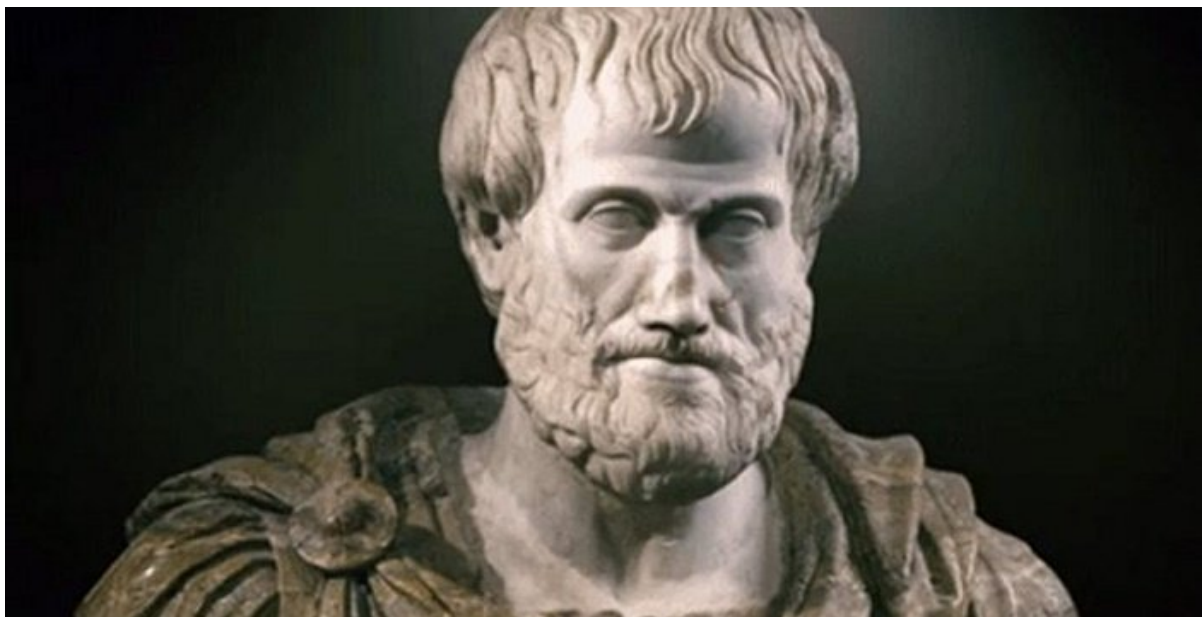# Markov chain Monte Carlo and Perfect simulation

45 minutes

Introduction {#sec-1}

Aristotle: "The more you know, the more you know you don't know."

History of MCMC {#sec-1}

Edward Teller (1908 – 2003). (Fermi once said that) Teller was the only monomaniac he knew who had several manias [@BrownMay-2004].

- The original Markov chain Monte Carlo method [@MetropolisRosenbluthRosenbluthTellerTeller-1953];
- Markov chain basics:
  - Transition probabilities, transition rates, unified view using exponential distribution;
  - Equilibrium, balance, detailed balance, reversibility;

- Aperiodicity, recurrence (and uniform and geometric recurrence);
  - Conditioning under detailed balance;
- MCMC: physicists almost always get there first!
- MCMC practicalities: Burn-in: what to do about it?
  - Theory tends to be much too pessimistic [example of @Zanella-2011 / @Zanella-2015b developing statistical methods for Anglo-Saxon history: *simplified* model converging in 10^5 steps, translating to about a week on a compute cluster, *versus* 10^9 steps in theory, translating to around 2 centuries …];
  - (a) One long run or (b) many short runs? (Option (b) assumes in effect that one can arrange for starts of short runs to be spread over the sample space — a task which in high-dimensional problems may be almost as hard as the problem one is trying to solve!)
  - Diagnostics? (Meta-theorem: for any diagnostic technique one can build a chain for which this technique is deceptive!)
  - Conclusion: effective MCMC does in fact require careful thought about appropriate length of run — the more one knows about the problem, the better! Can there ever be a better way?

Perfect Simulation {#sec-2}
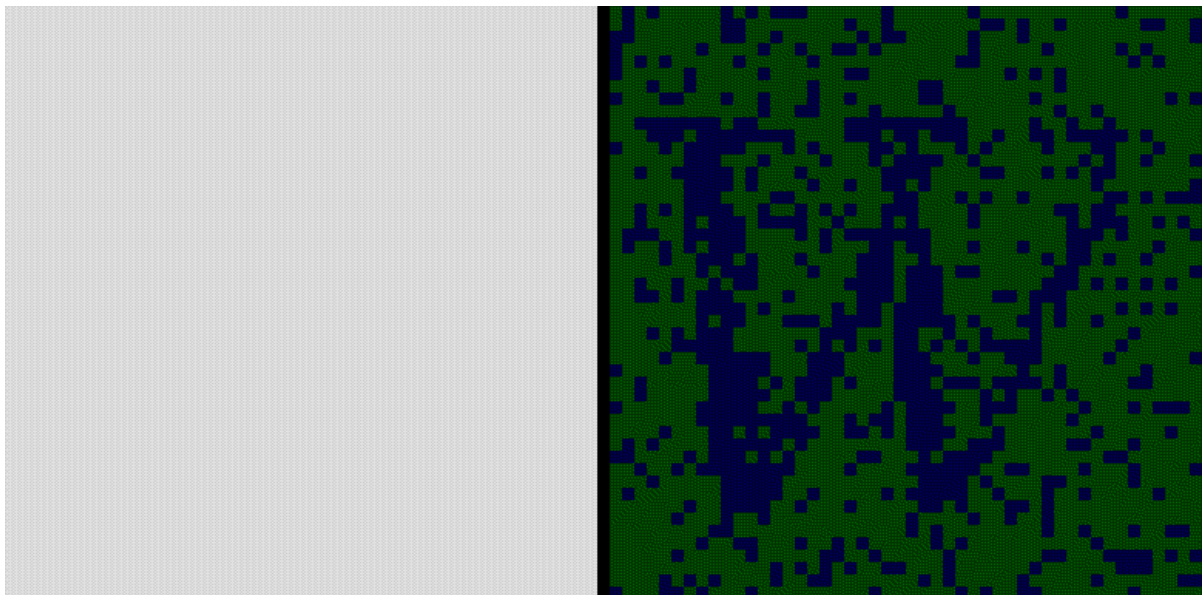
- The idea of exact simulation / Coupling from the Past (CFTP) / perfect simulation @ProppWilson-1996a (Persi Diaconis: "Like seeing the landscape of Mars for the first time");
- Ideas (of "*classic CFTP*"):
  - extending simulation *backwards* through time,
  - exploit monotonicity by coupling maximal and minimal processes,
  - seek coalescence;
- Details for *random-walk-CFTP*, which can be boosted as above to provide simple image reconstruction of an image using Ising model, @ProppWilson-1996a show how to vary a clever algorithm to get exact samples for **critical** Ising model (this is what impressed Diaconis);

- "Perfect simulation" [@Kendall-1998d]: because everyone knows it isn't going to be perfect, whereas people might imagine "exact simulation" somehow miraculously would defeat numerical approximation error :-).

An example and some theory {#sec-3}
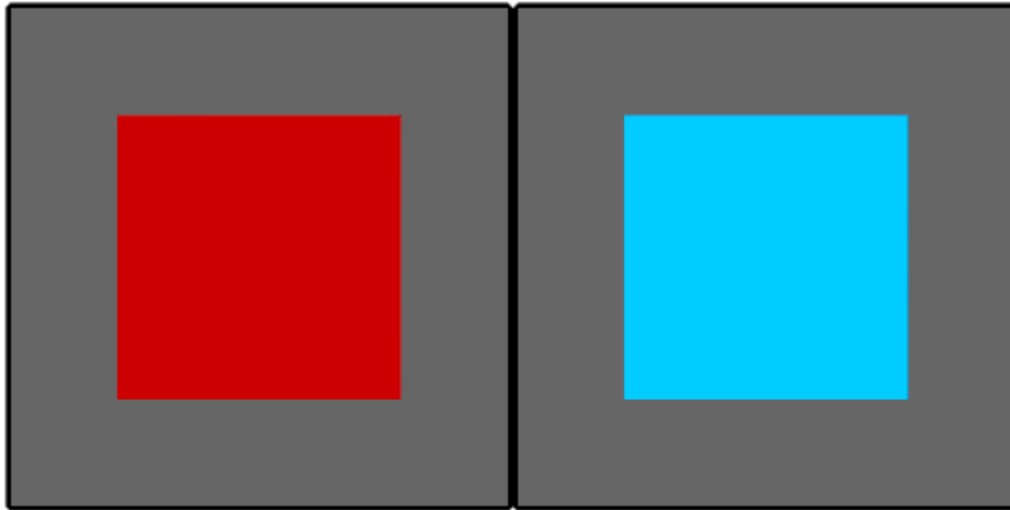
- An intensely visual example, which helps many people see intuitively what is going on here, is *DeadLeaves-CFTP* @KendallThoennes-1999 (technically, *Occluded CFTP*);

- What about cases where monotonicity fails? or there isn't a sensible "maximal" process? @Kendall-1998d:
    - cross-couple upper and lower envelope processes,
    - dominate by amenable "dominating process" (typically, time-reversible, can draw from equilibrium, can couple target processes to stay below dominating process);

- Theoretical limits:
    - *Classical CFTP* is *in principle* equivalent to uniform ergodicity [@FossTweedie-1997].
    - *Dominated CFTP* is *in principle* achievable if geometric ergodicity obtains[@Kendall-2004c].
    - It is possible to carry out Dominated CFTP in some **non**-geometrically ergodicity cases [@ConnorKendall-2007a; @ConnorKendall-2007b];

- We can use *Dominated CFTP* to carry out perfect simulation for stable point processes [@KendallMoeller-2000];

- For detailed exposition of much of this, see *eg* @Kendall-2005a, @Huber-2016. @Kendall-2014b introduces CFTP using R.

Queues and Epidemics {#sec-4}

PRACTISE SOCIAL DISTANCING

HELP STOP THE SPREAD!

2 Meters   2 Meters

https://covidposters.github.io/
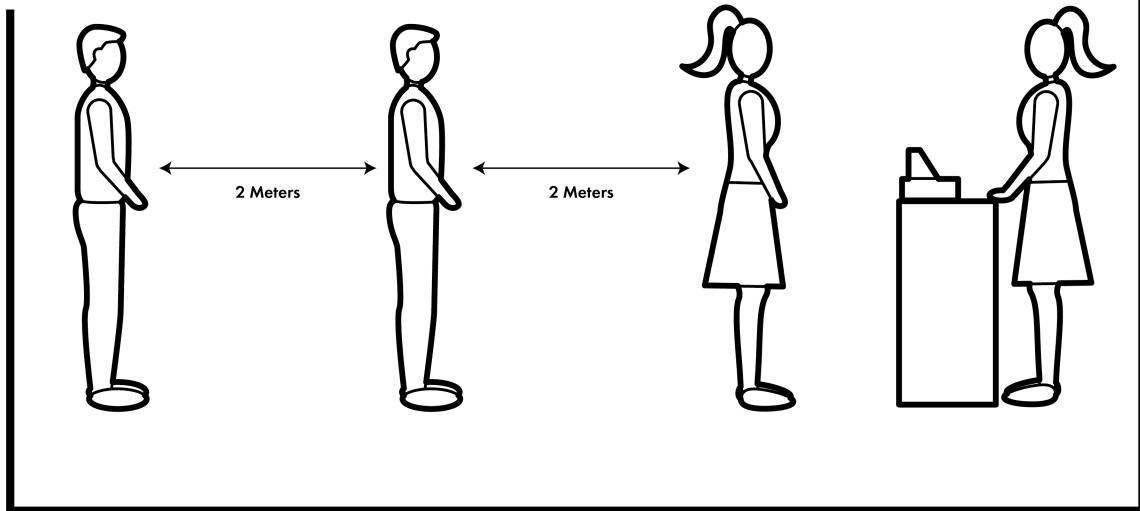
Queues

 The simplest queuing model (Poisson arrivals, exponential service times, single server) can be analyzed very throughly;

- Poisson arrivals are not unreasonable, but exponential service times are ludicrous. Fortunately we can analyze the case of general service time using the "embedded chain" (sample at instants of departure), so long as there is just one server;

- Multi-server case: computation of eg waiting-time distribution is theoretically out of reach so use simulation [and important ideas from @KieferWolfowitz-1955];

- @ConnorKendall-2014 extend @Sigman-2011, showing how Dominated CFTP can be applied to simulate (sub-critical!) queues perfectly (and this has now been generalized by others to the case of non-Poissonian inter-arrival times). Key idea: dominate the standard "First-come-first-served" (FCFS) discipline by "Random allocation" (RA), which allows us to factorize M/G/c to a product of c times M/G/1, then relate to a processor-sharing model which enables a time-reversibility argument. (Technical point: pathwise domination requires service times to be assigned in order of commencement of service!

Epidemics

- Even the simple case of deterministic S-I-R permits only *partial* closed-form solution;

- Suppose one can observe **only** the "removal times"? Can one do perfect simulation of the infection times? (Given basic parameters.)

- **YES**: work in progress by Connor and Kendall. Here is a GIF illustrating this for a real-life small-pox epidemic;
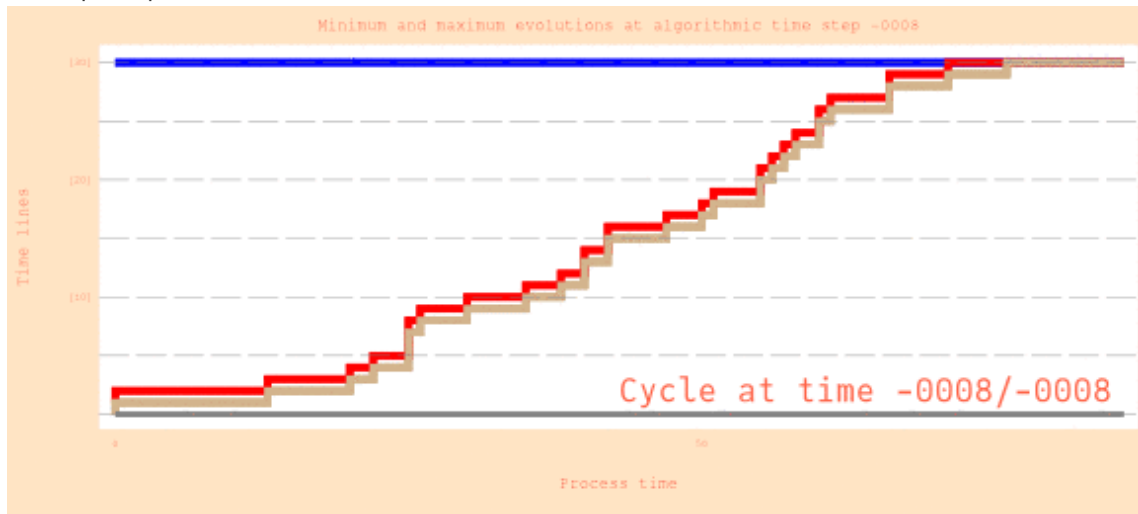
- (be clear about assumptions!)
- (indicate how the perfect simulation algorithm can be used as a high-dimensional integration device to enable simulation-based Bayesian inference!).

Conclusion {#sec-conclusion}

- You don't always have to put up with burn-in issues when doing MCMC;
- CFTP works even for significantly complex and relevant models of real-life phenomena;
- *Of course* really detailed models are still going to resist perfect simulation: but it will always be helpful to compare with a simpler model (using fewer parameters!).
- CFTP is clearly an important tool to be considered by the investigator seeking to do accurate and informative MCMC.

Figures

Aristotle	**http://brewminate.com/wp-content/uploads/2018/05/050218-23-Aristotle-Science-History-Philosophy.jpg**		
Edward Teller	https://upload.wikimedia.org/wikipedia/commons/b/b5/Edward_Teller_in_1958.jpg		
Perfect Ising	created by WSK		
Dead leaves	created by WSK		
Queues	https://covidposters.github.io/images/queues.png		
Epidemic	created by WSK		

Aristotle	http://brewminate.com/wp-content/uploads/2018/05/050218-23-Aristotle-Science-History-Philosophy.jpg		
	http://brewminate.com/wp-content/uploads/2018/05/050218-23-Aristotle-Science-History-Philosophy.jpg		