Stochastic roundings — coalescent proof

Suzie Brown

September 4, 2019

Definition 1. Let $X \geq 0$. A random variable $Y : \mathbb{R}_+ \to \mathbb{N}$ is a *stochastic rounding* of X if Y takes the values

$$Y = \begin{cases} \lfloor X \rfloor & \text{with probability } 1 - X + \lfloor X \rfloor \\ \lfloor X \rfloor + 1 & \text{with probability } X - \lfloor X \rfloor \end{cases}$$

Theorem 1. *Under the time scaling of Koskela et al. (2018, Theorem 1) and the conditions of Koskela et al. (2018, Lemma 3), genealogies of SMC algorithms with stochastic rounding-based resampling schemes converge to Kingman's n-coalescent in the sense of finite-dimensional distributions as $N \to \infty$.*

Proof. We need to show that there exists a deterministic sequence $(b_N)_{N\in\mathbb{N}}$ such that $\lim_{N\to\infty} b_N = 0$ and

$$\frac{\frac{1}{(N)_3}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_3 | \mathcal{F}_{t-1}]}{\frac{1}{(N)_2}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_2 | \mathcal{F}_{t-1}]} \leq b_N \tag{1}$$

for all $N \in \mathbb{N}$. Directly applying Definition 1, we find for the denominator

$$\begin{aligned}\mathbb{E}[(v_i)_2^{(r)} | w_i] &= \lfloor Nw_t^{(i)} \rfloor (\lfloor Nw_t^{(i)} \rfloor - 1)(1 - Nw_i + \lfloor Nw_t^{(i)} \rfloor) + (\lfloor Nw_t^{(i)} \rfloor + 1)\lfloor Nw_t^{(i)} \rfloor (Nw_i - \lfloor Nw_t^{(i)} \rfloor) \\ &= \lfloor Nw_t^{(i)} \rfloor \left(2(Nw_i - \lfloor Nw_t^{(i)} \rfloor) + \lfloor Nw_t^{(i)} \rfloor - 1 \right)\end{aligned}$$

For the numerator we find

$$\begin{aligned}\mathbb{E}[(\nu_t^{(i)})_3 | w_t^{(1:N)}] &= (\lfloor Nw_t^{(i)} \rfloor)_3 (1 - Nw_t^{(i)} + \lfloor Nw_t^{(i)} \rfloor) + (\lfloor Nw_t^{(i)} \rfloor + 1)_3 (Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor) \\ &= (\lfloor Nw_t^{(i)} \rfloor)_2 \left\{ (\lfloor Nw_t^{(i)} \rfloor - 2)(1 - Nw_t^{(i)} + \lfloor Nw_t^{(i)} \rfloor) + (\lfloor Nw_t^{(i)} \rfloor + 1)(Nw_t^{(i)} - \lfloor Nw_t^{(i)} \rfloor) \right\} \\ &= (\lfloor Nw_t^{(i)} \rfloor)_2 \left(3Nw_t^{(i)} - 2\lfloor Nw_t^{(i)} \rfloor - 2 \right)\end{aligned}$$

So for the ratio we have

$$
\begin{aligned}
\frac{\frac{1}{(N)_3}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_3|\mathcal{F}_{t-1}]}{\frac{1}{(N)_2}\sum_{i=1}^N \mathbb{E}[(\nu_t^{(i)})_2|\mathcal{F}_{t-1}]} &= \frac{1}{N-2}\frac{\sum_{i=1}^N \mathbb{E}[(\lfloor Nw_t^{(i)}\rfloor)_2(3Nw_t^{(i)}-2\lfloor Nw_t^{(i)}\rfloor-2)|\mathcal{F}_{t-1}]}{\sum_{i=1}^N \mathbb{E}[\lfloor Nw_t^{(i)}\rfloor(2Nw_t^{(i)}-\lfloor Nw_t^{(i)}\rfloor-1)|\mathcal{F}_{t-1}]} \\
&\leq \frac{1}{N-2}\frac{\sum_{i=1}^N \mathbb{E}[(\lfloor Nw_t^{(i)}\rfloor)_2 Nw_t^{(i)}|\mathcal{F}_{t-1}]}{\sum_{i=1}^N \mathbb{E}[\lfloor Nw_t^{(i)}\rfloor(Nw_t^{(i)}-1)|\mathcal{F}_{t-1}]} \qquad \text{since } Nw_t^{(i)}-\lfloor Nw_t^{(i)}\rfloor\in[0,1) \\
&\leq \frac{1}{N-2}\frac{\sum_{i=1}^N \mathbb{E}[(Nw_t^{(i)})_2 Nw_t^{(i)}|\mathcal{F}_{t-1}]}{\sum_{i=1}^N \mathbb{E}[(Nw_t^{(i)}-1)^2|\mathcal{F}_{t-1}]} \qquad \text{since } Nw_t^{(i)}-\lfloor Nw_t^{(i)}\rfloor\in[0,1) \\
&= \frac{1}{N-2}\frac{\sum_{i=1}^N \mathbb{E}[(Nw_t^{(i)})^3|\mathcal{F}_{t-1}]-\sum_{i=1}^N \mathbb{E}[(Nw_t^{(i)})^2|\mathcal{F}_{t-1}]}{N^2\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]-2N\sum_{i=1}^N \mathbb{E}[w_t^{(i)}|\mathcal{F}_{t-1}]+\sum_{i=1}^N \mathbb{E}[1|\mathcal{F}_{t-1}]} \\
&= \frac{1}{N-2}\frac{N^3\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^3|\mathcal{F}_{t-1}]-N^2\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]}{N^2\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]-2N+N} \qquad \text{since } \sum w_t^{(i)}=1 \\
&= \frac{1}{N-2}\frac{N^2\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^3|\mathcal{F}_{t-1}]-N\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]}{N\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]-1} \\
&= \frac{1}{N-2}\left[\frac{N^2\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^3|\mathcal{F}_{t-1}]}{N\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]-1}-\frac{N\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]}{N\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]-1}\right] \\
&= \frac{1}{N-2}\left[-1+\frac{N^2\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^3|\mathcal{F}_{t-1}]-1}{N\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]-1}\right]
\end{aligned}
$$

Then, using that $w_t^{(i)}=\Theta(N^{-1})$, the sum in the denominator is $O(N^{-1})$ and the sum in the numerator is $O(N^{-2})$. Hence the whole expression is $O(N^{-1})$ as $N\to\infty$, so we can find a suitable sequence $b_N \overset{N\to\infty}{\longrightarrow} 0$ to satisfy the conditions of Theorem (?) [the one with new assns]. □

The conclusion of the above proof is not rigorous... but we can use a simpler argument based on the bounded weights.

Proof. Directly applying Definition 1, we find firstly

$$
\begin{aligned}
\mathbb{E}[(v_i)_2^{(r)}|w_t^{(1:N)}] &= \lfloor Nw_t^{(i)}\rfloor(\lfloor Nw_t^{(i)}\rfloor-1)(1-Nw_i+\lfloor Nw_t^{(i)}\rfloor)+(\lfloor Nw_t^{(i)}\rfloor+1)\lfloor Nw_t^{(i)}\rfloor(Nw_i-\lfloor Nw_t^{(i)}\rfloor) \\
&= \lfloor Nw_t^{(i)}\rfloor\left(2(Nw_i-\lfloor Nw_t^{(i)}\rfloor)+\lfloor Nw_t^{(i)}\rfloor-1\right)
\end{aligned}
$$

In the same way,

$$
\begin{aligned}
\mathbb{E}[(\nu_t^{(i)})_3|w_t^{(1:N)}] &= (\lfloor Nw_t^{(i)}\rfloor)_3(1-Nw_t^{(i)}+\lfloor Nw_t^{(i)}\rfloor)+(\lfloor Nw_t^{(i)}\rfloor+1)_3(Nw_t^{(i)}-\lfloor Nw_t^{(i)}\rfloor) \\
&= (\lfloor Nw_t^{(i)}\rfloor)_2\left\{(\lfloor Nw_t^{(i)}\rfloor-2)(1-Nw_t^{(i)}+\lfloor Nw_t^{(i)}\rfloor)+(\lfloor Nw_t^{(i)}\rfloor+1)(Nw_t^{(i)}-\lfloor Nw_t^{(i)}\rfloor)\right\} \\
&= (\lfloor Nw_t^{(i)}\rfloor)_2\left(2Nw_t^{(i)}-\lfloor Nw_t^{(i)}\rfloor-1+(Nw_t^{(i)}-\lfloor Nw_t^{(i)}\rfloor-1)\right) \\
&\leq (\lfloor Nw_t^{(i)}\rfloor-1)\lfloor Nw_t^{(i)}\rfloor\left(2Nw_t^{(i)}-\lfloor Nw_t^{(i)}\rfloor-1\right).
\end{aligned}
$$

In the case where $\lfloor Nw_t^{(i)}\rfloor\geq 1$, we therefore have

$$
\begin{aligned}
\mathbb{E}[(\nu_t^{(i)})_3|w_t^{(1:N)}] &\leq (\lfloor Nw_t^{(i)}\rfloor-1)\mathbb{E}[(v_i)_2^{(r)}|w_t^{(1:N)}] \\
&\leq \mathbb{E}[(v_i)_2^{(r)}|w_t^{(1:N)}]\left(\frac{a^4}{\varepsilon^4}-1\right)
\end{aligned}
$$

with constants a and ε as before. The case $\lfloor Nw_t^{(i)}\rfloor<1$ is trivial because then $\mathbb{E}[(\nu_t^{(i)})_3|w_t^{(1:N)}]=0$ while $\mathbb{E}[(v_i)_2^{(r)}|w_t^{(1:N)}]\geq 0$, so b_N can be any non-negative sequence. □

References

Koskela, J., Jenkins, P. A., Johansen, A. M. and Spanò, D. (2018), 'Asymptotic genealogies of interacting particle systems with an application to sequential Monte Carlo', *arXiv preprint arXiv:1804.01811* .