}^{I} c_N(s_i) \right) \left(\prod_{j=I+1}^{l} D_N(\right.
\end{aligned} \tag{5.26}$$

where the last inequality just multiplies some positive terms by -1. Then (5.18)–(5.19) can be applied directly (noting that an upper bound on negative terms gives a lower bound overall):

$$-\sum_{I=0}^{l-1}\binom{l}{I}B^{l-I}\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\left(\prod_{i=1}^{I}c_N(s_i)\right)\left(\prod_{j=I+1}^{l}D_N(s_j)\right)\geq-\left(\sum_{s=1}^{\tau_N(t)}D_N(s)\right)(t{+}1)^{l-1}(1{+}B)^l \tag{5.27}$$

which concludes the proof. ■

5.2 Main components of weak convergence

Lemma 5.5 (Basis step)**.** *For any* $0 < t < \infty$,

$$\lim_{N\to\infty}\mathbb{E}\left[\prod_{r=1}^{\tau_N(t)}(1-p_r)\right]=e^{-\alpha_n t} \tag{5.28}$$

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

Proof. We start by showing that $\lim_{N\to\infty}\mathbb{E}\left[\prod_{r=1}^{\tau_N(t)}(1-p_r)\right]\leq e^{-\alpha_n t}$.
From Koskela et al. (2018, Lemma 1 Case 1), taking $\xi=\Delta$, we have for each r

$$1-p_r=p_{\Delta\Delta}(r)\leq 1-\alpha_n(1+O(N^{-1}))\left[c_N(r)-B_n'D_N(r)\right] \tag{5.29}$$

where the $O(N^{-1})$ term does not depend on r. When N is large enough, a sufficient condition to ensure the bound in (5.24) is non-negative is the event

$$E_r:=\left\{c_N(r)\leq\alpha_n^{-1}\right\} \tag{5.30}$$

and we define $E:=\bigcap_{r=1}^{\tau_N(t)}E_r$. Applying a multinomial expansion and then separating the positive and negative terms,

$$\begin{aligned}\prod_{r=1}^{\tau_N(t)}(1-p_r)&\leq 1+\sum_{l=1}^{\tau_N(t)}(-\alpha_n)^l(1+O(N^{-1}))\frac{1}{l!}\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\prod_{j=1}^{l}\left[c_N(s_j)-B_n'D_N(s_j)\right]\mathbb{1}_E\\&=1+\sum_{\substack{l=2\\\text{even}}}^{\tau_N(t)}\alpha_n^l(1+O(N^{-1}))\frac{1}{l!}\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\prod_{j=1}^{l}\left[c_N(s_j)-B_n'D_N(s_j)\right]\mathbb{1}_E\\&\quad-\sum_{\substack{l=1\\\text{odd}}}^{\tau_N(t)}\alpha_n^l(1+O(N^{-1}))\frac{1}{l!}\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\prod_{j=1}^{l}\left[c_N(s_j)-B_n'D_N(s_j)\right]\mathbb{1}_E.\end{aligned} \tag{5.31}$$

This is further bounded by applying Lemma 5.3 and then both bounds of Lemma 5.1:

$$
\begin{aligned}
\prod_{r=1}^{\tau_N(t)} (1-p_r) \le 1 + \Bigg\{ & \sum_{\substack{l=2\\ \text{even}}}^{\tau_N(t)} \alpha_n^l (1+O(N^{-1})) \frac{1}{l!} \sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) \\
& - \sum_{\substack{l=1\\ \text{odd}}}^{\tau_N(t)} \alpha_n^l (1+O(N^{-1})) \frac{1}{l!} \left[\sum_{s_1 \neq \cdots \neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) - \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right) (t+1)^{l-1} (1+B_n')^l \right] \\
\le 1 + \Bigg\{ & \sum_{\substack{l=2\\ \text{even}}}^{\tau_N(t)} \alpha_n^l (1+O(N^{-1})) \frac{1}{l!} \left\{ t^l + c_N(\tau_N(t))(t+1)^l \right\} \\
& - \sum_{\substack{l=1\\ \text{odd}}}^{\tau_N(t)} \alpha_n^l (1+O(N^{-1})) \frac{1}{l!} \left[t^l - \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \right) \binom{l}{2} (t+1)^{l-2} - \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right) (t+ \right.
\end{aligned}
\tag{5.32}
$$

Collecting some terms,

$$
\begin{aligned}
\prod_{r=1}^{\tau_N(t)} (1-p_r) \le 1 &+ \sum_{l=1}^{\tau_N(t)} (-\alpha_n)^l (1+O(N^{-1})) \frac{1}{l!} t^l \mathbb{1}_E + c_N(\tau_N(t)) \sum_{\substack{l=2\\ \text{even}}}^{\tau_N(t)} \alpha_n^l (1+O(N^{-1})) \frac{1}{l!} (t+1)^l \\
&+ \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \right) \sum_{\substack{l=1\\ \text{odd}}}^{\tau_N(t)} \alpha_n^l (1+O(N^{-1})) \frac{1}{l!} \binom{l}{2} (t+1)^{l-2} \\
&+ \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right) \sum_{\substack{l=1\\ \text{odd}}}^{\tau_N(t)} \alpha_n^l (1+O(N^{-1})) \frac{1}{l!} (t+1)^{l-1} (1+B_n')^l \\
\le 1 &+ \sum_{l=1}^{\infty} (-\alpha_n)^l (1+O(N^{-1})) \frac{1}{l!} t^l \mathbb{1}_{\{\tau_N(t) \ge l\}} \mathbb{1}_E + c_N(\tau_N(t)) \exp[\alpha_n (1+O(N^{-1}))(t+1)] \\
&+ \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \right) \frac{1}{2} \alpha_n^2 \exp[\alpha_n (1+O(N^{-1}))(t+1)] \\
&+ \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right) \exp[\alpha_n (1+O(N^{-1}))(t+1)(1+B_n')].
\end{aligned}
\tag{5.33}
$$

Now, taking the expectation and limit, then applying Brown et al. (2021, Equations (3.3)–

(3.5)), and Lemmata 5.9 and 5.11 to deal with the indicators,

$$\lim_{N\to\infty}\mathbb{E}\left[\prod_{r=1}^{\tau_N(t)}(1-p_r)\right] \leq 1+\sum_{l=1}^{\infty}(-\alpha_n)^l\frac{1}{l!}t^l\lim_{N\to\infty}\mathbb{P}\left[\{\tau_N(t)\geq l\}\cap E\right]+\lim_{N\to\infty}\mathbb{E}\left[c_N(\tau_N(t))\right]\exp[\alpha_n(t+1)]$$

$$+\lim_{N\to\infty}\mathbb{E}\left[\sum_{s=1}^{\tau_N(t)}c_N(s)^2\right]\frac{1}{2}\alpha_n^2\exp[\alpha_n(t+1)]$$

$$+\lim_{N\to\infty}\mathbb{E}\left[\sum_{s=1}^{\tau_N(t)}D_N(s)\right]\exp[\alpha_n(t+1)(1+B_n')]$$

$$=1+\sum_{l=1}^{\infty}(-\alpha_n)^l\frac{1}{l!}t^l=e^{-\alpha_n t}. \tag{5.34}$$

It remains to show the corresponding lower bound $\lim_{N\to\infty}\mathbb{E}\left[\prod_{r=1}^{\tau_N(t)}(1-p_r)\right]\geq e^{-\alpha_n t}$. From Brown et al. (2021, Equation (3.14)), taking $\xi=\Delta$, we have

$$1-p_t=p_{\Delta\Delta}(t)\geq 1-\alpha_n(1+O(N^{-1}))\left[c_N(t)+B_nD_N(t)\right] \tag{5.35}$$

where $B_n>0$ and the $O(N^{-1})$ term does not depend on t. In particular,

$$1-p_t=p_{\Delta\Delta}(t)\geq 1-\frac{N^{n-2}}{(N-2)_{n-2}}\alpha_n[c_N(t)+B_nD_N(t)]. \tag{5.36}$$

Since $D_N(s)\leq c_N(s)$ for all s (Koskela et al. 2018, p.9), a sufficient condition for this bound to be non-negative is

$$E_r:=\left\{c_N(r)\leq\frac{(N-2)_{n-2}}{N^{n-2}}\alpha_n^{-1}(1+B_n)^{-1}\right\}, \tag{5.37}$$

and we again define $E:=\bigcap_{r=1}^{\tau_N(t)}E_r$. We now apply a multinomial expansion to the product, and split into positive and negative terms:

$$\prod_{r=1}^{\tau_N(t)}(1-p_r)\geq\left\{1+\sum_{l=1}^{\tau_N(t)}(-\alpha_n)^l(1+O(N^{-1}))\frac{1}{l!}\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\prod_{j=1}^{l}\left[c_N(s_j)+B_nD_N(s_j)\right]\right\}\mathbb{1}_E$$

$$=\left\{1+\sum_{\substack{l=2\\ \text{even}}}^{\tau_N(t)}\alpha_n^l(1+O(N^{-1}))\frac{1}{l!}\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\prod_{j=1}^{l}\left[c_N(s_j)+B_nD_N(s_j)\right]\right.$$

$$\left.-\sum_{\substack{l=1\\ \text{odd}}}^{\tau_N(t)}\alpha_n^l(1+O(N^{-1}))\frac{1}{l!}\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\prod_{j=1}^{l}\left[c_N(s_j)+B_nD_N(s_j)\right]\right\}\mathbb{1}_E \tag{5.38}$$

This is further bounded by applying Lemma 5.2 and both bounds in Lemma 5.1:

$$\begin{aligned}
\prod_{r=1}^{\tau_N(t)}(1-p_r) \geq \Bigg\{ 1 &+ \sum_{\substack{l=2\\ \text{even}}}^{\tau_N(t)} \alpha_n^l (1+O(N^{-1}))\frac{1}{l!} \sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) \\
&- \sum_{\substack{l=1\\ \text{odd}}}^{\tau_N(t)} \alpha_n^l (1+O(N^{-1}))\frac{1}{l!}\left[\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)} \prod_{j=1}^{l} c_N(s_j) + \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right)(t+1)^{l-1}(1+B_n)^l \right] \\
\geq \Bigg\{ 1 &+ \sum_{\substack{l=2\\ \text{even}}}^{\tau_N(t)} \alpha_n^l (1+O(N^{-1}))\frac{1}{l!}\left[t^l - \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right)\binom{l}{2}(t+1)^{l-2}\right] \\
&- \sum_{\substack{l=1\\ \text{odd}}}^{\tau_N(t)} \alpha_n^l (1+O(N^{-1}))\frac{1}{l!}\Bigg[t^l + c_N(\tau_N(t))(t+1)^l + \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right)(t+1)^{l-1}(1+B_n
\end{aligned} \tag{5.39}$$

Collecting terms,

$$\begin{aligned}
\prod_{r=1}^{\tau_N(t)}(1-p_r) &\geq \sum_{l=0}^{\tau_N(t)} (-\alpha_n)^l (1+O(N^{-1}))\frac{1}{l!} t^l \mathbb{1}_E - \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right) \sum_{\substack{l=2\\ \text{even}}}^{\tau_N(t)} \alpha_n^l (1+O(N^{-1}))\frac{1}{l!}\binom{l}{2}(t+1)^{l-2} \\
&\quad - c_N(\tau_N(t)) \sum_{\substack{l=1\\ \text{odd}}}^{\tau_N(t)} \alpha_n^l (1+O(N^{-1}))\frac{1}{l!}(t+1)^l \\
&\quad - \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right) \sum_{\substack{l=1\\ \text{odd}}}^{\tau_N(t)} \alpha_n^l (1+O(N^{-1}))\frac{1}{l!}(t+1)^{l-1}(1+B_n)^l \\
&\geq \sum_{l=0}^{\infty} (-\alpha_n)^l (1+O(N^{-1}))\frac{1}{l!} t^l \mathbb{1}_E \mathbb{1}_{\{\tau_N(t)\geq l\}} - \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right)\frac{1}{2}\alpha_n^2 \exp[\alpha_n(1+O(N^{-1}))(t \\
&\quad - c_N(\tau_N(t)) \exp[\alpha_n(1+O(N^{-1}))(t+1)] \\
&\quad - \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right) \exp[\alpha_n(1+O(N^{-1}))(t+1)(1+B_n)].
\end{aligned} \tag{5.40}$$

Now, taking the expectation and limit, and applying Brown et al. (2021, Equations (3.3)–(3.5)) to show that all but the first sum vanish, and Lemmata 5.9 and 5.8 to show that

$\lim_{N\to\infty}\mathbb{P}[\{\tau_N(t)\ge l\}\cap E]=1$,

$$\begin{aligned}\lim_{N\to\infty}\mathbb{E}\left[\prod_{r=1}^{\tau_N(t)}(1-p_r)\right] &\ge \sum_{l=0}^{\infty}(-\alpha_n)^l(1+O(N^{-1}))\frac{1}{l!}t^l\lim_{N\to\infty}\mathbb{P}\left[\{\tau_N(t)\ge l\}\cap E\right]\\ &\quad -\lim_{N\to\infty}\mathbb{E}\left[\sum_{s=1}^{\tau_N(t)}c_N(s)^2\right]\frac{1}{2}\alpha_n^2\exp[\alpha_n(t+1)]\\ &\quad -\lim_{N\to\infty}\mathbb{E}\left[c_N(\tau_N(t))\right]\exp[\alpha_n(t+1)]\\ &\quad -\lim_{N\to\infty}\mathbb{E}\left[\sum_{s=1}^{\tau_N(t)}D_N(s)\right]\exp[\alpha_n(t+1)(1+B_n)]\\ &=\sum_{l=0}^{\infty}(-\alpha_n)^l\frac{1}{l!}t^l=e^{-\alpha_n t}.\end{aligned}\tag{5.41}$$

Combining the upper and lower bounds in (5.29) and (5.36) respectively concludes the proof. ■

Lemma 5.6 (Induction step upper bound)**.** *Fix $k\in\mathbb{N}$, $i_0:=0$, $i_k:=k$. For any sequence of times $0=t_0\le t_1\le\cdots\le t_k\le t$,*

$$\lim_{N\to\infty}\mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}}\left(\prod_{i=1}^{k}p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\ldots,r_k\}}}^{\tau_N(t)}(1-p_r)\right)\right]\le\alpha_n^k\sum_{l=0}^{\infty}\frac{1}{l!}(-\alpha_n)^l t^l\sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\ge j\forall j}}\prod_{j=1}^{k}\frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}\tag{5.42}$$

Proof. We use the bound on $(1-p_r)$ from (5.24) and apply a multinomial expansion, defining as in (5.25) the event E which ensures the bound is non-negative:

$$\begin{aligned}\prod_{\substack{r=1\\ \notin\{r_1,\ldots,r_k\}}}^{\tau_N(t)}(1-p_r) &\le \prod_{\substack{r=1\\ \notin\{r_1,\ldots,r_k\}}}^{\tau_N(t)}\left\{1-\alpha_n(1+O(N^{-1}))[c_N(r)-B_n'D_N(r)]\mathbb{1}_E\right\}\\ &=1+\sum_{l=1}^{\tau_N(t)-k}(-\alpha_n)^l(1+O(N^{-1}))\frac{1}{l!}\sum_{\substack{s_1\neq\cdots\neq s_l\\ \notin\{r_1,\ldots,r_k\}}}^{\tau_N(t)}\prod_{j=1}^{l}[c_N(s_j)-B_n'D_N(s_j)]\mathbb{1}_E\\ &=1+\sum_{l=1}^{\tau_N(t)-k}(-\alpha_n)^l(1+O(N^{-1}))\frac{1}{l!}\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\prod_{j=1}^{l}[c_N(s_j)-B_n'D_N(s_j)]\mathbb{1}_E\\ &\quad-\sum_{l=1}^{\tau_N(t)-k}(-\alpha_n)^l(1+O(N^{-1}))\frac{1}{l!}\sum_{\substack{s_1\neq\cdots\neq s_l\\ \exists i,i':s_i=r_{i'}}}\prod_{j=1}^{l}[c_N(s_j)-B_n'D_N(s_j)]\mathbb{1}_E.\end{aligned}\tag{5.43}$$

The penultimate line above is exactly the expansion we had in the basis step (5.26), except for the limit on l, and as such following the same arguments gives a bound like that in (5.28):

$$
\begin{aligned}
1+\sum_{l=1}^{\tau_N(t)-k}(-\alpha_n)^l(1+O(N^{-1}))\frac{1}{l!}\sum_{s_1\neq\cdots\neq s_l}^{\tau_N(t)}\prod_{j=1}^{l}[c_N(s_j)-B_n'D_N(s_j)]\mathbb{1}_E& \\
\leq 1+\sum_{l=1}^{\tau_N(t)-k}(-\alpha_n)^l(1+O(N^{-1}))\frac{1}{l!}t^l\mathbb{1}_E+c_N(\tau_N(t))\exp[\alpha_n(1+O(&\\
+\left(\sum_{s=1}^{\tau_N(t)}c_N(s)^2\right)\frac{1}{2}\alpha_n^2\exp[\alpha_n(1+O(N^{-1}))(t+1)]&\\
+\left(\sum_{s=1}^{\tau_N(t)}D_N(s)\right)\exp[\alpha_n(1+O(N^{-1}))(t+1)(1+B_n')].&
\end{aligned}
\tag{5.44}
$$

For the last line of (5.38),

$$
\begin{aligned}
-\sum_{l=1}^{\tau_N(t)-k}(-\alpha_n)^l(1+O(N^{-1}))\frac{1}{l!}\sum_{\substack{s_1\neq\cdots\neq s_l\\ \exists i,i':s_i=r_{i'}}}\prod_{j=1}^{l}\{c_N(s_j)-B_n'D_N(s_j)\}\mathbb{1}_E&\\
\leq\sum_{l=1}^{\tau_N(t)-k}\alpha_n^l(1+O(N^{-1}))\frac{1}{l!}\sum_{\substack{s_1\neq\cdots\neq s_l\\ \exists i,i':s_i=r_{i'}}}\prod_{j=1}^{l}\{c_N(s_j)+B_n'D_N(s_j)\}&\\
\leq\sum_{l=1}^{\tau_N(t)-k}\alpha_n^l(1+O(N^{-1}))\frac{1}{l!}\sum_{\substack{s_1\neq\cdots\neq s_l\\ \exists i,i':s_i=r_{i'}}}(1+B_n')^l\prod_{j=1}^{l}c_N(s_j)&\\
\leq\sum_{l=1}^{\tau_N(t)-k}\alpha_n^l(1+O(N^{-1}))\frac{1}{(l-1)!}\sum_{s_1\in\{r_1,\dots,r_k\}}\sum_{s_2\neq\cdots\neq s_l}^{\tau_N(t)}(1+B_n')^l\prod_{j=1}^{l}c_N(&\\
=\sum_{s\in\{r_1,\dots,r_k\}}c_N(s)\sum_{l=1}^{\tau_N(t)-k}\alpha_n^l(1+O(N^{-1}))\frac{1}{(l-1)!}(1+B_n')^l\sum_{s_1\neq\cdots\neq s_{l-1}}^{\tau_N(t)}&\\
\leq\sum_{j=1}^{k}c_N(r_j)\sum_{l=1}^{\tau_N(t)-k}\alpha_n^l(1+O(N^{-1}))\frac{1}{(l-1)!}(1+B_n')^l(t+1)^{l-1}&\\
\leq\left(\sum_{j=1}^{k}c_N(r_j)\right)\alpha_n(1+B_n')\exp[\alpha_n(1+O(N^{-1}))(1+B_n')(t+1)].&
\end{aligned}
\tag{5.45}
$$

Putting these together, we have

$$\prod_{\substack{r=1\\ \notin\{r_1,\dots,r_k\}}}^{\tau_N(t)} (1-p_r) \leq 1 + \sum_{l=1}^{\tau_N(t)-k} (-\alpha_n)^l (1+O(N^{-1}))\frac{1}{l!}t^l \mathbb{1}_E + c_N(\tau_N(t)) \exp[\alpha_n(1+O(N^{-1}))(t+1)]$$
$$+ \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right) \frac{1}{2}\alpha_n^2 \exp[\alpha_n(1+O(N^{-1}))(t+1)]$$
$$+ \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right) \exp[\alpha_n(1+O(N^{-1}))(t+1)(1+B_n')]$$
$$+ \left(\sum_{j=1}^{k} c_N(r_j)\right) \alpha_n(1+B_n') \exp[\alpha_n(1+O(N^{-1}))(1+B_n')(t+1)]. \tag{5.46}$$

Meanwhile, using the bound on p_r from (5.30) then applying a modification of Lemma 5.2,

$$\sum_{\substack{r_1<\dots<r_k:\\ r_i\leq\tau_N(t_i)\forall i}} \prod_{i=1}^{k} p_{r_i} \leq \alpha_n^k(1+O(N^{-1})) \sum_{\substack{r_1<\dots<r_k:\\ r_i\leq\tau_N(t_i)\forall i}} \prod_{i=1}^{k} [c_N(r_i) + B_n D_N(r_i)]$$
$$\leq \alpha_n^k(1+O(N^{-1})) \sum_{\substack{r_1<\dots<r_k:\\ r_i\leq\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i) + \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right) \alpha_n^k(1+O(N^{-1}))(t+1)^{k-1}(1 \tag{5.47}$$

A more liberal (but simpler) bound can be arrived at thus:

$$\prod_{i=1}^{k} p_{r_i} \leq \alpha_n^k(1+O(N^{-1})) \prod_{i=1}^{k} [c_N(r_i) + B_n D_N(r_i)]$$
$$\leq \alpha_n^k(1+O(N^{-1})) \prod_{i=1}^{k} c_N(r_i)(1+B_n)$$
$$\leq \alpha_n^k(1+O(N^{-1}))(1+B_n)^k \prod_{i=1}^{k} c_N(r_i) \tag{5.48}$$

which also leads to the deterministic bound

$$\sum_{\substack{r_1<\dots<r_k:\\ r_i\leq\tau_N(t_i)\forall i}} \prod_{i=1}^{k} p_{r_i} \leq \alpha_n^k(1+O(N^{-1}))(1+B_n)^k \sum_{\substack{r_1<\dots<r_k:\\ r_i\leq\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)$$
$$\leq \alpha_n^k(1+O(N^{-1}))(1+B_n)^k \frac{1}{k!} \sum_{r_1\neq\dots\neq r_k}^{\tau_N(t)} \prod_{i=1}^{k} c_N(r_i)$$
$$\leq \alpha_n^k(1+O(N^{-1}))(1+B_n)^k \frac{1}{k!}(t+1)^k. \tag{5.49}$$

Combining (5.41) with the other product, the expression inside the expectation in (5.37) is bounded above by

$$
\begin{aligned}
&\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\dots,r_k\}}}^{\tau_N(t)} (1-p_r)\right) \le \left\{1+\sum_{l=1}^{\tau_N(t)-k} (-\alpha_n)^l (1+O(N^{-1}))\frac{1}{l!}t^l \mathbb{1}_E\right\} \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^{k} p_{r_i} \\
&\qquad + \left\{c_N(\tau_N(t))\exp[\alpha_n(1+O(N^{-1}))(t+1)] + \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right)\frac{1}{2}\alpha_n^2 \exp[\alpha_n(1+O(N^{-1}))(\right. \\
&\qquad\qquad \left. + \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right)\exp[\alpha_n(1+O(N^{-1}))(t+1)(1+B_n')]\right\} \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1} \\
&\qquad + \exp[\alpha_n(1+O(N^{-1}))(1+B_n')(t+1)]\alpha_n(1+B_n') \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \sum_{j=1}^{k} c_N(r_j)\prod_{i=1}^{k} p_{r_i}.
\end{aligned}
\tag{5.50}
$$

Applying the various bounds (5.42)–(5.44), we have

$$
\begin{aligned}
&\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\dots,r_k\}}}^{\tau_N(t)} (1-p_r)\right) \le \alpha_n^k(1+O(N^{-1}))\left\{1+\sum_{l=1}^{\tau_N(t)-k} (-\alpha_n)^l (1+O(N^{-1}))\frac{1}{l!}t^l \mathbb{1}_E\right\} \\
&\qquad + \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right)\alpha_n^k(1+O(N^{-1}))(t+1)^{k-1}(1+B_n)^k \sum_{l=0}^{\tau_N(t)} (\alpha_n)^l(1+O(N^{-1}))\frac{1}{l!}t^l \\
&\qquad + \left\{c_N(\tau_N(t))\exp[\alpha_n(1+O(N^{-1}))(t+1)] + \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right)\frac{1}{2}\alpha_n^2 \exp[\alpha_n(1+O(N^{-1}))(\right. \\
&\qquad\qquad \left. + \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right)\exp[\alpha_n(1+O(N^{-1}))(t+1)(1+B_n')]\right\}\alpha_n^k(1+O(N^{-1}))(1+B \\
&\qquad + \exp[\alpha_n(1+B_n')(t+1)]\alpha_n(1+B_n')\alpha_n^k(1+O(N^{-1}))(1+B_n)^k \\
&\qquad\qquad \times \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \sum_{j=1}^{k} c_N(r_j)\prod_{i=1}^{k} c_N(r_i).
\end{aligned}
\tag{5.51}
$$

Upon taking the expectation and limit, we have

$$\begin{aligned}
\lim_{N\to\infty}\mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}}\left(\prod_{i=1}^{k}p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\dots,r_k\}}}^{\tau_N(t)}(1-p_r)\right)\right] &\le \alpha_n^k\lim_{N\to\infty}\mathbb{E}\left[\left(1+\sum_{l=1}^{\tau_N(t)-k}(-\alpha_n)^l\frac{1}{l!}t^l\mathbb{1}_E\right)\sum_{\substack{r_1<\cdots<r\\ r_i\le\tau_N(t_i}}\right.\\
&+\lim_{N\to\infty}\mathbb{E}\left[\sum_{s=1}^{\tau_N(t)}D_N(s)\right]\alpha_n^k(t+1)^{k-1}(1+B_n)^k\exp[\alpha_n t]\\
&+\left\{\lim_{N\to\infty}\mathbb{E}\left[c_N(\tau_N(t))\right]\exp[\alpha_n(t+1)]+\lim_{N\to\infty}\mathbb{E}\left[\sum_{s=1}^{\tau_N(t)}c_N(s)^2\right]\frac{1}{2}\alpha_n^2\exp[\alpha_n(t+1)]\right.\\
&\left.+\lim_{N\to\infty}\mathbb{E}\left[\sum_{s=1}^{\tau_N(t)}D_N(s)\right]\exp[\alpha_n(t+1)(1+B_n')]\right\}\alpha_n^k(1+B_n)^k\frac{1}{k!}(t+1)^k\\
&+\exp[\alpha_n(1+B_n')(t+1)]\alpha_n^{k+1}(1+B_n')(1+B_n)^k\lim_{N\to\infty}\mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}}\sum_{j=1}^{k}c_N(r_j)\prod_{i=1}^{k}c_N(r_i)\right]
\end{aligned}\tag{5.52}$$

The middle terms vanish due to Brown et al. (2021, Equations (3.3)–(3.5)) and the expression becomes

$$\begin{aligned}
\lim_{N\to\infty}\mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}}\left(\prod_{i=1}^{k}p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\dots,r_k\}}}^{\tau_N(t)}(1-p_r)\right)\right] &\le \alpha_n^k\lim_{N\to\infty}\mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}}\prod_{i=1}^{k}c_N(r_i)\right]\\
&+\alpha_n^k\sum_{l=1}^{\infty}(-\alpha_n)^l\frac{1}{l!}t^l\lim_{N\to\infty}\mathbb{E}\left[\mathbb{1}_{\{\tau_N(t)\ge k+l\}}\mathbb{1}_E\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}}\prod_{i=1}^{k}c_N(r_i)\right]\\
&+\exp[\alpha_n(1+B_n')(t+1)]\alpha_n^{k+1}(1+B_n')(1+B_n)^k\lim_{N\to\infty}\mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}}\sum_{j=1}^{k}c_N(r_j)\prod_{i=1}^{k}c_N(r_i)\right]
\end{aligned}\tag{5.53}$$

To simplify the last line,

$$\begin{aligned}
\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \sum_{j=1}^{k} c_N(r_j) \prod_{i=1}^{k} c_N(r_i) &\le \frac{1}{k!} \sum_{r_1\neq\cdots\neq r_k}^{\tau_N(t)} \sum_{j=1}^{k} c_N(r_j) \prod_{i=1}^{k} c_N(r_i) \\
&= \frac{1}{k!} \sum_{r_1\neq\cdots\neq r_k}^{\tau_N(t)} \sum_{j=1}^{k} c_N(r_j)^2 \prod_{i\neq j} c_N(r_i) \\
&\le \frac{1}{k!} \sum_{j=1}^{k} \sum_{s=1}^{\tau_N(t)} c_N(s)^2 \sum_{r_1\neq\cdots\neq r_{k-1}}^{\tau_N(t)} \prod_{i=1}^{k-1} c_N(r_i) \\
&\le \frac{1}{(k-1)!} \sum_{s=1}^{\tau_N(t)} c_N(s)^2 (t+1)^{k-1} \qquad (5.54)
\end{aligned}$$

hence

$$\lim_{N\to\infty} \mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \sum_{s\in\{r_1,\ldots,r_k\}} c_N(s) \prod_{i=1}^{k} c_N(r_i) \right] \le \frac{1}{(k-1)!}(t+1)^{k-1} \lim_{N\to\infty} \mathbb{E}\left[\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \right] = 0 \qquad (5.55)$$

by Brown et al. (2021, Equation (3.5)). By Lemmata 5.9 and 5.8, $\lim_{N\to\infty} \mathbb{P}[\{\tau_N(t) \ge k+l\} \cap E] = 1$, so we can apply Lemma 5.7 to the remaining expectations in (5.48), yielding

$$\begin{aligned}
\lim_{N\to\infty} \mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right) \left(\prod_{\substack{r=1\\ \notin\{r_1,\ldots,r_k\}}}^{\tau_N(t)} (1-p_r) \right) \right] &\le \alpha_n^k \sum_{l=0}^{\infty} (-\alpha_n)^l \frac{1}{l!} t^l \sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\ge j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!} \\
&= \alpha_n^k e^{-\alpha_n t} \sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\ge j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}
\end{aligned} \qquad (5.56)$$

as required. ■

Lemma 5.7 (Induction step lower bound)**.** *Fix $k \in \mathbb{N}$, $i_0 := 0$, $i_k := k$. For any sequence of times $0 = t_0 \le t_1 \le \cdots \le t_k \le t$,*

$$\lim_{N\to\infty} \mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right) \left(\prod_{\substack{r=1\\ \notin\{r_1,\ldots,r_k\}}}^{\tau_N(t)} (1-p_r) \right) \right] \ge \alpha_n^k \sum_{l=0}^{\infty} \frac{1}{l!} (-\alpha_n)^l t^l \sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\ge j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!} \qquad (5.57)$$

Proof. Firstly,

$$\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\ldots,r_k\}}}^{\tau_N(t)} (1-p_r)\right) \ge \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right)\left(\prod_{r=1}^{\tau_N(t)} (1-p_r)\right). \quad (5.58)$$

Now the second product does not depend on $r_1,\ldots,r_k$, and we can use the lower bound from (5.35):

$$\begin{aligned}\prod_{r=1}^{\tau_N(t)}(1-p_r) \ge \sum_{l=0}^{\tau_N(t)}(-\alpha_n)^l(1+O(N^{-1}))\frac{1}{l!}t^l\mathbb{1}_E - \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right)\frac{1}{2}\alpha_n^2\exp[\alpha_n(1+O(N^{-1}))(t+1)]\\ - c_N(\tau_N(t))\exp[\alpha_n(1+O(N^{-1}))(t+1)]\\ - \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right)\exp[\alpha_n(1+O(N^{-1}))(t+1)(1+B_n)] \qquad (5.59)\end{aligned}$$

where E is defined as in (5.32). We will also need an upper bound on this product, which is formed from (5.28) with a further deterministic bound:

$$\begin{aligned}\prod_{r=1}^{\tau_N(t)}(1-p_r) &\le \sum_{l=0}^{\tau_N(t)}(-\alpha_n)^l(1+O(N^{-1}))\frac{1}{l!}t^l + c_N(\tau_N(t))\exp[\alpha_n(1+O(N^{-1}))(t+1)]\\ &\quad + \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2\right)\frac{1}{2}\alpha_n^2\exp[\alpha_n(1+O(N^{-1}))(t+1)]\\ &\quad + \left(\sum_{s=1}^{\tau_N(t)} D_N(s)\right)\exp[\alpha_n(1+O(N^{-1}))(t+1)(1+B'_n)]\\ &\le \exp[\alpha_n(1+O(N^{-1}))t] + \exp[\alpha_n(1+O(N^{-1}))(t+1)]\\ &\quad + \frac{1}{2}\alpha_n^2(t+1)\exp[\alpha_n(1+O(N^{-1}))(t+1)] + (t+1)\exp[\alpha_n(1+O(N^{-1}))(t+1)(1+ B\\ &\le \left(2+\frac{\alpha_n^2(t+1)}{2}\right)\exp[\alpha_n(1+O(N^{-1}))(t+1)] + (t+1)\exp[\alpha_n(1+O(N^{-1}))(t+1)(1+\end{aligned}$$

$$(5.60)$$

Now let us consider the remaining sum-product on the RHS of (5.53). We use the same bound on p_r as in (5.24):

$$p_r = 1 - p_{\Delta\Delta}(r) \ge \alpha_n(1+O(N^{-1}))\left[c_N(r) - B'_n D_N(r)\right] \qquad (5.61)$$

where the $O(N^{-1})$ term does not depend on r. When N is large enough for the factor of $(1+O(N^{-1}))$ to be non-negative, a sufficient condition to ensure the bound in (5.56) is non-negative is the event

$$E'_r := \left\{c_N(r) \ge B'_n D_N(r)\right\} \qquad (5.62)$$

and we define $E' := \bigcap_{r=1}^{\tau_N(t)} E'_r$. Then

$$\prod_{i=1}^{k} p_{r_i} \geq \alpha_n^k (1 + O(N^{-1})) \prod_{i=1}^{k} \left[c_N(r_i) - B'_n D_N(r_i) \right] \mathbb{1}_{E'}. \tag{5.63}$$

Applying a modification of Lemma 5.3,

$$\begin{aligned} \sum_{\substack{r_1 < \cdots < r_k: \\ r_i \leq \tau_N(t_i) \forall i}} \left(\prod_{i=1}^{k} p_{r_i} \right) &\geq \alpha_n^k (1 + O(N^{-1})) \sum_{\substack{r_1 < \cdots < r_k: \\ r_i \leq \tau_N(t_i) \forall i}} \prod_{i=1}^{k} \left[c_N(r_i) - B'_n D_N(r_i) \right] \mathbb{1}_{E'} \\ &\geq \alpha_n^k (1 + O(N^{-1})) \left\{ \sum_{\substack{r_1 < \cdots < r_k: \\ r_i \leq \tau_N(t_i) \forall i}} \prod_{i=1}^{k} c_N(r_i) \mathbb{1}_{E'} - \frac{1}{k!} \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right) (t+1)^{k-1} (1 + B \right. \end{aligned} \tag{5.64}$$

The above expression is already split into positive and negative terms; a lower bound on (5.53) can be formed by multiplying the positive terms by the lower bound (5.54) and the negative terms by the upper bound (5.55). Thus

$$\begin{aligned} \sum_{\substack{r_1 < \cdots < r_k: \\ r_i \leq \tau_N(t_i) \forall i}} \left(\prod_{i=1}^{k} p_{r_i} \right) \left(\prod_{\substack{r=1 \\ \notin \{r_1, \ldots, r_k\}}}^{\tau_N(t)} (1 - p_r) \right) \geq \alpha_n^k (1 + O(N^{-1})) & \sum_{\substack{r_1 < \cdots < r_k: \\ r_i \leq \tau_N(t_i) \forall i}} \prod_{i=1}^{k} c_N(r_i) \mathbb{1}_{E'} \Bigg\{ \\ & \sum_{l=0}^{\tau_N(t)} (-\alpha_n)^l (1 + O(N^{-1})) \frac{1}{l!} t^l \mathbb{1}_E \\ & - \left(\sum_{s=1}^{\tau_N(t)} c_N(s)^2 \right) \frac{1}{2} \alpha_n^2 \exp[\alpha_n (1 + O(N^{-1}))(t+1)] \\ & - c_N(\tau_N(t)) \exp[\alpha_n (1 + O(N^{-1}))(t+1)] \\ & - \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right) \exp[\alpha_n (1 + O(N^{-1}))(t+1)(1 + B_n)] \Bigg\} \\ - \left(\sum_{s=1}^{\tau_N(t)} D_N(s) \right) \alpha_n^k (1 + O(N^{-1})) \frac{1}{k!} (t+1)^{k-1} (1 + B'_n)^k & \Bigg\{ \\ & \left(2 + \frac{\alpha_n^2 (t+1)}{2} \right) \exp[\alpha_n (1 + O(N^{-1}))(t+1)] \\ & + (t+1) \exp[\alpha_n (1 + O(N^{-1}))(t+1)(1 + B'_n)] \Bigg\}. \end{aligned} \tag{5.65}$$

Due to Brown et al. (2021, Equations (3.3)–(3.5)), all but the first two lines in the above have vanishing expectation, leaving

$$
\begin{aligned}
\lim_{N\to\infty} \mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\ldots,r_k\}}}^{\tau_N(t)} (1-p_r)\right)\right] \\
\ge \lim_{N\to\infty} \mathbb{E}\left[\alpha_n^k(1+O(N^{-1})) \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\mathbb{1}_{E'} \sum_{l=0}^{\tau_N(t)} (-\alpha_n)^l(1+O(N^{-1}))\frac{1}{l!}t^l \mathbb{1}_{I}\right. \\
= \alpha_n^k \sum_{l=0}^{\infty}(-\alpha_n)^l\frac{1}{l!}t^l \lim_{N\to\infty} \mathbb{E}\left[\mathbb{1}_{\{\tau_N(t)\ge l\}}\mathbb{1}_{E\cap E'} \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\right]
\end{aligned}
\tag{5.66}
$$

Lemmata 5.8 and 5.11 establish that $\lim_{N\to\infty}\mathbb{P}[E\cap E'] = 1$ and Lemma 5.9 deals with the other indicator. We can therefore apply Lemma 5.7 to conclude that

$$
\begin{aligned}
\lim_{N\to\infty} \mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \left(\prod_{i=1}^{k} p_{r_i}\right)\left(\prod_{\substack{r=1\\ \notin\{r_1,\ldots,r_k\}}}^{\tau_N(t)} (1-p_r)\right)\right] &\ge \alpha_n^k \sum_{l=0}^{\infty}(-\alpha_n)^l\frac{1}{l!}t^l \sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\ge j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!} \\
&= \alpha_n^k e^{-\alpha_n t} \sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\ge j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}
\end{aligned}
\tag{5.67}
$$

as required. ∎

Lemma 5.8. *Fix* $k\in\mathbb{N}$, $i_0 := 0$, $i_k := k$. *Let* E *be any event independent of* $r_1,\ldots,r_k$ *such that* $\lim_{N\to\infty}\mathbb{P}[E] = 1$. *Then for any sequence of times* $0 = t_0 \le t_1 \le \cdots \le t_k \le t$,

$$
\lim_{N\to\infty} \mathbb{E}\left[\mathbb{1}_E \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\right] = \sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\ge j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}. \tag{5.68}
$$

Proof. As pointed out by Möhle (1999, p. 460), the sum-product on the left hand side can be expanded as

$$\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}}\prod_{i=1}^{k}c_N(r_i)=\sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\ge j\forall j}}\prod_{j=1}^{k}\sum_{\substack{r_{i_{j-1}+1}<\cdots<r_{i_j}\\ =\tau_N(t_{j-1})+1}}^{\tau_N(t_j)}\prod_{i=i_{j-1}+1}^{i_j}c_N(r_i). \tag{5.69}$$

Moreover,

$$\sum_{\substack{r_{i_{j-1}+1}<\cdots<r_{i_j}\\ =\tau_N(t_{j-1})+1}}^{\tau_N(t_j)}\prod_{i=i_{j-1}+1}^{i_j}c_N(r_i)=\frac{1}{(i_j-i_{j-1})!}\sum_{\substack{r_{i_{j-1}+1}\neq\cdots\neq r_{i_j}\\ =\tau_N(t_{j-1})+1}}^{\tau_N(t_j)}\prod_{i=i_{j-1}+1}^{i_j}c_N(r_i). \tag{5.70}$$

By a modification of the upper bound in Lemma 5.1,

$$\begin{aligned}\sum_{\substack{r_{i_{j-1}+1}\neq\cdots\neq r_{i_j}\\ =\tau_N(t_{j-1})+1}}^{\tau_N(t_j)}\prod_{i=i_{j-1}+1}^{i_j}c_N(r_i)&\le(t_j-t_{j-1})^{i_j-i_{j-1}}+c_N(\tau_N(t_j))(t_j-t_{j-1}+1)^{i_j-i_{j-1}}\\ &\le(t_j-t_{j-1})^{i_j-i_{j-1}}+c_N(\tau_N(t_j))(t_j-t_{j-1}+1)^{k}. \end{aligned}\tag{5.71}$$

Now, taking the product on the outside,

$$\begin{aligned}\prod_{j=1}^{k}\sum_{\substack{r_{i_{j-1}+1}<\cdots<r_{i_j}\\ =\tau_N(t_{j-1})+1}}^{\tau_N(t_j)}\prod_{i=i_{j-1}+1}^{i_j}c_N(r_i)&\le\prod_{j=1}^{k}\left\{\frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}+c_N(\tau_N(t_j))\frac{(1+t_j-t_{j-1})^k}{(i_j-i_{j-1})!}\right\}\\
&\le\prod_{j=1}^{k}\left\{\frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}+c_N(\tau_N(t_j))(1+t_j-t_{j-1})^k\right\}\\
&=\sum_{\mathcal{I}\subseteq[k]}\left(\prod_{j\in\mathcal{I}}\frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}\right)\left(\prod_{j\notin\mathcal{I}}c_N(\tau_N(t_j))(1+t_j-t_{j-1})^k\right)\\
&=\prod_{j=1}^{k}\frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}\\
&\quad+\sum_{\mathcal{I}\subset[k]}\left(\prod_{j\in\mathcal{I}}\frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}\right)\left(\prod_{j\notin\mathcal{I}}c_N(\tau_N(t_j))(1+t_j-t_{j-1})^k\right)\\
&\le\prod_{j=1}^{k}\frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}\\
&\quad+\sum_{\mathcal{I}\subset[k]}c_N(\tau_N(t_{j^*}))\left(\prod_{j\in\mathcal{I}}\frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}\right)(1+t_j-t_{j-1})^{k^2}\end{aligned}\tag{5.72}$$

where, say, $j^* := \min\{j \notin \mathcal{I}\}$. Now we are in a position to evaluate the limit in (5.63):

$$
\begin{aligned}
\lim_{N\to\infty} \mathbb{E}\left[\mathbb{1}_E \sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\right] &\le \lim_{N\to\infty} \mathbb{E}\left[\sum_{\substack{r_1<\cdots<r_k:\\ r_i\le\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\right] \\
&\le \sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\dots,k\}:\\ i_j\ge j\forall j}} \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \\
&\quad + \sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\dots,k\}:\\ i_j\ge j\forall j}} \sum_{\mathcal{I}\subset[k]} \lim_{N\to\infty} \mathbb{E}\left[c_N(\tau_N(t_{j^*}))\right] \left(\prod_{j\in\mathcal{I}} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!}\right) (1 + t_j - t_{j-1})^{k^2} \\
&= \sum_{\substack{i_1\le\cdots\le i_{k-1}\\ \in\{0,\dots,k\}:\\ i_j\ge j\forall j}} \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \qquad (5.73)
\end{aligned}
$$

using Brown et al. (2021, Equation (3.3)).

For the corresponding lower bound, by a modification of the lower bound in Lemma 5.1,

$$
\begin{aligned}
\sum_{\substack{r_{i_{j-1}+1}\ne\cdots\ne r_{i_j}\\ =\tau_N(t_{j-1})+1}}^{\tau_N(t_j)} \prod_{i=i_{j-1}+1}^{i_j} c_N(r_i) &\ge (t_j - t_{j-1})^{i_j - i_{j-1}} - \binom{i_j - i_{j-1}}{2} \left(\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2\right) (t_j - t_{j-1} + 1)^{i_j-} \\
&\ge (t_j - t_{j-1})^{i_j - i_{j-1}} - (i_j - i_{j-1})! \left(\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2\right) (t_j - t_{j-1} + 1)^{k-1}
\end{aligned}
\qquad (5.74)
$$

Define the event

$$
E_j^\star = \left\{\left(\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2\right) \le \frac{(t_j - t_{j-1})^{i_j - i_{j-1} - k + 1}}{(i_j - i_{j-1})!}\right\}, \qquad (5.75)
$$

which is sufficient to ensure the j^{th} term in the following product is non-negative, and define $E^\star := \bigcap_{j=1}^k E_j^\star$. Now, taking a product over j,

$$
\begin{aligned}
\prod_{j=1}^{k} \sum_{\substack{r_{i_{j-1}+1}<\cdots<r_{i_j} \\ =\tau_N(t_{j-1})+1}}^{\tau_N(t_j)} \prod_{i=i_{j-1}+1}^{i_j} c_N(r_i) &\geq \prod_{j=1}^{k} \left\{ \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} - \left(\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2 \right) (t_j - t_{j-1} + 1)^{k-1} \right\} \\
&= \sum_{\mathcal{I}\subseteq[k]} (-1)^{k-|\mathcal{I}|} \left(\prod_{j\in\mathcal{I}} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \right) \left(\prod_{j\notin\mathcal{I}} \left(\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2 \right) (t_j - t_{j-1} + 1 \right. \\
&= \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \mathbb{1}_{E^\star} \\
&\quad + \sum_{\mathcal{I}\subset[k]} (-1)^{k-|\mathcal{I}|} \left(\prod_{j\in\mathcal{I}} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \right) \left(\prod_{j\notin\mathcal{I}} \left(\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2 \right) (t_j - t_{j-1} \right. \\
&\geq \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \mathbb{1}_{E^\star} \\
&\quad - \sum_{\mathcal{I}\subset[k]} \left(\prod_{j\in\mathcal{I}} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \right) \left(\prod_{j\notin\mathcal{I}} \left(\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2 \right) (t_j - t_{j-1} + 1)^{k-1} \right) \\
&\geq \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \mathbb{1}_{E^\star} \\
&\quad - \sum_{\mathcal{I}\subset[k]} \left(\sum_{s=\tau_N(t_{j^*-1})+1}^{\tau_N(t_{j^*})} c_N(s)^2 \right) \left(\prod_{j\in\mathcal{I}} (t_j - t_{j-1})^k \right) \left(\prod_{j\notin\mathcal{I}} (t_j - t_{j-1} + 1)^k \right) \\
&\geq \prod_{j=1}^{k} \frac{(t_j - t_{j-1})^{i_j - i_{j-1}}}{(i_j - i_{j-1})!} \mathbb{1}_{E^\star} - \sum_{\mathcal{I}\subset[k]} \left(\sum_{s=\tau_N(t_{j^*-1})+1}^{\tau_N(t_{j^*})} c_N(s)^2 \right) (t_j - t_{j-1} + 1)^{k^2},
\end{aligned}
\tag{5.76}
$$

where again we have arbitrarily set $j^\star := \min\{j \notin \mathcal{I}\}$. We can now evaluate the limit:

$$
\begin{aligned}
\lim_{N\to\infty} \mathbb{E}\left[\mathbb{1}_E \sum_{\substack{r_1<\cdots<r_k:\\ r_i\leq\tau_N(t_i)\forall i}} \prod_{i=1}^{k} c_N(r_i)\right] \geq{}& \lim_{N\to\infty} \mathbb{E}\left[\mathbb{1}_{E\cap E^\star} \sum_{\substack{i_1\leq\cdots\leq i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\geq j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}\right] \\
&- \lim_{N\to\infty} \mathbb{E}\left[\mathbb{1}_E \sum_{\substack{i_1\leq\cdots\leq i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\geq j\forall j}} \sum_{\mathcal{I}\subset[k]} \left(\sum_{s=\tau_N(t_{j^*-1})+1}^{\tau_N(t_{j^*})} c_N(s)^2\right) (t_j - t_{j-1} \right. \\
\geq{}& \sum_{\substack{i_1\leq\cdots\leq i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\geq j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!} \lim_{N\to\infty} \mathbb{E}[\mathbb{1}_{E\cap E^\star}] \\
&- \lim_{N\to\infty} \mathbb{E}\left[\sum_{\substack{i_1\leq\cdots\leq i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\geq j\forall j}} \sum_{\mathcal{I}\subset[k]} \left(\sum_{s=\tau_N(t_{j^*-1})+1}^{\tau_N(t_{j^*})} c_N(s)^2\right) (t_j - t_{j-1} + 1 \right. \\
={}& \sum_{\substack{i_1\leq\cdots\leq i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\geq j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!} \lim_{N\to\infty} \mathbb{P}[E\cap E^\star] \\
&- \sum_{\substack{i_1\leq\cdots\leq i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\geq j\forall j}} \sum_{\mathcal{I}\subset[k]} \lim_{N\to\infty} \mathbb{E}\left[\sum_{s=\tau_N(t_{j^*-1})+1}^{\tau_N(t_{j^*})} c_N(s)^2\right] (t_j - t_{j-1} + 1)^{k^2} \\
={}& \sum_{\substack{i_1\leq\cdots\leq i_{k-1}\\ \in\{0,\ldots,k\}:\\ i_j\geq j\forall j}} \prod_{j=1}^{k} \frac{(t_j-t_{j-1})^{i_j-i_{j-1}}}{(i_j-i_{j-1})!}
\end{aligned} \tag{5.77}
$$

where for the last equality we use Brown et al. (2021, Equation (3.5)) to show that the second sum vanishes and Lemma 5.10 to show that $\lim_{N\to\infty}\mathbb{P}[E\cap E^\star] = 1$. We have shown that the upper and lower bounds coincide, so the result follows. ∎

5.3 Indicators

Lemma 5.9. *Let K be a constant which may depend on n, N but not on r, such that $K^{-2} = O(1)$ as $N \to \infty$. Define the events $E_r := \{c_N(r) < K\}$ and denote $E := \bigcap_{r=1}^{\tau_N(t)} E_r$. Then $\lim_{N\to\infty}\mathbb{P}[E] = 1$.*

Proof.

$$\mathbb{P}[E] = 1 - \mathbb{P}[E^c] = 1 - \mathbb{P}\left[\bigcup_{r=1}^{\tau_N(t)} E_r^c\right] = 1 - \mathbb{E}\left[\mathbb{1}_{\bigcup E_r^c}\right] \geq 1 - \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{1}_{E_r^c}\right]$$
$$= 1 - \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{E}\left[\mathbb{1}_{E_r^c} \mid \mathcal{F}_{r-1}\right]\right] = 1 - \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{P}\left[E_r^c \mid \mathcal{F}_{r-1}\right]\right] \tag{5.78}$$

where for the second line we apply Lemma 5.13 with $f(r) = \mathbb{1}_{E_r^c}$. By the generalised Markov inequality,

$$\mathbb{P}[E_r^c \mid \mathcal{F}_{r-1}] = \mathbb{P}[c_N(r) \geq K \mid \mathcal{F}_{r-1}] \leq K^{-2}\mathbb{E}[c_N(r)^2 \mid \mathcal{F}_{r-1}]. \tag{5.79}$$

Substituting this into (5.73) and applying Lemma 5.13 again, this time with $f(r) = c_N(r)^2$,

$$\mathbb{P}[E] \geq 1 - K^{-2}\mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{E}[c_N(r)^2 \mid \mathcal{F}_{r-1}]\right] = 1 - K^{-2}\mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} c_N(r)^2\right]. \tag{5.80}$$

Applying Brown et al. (2021, Equation (3.5)), the limit is

$$\lim_{N\to\infty} \mathbb{P}[E] = 1 - O(1) \times 0 = 1 \tag{5.81}$$

as required. ■

Lemma 5.10. *Fix $t > 0$. For any $l \in \mathbb{N}$, $\lim_{N\to\infty} \mathbb{P}[\tau_N(t) \geq l] = 1$.*

Proof. We can replace the event $\{\tau_N(t) \geq\}$ with an event of the form of E in Lemma 5.8:

$$\{\tau_N(t) \geq l\} = \left\{\min\left\{s \geq 1 : \sum_{r=1}^{s} c_N(r) \geq t\right\} \geq l\right\} = \left\{\sum_{r=1}^{l-1} c_N(r) < t\right\} \supseteq \bigcap_{r=1}^{l-1}\left\{c_N(r) < \frac{t}{l}\right\} \supseteq \bigcap_{r=1}^{\tau_N(t)}\Big\{c_N \tag{5.82}$$

Hence

$$\lim_{N\to\infty} \mathbb{P}[\tau_N(t) \geq l] \geq \lim_{N\to\infty} \mathbb{P}\left[\bigcap_{r=1}^{\tau_N(t)} \left\{c_N(r) < \frac{t}{l}\right\}\right] = 1 \tag{5.83}$$

by applying Lemma 5.8 with $K = t/l$. ■

Lemma 5.11. *Fix $k \in \mathbb{N}$, a sequence of times $0 = t_0 \leq t_1 \leq \cdots \leq t_k \leq t$, and let $K_1, \dots, K_k$ be constants such that for each j, $K_j^{-1} = O(1)$ as $N \to \infty$. Define the event*

$$E^\star := \bigcap_{j=1}^{k} \left\{ \sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2 \leq K_j \right\}. \tag{5.84}$$

Then $\lim_{N\to\infty} \mathbb{P}[E^\star] = 1$.

Proof.

$$\mathbb{P}[E^\star] = 1-\mathbb{P}[(E^\star)^c] = 1-\mathbb{P}\left[\bigcup_{j=1}^{k} \left\{ \sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2 > K_j \right\}\right] \geq 1-\sum_{j=1}^{k} \mathbb{P}\left[\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2 \geq K_j \right] \tag{5.85}$$

Applying Markov's inequality,

$$\mathbb{P}[E^\star] \geq 1 - \sum_{j=1}^{k} K_j^{-1} \mathbb{E}\left[\sum_{s=\tau_N(t_{j-1})+1}^{\tau_N(t_j)} c_N(s)^2 \right] \underset{N\to\infty}{\longrightarrow} 1 - \sum_{j=1}^{k} O(1) \times 0 = 1 \tag{5.86}$$

by Brown et al. (2021, Equation (3.5)). The statement of (3.5) is slightly less general than we need here: the relevant statement can be found in Koskela et al. (2018). ∎

Lemma 5.12. *Fix $t > 0$. Let K be a constant not depending on N, r, but which may depend on n.*

$$\lim_{N\to\infty} \mathbb{P}\left[\bigcap_{r=1}^{\tau_N(t)} \{c_N(r) \geq K D_N(r)\} \right] = 1. \tag{5.87}$$

Proof.

$$\begin{aligned}
\mathbb{P}\left[\bigcap_{r=1}^{\tau_N(t)} \{c_N(r) \geq K D_N(r)\} \right] &\geq \mathbb{P}\left[\bigcap_{r=1}^{\tau_N(t)} \{c_N(r) > K D_N(r)\} \right] \\
&= 1 - \mathbb{P}\left[\bigcup_{r=1}^{\tau_N(t)} \{c_N(r) \leq K D_N(r)\} \right] \\
&= 1 - \mathbb{E}\left[\mathbb{1}_{\bigcup\{c_N(r) \leq K D_N(r)\}} \right] \\
&\geq 1 - \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{1}_{\{c_N(r) \leq K D_N(r)\}} \right] \\
&= 1 - \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{P}[c_N(r) \leq K D_N(r) \mid \mathcal{F}_{r-1}] \right]
\end{aligned} \tag{5.88}$$

where the final inequality is an application of Lemma 5.13 with $f(r) = \mathbb{1}_{\{c_N(r) \leq K D_N(r)\}}$.

Fix $0 < \varepsilon < K^{-1}/2$ and assume $N > \max\{\varepsilon^{-1}, (\binom{n-2}{2} - 2\varepsilon)^{-1}\}$. For each r, i define the event $A_i(r) := \{\nu_r^{(i)} \leq N\varepsilon\}$. Conditional on $\mathcal{F}_{r-1}$, we have

$$D_N(r) = \frac{1}{N(N)_2}\sum_{i=1}^{N}(\nu_r^{(j)})_2 \left[\nu_r^{(i)} + \frac{1}{N}\sum_{j\neq i}(\nu_r^{(i)})^2\right]\mathbb{1}_{A_i^c(r)} + \frac{1}{N(N)_2}\sum_{i=1}^{N}(\nu_r^{(i)})_2 \left[\nu_r^{(i)} + \frac{1}{N}\sum_{j\neq i}(\nu_r^{(j)})^2\right]\mathbb{1}_{A_i(} \tag{5.89}$$

For the first term,

$$\frac{1}{N(N)_2}\sum_{i=1}^{N}(\nu_r^{(i)})_2 \left[\nu_r^{(i)} + \frac{1}{N}\sum_{j\neq i}(\nu_r^{(j)})^2\right]\mathbb{1}_{A_i^c(r)} \leq \sum_{i=1}^{N}\mathbb{1}_{A_i^c(r)}. \tag{5.90}$$

For the second term,

$$\begin{aligned}
\frac{1}{N(N)_2}\sum_{i=1}^{N}(\nu_r^{(i)})_2 \left[\nu_r^{(i)} + \frac{1}{N}\sum_{j\neq i}(\nu_r^{(j)})^2\right]\mathbb{1}_{A_i(r)} &\leq \frac{1}{N(N)_2}\sum_{i=1}^{N}(\nu_r^{(i)})_2\nu_r^{(i)}\mathbb{1}_{A_i(r)} + \frac{1}{N^2(N)_2}\sum_{i=1}^{N}(\nu_r^{(i)})_2\sum_{j=1}^{N}(\\
&\leq \frac{1}{N}c_N(r)N\varepsilon + \frac{1}{N^2(N)_2}\sum_{i=1}^{N}(\nu_r^{(i)})_2\sum_{j=1}^{N}(\nu_r^{(j)})_2\mathbb{1}_{A_i(r)} \\
&\quad + \frac{1}{N^2(N)_2}\sum_{i=1}^{N}(\nu_r^{(i)})_2\sum_{j=1}^{N}(\nu_r^{(j)})\mathbb{1}_{A_i(r)} \\
&\leq \varepsilon c_N(r) + \frac{1}{N^2}\sum_{i=1}^{N}\nu_r^{(i)}N\varepsilon c_N(r) + \frac{1}{N^2}c_N(r)N \\
&= c_N(r)\left(2\varepsilon + \frac{1}{N}\right).
\end{aligned} \tag{5.91}$$

Altogether we have

$$D_N(r) \leq c_N(r)\left(2\varepsilon + \frac{1}{N}\right) + \sum_{i=1}^{N}\mathbb{1}_{A_i^c(r)}. \tag{5.92}$$

Hence, still conditional on $\mathcal{F}_{r-1}$,

$$\begin{aligned}
\{c_N(r) \leq KD_N(r)\} &\subseteq \left\{c_N(r) \leq Kc_N(r)(2\varepsilon + N^{-1}) + K\sum_{i=1}^{N}\mathbb{1}_{A_i^c(r)}\right\} \\
&= \left\{K^{-1} - 2\varepsilon - \frac{1}{N} \leq \sum_{i=1}^{N}\frac{\mathbb{1}_{A_i^c(r)}}{c_N(r)}\right\}
\end{aligned} \tag{5.93}$$

where the ratio $\mathbb{1}_{A_i^c(r)}/c_N(r)$ is well-defined because

$$A_i^c(r) \Rightarrow c_N(r) := \frac{1}{(N)_2}\sum_{j=1}^{N}(\nu_r^{(j)})_2 \geq \frac{1}{(N)_2}(\nu_r^{(i)})_2 \geq \frac{\varepsilon(N\varepsilon - 1)}{N-1} \geq \varepsilon\left(\varepsilon - \frac{1}{N}\right) > 0. \tag{5.94}$$

Hence by Markov's inequality (the conditions on ε, N ensuring the constant is always strictly positive),

$$
\begin{aligned}
\mathbb{P}\left[c_N(r) \le KD_N(r) \mid \mathcal{F}_{r-1}\right] &\le \mathbb{P}\left[\sum_{i=1}^{N} \mathbb{1}_{A_i^c(r)} \ge \left(K^{-1} - 2\varepsilon - \frac{1}{N}\right)\varepsilon\left(\varepsilon - \frac{1}{N}\right)\middle|\mathcal{F}_{r-1}\right] \\
&\le \frac{1}{\left(K^{-1} - 2\varepsilon - \frac{1}{N}\right)\varepsilon\left(\varepsilon - \frac{1}{N}\right)} \mathbb{E}\left[\sum_{i=1}^{N} \mathbb{1}_{A_i^c(r)}\middle|\mathcal{F}_{r-1}\right] \\
&\le \frac{1}{\left(K^{-1} - 2\varepsilon - \frac{1}{N}\right)\varepsilon\left(\varepsilon - \frac{1}{N}\right)} \mathbb{E}\left[\sum_{i=1}^{N} \frac{(\nu_r^{(i)})_3}{(N\varepsilon)_3}\middle|\mathcal{F}_{r-1}\right] \\
&\le \frac{1}{\left(K^{-1} - 2\varepsilon - \frac{1}{N}\right)\varepsilon\left(\varepsilon - \frac{1}{N}\right)} \mathbb{E}\left[\frac{N(N)_2}{(N\varepsilon)_3} D_N(r)\middle|\mathcal{F}_{r-1}\right].
\end{aligned}
\tag{5.95}
$$

Applying Lemma 5.13 once more, with $f(r) = D_N(r)$,

$$
\begin{aligned}
\mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{P}[c_N(r) \le KD_N(r) \mid \mathcal{F}_{r-1}]\right] &\le \frac{1}{\left(K^{-1} - 2\varepsilon - \frac{1}{N}\right)\varepsilon\left(\varepsilon - \frac{1}{N}\right)} \frac{N(N)_2}{(N\varepsilon)_3} \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{E}[D_N(r) \mid \mathcal{F}_{r-1}]\right] \\
&= \frac{1}{\left(K^{-1} - 2\varepsilon - \frac{1}{N}\right)\varepsilon\left(\varepsilon - \frac{1}{N}\right)} \frac{N(N)_2}{(N\varepsilon)_3} \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} D_N(r)\right] \\
&\xrightarrow[N\to\infty]{} \frac{1}{(K^{-1} - 2\varepsilon)\varepsilon^5} \times 0 = 0.
\end{aligned}
\tag{5.96}
$$

Substituting this back into (5.83) concludes the proof. ∎

5.4 Other useful results

The following Lemma is taken from Koskela et al. (2018, Lemma 2), where the function is set to $f(r) = c_N(r)$, but the authors remark that the result holds for other choices of function.

Lemma 5.13. *Fix $t > 0$. Let $(\mathcal{F}_r)$ be the backwards-in-time filtration generated by the offspring counts $\nu_r^{(1:N)}$ at each generation r, and let $f(r)$ be any deterministic function of $\nu_r^{(1:N)}$ that is non-negative and bounded. In particular, for all r there exists $B < \infty$ such that $0 \le f(r) \le B$. Then*

$$
\mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} f(r)\right] = \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{E}[f(r) \mid \mathcal{F}_{r-1}]\right]. \tag{5.97}
$$

Proof. Define

$$
M_s := \sum_{r=1}^{s} \left\{f(r) - \mathbb{E}[f(r) \mid \mathcal{F}_{r-1}]\right\}. \tag{5.98}
$$

It is easy to establish that (M_s) is a martingale with respect to $(\mathcal{F}_s)$, and $M_0 = 0$. Now fix $K \geq 1$ and note that $\tau_N(t) \wedge K$ is a bounded $\mathcal{F}_t$-stopping time. Hence we can apply the optional stopping theorem:

$$\mathbb{E}[M_{\tau_N(t)\wedge K}] = \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)\wedge K} \{f(r) - \mathbb{E}[f(r) \mid \mathcal{F}_{r-1}]\}\right] = \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)\wedge K} f(r)\right] - \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)\wedge K} \mathbb{E}[f(r) \mid \mathcal{F}_{r-1}]\right] = \tag{5.99}$$

Since this holds for all $K \geq 1$,

$$\lim_{K\to\infty} \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)\wedge K} f(r)\right] = \lim_{K\to\infty} \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)\wedge K} \mathbb{E}[f(r) \mid \mathcal{F}_{r-1}]\right]. \tag{5.100}$$

The monotone convergence theorem allows the limit to pass inside the expectation on each side (since increasing K can only increase each sum, by possibly adding non-negative terms). Hence

$$\mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} f(r)\right] = \mathbb{E}\left[\lim_{K\to\infty} \sum_{r=1}^{\tau_N(t)\wedge K} f(r)\right] = \mathbb{E}\left[\lim_{K\to\infty} \sum_{r=1}^{\tau_N(t)\wedge K} \mathbb{E}[f(r) \mid \mathcal{F}_{r-1}]\right] = \mathbb{E}\left[\sum_{r=1}^{\tau_N(t)} \mathbb{E}[f(r) \mid \mathcal{F}_{r-1}]\right] \tag{5.101}$$

which concludes the proof. ■

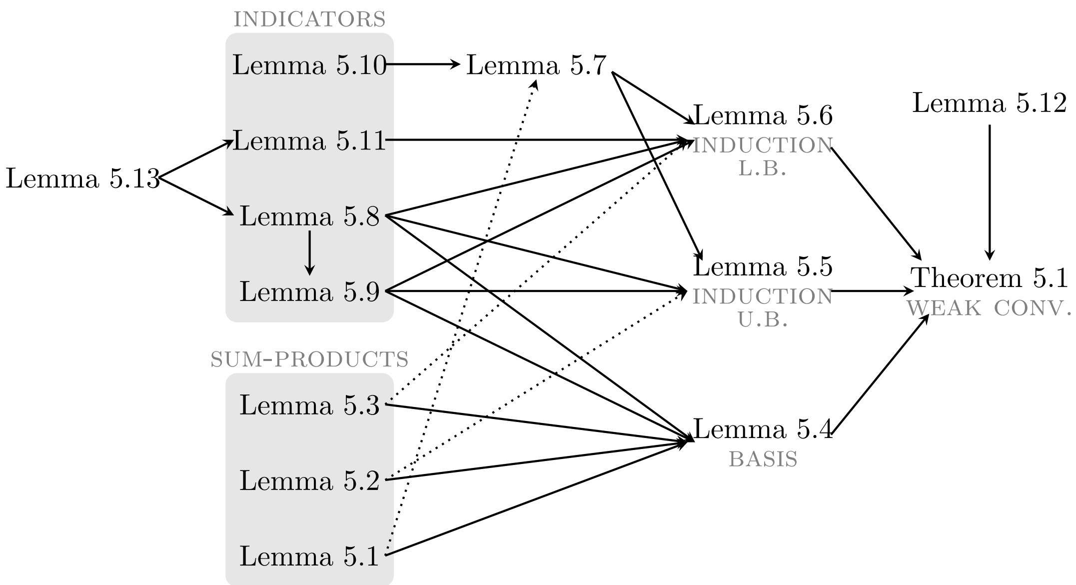

Figure 5.1: Graph showing dependencies between the lemmata used to prove weak convergence. Dotted arrows indicate dependence via a slight modification of the preceding lemma. Add FDD convergence theorem as another precedent of weak convergence theorem.

6 Discussion

Bibliography

[1] Suzie Brown et al. “Simple Conditions for Convergence of Sequential Monte Carlo Genealogies with Applications”. In: *Electronic Journal of Probability* 26.1 (2021), pp. 1–22. ISSN: 1083-6489. DOI: 10.1214/20-EJP561.

[2] Rick Durrett. *Probability: Theory and Examples*. 5th ed. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2019. DOI: 10.1017/9781108591034.

[3] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence*. John Wiley & Sons, 2009.

[4] Jere Koskela et al. *Asymptotic genealogies of interacting particle systems with an application to sequential Monte Carlo*. Mathematics e-print 1804.01811. ArXiv, 2018.

[5] Martin Möhle. “Weak Convergence to the Coalescent in Neutral Population Models”. In: *Journal of Applied Probability* 36.2 (1999), pp. 446–460.

[6] James E. Mosimann. “On the Compound Multinomial Distribution, the Multivariate β-Distribution, and Correlations among Proportions”. In: *Biometrika* 49.1/2 (1962), pp. 65–82.