Conditional SMC with Multinomial Resampling

Suzie Brown

December 4, 2018

In standard SMC with multinomial resampling, the marginal offspring distributions, conditioned on the filtration $\mathcal{F}_{t-1}$ generated by the previous offspring counts, are

$$v_t^{(i)} \overset{d}{=} \mathrm{Bin}(N, w_t^{(i)}), \qquad i = 1, \ldots, N$$

where $v_t^{(i)}$ is the number of offspring in generation $t+1$ of the ith particle in generation t, N is the number of particles and $w_t^{(i)}$ is the weight associated with the ith particle in generation t.

In conditional SMC we condition on the immortal line surviving each resampling step. By exchangeability we can set without loss of generality that the immortal line consists of particle 1 in each generation. At each resampling step, particle 1 must therefore choose particle 1 as its parent, while the remaining $N-1$ offspring are assigned multinomially to the N possible parents. The marginal offspring distributions are then

$$\tilde{v}_t^{(1)} \overset{d}{=} 1 + \mathrm{Bin}(N-1, w_t^{(1)})$$
$$\tilde{v}_t^{(i)} \overset{d}{=} \mathrm{Bin}(N-1, w_t^{(i)}), \qquad i = 2, \ldots, N$$

where we write $\tilde{v}_t^{(i)}$ to distinguish from $v_t^{(i)}$. Similarly, any other quantities written with tilde are the conditional SMC analogues of the corresponding untilded quantities.

In the following we will make extensive use of the formula for factorial moments of the multinomial distribution given in Mosimann (1962, p.67), which has the convenient form

$$\mathbb{E}[(X_i)_a (X_j)_b] = (n)_{a+b}\, p_i^a p_j^b$$

where $(X_1, \ldots X_k) \sim \mathrm{Multinomial}(n, \mathbf{p})$. To take advantage of this convenient expression we will convert any regular moments appearing in the calculations into factorial moments, using the identities summarised in Table 1.

The coalescence rate is given by

$$c_N(t) := \frac{1}{(N)_2} \sum_{i=1}^{N} \mathbb{E}\left[(v_t^{(i)})_2\right].$$

In the following, all expectations are conditional on $\mathcal{F}_{t-1}$ but the conditioning is omitted for notational brevity. In the standard case we have, using the tower rule and the multinomial moments,

$$\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}] = \frac{1}{(N)_2} \sum_{i=1}^{N} \mathbb{E}\left[\mathbb{E}[(v_t^{(i)})_2]|\mathcal{F}_{t-1}\right] = \frac{1}{(N)_2} \sum_{i=1}^{N} \mathbb{E}\left[(N)_2 (w_t^{(i)})^2|\mathcal{F}_{t-1}\right] = \sum_{i=1}^{N} \mathbb{E}\left[(w_t^{(i)})^2|\mathcal{F}_{t-1}\right]$$

x	$=$	$(x)_1$
x^2	$=$	$(x)_2 + (x)_1$
x^3	$=$	$(x)_3 + 3(x)_2 + (x)_1$
x^4	$=$	$(x)_4 + 6(x)_3 + 7(x)_2 + (x)_1$
xy	$=$	$(x)_1(y)_1$
$x^2 y$	$=$	$(x)_2(y)_1 + (x)_1(y)_1$
xy^2	$=$	$(x)_1(y)_2 + (x)_1(y)_1$
$x^2 y^2$	$=$	$(x)_2(y)_2 + (x)_2(y)_1 + (x)_1(y)_2 + (x)_1(y)_1$

Table 1: Conversion of regular powers into falling factorial powers

The corresponding quantity in the conditional case is

$$\tilde{c}_N(t) = \frac{1}{(N)_2} \sum_{i=1}^{N} \mathbb{E}\left[(\tilde{v}_t^{(i)})_2\right] = \frac{1}{(N)_2}\mathbb{E}\left[(\tilde{v}_t^{(1)})_2\right] + \frac{1}{(N)_2}\sum_{i=2}^{N}\mathbb{E}\left[(\tilde{v}_t^{(i)})_2\right]$$

Then applying the tower rule and factorial moments as before we obtain

$$\mathbb{E}[\tilde{c}_N(t)|\mathcal{F}_{t-1}] = \frac{1}{(N)_2}\left\{(N-1)_2\mathbb{E}[(w_t^{(1)})^2|\mathcal{F}_{t-1}] + 2(N-1)\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}] + \sum_{i=2}^{N}(N-1)_2\mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]\right\}$$

$$= \frac{(N-1)_2}{(N)_2}\sum_{i=1}^{N}\mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}] + \frac{2(N-1)}{(N)_2}\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}]$$

$$= \frac{N-2}{N}\mathbb{E}[c_N(t)|\mathcal{F}_{t-1}] + \frac{2}{N}\mathbb{E}[w_t^{(1)}|\mathcal{F}_{t-1}]$$

The rate of super-binary mergers is bounded above by

$$D_N(t) := \frac{1}{N(N)_2}\sum_{i=1}^{N}(v_t^{(i)})_2\left(v_t^{(i)} + \frac{1}{N}\sum_{j\neq i}(v_t^{(j)})^2\right)$$

$$= \frac{1}{N(N)_2}\sum_{i=1}^{N}\left\{(v_t^{(i)})_2 v_t^{(i)} + \frac{1}{N}\sum_{j\neq i}(v_t^{(i)})_2(v_t^{(j)})^2\right\}$$

$$= \frac{1}{N(N)_2}\sum_{i=1}^{N}\left\{(v_t^{(i)})_3 + 2(v_t^{(i)})_2 + \frac{1}{N}\sum_{j\neq i}\left((v_t^{(i)})_2(v_t^{(j)})_2 + (v_t^{(i)})_2 v_t^{(j)}\right)\right\}$$

We therefore find

$$\mathbb{E}[D_N(t)|\mathcal{F}_{t-1}] = \frac{1}{N(N)_2}\sum_{i=1}^{N}\left\{(N)_3\mathbb{E}[(w_t^{(i)})^3|\mathcal{F}_{t-1}] + 2(N)_2\mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}]\right\}$$

$$+ \frac{1}{N^2(N)_2}\sum_{i=1}^{N}\sum_{j\neq i}\left\{(N)_4\mathbb{E}[(w_t^{(i)})^2(w_t^{(j)})^2|\mathcal{F}_{t-1}] + (N)_3\mathbb{E}[(w_t^{(i)})^2 w_t^{(j)}|\mathcal{F}_{t-1}]\right\}$$

For the conditional case, we separate the terms involving particle 1:

$$\tilde{D}_N(t) = \frac{1}{N(N)_2}(\tilde{v}_t^{(1)})_2\left(\tilde{v}_t^{(1)} + \frac{1}{N}\sum_{j\neq 1}(\tilde{v}_t^{(j)})^2\right) + \frac{1}{N(N)_2}\sum_{i\neq 1}(\tilde{v}_t^{(i)})_2\left(\tilde{v}_t^{(i)} + \frac{1}{N}(\tilde{v}_t^{(1)})^2 + \frac{1}{N}\sum_{1\neq j\neq i}(\tilde{v}_t^{(j)})^2\right)$$

$$= \frac{1}{N(N)_2}\left\{(v_t^{(1)})_2 v_t^{(1)} + \frac{1}{N}\sum_{j\neq 1}(v_t^{(1)})_2(v_t^{(j)})^2 + \frac{1}{N}\sum_{i\neq 1}(v_t^{(i)})_2(v_t^{(1)})^2\right\}$$

$$+ \frac{1}{N(N)_2}\sum_{i\neq 1}\left\{(v_t^{(i)})_2 v_t^{(i)} + \frac{1}{N}\sum_{1\neq j\neq i}(v_t^{(i)})_2(v_t^{(j)})^2\right\}$$

$$= \frac{1}{N(N)_2}\left\{(v_t^{(1)})_2 v_t^{(1)} + \frac{1}{N}\sum_{i\neq 1}\left((v_t^{(1)})_2(v_t^{(i)})^2 + (v_t^{(i)})_2(v_t^{(1)})^2\right)\right\}$$

$$+ \frac{1}{N(N)_2}\sum_{i\neq 1}\left\{(v_t^{(i)})_3 + 2(v_t^{(i)})_2 + \frac{1}{N}\sum_{1\neq j\neq i}\left((v_t^{(i)})_2(v_t^{(j)})_2 + (v_t^{(i)})_2 v_t^{(j)}\right)\right\}$$

and we find the expectation..... (below not corrected yet)

$$\mathbb{E}[\tilde{D}_N(t)] = \frac{N-2}{N^2} \sum_{i=1}^{N} \left((N-3)\mathbb{E}[(w_t^{(i)})^3] + \left(2 + \frac{1}{N}\right) \mathbb{E}[(w_t^{(i)})^2] + \frac{3}{N-2}\mathbb{E}[w_t^{(i)}] \right)$$

$$+ \frac{(N-2)(N-3)}{N^3} \sum_{i=1}^{N}\sum_{j\neq i} \left((N-4)\mathbb{E}[(w_t^{(i)})^2(w_t^{(j)})^2] + \mathbb{E}[(w_t^{(i)})^2 w_t^{(j)}] \right)$$

$$+ \frac{1}{N^2}\mathbb{E}[w_t^{(1)}] + \left(3 - \frac{1}{N}\right)\frac{(N-2)}{N^2}\mathbb{E}[(w_t^{(1)})^2] + \frac{2(N-2)}{N^3}\sum_{i=2}^{N}\mathbb{E}[w_t^{(1)}w_t^{(i)}] + \frac{4(N-2)(N-3)}{N^3}\sum_{i=2}^{N}\mathbb{E}[w_t^{(1)}(w_t^{(i)})^2]$$

$$\leq \mathbb{E}[D_N(t)] + \frac{1}{N^3}\sum_{i=1}^{N}\mathbb{E}[(w_t^{(i)})^2] + \frac{1}{N^2}\mathbb{E}[w_t^{(1)}] + \frac{3}{N}\mathbb{E}[(w_t^{(1)})^2] + \frac{2}{N^2}\sum_{i=2}^{N}\mathbb{E}[w_t^{(1)}w_t^{(i)}] + \frac{4}{N}\sum_{i=2}^{N}\mathbb{E}[w_t^{(1)}(w_t^{(i)})^2].$$

We also require a bound on the squared coalescence rate, obtained by a similar separation of terms involving particle 1.

$$\mathbb{E}[c_N(t)^2] = \frac{1}{(N)_2^2}\mathbb{E}\left[\left(\sum_{i=1}^{N}(v_t^{(i)})_2\right)^2\right]$$

$$= \frac{1}{(N)_2^2}\left\{\sum_{i=1}^{N}\mathbb{E}[(v_t^{(i)})_2^2] + \sum_{i=1}^{N}\sum_{j\neq i}\mathbb{E}[(v_t^{(i)})_2(v_t^{(j)})_2]\right\}$$

$$= \frac{1}{(N)_2^2}\left\{(N)_4\sum_{i=1}^{N}\mathbb{E}[(w_t^{(i)})^4] + 4(N)_3\sum_{i=1}^{N}\mathbb{E}[(w_t^{(i)})^3] + 2(N)_2\sum_{i=1}^{N}\mathbb{E}[(w_t^{(i)})^2]\right\}$$

$$+ \frac{1}{(N)_2^2}(N)_4\sum_{i=1}^{N}\sum_{j\neq i}\mathbb{E}[(w_t^{(i)})^2(w_t^{(j)})^2]$$

and

$$\mathbb{E}[\tilde{c}_N(t)^2] = \frac{1}{(N)_2^2}\left\{\sum_{i=1}^{N}\mathbb{E}[(\tilde{v}_t^{(i)})_2^2] + \sum_{i=1}^{N}\sum_{j\neq i}\mathbb{E}[(\tilde{v}_t^{(i)})_2(\tilde{v}_t^{(j)})_2]\right\}$$

$$= \frac{1}{(N)_2^2}\left\{\sum_{i=2}^{N}\mathbb{E}[(\tilde{v}_t^{(i)})_2^2] + \mathbb{E}[(\tilde{v}_t^{(1)})_2^2] + \sum_{i=2}^{N}\sum_{1\neq j\neq i}\mathbb{E}[(\tilde{v}_t^{(i)})_2(\tilde{v}_t^{(j)})_2] + 2\sum_{i=1}^{N}\mathbb{E}[(\tilde{v}_t^{(1)})_2(\tilde{v}_t^{(i)})_2]\right\}$$

$$= \frac{1}{(N)_2^2}\left\{(N-1)_4\sum_{i=1}^{N}\mathbb{E}[(w_t^{(i)})^4] + 4(N-1)_3\sum_{i=1}^{N}\mathbb{E}[(w_t^{(i)})^3] + 2(N-1)_2\sum_{i=1}^{N}\mathbb{E}[(w_t^{(i)})^2]\right\}$$

$$+ \frac{1}{(N)_2^2}\left\{(N-1)_4\sum_{i=1}^{N}\sum_{j\neq i}\mathbb{E}[(w_t^{(i)})^2(w_t^{(j)})^2] + 4(N-1)_3\mathbb{E}[(w_t^{(1)})^3] + 12(N-1)_2\mathbb{E}[(w_t^{(1)})^2] + 4(N-1)\mathbb{E}[w_t^{(1)}]\right\}$$

$$+ \frac{1}{(N)_2^2}\left\{2(N-1)_3\sum_{i=2}^{N}\mathbb{E}[w_t^{(1)}(w_t^{(i)})^2]\right\}$$

$$\leq \mathbb{E}[c_N(t)^2] + \frac{1}{(N)_2^2}\left\{4(N-1)_3\mathbb{E}[(w_t^{(1)})^3] + 12(N-1)_2\mathbb{E}[(w_t^{(1)})^2] + 4(N-1)\mathbb{E}[w_t^{(1)}] + 2(N-1)_3\sum_{i=2}^{N}\mathbb{E}[w_t^{(1)}(w_t^{(i)})^2]\right\}$$

We can modify the proof of Koskela et al. (2018, Lemma 3) by replacing the untilded quantities with these modified expressions. We still assume the conditions (18) and (19) of the lemma, which control the weights and allow us to find the following asymptotic expressions:

$$\mathbb{E}[\tilde{c}_N(t)] \leq \mathbb{E}[c_N(t)] + O(N^{-2})$$
$$\mathbb{E}[\tilde{c}_N(t)^2] \leq \mathbb{E}[c_N(t)^2] + O(N^{-3})$$
$$\mathbb{E}[\tilde{D}_N(t)] \leq \mathbb{E}[D_N(t)] + O(N^{-3})$$

Therefore we get the upper bound

$$
\begin{aligned}
\mathbb{E}[\tilde{c}_N(t)] &= \frac{N-2}{N}\mathbb{E}[c_N(t)] + \frac{2}{N}\mathbb{E}[w_t^{(1)}] \\
&\leq \mathbb{E}[c_N(t)] + \frac{2}{N}\mathbb{E}[w_t^{(1)}] \\
&\leq \frac{a^4}{N\varepsilon^4} + \frac{2}{N}\mathbb{E}[w_t^{(1)}] \\
&= \frac{a^4}{N\varepsilon^4} + O(N^{-2})
\end{aligned}
$$

and lower bound

$$
\begin{aligned}
\mathbb{E}[\tilde{c}_N(t)] &= \frac{N-2}{N}\mathbb{E}[c_N(t)] + \frac{2}{N}\mathbb{E}[w_t^{(1)}] \\
&\geq \frac{N-2}{N}\frac{\varepsilon^4}{Na^4} + O(N^{-2}) \\
&= \frac{\varepsilon^4}{Na^4} - \frac{2\varepsilon^4}{N^2a^4} + O(N^{-2}) \\
&= \frac{\varepsilon^4}{Na^4} + O(N^{-2})
\end{aligned}
$$

as in Koskela et al. (2018), with the addition of an error term. Using that (22) in the lemma implies $\mathbb{E}[c_N(t)] = O(N^{-1})$, we also have the reverse inequality

$$
\mathbb{E}[\tilde{c}_N(t)] = \mathbb{E}[c_N(t)] - \frac{2}{N}\mathbb{E}[c_N(t)] + O(N^{-2}) \geq \mathbb{E}[c_N(t)] + O(N^{-2}).
$$

This allows us to deduce the following upper bounds:

$$
\begin{aligned}
\mathbb{E}[\tilde{D}_N(t)] &\leq \mathbb{E}[D_N(t)] + O(N^{-3}) \\
&\leq \frac{C}{N}\mathbb{E}[c_N(t)] + O(N^{-3}) \\
&= \frac{C}{N}\mathbb{E}[\tilde{c}_N(t)] + O(N^{-3})
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}[\tilde{c}_N(t)^2] &\leq \mathbb{E}[c_N(t)^2] + O(N^{-3}) \\
&\leq \frac{C}{N}\mathbb{E}[c_N(t)] + O(N^{-3}) \\
&= \frac{C}{N}\mathbb{E}[\tilde{c}_N(t)] + O(N^{-3})
\end{aligned}
$$

which correspond to (20) and (21) respectively in the statement of the lemma, again with additional error terms. C indicates an arbitrary constant which may change from one line to the next.

We conclude an analogous result to Koskela et al. (2018, Lemma 3) for conditional SMC, involving the tilded quantities and with the addition of some vanishing errors that should not affect the further analysis.

References

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

Mosimann, J. E. (1962), 'On the compound multinomial distribution, the multivariate β-distribution, and correlations among proportions', *Biometrika* **49**(1/2), 65–82.