Some updated calculations for conditional SMC

Suzie Brown

March 18, 2019

Expected coalescence rates

Standard SMC with multinomial resampling has marginal offspring distributions

$$v_t^{(i)} \overset{d}{=} \text{Bin}(N, w_t^{(i)}), \qquad i = 1, \dots, N.$$

The coalescence rate is given by

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

In the case of multinomial resampling we have

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

In the conditional SMC case, to ensure the immortal line survives, individual 1 in each time step necessarily produces at least one offspring. (Exchangeability means we can label the immortal particle as particle 1 in each generation). It is straightforward to check that under this conditioning, the remaining $N-1$ offspring are assigned multinomially to the N possible parents as usual, giving the following offspring distributions:

$$\begin{aligned}
\tilde{v}_t^{(1)} &\overset{d}{=} 1 + \text{Bin}(N-1, w_t^{(1)}) \\
\tilde{v}_t^{(i)} &\overset{d}{=} \text{Bin}(N-1, w_t^{(i)}), \qquad i = 2, \dots, N.
\end{aligned}$$

and we can derive the altered coalescence rate:

$$\begin{aligned}
\mathbb{E}[\tilde{c}_N(t)] &= \frac{1}{(N)_2} \sum_{i=1}^{N} \mathbb{E}\left[(\tilde{v}_t^{(i)})_2 \right] \\
&= \frac{1}{(N)_2} \mathbb{E}\left[(\tilde{v}_t^{(1)})^2 - \tilde{v}_t^{(1)} \right] + \frac{1}{(N)_2} \sum_{i=2}^{N} \mathbb{E}\left[(\tilde{v}_t^{(i)})^2 - \tilde{v}_t^{(i)} \right] \\
&= \frac{1}{(N)_2} \left[(N-1)(N-2)\mathbb{E}[(w_t^{(1)})^2] + 2(N-1)\mathbb{E}[w_t^{(1)}] \right] + \frac{1}{(N)_2} \sum_{i=2}^{N} (N-1)(N-2)\mathbb{E}[(w_t^{(i)})^2] \\
&= \frac{1}{(N)_2} \sum_{i=1}^{N} (N-1)(N-2)\mathbb{E}[(w_t^{(i)})^2] + \frac{1}{(N)_2} 2(N-1)\mathbb{E}[w_t^{(1)}] \\
&= \frac{N-2}{N} \mathbb{E}[c_N(t)] + \frac{2}{N} \mathbb{E}[w_t^{(1)}]
\end{aligned} \tag{2}$$

The rate of super-binary mergers is bounded above by

$$D_N(t) := \frac{1}{N(N)_2} \sum_{i=1}^{N} (v_t^{(i)})_2 \left(v_t^{(i)} + \frac{1}{N} \sum_{j \neq i} (v_t^{(j)})^2 \right). \tag{3}$$

From the definition, first separate the terms involving particle 1 (which is special in the conditional model).

$$
\begin{aligned}
D_N(t) &:= \frac{1}{N(N)_2}\sum_{i=1}^{N}(v_t^{(i)})_2\left(v_t^{(i)}+\frac{1}{N}\sum_{j\neq i}(v_t^{(j)})^2\right)\\
&= \frac{1}{N(N)_2}(v_t^{(1)})_2\left(v_t^{(1)}+\frac{1}{N}\sum_{j\neq 1}(v_t^{(j)})^2\right)\\
&\quad + \frac{1}{N(N)_2}\sum_{i\neq 1}(v_t^{(i)})_2\left(v_t^{(i)}+\frac{1}{N}(v_t^{(1)})^2+\frac{1}{N}\sum_{1\neq j\neq i}(v_t^{(j)})^2\right)\\
&= \frac{1}{N(N)_2}\left((v_t^{(1)})^3-(v_t^{(1)})^2\right)+\frac{1}{N(N)_2}\sum_{i\neq 1}\left(\frac{1}{N}(v_t^{(1)})^2(v_t^{(i)})^2-\frac{1}{N}v_t^{(1)}(v_t^{(i)})^2\right)\\
&\quad + \frac{1}{N(N)_2}\sum_{i\neq 1}\left((v_t^{(i)})^3-(v_t^{(i)})^2+\frac{1}{N}(v_t^{(i)})^2(v_t^{(1)})^2-\frac{1}{N}(v_t^{(1)})^2v_t^{(i)}\right)\\
&\quad + \frac{1}{N^2(N)_2}\sum_{i\neq 1}\sum_{1\neq j\neq i}\left((v_t^{(i)})^2(v_t^{(j)})^2-v_t^{(i)}(v_t^{(j)})^2\right) \qquad (4)
\end{aligned}
$$

Let us consider the terms separately:

$$
\begin{aligned}
A &:= (v_t^{(1)})^3-(v_t^{(1)})^2\\
B &:= \frac{1}{N}(v_t^{(1)})^2(v_t^{(i)})^2-\frac{1}{N}v_t^{(1)}(v_t^{(i)})^2 && i>1\\
C &:= (v_t^{(i)})^3-(v_t^{(i)})^2 && i>1\\
D &:= \frac{1}{N}(v_t^{(i)})^2(v_t^{(1)})^2-\frac{1}{N}(v_t^{(1)})^2v_t^{(i)} && i>1\\
E &:= (v_t^{(i)})^2(v_t^{(j)})^2-v_t^{(i)}(v_t^{(j)})^2 && i,j>1; i\neq j
\end{aligned}
$$

Notice that A depends only on particle 1; C and E do not depend on particle 1; and B and D depend on particle 1 and others. The terms C and E will be the same in the standard and conditional cases, except that N is replaced by $N-1$ for conditional resampling. In the standard case, expressions involving particle 1 will be the same as the corresponding terms for other particles, but this is not the case for conditional resampling.

Let $X_1,\dots,X_k \sim \mathrm{MN}(n,(p_1,\dots,p_k))$. Then we have the following moments (due to REF):

$$
\begin{aligned}
\mathbb{E}[X_i] &= np_i\\
\mathbb{E}[X_i^2] &= np_i((n-1)p_i+1)\\
\mathbb{E}[X_i^3] &= np_i((n-1)(n-2)p_i^2+3(n-1)p_i+4)\\
\mathbb{E}[X_iX_j] &= n(n-1)p_ip_j\\
\mathbb{E}[X_i^2X_j] &= n(n-1)p_ip_j((n-2)p_i+1)\\
\mathbb{E}[X_i^2X_j^2] &= n(n-1)p_ip_j((n-2)(n-3)p_ip_j+(n-2)(p_i+p_j)+1)
\end{aligned}
$$

We can now calculate the quantities A–E above (first in the standard SMC case):

$$
\begin{aligned}
\mathbb{E}[A] &= N\left((N-1)(N-2)\mathbb{E}[(w_t^{(1)})^3]+2(N-1)\mathbb{E}[(w_t^{(1)})^2]+3\mathbb{E}[w_t^{(1)}]\right)\\
\mathbb{E}[B] &= (N-1)(N-2)\left((N-3)\mathbb{E}[(w_t^{(1)})^2(w_t^{(i)})^2]+\mathbb{E}[(w_t^{(1)})^2w_t^{(i)}]\right)\\
\mathbb{E}[C] &= N\left((N-1)(N-2)\mathbb{E}[(w_t^{(i)})^3]+2(N-1)\mathbb{E}[(w_t^{(i)})^2]+3\mathbb{E}[w_t^{(i)}]\right)\\
\mathbb{E}[D] &= (N-1)(N-2)\left((N-3)\mathbb{E}[(w_t^{(1)})^2(w_t^{(i)})^2]+\mathbb{E}[w_t^{(1)}(w_t^{(i)})^2]\right)\\
\mathbb{E}[E] &= N(N-1)(N-2)\left((N-3)\mathbb{E}(w_t^{(i)})^2(w_t^{(j)})^2]+\mathbb{E}[(w_t^{(i)})^2w_t^{(j)}]\right)
\end{aligned}
$$

We find therefore that

$$
\begin{aligned}
\mathbb{E}[D_N(t)] &= \frac{1}{N}\sum_{i=1}^{N}\left((N-2)\mathbb{E}[(w_t^{(i)})^3] + 2\mathbb{E}[(w_t^{(i)})^2] + \frac{3}{N-1}\mathbb{E}[w_t^{(i)}]\right) \\
&\quad + \frac{N-2}{N^2}\sum_{i=1}^{N}\sum_{j\neq i}\left((N-3)\mathbb{E}[(w_t^{(i)})^2(w_t^{(j)})^2] + \mathbb{E}[(w_t^{(i)})^2 w_t^{(j)}]\right)
\end{aligned}
$$

Let us calculate the expectations of A–E in the conditional case:

$$
\begin{aligned}
\mathbb{E}[\tilde{A}] &= (N-1)\left((N-2)(N-3)\mathbb{E}[(w_t^{(1)})^3] + 5(N-2)\mathbb{E}[(w_t^{(1)})^2] + 4\mathbb{E}[w_t^{(1)}]\right) \\
\mathbb{E}[\tilde{B}] &= \frac{1}{N}(N-1)(N-2)\left((N-3)(N-4)\mathbb{E}[(w_t^{(1)})^2(w_t^{(i)})^2] + (N-3)\mathbb{E}[(w_t^{(1)})^2 w_t^{(i)}] + 2(N-3)\mathbb{E}[w_t^{(1)}(w_t^{(i)})^2] + 2\mathbb{E}[w_t^{(1)}w_t^{(i)}]\right) \\
\mathbb{E}[\tilde{C}] &= (N-1)\left((N-2)(N-3)\mathbb{E}[(w_t^{(i)})^3] + 2(N-2)\mathbb{E}[(w_t^{(i)})^2] + 3\mathbb{E}[w_t^{(i)}]\right) \\
\mathbb{E}[\tilde{D}] &= \frac{(N-1)(N-2)}{N}\left((N-3)(N-4)\mathbb{E}[(w_t^{(1)})^2(w_t^{(i)})^2] + 3(N-3)\mathbb{E}[w_t^{(1)}(w_t^{(i)})^2] + \mathbb{E}[(w_t^{(i)})^2]\right) \\
\mathbb{E}[\tilde{E}] &= (N-1)(N-2)(N-3)\left((N-4)\mathbb{E}[(w_t^{(i)})^2(w_t^{(j)})^2] + \mathbb{E}[(w_t^{(i)})^2 w_t^{(j)}]\right)
\end{aligned}
$$

We find therefore that

$$
\begin{aligned}
\mathbb{E}[\tilde{D}_N(t)] &= \frac{N-2}{N^2}\sum_{i=1}^{N}\left((N-3)\mathbb{E}[(w_t^{(i)})^3] + \left(2+\frac{1}{N}\right)\mathbb{E}[(w_t^{(i)})^2] + \frac{3}{N-2}\mathbb{E}[w_t^{(i)}]\right) \\
&\quad + \frac{(N-2)(N-3)}{N^3}\sum_{i=1}^{N}\sum_{j\neq i}\left((N-4)\mathbb{E}[(w_t^{(i)})^2(w_t^{(j)})^2] + \mathbb{E}[(w_t^{(i)})^2 w_t^{(j)}]\right) \\
&\quad + \frac{1}{N^2}\mathbb{E}[w_t^{(1)}] + \left(3-\frac{1}{N}\right)\frac{(N-2)}{N^2}\mathbb{E}[(w_t^{(1)})^2] + \frac{2(N-2)}{N^3}\sum_{i=2}^{N}\mathbb{E}[w_t^{(1)}w_t^{(i)}] + \frac{4(N-2)(N-3)}{N^3}\sum_{i=2}^{N}\mathbb{E}[w_t^{(1)}(w_t^{(i)})^2]
\end{aligned}
$$

Using repeatedly that $(N-k)/N \leq 1)$, we find the upper bound in terms of $D_N(t)$:

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

For all but the last term it is enough that the weights are bounded in $[0,1]$ for them to vanish as $N\to\infty$. We can ensure the last term also vanishes by applying the strong but standard assumptions of KJJS Lemma 3. We then have in the limit as $N\to\infty$

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

We also need an upper bound on the squared coalescence rate...

$$
\begin{aligned}
\mathbb{E}[c_N(t)^2] &= \frac{1}{(N)_2^2}\mathbb{E}\left[\left(\sum_{i=1}^{N}(v_t^{(i)})_2\right)^2\right] \\
&= \frac{1}{(N)_2^2}\left\{\sum_{i=1}^{N}\mathbb{E}[(v_t^{(i)})_2^2] + \sum_{i=1}^{N}\sum_{j\neq i}\mathbb{E}[(v_t^{(i)})_2(v_t^{(j)})_2]\right\} \\
&= \frac{1}{(N)_2^2}\left\{(N)_4\sum_{i=1}^{N}\mathbb{E}[(w_t^{(i)})^4] + 4(N)_3\sum_{i=1}^{N}\mathbb{E}[(w_t^{(i)})^3] + 11(N)_2\sum_{i=1}^{N}\mathbb{E}[(w_t^{(i)})^2] - 3N\sum_{i=1}^{N}\mathbb{E}[w_t^{(i)}]\right\} \\
&\quad + \frac{1}{(N)_2^2}(N)_4\sum_{i=1}^{N}\sum_{j\neq i}\mathbb{E}[(w_t^{(i)})^2(w_t^{(j)})^2]
\end{aligned}
$$

In the conditional case we have...

$$\begin{aligned}
\mathbb{E}[\tilde{c}_N(t)^2] &= \frac{1}{(N)_2^2}\left\{\sum_{i=1}^N \mathbb{E}[(\tilde{v}_t^{(i)})_2^2] + \sum_{i=1}^N \sum_{j\neq i} \mathbb{E}[(\tilde{v}_t^{(i)})_2(\tilde{v}_t^{(j)})_2]\right\} \\
&= \frac{1}{(N)_2^2}\left\{\sum_{i=2}^N \mathbb{E}[(\tilde{v}_t^{(i)})_2^2] + \mathbb{E}[(\tilde{v}_t^{(1)})_2^2] + \sum_{i=2}^N \sum_{1\neq j\neq i} \mathbb{E}[(\tilde{v}_t^{(i)})_2(\tilde{v}_t^{(j)})_2] + 2\sum_{i=1}^N \mathbb{E}[(\tilde{v}_t^{(1)})2(\tilde{v}_t^{(i)})2]\right\} \\
&= \frac{1}{(N)_2^2}\left\{(N-1)_4\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^4] + 4(N-1)_3\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^3] + 11(N-1)_2\sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2] - 3(N-1)\sum_{i=1}^N \mathbb{E}[w_t^{(i)}]\right\} \\
&\quad + \frac{1}{(N)_2^2}\left\{(N-1)_4\sum_{i=1}^N \sum_{j\neq i} \mathbb{E}[(w_t^{(i)})^2(w_t^{(j)})^2] + 4(N-1)_3\mathbb{E}[(w_t^{(1)})^3] + 13(N-1)_2\mathbb{E}[(w_t^{(1)})^2] + 17(N-1)\mathbb{E}[w_t^{(1)}]\right\} \\
&\quad + \frac{1}{(N)_2^2}\left\{4(N-1)_3\sum_{i=2}^N \mathbb{E}[w_t^{(1)}(w_t^{(i)})^2] + 4(N-1)_2\sum_{i=2}^N \mathbb{E}[w_t^{(1)}w_t^{(i)}]\right\} \\
&\leq \mathbb{E}[c_N(t)^2] + \frac{1}{(N)_2^2}\left\{4(N-1)_3\mathbb{E}[(w_t^{(1)})^3] + 13(N-1)_2\mathbb{E}[(w_t^{(1)})^2] + 17(N-1)\mathbb{E}[w_t^{(1)}]\right\} \\
&\quad + \frac{1}{(N)_2^2}\left\{4(N-1)_3\sum_{i=2}^N \mathbb{E}[w_t^{(1)}(w_t^{(i)})^2] + 4(N-1)_2\sum_{i=2}^N \mathbb{E}[w_t^{(1)}w_t^{(i)}]\right\} \\
&= \mathbb{E}[c_N(t)^2] + O(N^{-3})
\end{aligned}$$

Proof of Lemma 3

We still assume the conditions (18) and (19) from KJJS. The conditional independence structure of the process (with time labelled backwards) gives us that, for any integrable function f,

$$\mathbb{E}[f(\mathbf{a}_t) \mid \mathcal{F}_{t-1}] = \mathbb{E}[\mathbb{E}[f(\mathbf{a}_t) \mid \mathbf{a}_{t+1}, \mathbf{X}_{t+1}, \mathbf{X}_t, \mathbf{X}_{t-1}, \mathbf{w}_{t-1}] \mid \mathcal{F}_{t-1}]$$

as in KJJS. In the conditional case with multinomial resampling, we have

$$\mathbb{P}(\mathbf{a}_t = \mathbf{a} \mid \mathbf{a}_{t+1}, \mathbf{X}_{t+1}, \mathbf{X}_t, \mathbf{X}_{t-1}, \mathbf{w}_{t-1}) \propto \mathbb{I}\{a_1 = 1\} \prod_{i=2}^N g_t(X_{t+1}^{(a_{t+1}^{a_i})}, X_t^{(a_i)}) q_{t-1}(X_t^{(a_i)}, X_{t-1}^{(i)})$$

That is,

$$\begin{aligned}
&a_t^{(1)} \mid \mathbf{a}_{t+1}, \mathbf{X}_{t+1}, \mathbf{X}_t, \mathbf{X}_{t-1}, \mathbf{w}_{t-1} = 1 \\
&a_t^{(i)} \mid \mathbf{a}_{t+1}, \mathbf{X}_{t+1}, \mathbf{X}_t, \mathbf{X}_{t-1}, \mathbf{w}_{t-1} \sim \text{Categorical}\left(g_t(X_{t+1}^{(a_{t+1}^{a_i})}, X_t^{(1)}) q_{t-1}(X_t^{(1)}, X_{t-1}^{(i)}), \dots, g_t(X_{t+1}^{(a_{t+1}^{a_i})}, X_t^{(N)}) q_{t-1}(X_t^{(N)}, X_{t-1}^{(i)})\right) \\
&\qquad \text{for } i = 2, \dots, N
\end{aligned}$$

By formulating the definition of I-increasing in terms of the modified (conditional) ancestral process, we still have that $f_i(\mathbf{a}_t) := (v_t^{(i)})_2$ is $\{i\}$-increasing for all i, but we need to modify the consequent result.

To get a result of the form $\mathbb{E}[f(\mathbf{a}_t) \mid \mathbf{a}_{t+1}, \mathbf{X}_{t+1}, \mathbf{X}_t, \mathbf{X}_{t-1}, \mathbf{w}_{t-1}] \leq \mathbb{E}[f(\tilde{\mathbf{a}}_t)]$, we need to modify the distribution of $\tilde{\mathbf{a}}_t$. We require that the probability of assigning to a parent i is higher under $\tilde{\mathbf{a}}_t$ than under $\mathbf{a}_t$ if and only if $i \in I$. So, by the same balls-in-bins argument of KJJS, the following distribution will work (using a and ε from the bounds (18) and (19)):

$$\begin{aligned}
\tilde{a}_t^{(1)} &= 1 \\
\tilde{a}_t^{(i)} &= \text{Categorical}\left(\left(\frac{a}{\varepsilon}\right)^{\mathbb{I}\{1\in I\}-\mathbb{I}\{1\notin I\}}, \dots, \left(\frac{a}{\varepsilon}\right)^{\mathbb{I}\{N\in I\}-\mathbb{I}\{N\notin I\}}\right), \qquad i = 2, \dots, N
\end{aligned}$$

Can't write that as a multinomial offspring distribution, which is a bit troublesome... Anyway, carrying this through we have

$$
\begin{aligned}
\mathbb{E}[\tilde{c}_N(t) \mid \mathcal{F}_{t-1}] &= \frac{1}{(N)_2} \sum_{i=1}^{N} \mathbb{E}[(v_t^{(i)})_2 \mid \mathcal{F}_{t-1}] \\
&=: \frac{1}{(N)_2} \sum_{i=1}^{N} \mathbb{E}[f_i(\mathbf{a}_t) \mid \mathcal{F}_{t-1}] \\
&\leq \frac{1}{(N)_2} \sum_{i=1}^{N} \mathbb{E}[f_i(\tilde{\mathbf{a}}_t)] \\
&= \frac{1}{(N)_2} \mathbb{E}[f_i(\tilde{a}_t^{(1)})] + \frac{1}{(N)_2} \sum_{i=2}^{N} \mathbb{E}[f_i(\tilde{a}_t^{(i)})] \\
&= \ldots
\end{aligned}
$$

Actually, we can probably just plug in the results from (2) and such above:

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

And for the lower bound:

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

Now for the bound on $D_N(t)$:

$$
\begin{aligned}
\mathbb{E}[\tilde{D}_N(t) \mid \mathcal{F}_{t-1}] &\leq \mathbb{E}[D_N(t) \mid \mathcal{F}_{t-1}] + O(N^{-3}) \\
&\leq \frac{C}{N} \mathbb{E}[c_N(t) \mid \mathcal{F}_{t-1}] + O(N^{-3}) \\
&\sim \frac{C}{N} \mathbb{E}[\tilde{c}_N(t) \mid \mathcal{F}_{t-1}]
\end{aligned}
$$

And the bound on $(c_N(t))^2$:

$$
\begin{aligned}
\mathbb{E}[\tilde{c}_N^2(t) \mid \mathcal{F}_{t-1}] &\leq \mathbb{E}[c_N^2(t) \mid \mathcal{F}_{t-1}] + O(N^{-3}) \\
&\leq \frac{C}{N} \mathbb{E}[c_N(t) \mid \mathcal{F}_{t-1}] + O(N^{-3}) \\
&\sim \frac{C}{N} \mathbb{E}[\tilde{c}_N(t) \mid \mathcal{F}_{t-1}]
\end{aligned}
$$