Possible future research topics

Suzie Brown

1 July 2021

Extending the theoretical results

- Relax condition that potentials are bounded below (e.g. see Cappé, Moulines, Rydén "Inference in HMMs", Chapter 9)
- Show that the condition of BJJK Theorem 1 is necessary as well as sufficient (e.g. see Möhle & Sagitov 1998, 2001, 2003?)
- What is the effect of adaptive resampling?
- Corollaries also for residual-multinomial and residual-star resampling (although it might be a negative result in the case of residual-star).
- Rates of convergence (modifying more of Möhle's work).
- Could badly-behaved potentials produce a non-Kingman Lambda-coalescent? For example, use for potentials some heavy-tailed fitnesses like in Schweinsberg model. (This is probably not interesting from an SMC point of view.)
- A way to estimate coalescent rates / time scale a priori, for some specific tractable class of models, say.
- Do our results apply to cloning models, or other continuous-time models?
- See if Jacob & Rubenthaler's ("path storage in the PF") brute-force technique can be adapted for use in Conditional SMC, to UB tree height
- Does CSMC with an "unlikely" immortal line converge to a structured coalescent?

Resampling

- Derive expressions for the one-step Monte Carlo variance with SSP resampling and residual-systematic resampling, (and systematic resampling?). Is there another ordering result there?
- Compare residual-systematic vs. systematic resampling: are they equivalent? What if the weights are sorted? Maybe start by coding up an exploratory experiment.
- Explore more generally the effect of pre-sorting the weights. How does this link with results of Gerber Chopin Whiteley (where they sort by particle position)?
- Compare theoretical computation/storage costs and parallelisability between the different resampling schemes.
- Can the comparison of time-scales for different schemes be wrangled into/ related to a direct comparison of the variances? Is this even useful?
- (See marked notebook 3 page circa Jun/Jul 2020 for some more thoughts/ideas.)
- The difference Δ_i, as defined in `phd/latex/randomised_rounding/randomised_rounding.pdf`, seems to tend to a quadratic in the weight as $N \to \infty$. Prove it?

- (Just for fun:) in stochastic rounding, how many possible ways are there to assign the offspring counts? Consider that each of the N counts takes one of two possible values, but this will overcount a lot because we are also constrained by offspring counts summing to N. It's $\binom{N}{R}$ isn't it? Where R as in residual resampling. Give each parent its minimum number of offspring, and you'll be left with R unassigned offspring that have to be given to R distinct parents among N. $\binom{N}{R}$ is maximised if $R \simeq N/2$, in which case $\binom{N}{R} \simeq$??.
- (Conjecture:) pre-sorting of weights reduces resampling variance, but increases the coalescence rate. Intuition: when weights are sorted, small weights that sum to less than 1/N are grouped together so only one of them can have a child. (It may be that this effect is entirely cancelled by the reduction in variance elsewhere.)

Simulation experiments

- Now we have weak convergence, redo similar experiments to those at end of KJJS, but without having to fudge it.
- Code up all the different sampling schemes (and sorted/unsorted variants etc.) and come up with some useful (if only illustrative) experiments comparing them. (I started doing this by plotting $\tau_N(t)$ against t, and got some nice results.)
- Decide which functions would be most interesting to illustrate the performance of resampling schemes, say using ternary diagrams. (See marked notebook 3 page at 8/7/20.)
- Explore pre-asymptotic behaviour of CSMC for example. I did some work on this previously.
- Make edits to my ternary plot "library" (see notebook 3 marked page at 8/7/20).