

Introduction

- Delayed-acceptance (Christen and Fox (2005), DA) can reduce computational effort for expensive likelihoods within Metropolis-Hastings algorithm (MH).
- A surrogate approximates the expensive likelihood.
- Acceptance of proposals are delayed until they passed the surrogate likelihood.
- Delayed-acceptance reduces the cost of MH whilst preserving the invariant distribution.

Objectives

- Incorporate DA into sequential Monte Carlo.
- Tune surrogate likelihood using particles.
- Control the computational cost of DA-SMC.

Sequential Monte Carlo

- Choose a schedule to connect a tractable starting distribution $p_0(\theta)$ to the posterior $p_T(\theta) = p(\theta | \mathbf{y})$, e.g.

$$p_t(\theta) = p_0(\theta)^{1-\gamma_t} p(\theta | \mathbf{y})^{\gamma_t}$$

$$0 = \gamma_0 < \gamma_1 < \dots < \gamma_T = 1$$

- Draw a sample from starting distribution $p_0(\theta)$ particles, then iterate reweight, resample, and refresh steps as in Figure 1.

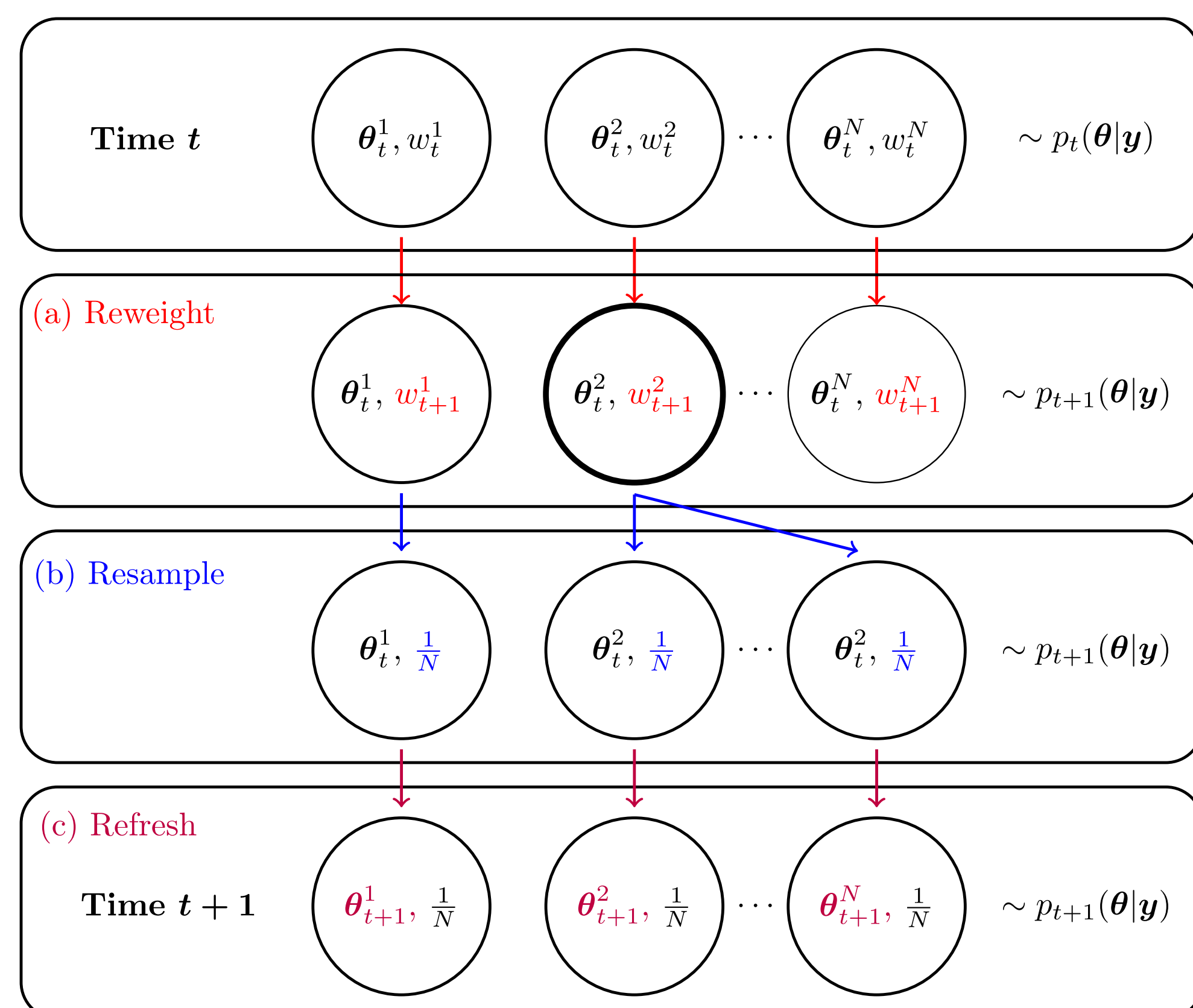


Figure 1: Visualisation of SMC algorithm: iteration $t \rightarrow t+1$

Delayed-acceptance

SMC can use MH to perform the refresh. For simplicity consider Metropolis algorithm (symmetric proposal).

Let $\tilde{p}_t(\theta_t | \mathbf{y})$ be a cheap approximation of the full posterior $p_t(\theta_t | \mathbf{y})$.

- From θ_{t-1} propose new $\theta' \sim k(\theta | \theta_{t-1})$.
- Calculate surrogate acceptance probability

$$\alpha^1 = \min \left(1, \frac{\tilde{p}_t(\theta' | \mathbf{y})}{\tilde{p}_t(\theta_{t-1} | \mathbf{y})} \right)$$

and provisionally accept with probability α^1 .

- If rejected: $\theta_t \leftarrow \theta_{t-1}$.

- If provisionally accepted, calculate

$$\alpha^2 = \min \left(1, \frac{\tilde{p}(\theta_{t-1} | \mathbf{y})}{\tilde{p}(\theta' | \mathbf{y})} \frac{p(\theta' | \mathbf{y})}{p(\theta_{t-1} | \mathbf{y})} \right)$$

and accept with probability α^2 .

- If accepted: $\theta_t \leftarrow \theta'$.
- If rejected: $\theta_t \leftarrow \theta_{t-1}$.

Surrogate tuning

- Correct for differences in surrogate and full likelihood with tuning using SMC particles.
- Record every location (and value) where the expensive likelihood has been evaluated (cache with a hash function).

At each SMC iteration we can minimise a distance (or discrepancy) to find ξ^* , the optimal transformation, with

$$\min_{\xi} \sum_{\theta \in H(L)} d[L(\mathbf{y} | \theta), \tilde{L}(\mathbf{y} | T_{\xi}(\theta))]$$

- $L(\mathbf{y} | \theta)$ and $\tilde{L}(\mathbf{y} | \theta)$ are full approximate likelihoods.
- T_{ξ} is a pre-specified transformation.

The expensive likelihood does not need to be evaluated since we use the cache $H(L)$. In fact, since $H(L)$ grows each iteration, we can use a subset of values.

Optimal computation time

DA MH is most efficient when big steps are taken with lots of rejections, but a few big accepted jumps. We characterise this as an optimisation problem.

$$\min_{k \in \mathbb{Z}^+, \phi \in \Phi} C(k, \phi) \quad \text{s.t.} \quad D(k, \phi) \geq d$$

- k is number of MCMC iterations.
- ϕ is tuning parameters of proposal kernel.
- $C(k, \phi)$ is cost of k MCMC iterations.
- $D(k, \phi)$ is a diversification condition, to ensure the particles are sufficiently mutated after resampling.

In DA-MH, the cost is

$$C(k, \phi) = k(\tilde{L}_C + \alpha^1(\phi)L_C)$$

A good diversification condition can be the expected squared jumping distance (ESJD) (Pasarica and Gelman 2010):

$$\text{median} \left\{ \sum_{i=1}^k J_i(\phi) \right\} > D$$

where J_i is the ESJD for particle i .

Approximate annealing

Another way to increase the speed is to anneal to the approximate posterior first. Instead of moving from prior to posterior, we move from prior to approximate posterior, then to full posterior.

$$\tilde{p}_t(\theta) = p_0(\theta)^{1-\gamma_t} \tilde{p}(\theta | \mathbf{y})^{\gamma_t}, \quad 0 = \gamma_0 < \dots < \gamma_S = 1$$

followed by (using $\tilde{p}_S(\theta | \mathbf{y})$ from first SMC run)

$$p_t(\theta) = \tilde{p}_S(\theta | \mathbf{y})^{1-\zeta_t} p(\theta | \mathbf{y})^{\zeta_t}, \quad 0 = \zeta_0 < \dots < \zeta_T = 1$$

Some care is needed to ensure the approximate posterior has heavier tails than the full posterior.

Simulation

Simple test of DA-SMC framework using 2000 particles for linear regression.

- Likelihood: Normal or t-distribution with artificial delay.
- Surrogate model: Biased normal.
- Proposal: Adaptive multivariate normal.
- Surrogate tuning: Least-squares on log L with T_{η} univariate shift and scale.
- Full to surrogate cost ratios: $L_C/\tilde{L}_C \in \{10^2, 10^3, 10^4, 10^5\}$

- Computation time optimisation: Approximation of ESJD and costs using pilot run over grid on proposal step-size.
- ESJD Model, to estimate k for diversification to be met;
 - Bootstrap using pilot run draws of ESJD, or
 - Assume ESJD follows Gamma distribution estimated from pilot run.

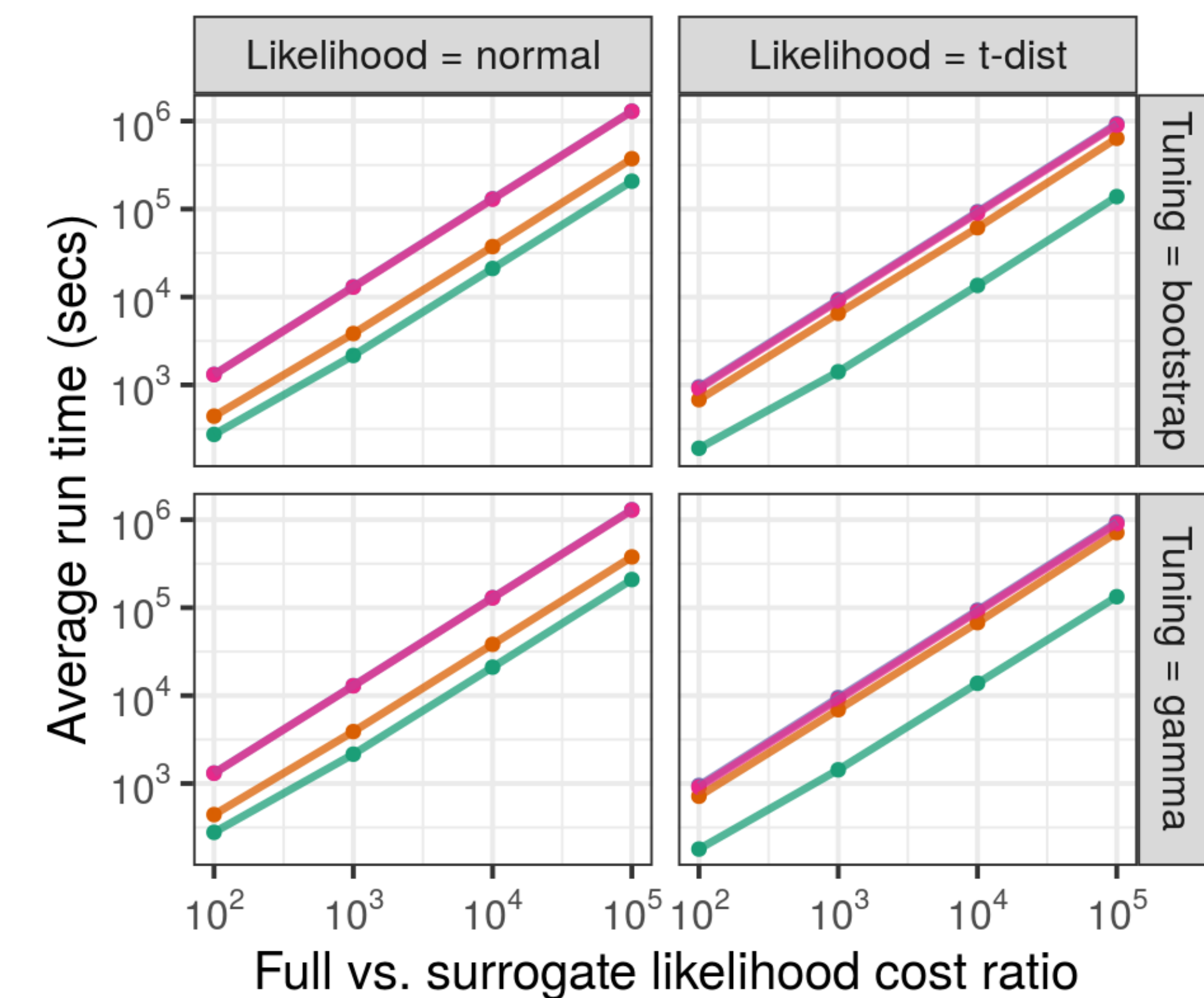


Figure 2: Average simulation computation times (secs)

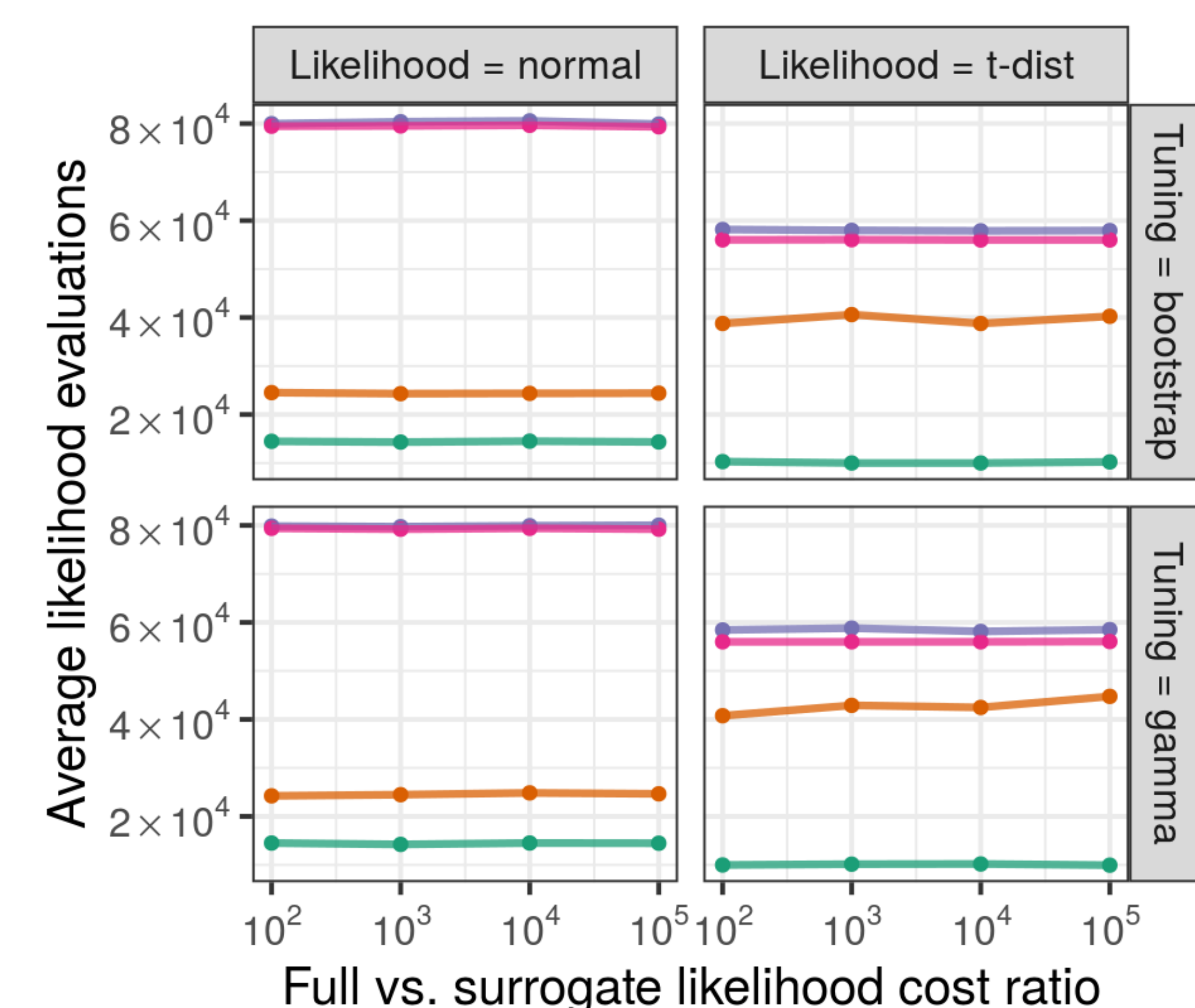


Figure 3: Average full likelihood evaluations

Conclusion

- Longest run time reduction 10 days to 1.5 days with approximate annealing and DA.
- Mean squared error was near-identical for all methods.
- Pilot median method works well for standard MH in SMC (Salomone et al. 2018).
- Approximations to optimal computation time problem for plain MH correspond to choices used in the literature.

References

Poster made with [posterdown](#) (Thorne 2019).

Christen, J Andr s, and Colin Fox. 2005. “Markov Chain Monte Carlo Using an Approximation.” *Journal of Computational and Graphical Statistics* 14 (4): 795–810.

Pasarica, Cristian, and Andrew Gelman. 2010. “Adaptively Scaling the Metropolis Algorithm Using Expected Squared Jumped Distance.” *Statistica Sinica* 20 (1): 343–64.

Salomone, Robert, Leah F South, Christopher C Drovandi, and Dirk P Kroese. 2018. “Unbiased and Consistent Nested Sampling via Sequential Monte Carlo.” *arXiv Preprint arXiv:1805.03924*. <https://arxiv.org/abs/1805.03924>.

Thorne, Brent. 2019. *Posterdown: Generate PDF Conference Posters Using R Markdown*. <https://CRAN.R-project.org/package=posterdown>.

