

Major changes

Structure

- Took likelihood calculations out of objective section and moved them into research chapter (sec. 2.1.4 "Justification for approach").
- Instead of displaying many posterior distributions from synthetic data in the appendices, only show two in the main text: one "good" (ie. low error rate) and one "bad".
- Structure of chapter 2 has been rearranged: one section about method, one about synthetic data experiments, one about real data, and then the discussion.

Analysis

- It turns out that only networks with $\alpha = 1$ have a power law distribution (p. 12). However power laws still capture the salient features of the degree distributions (fig. A26) so I continue to report them.
- Throughout, maximum a posteriori estimates were replaced with posterior means, which are more conventional and give better results.
- Removed ad-hoc statistical tests from results on grid search and ABC. For grid search, only error distributions were reported numerically. Qualitative statements were made about kernel score distributions. For ABC, effects of each parameter on error rates was analysed with a p-value corrected GLM (p. 52-53).
- Two extra analyses to address possible sources of ABC error (a reviewer asked for this) described on p. 46, results on p. 62-63.
- Show some 2-dimensional marginals (p. 55, 57).
- The real data analysis has been significantly expanded to include six more datasets.
- Analysed three datasets with both env and gag genes and compare results (fig. 2.17 on p. 65).
- Found several published estimates of power law exponent gamma from literature (table 2.9 on p. 74).

Writing

- Added some introduction to SMC section to clarify what the relationship to our problem is, ie. we are trying to make a particle approximation to the posterior.
- Same thing for ABC section (p. 21-23).
- Added summary of chapter 1 (p. 26-27).
- Added introductory paragraph to chapter 2 (p. 28).
- Expanded on why ABC is needed in this setting (p. 33-36), why we should study the BA model (p. 36-37).
- Wrote high-level summary and motivation of all synthetic data experiments (p. 38-39, 40, 44).

- Results section for real data experiments has been entirely rewritten.
- Discussion of real data results is also entirely new. I relate our results to a couple of papers looking at preferential attachment (PA) networks, one which found sub-linear PA for real world data (which we also found), and one which claimed IDU networks were more PA-ish than sexual networks (which we also found too).

Minor changes

- Added text to abstract (UBC limit is 350 words).
- Better definitions for "production rule", "decay factor", and "ladderize" (p. 8-9).
- More conventional definition of importance sampling (p. 16).
- Added subsection under SMC for actual SMC (SIS with resampling), plus discussion of particle degeneracy (p. 18).
- Throughout, multi-step mathematical derivations have had justifications added beside each step (eg. p. 20).
- Added a paragraph about the importance of the distance function for ABC (p. 23-24).
- Changed " α " from Del Moral paper to " α_{ESS} " to avoid confusion (p. 32).
- Fig. 2.7 is now in color.
- Added notches for 50% highest density intervals (similar to IQR) in fig. 2.8.
- Reported relative errors for I and N in grid search.
- Refer to datasets as eg. "IDU/Romania data" instead of "Niculescu et al. data" for easier interpretation.
- Proper scaling and legends for fig. 2.18 (p. 69).