

Asymptotic genealogies of sequential Monte Carlo algorithms

Suzie Brown

University of Warwick

with Paul Jenkins, Adam Johansen & Jere Koskela

18 November 2020

Outline

1. Sequential Monte Carlo
2. Resampling and degeneracy
3. SMC genealogies
4. Examples

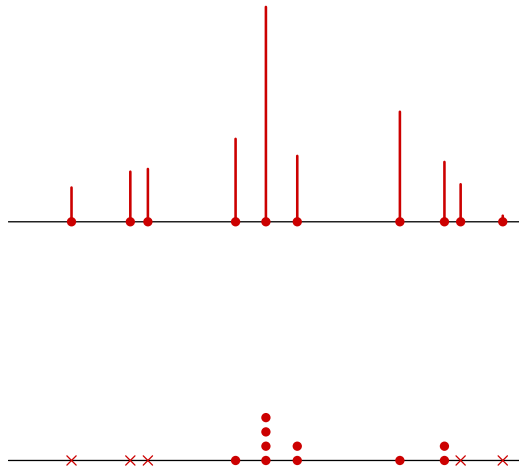
Sequential Monte Carlo

Iterate these steps:

1. **Mutate** particles via Markov transition density q_t
2. **Weight** particles by potential function g_t
3. **Resample** particles in proportion to their weights

Resampling

Stochastically map continuous weights $(w_t^{(1)}, \dots, w_t^{(N)})$ to discrete offspring counts $(\nu_t^{(1)}, \dots, \nu_t^{(N)})$



Resampling

Stochastically map continuous weights $(w_t^{(1)}, \dots, w_t^{(N)})$ to discrete offspring counts $(\nu_t^{(1)}, \dots, \nu_t^{(N)})$

Properties:

- ▶ Number of particles constant

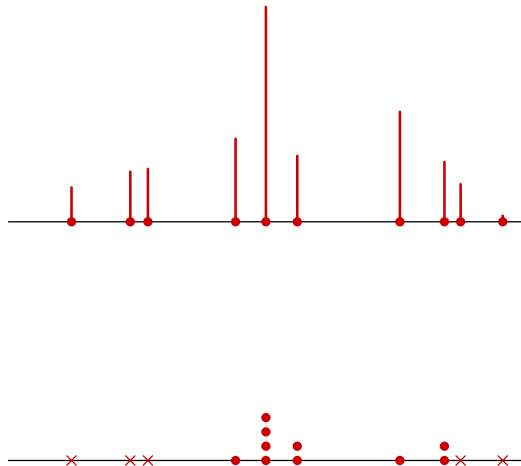
$$\sum_{i=1}^N \nu_t^{(i)} = N$$

- ▶ Equal weights after resampling

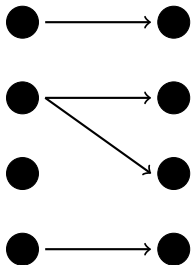
$$w_{t+}^{(i)} = 1/N$$

- ▶ Unbiased

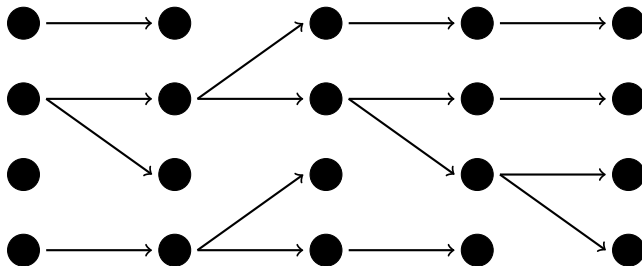
$$\mathbb{E}[\nu_t^{(i)} | w_t^{(i)}] = N w_t^{(i)}$$



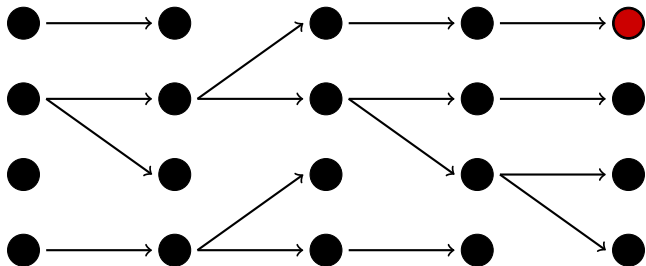
Resampling induces a genealogy



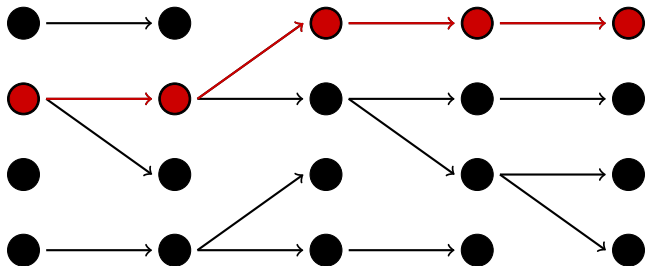
Resampling induces a genealogy



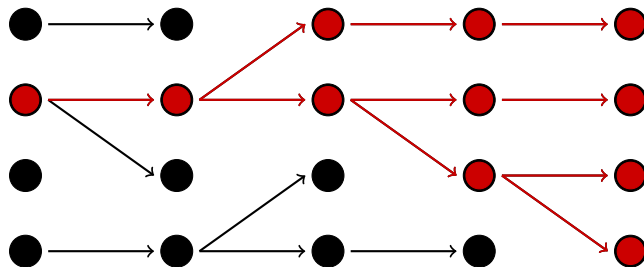
Resampling induces a genealogy



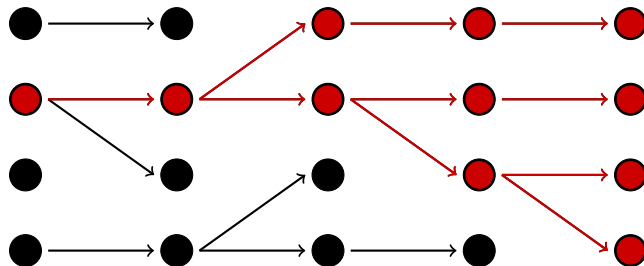
Resampling induces a genealogy



Resampling induces a genealogy



Resampling induces a genealogy

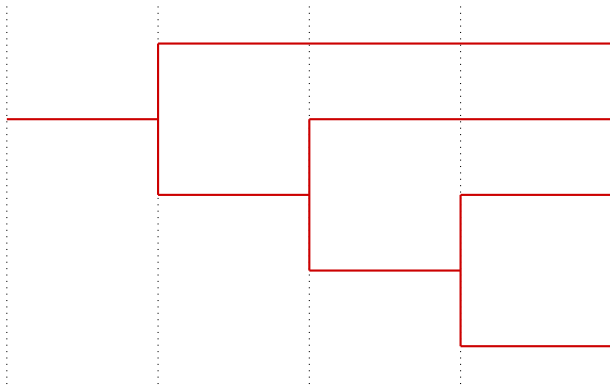


Ancestral degeneracy: for $t \ll T$, few distinct samples are available

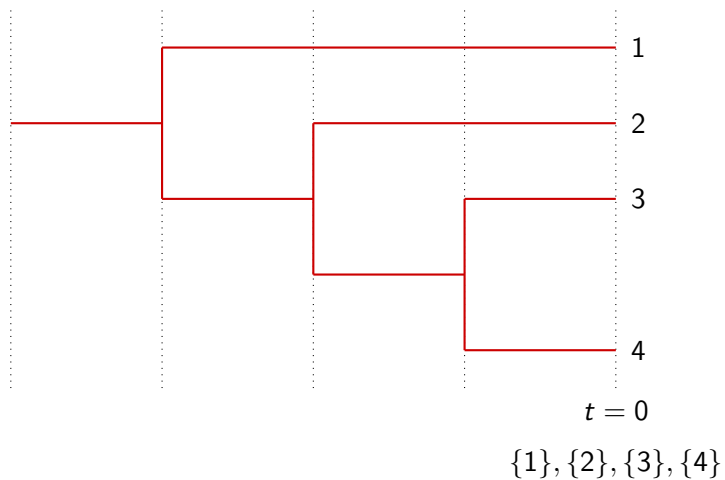
Encoding genealogies

- ▶ Label time in reverse
- ▶ Population of N particles
- ▶ Sample $n \leq N$ terminal particles
- ▶ Describe genealogy by stochastic process $(G_t^{(n,N)})_{t \in \mathbb{N}_0}$ on space of partitions of $\{1, \dots, n\}$
- ▶ Elements i, j are in the same block of the partition $G_t^{(n,N)}$ iff particles i and j share a common ancestor at time t
- ▶ Initially $G_0^{(n,N)} = \{\{1\}, \dots, \{n\}\}$
- ▶ The only possible non-identity transitions are those that merge blocks
- ▶ The trivial partition $\{\{1, \dots, n\}\}$ is an absorbing state

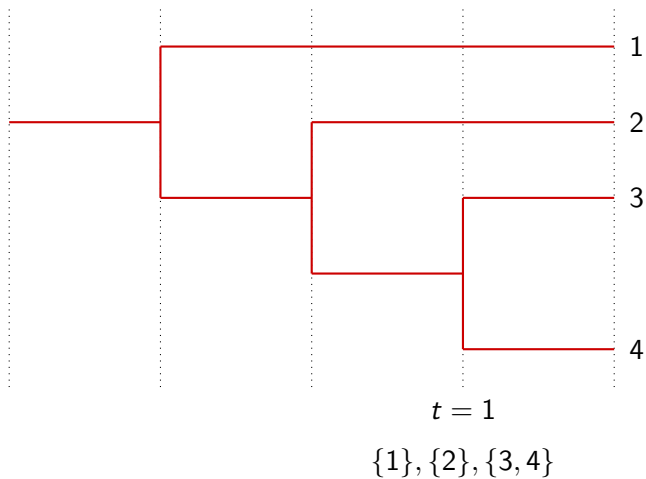
Encoding genealogies



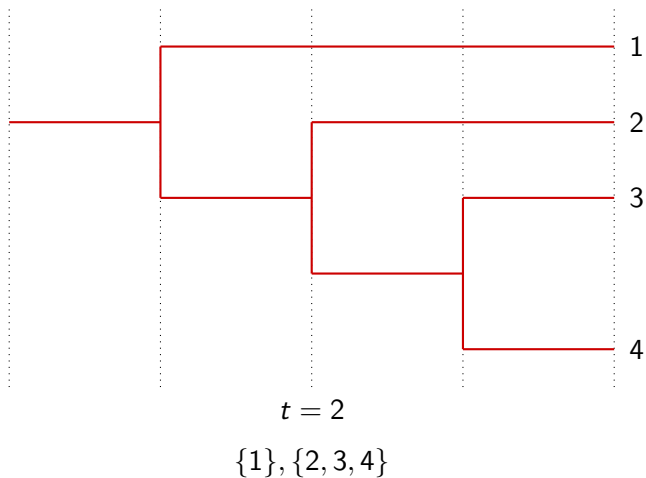
Encoding genealogies



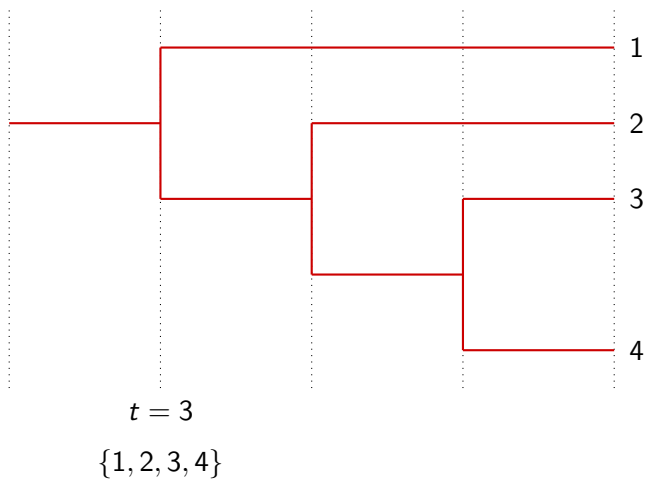
Encoding genealogies



Encoding genealogies



Encoding genealogies



Kingman's n -coalescent

- ▶ Continuous-time Markov chain on the space of partitions of $\{1, \dots, n\}$
- ▶ Single pair mergers only
- ▶ Each pair merges independently at rate 1 (total merge rate $\binom{k}{2}$ while there are k distinct lineages)

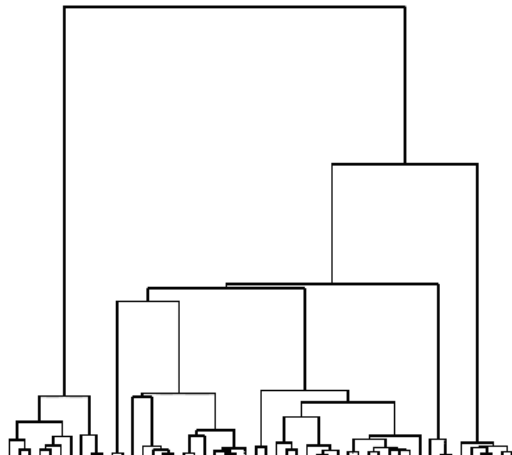


image: Wikimedia commons

Time scale

The probability that a randomly chosen pair of particles at generation t share a common ancestor at generation $(t - 1)$, conditional on offspring counts, is

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

To get an n -coalescent, this should converge to 1 (the required pair merger rate), so we rescale time by the inverse

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

Main theorem

- ▶ Offspring counts $(\nu_t^{(1)}, \dots, \nu_t^{(N)})$

Main theorem

- ▶ Offspring counts $(\nu_t^{(1)}, \dots, \nu_t^{(N)})$
- ▶ Parent-offspring assignments are uniform given offspring counts
- ▶ Time scale does not explode (i.e. $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all finite t)
- ▶ There exists a sequence (b_N) such that $\lim_{N \rightarrow \infty} b_N = 0$ and

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

Main theorem

- ▶ Offspring counts $(\nu_t^{(1)}, \dots, \nu_t^{(N)})$
- ▶ Parent-offspring assignments are uniform given offspring counts
- ▶ Time scale does not explode (i.e. $\mathbb{P}[\tau_N(t) = \infty] = 0$ for all finite t)
- ▶ There exists a sequence (b_N) such that $\lim_{N \rightarrow \infty} b_N = 0$ and

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

Then the finite-dimensional distributions of the time-rescaled genealogies converge to Kingman's n -coalescent as $N \rightarrow \infty$.

We established the theorem in these cases:

- ▶ Multinomial resampling
- ▶ Stochastic rounding
- ▶ Conditional SMC with multinomial resampling

Multinomial resampling

- ▶ Offspring counts are sampled from Multinomial distribution parametrised by weights
- ▶ Easy to analyse, but doesn't perform well
- ▶ (For the population geneticists): different from Wright-Fisher model because not neutral

Corollary: multinomial resampling

Consider an SMC algorithm with potential g and transition density q , satisfying

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

for constants $0 < \varepsilon \leq 1 \leq a < \infty$, and probability distribution $h(\cdot)$.

Under multinomial resampling, the rescaled genealogies converge to Kingman's coalescent.

Stochastic rounding

$\mathbf{Y} : \mathbb{R}_+^N \rightarrow \mathbb{N}^N$ is a *stochastic rounding* of \mathbf{X} if for $i = 1, \dots, N$

$$Y_i \mid X_i = \begin{cases} \lfloor X_i \rfloor & \text{with probability } 1 - X_i + \lfloor X_i \rfloor \\ \lfloor X_i \rfloor + 1 & \text{with probability } X_i - \lfloor X_i \rfloor \end{cases}$$

- ▶ We can construct low-variance resampling schemes using stochastic rounding
- ▶ Take $X_i = Nw_t^{(i)}$ and $Y_i = \nu_t^{(i)}$
- ▶ Require further constraint $Y_1 + \dots + Y_N = N$
- ▶ **Examples:** systematic, residual-stratified, branching system¹, SSP resampling²

¹Crisan, Lyons. *Prob Theory & Related Fields*, 1997

²Gerber, Whiteley, Chopin. *Ann Stat*, 2019.

Corollary: stochastic rounding

Consider an SMC algorithm with potential g and transition density q , satisfying

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

for constants $0 < \varepsilon \leq 1 \leq a < \infty$, and probability distribution $h(\cdot)$.

Under stochastic rounding -based resampling, the rescaled genealogies converge to Kingman's coalescent.

Conditional SMC

- ▶ Used for SMC updates in particle MCMC³
- ▶ One “immortal lineage” is conditioned to survive all mutation and resampling steps
- ▶ Resampling algorithm must deterministically propagate this “immortal lineage”
- ▶ For example: fix one offspring to immortal lineage, multinomial sampling for the remaining counts

³Andrieu, Doucet, Holenstein. *JRSSB*, 2010.

Conditional SMC

- ▶ Used for SMC updates in particle MCMC³
- ▶ One “immortal lineage” is conditioned to survive all mutation and resampling steps
- ▶ Resampling algorithm must deterministically propagate this “immortal lineage”
- ▶ For example: fix one offspring to immortal lineage, multinomial sampling for the remaining counts
- ▶ Conjecture: theorem also applies to conditional SMC with stochastic rounding

³Andrieu, Doucet, Holenstein. *JRSSB*, 2010.

Corollary: conditional SMC

Consider a conditional SMC algorithm with potential g and transition density q , satisfying

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

for constants $0 < \varepsilon \leq 1 \leq a < \infty$, and probability distribution $h(\cdot)$.

Under multinomial resampling, the rescaled genealogies converge to Kingman's coalescent.

Summary

- ▶ We consider genealogies of SMC algorithms
- ▶ We provide simple conditions for asymptotically Kingman genealogies
- ▶ Satisfied by stochastic rounding -based resampling
- ▶ Satisfied by conditional SMC with multinomial resampling...
- ▶ ... but this hides some interesting pre-limiting behaviour

Open questions

- ▶ Verify theorem for other important resampling schemes (stratified, residual-multinomial)
- ▶ How to estimate the time scale c_N a priori (since it depends on offspring counts)
- ▶ Weak convergence so we can say more about convergence of expectations
- ▶ Rates of convergence
- ▶ Finite- N behaviour

In conclusion...

- ▶ Genealogies can help us to analyse performance of smoothing algorithms which suffer ancestral degeneracy
- ▶ We have simple conditions under which these genealogies converge to n -coalescent
- ▶ These conditions are verified for some important classes of SMC algorithms

In conclusion...

- ▶ Genealogies can help us to analyse performance of smoothing algorithms which suffer ancestral degeneracy
- ▶ We have simple conditions under which these genealogies converge to n -coalescent
- ▶ These conditions are verified for some important classes of SMC algorithms

Open questions

- ▶ Verify theorem for other important resampling schemes (stratified, residual-multinomial)
- ▶ How to estimate the time scale c_N a priori (since it depends on offspring counts)
- ▶ Weak convergence so we can say more about convergence of expectations
- ▶ Rates of convergence
- ▶ Finite- N behaviour