Residual resampling in SMC

Suzie Brown

3 April 2019

THIS DOCUMENT IS OBSOLETE. The calculations in section 1 are flawed because the definition of c_N used is not correct. Sections 2 and 3 have been rendered irrelevant since we simplified the theorem conditions to involve $\mathbb{E}_t[(\nu_i)_2]$ and $\mathbb{E}_t[(\nu_i)_3]$ only.

The offspring counts are sampled according to:

$$v_t^{(i)} = \lfloor N w_t^{(i)} \rfloor + X_i$$
$$X_i \sim \text{Multinomial}(N-k, (\bar{w}_t^{(1)}, \dots, \bar{w}_t^{(N)}))$$

where $k := \sum_{i=1}^N \lfloor N w_t^{(i)} \rfloor$ is the number of offspring assigned deterministically, and $\bar{w}_t^{(i)} := \frac{N w_t^{(i)} - \lfloor N w_t^{(i)} \rfloor}{N-k}$ are the residual weights. Let us also define the residuals $r_i := N w_t^{(i)} - \lfloor N w_t^{(i)} \rfloor$.

1 Coalescence rate

We shall calculate the expected coalescence rate in the case of residual resampling. The coalescence rate is defined in Koskela et al. (2018) as

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

Edit 27 May 2021: the above expression is not even correct; in our setting c_N is random and does not have the expectation in its definition.

The inner expectation comes out as

$$\begin{aligned}
\mathbb{E}[(v_t^{(i)})_2] &= \mathbb{E}[(v_t^{(i)})^2] - \mathbb{E}[v_t^{(i)}] \\
&= \lfloor N w_t^{(i)} \rfloor^2 + 2 \lfloor N w_t^{(i)} \rfloor r_i + r_i \left(1 - \frac{r_i}{N-k} + r_i \right) - N w_t^{(i)} \\
&= \lfloor N w_t^{(i)} \rfloor^2 - \lfloor N w_t^{(i)} \rfloor + 2 \lfloor N w_t^{(i)} \rfloor r_i + r_i^2 \left(1 - \frac{1}{N-k} \right) \\
&= (N w_t^{(i)})^2 - \lfloor N w_t^{(i)} \rfloor - \frac{r_i^2}{N-k}
\end{aligned}$$

so we get

$$\begin{aligned}
\mathbb{E}[c_N^r(t) | \mathcal{F}_{t-1}] &= \frac{1}{(N)_2} \mathbb{E}\left[\sum_{i=1}^N \mathbb{E}[(v_t^{(i)})_2] | \mathcal{F}_{t-1} \right] \\
&= \frac{N}{N-1} \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2 | \mathcal{F}_{t-1}] - \frac{1}{(N)_2} \sum_{i=1}^N \mathbb{E}\left[\frac{r_i^2}{N-k} | \mathcal{F}_{t-1} \right] - \frac{1}{(N)_2} \mathbb{E}[k | \mathcal{F}_{t-1}] \\
&= \mathbb{E}[c_N^m(t) | \mathcal{F}_{t-1}] \left(1 + \frac{1}{N-1} \right) - \frac{1}{(N)_2} \mathbb{E}\left[\frac{\sum_{i=1}^N (N w_t^{(i)} - \lfloor N w_t^{(i)} \rfloor)^2}{\sum_{j=1}^N (N w_t^{(j)} - \lfloor N w_t^{(j)} \rfloor)} | \mathcal{F}_{t-1} \right] \\
&\quad - \frac{1}{(N)_2} \mathbb{E}\left[\sum_{i=1}^N \lfloor N w_t^{(i)} \rfloor | \mathcal{F}_{t-1} \right]
\end{aligned}$$

Sanity check:
When the weights are all equal, $w_t^{(i)} \equiv 1/N$, we should have $\mathbb{E}[c_N^r(t)|\mathcal{F}_{t-1}] = 0$ since each particle will have exactly one offspring so it is impossible for any lineages to coalesce. In this case we have $\mathbb{E}[c_N^m(t)|\mathcal{F}_{t-1}] = \sum_{i=1}^N \mathbb{E}[(w_t^{(i)})^2|\mathcal{F}_{t-1}] = 1/N$ for multinomial resampling. We also have that $Nw_t^{(i)} \equiv \lfloor Nw_t^{(i)} \rfloor \equiv 1$ and hence $r_i = 0$ and $k = N$. Thus the RHS comes out as

$$\frac{1}{N}\frac{N}{N-1} - 0 - \frac{1}{(N)_2}N = \frac{1}{N-1} - \frac{1}{N-1} = 0$$

as expected.

2 Squared coalescence rate and other awful expressions

First let's write down various moments that will be needed (where $i \neq j$).

$$\begin{aligned}
\mathbb{E}[(v_t^{(i)})^2|\mathcal{F}_{t-1}] &= \lfloor Nw_t^{(i)} \rfloor^2 + 2\lfloor Nw_t^{(i)} \rfloor Nw_t^{(i)} + (N)_2(w_t^{(i)})^2 \\
\mathbb{E}[(v_t^{(i)})^3|\mathcal{F}_{t-1}] &= \lfloor Nw_t^{(i)} \rfloor^3 + 3\lfloor Nw_t^{(i)} \rfloor^2 Nw_t^{(i)} + 3\lfloor Nw_t^{(i)} \rfloor (N)_2(w_t^{(i)})^2 + (N)_3(w_t^{(i)})^3 \\
\mathbb{E}[(v_t^{(i)})^4|\mathcal{F}_{t-1}] &= \lfloor Nw_t^{(i)} \rfloor^4 + 4\lfloor Nw_t^{(i)} \rfloor^3 Nw_t^{(i)} + 6\lfloor Nw_t^{(i)} \rfloor^2 (N)_2(w_t^{(i)})^2 + 4\lfloor Nw_t^{(i)} \rfloor (N)_3(w_t^{(i)})^3 + (N)_4(w_t^{(i)})^4 \\
\mathbb{E}[v_t^{(i)}v_t^{(j)}|\mathcal{F}_{t-1}] &= \lfloor Nw_t^{(i)} \rfloor \lfloor Nw_t^{(j)} \rfloor + \lfloor Nw_t^{(i)} \rfloor Nw_t^{(j)} + \lfloor Nw_t^{(j)} \rfloor Nw_t^{(i)} + (N)_2 w_t^{(i)} w_t^{(j)} \\
\mathbb{E}[(v_t^{(i)})^2 v_t^{(j)}|\mathcal{F}_{t-1}] &= \lfloor Nw_t^{(i)} \rfloor^2 \lfloor Nw_t^{(j)} \rfloor + \lfloor Nw_t^{(i)} \rfloor^2 Nw_t^{(j)} + 2\lfloor Nw_t^{(i)} \rfloor \lfloor Nw_t^{(j)} \rfloor Nw_t^{(i)} + 2\lfloor Nw_t^{(i)} \rfloor (N)_2 w_t^{(i)} w_t^{(j)} \\
&\quad + \lfloor Nw_t^{(j)} \rfloor (N)_2 (w_t^{(i)})^2 + (N)_3 (w_t^{(i)})^2 w_t^{(j)} \\
\mathbb{E}[(v_t^{(i)})^2 (v_t^{(j)})^2|\mathcal{F}_{t-1}] &= \lfloor Nw_t^{(i)} \rfloor^2 \lfloor Nw_t^{(j)} \rfloor^2 + 2\lfloor Nw_t^{(i)} \rfloor^2 \lfloor Nw_t^{(j)} \rfloor Nw_t^{(j)} + 2\lfloor Nw_t^{(i)} \rfloor \lfloor Nw_t^{(j)} \rfloor^2 Nw_t^{(i)} \\
&\quad + \lfloor Nw_t^{(i)} \rfloor^2 (N)_2 (w_t^{(j)})^2 + 4\lfloor Nw_t^{(i)} \rfloor \lfloor Nw_t^{(j)} \rfloor (N)_2 w_t^{(i)} w_t^{(j)} + \lfloor Nw_t^{(j)} \rfloor^2 (N)_2 (w_t^{(i)})^2 \\
&\quad + 2\lfloor Nw_t^{(i)} \rfloor (N)_3 w_t^{(i)} (w_t^{(j)})^2 + 2\lfloor Nw_t^{(j)} \rfloor (N)_3 (w_t^{(i)})^2 w_t^{(j)} + (N)_4 (w_t^{(i)})^2 (w_t^{(j)})^2
\end{aligned}$$

For the squared coalescence rate, expanding the falling factorials appropriately, we get

$$\begin{aligned}
\mathbb{E}[(c_N^r(t))^2|\mathcal{F}_{t-1}] &= \frac{1}{(N)_2^2}\left(\sum_{i=1}^N \mathbb{E}[(v_t^{(i)})_2^2|\mathcal{F}_{t-1}] + \sum_{i=1}^N \sum_{j\neq i} \mathbb{E}[(v_t^{(i)})_2 (v_t^{(j)})_2|\mathcal{F}_{t-1}]\right) \\
&= \frac{1}{(N)_2^2}\sum_{i=1}^N \left(\mathbb{E}[(v_t^{(i)})^4|\mathcal{F}_{t-1}] - 2\mathbb{E}[(v_t^{(i)})^3|\mathcal{F}_{t-1}] + \mathbb{E}[(v_t^{(i)})^2|\mathcal{F}_{t-1}]\right) \\
&\quad + \frac{1}{(N)_2^2}\sum_{i=1}^N \sum_{j\neq i}\left(\mathbb{E}[(v_t^{(i)})^2(v_t^{(j)})^2|\mathcal{F}_{t-1}] - \mathbb{E}[(v_t^{(i)})^2 v_t^{(j)}|\mathcal{F}_{t-1}] - \mathbb{E}[v_t^{(i)}(v_t^{(j)})^2|\mathcal{F}_{t-1}] + \mathbb{E}[v_t^{(i)}v_t^{(j)}|\mathcal{F}_{t-1}]\right)
\end{aligned}$$

Then we can try plugging in the expressions derived above and find that none of the terms cancel. Maybe if we're clever we can factorise it or something.

3 Mega-merger rate

Now the rate of super0binary mergers...

$$\begin{aligned}
\mathbb{E}[D_N(t)|\mathcal{F}_{t-1}] &= \frac{1}{N(N)_2}\sum_{i=1}^N \left(\mathbb{E}[(v_t^{(i)})^3|\mathcal{F}_{t-1}] - \mathbb{E}[(v_t^{(i)})^2|\mathcal{F}_{t-1}]\right) \\
&\quad + \frac{1}{N(N)_2}\sum_{i=1}^N \sum_{j\neq i}\left(\mathbb{E}[(v_t^{(i)})^2(v_t^{(j)})^2|\mathcal{F}_{t-1}] - \mathbb{E}[v_t^{(i)}(v_t^{(j)})^2|\mathcal{F}_{t-1}]\right)
\end{aligned}$$

References

Douc, R. and Cappé, O. (2005), Comparison of resampling schemes for particle filtering, *in* 'Image and Signal Processing and Analysis, 2005. ISPA 2005. Proceedings of the 4th International Symposium on', IEEE, pp. 64–69.

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* .

Liu, J. S. and Chen, R. (1998), 'Sequential monte carlo methods for dynamic systems', *Journal of the American statistical association* **93**(443), 1032–1044.