Perfect Epidemics

Applied Probability Seminar
Department of Statistics, University of Warwick

W S Kendall   S B Connor
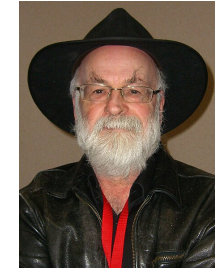
Warwick, York

24 January 2025

Introduction

"Maybe the only significant difference between a really smart simulation and a human being was the noise they made when you punched them."
(The Long Earth, Pratchett & Baxter, 2012)

Handout is on the web: use the `QR-code` or visit `wilfridskendall.github.io/talks/PerfectEpidemics`.

This is initial work on using perfect simulation (CFTP) for epidemics.
WSK acknowledges the support of UK EPSRC grant EP/R022100.

---

1. Introduction to Perfect Simulation

- Propp & Wilson (1996) invented exact simulation / Coupling from the Past (CFTP) / perfect simulation;

- The term "perfect simulation" (WSK, 1998) should make you suspicious: perfection is never achieved. This is why the term was chosen!

- Key ideas of "*classic CFTP*":
  - ▶ extend simulation *backwards* through time not forwards;
  - ▶ exploit monotonicity (*couple* maximal and minimal processes);
  - ▶ seek coalescence.

- Simplest possible example: *random-walk-CFTP*
(can boost to use Ising model to do simple image reconstruction).

2025-01-24

Perfect Epidemics

└─ 1. Introduction to Perfect Simulation

Here's a very brief summary of CFTP / perfect simulation.

1. Jim Propp described the discovery of CFTP as like walking down the street and suddenly noticing a 50$ bill lying on the ground.

2. In particular, "exact simulation" cannot somehow miraculously defeat numerical approximation error `:-)`.

3. We will cover these key ideas in a single very specific example.

4. Propp & Wilson (1996) show how to vary random walk CFTP to get exact samples for a critical Ising model (Persi Diaconis: "Like seeing the landscape of Mars for the first time"). The Ising model with an external field can be used to model images, hence CFTP can be used for image reconstruction.
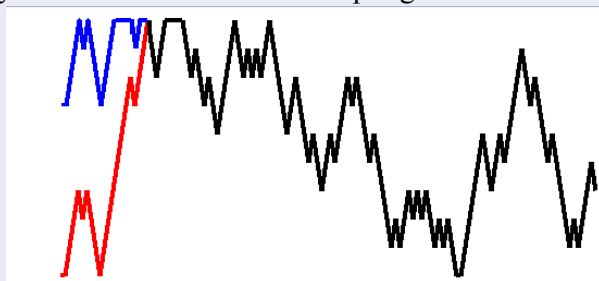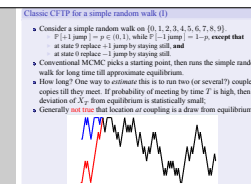
Classic CFTP for a simple random walk (I)

- Consider a simple random walk on $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.
  - $\mathbb{P}[+1 \text{ jump}] = p \in (0, 1)$, while $\mathbb{P}[-1 \text{ jump}] = 1-p$, **except that**
  - at state 9 replace $+1$ jump by staying still, **and**
  - at state 0 replace -1 jump by staying still.
- Conventional MCMC picks a starting point, then runs the simple random walk for long time till approximate equilibrium.
- How long? One way to *estimate* this is to run two (or several?) coupled copies till they meet. If probability of meeting by time T is high, then deviation of X_T from equilibrium is statistically small;
- Generally not true that location *at* coupling is a draw from equilibrium.
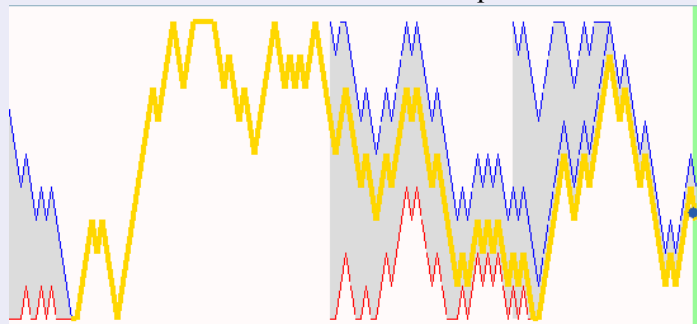
Perfect Epidemics

This random process is asymmetric simple random walk on the integers, but forbidding any transitions outside the specified state-space.

1. The process is irreducible _and_ aperiodic, so equilibrium is a truncated Geometric distribution (hint: use reversibility to check this!).

2. Conventional MCMC is overkill here, of course: we could simulate directly from the equilibrium distribution.

3. This observation was originally formalized by Aldous. Here we use 'synchronous coupling – variants are possible but not particularly useful.

4. Clearly not true in our case: coupling can only occur at 0 or 9!

5. In this case the blue and red evolutions can only couple at a boundary, so self-evidently the value at time of coupling cannot be a draw from the equilibrium!
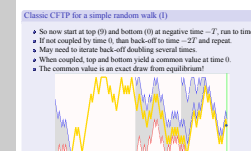
Classic CFTP for a simple random walk (I)

- So now start at top (9) and bottom (0) at negative time $-T$, run to time 0.
- If not coupled by time 0, than back-off to time $-2T$ and repeat.
- May need to iterate back-off doubling several times.
- When coupled, top and bottom yield a common value at time 0.
- The common value is an exact draw from equilibrium!

Perfect Epidemics

1. Ideally one needs to choose T neither too small nor too large. But the result is not particularly sensitive to this.
2. Very important that one doesn't introduce different jumps for the same time $-t$ in this binary back-off! We must couple to re-use randomness.
3. The binary back-off procedure means, if T is initially set to be too small then the extra work that is required is only ever a factor of 4!
4. Re-use of randomness means there is now no point in continuing: the common value at time 0 will be the same however far we extend into the past with further back-offs.
5. Why is the common value an exact draw from equilibrium? Informally, because one would get the same result however far one backed-off: therefore the draw is effectively a draw from time $-\infty$. The golden thread can be viewed as a perfect draw from the last segment of such a simulation (this sort of device is very well-known to ergodic theorists). Remarkably, this can easily be converted into a fully rigorous proof!

A little theory about CFTP

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

- Theoretical limits: *in principle* we can show that
  - ▶ *Classical CFTP* equivalent to uniform ergodicity (Foss & Tweedie, 1998);
  - ▶ *Dominated CFTP* achievable under geometric ergodicity (WSK, 2004);
  - ▶ Dominated CFTP can work in some **non**-geometrically ergodicity cases (SBC & WSK, 2007a; *nb* corrigendum SBC & WSK, 2007b).

- *Dominated CFTP* can be used to carry out perfect simulation for stable point processes (WSK & Møller, 2000);

- Detailed expositions: WSK (2005), Huber (2015).
  (Want to implement CFTP in R? see WSK, 2015.)

---

1. To be computationally effective, the WSK (1998) fixes still require some associated notion of partial order and (perhaps weak) amenability of the dominating process.
2. Basic ideas: use the notion of regenerative sets ("small sets"), and (Foster-)Lyapunov arguments. Note that the recipes given for these tend not to be computationally practical: they simply show the possibility of (possibly computationally infeasible) CFTP. They are intended as a challenge: when can one find *practical* methods?
3. (Locally finite) point processes cannot in general be treated using Classic CFTP, because the "top" pattern would have points *everywhere*. More generally, the "in principle" results encourage us to try CFTP out on a wide variety of models (For example, for $M/G/k$ multiserver queues, SBC & WSK, 2015; SBC, 2020).
4. This is the end of the visual introduction to CFTP. == People who are interested in CFTP from a practical point of view may find it useful to work through WSK (2015).

---

2. Perfect Epidemics: a challenge problem for CFTP

Many important inferential questions (Cori & Kucharski, 2024).

Simplest models (versus UK model with 10^6 agents!, Fraser & Others, 2023):

S-I-R deterministic epidemic: susceptibles s, infectives i, removals r (constant total population $s + i + r = n$):

$$
\begin{aligned}
s' &= -\alpha\, s\, i\,, \\
i' &= (\alpha\, s - \beta)\, i\,, \\
r' &= \beta\, i\,.
\end{aligned}
$$

S-I-R stochastic epidemic: a Markov chain (S, I, R) with transitions

Infection:	$S \to S - 1\,,$	$I \to I + 1$	at rate	$\alpha\, S I\,,$
Removal:	$I \to I - 1\,,$	$R \to R + 1$	at rate	$\beta\, I\,.$

Both models make an unrealistic assumption: homogeneous mixing.

Cori & Kucharski (2024) provide a broad perspective on statistical challenges from a very practical point of view!

Here are two classical models (one deterministic, one stochastic) from mathematical epidemiology. Even the exceedingly simple case of deterministic S-I-R permits only *partial* closed-form solution.

The assumption of homogeneous mixing of population greatly simplifies the mathematics *and* statistics. Without this assumption one ends up with huge numbers of parameters, which is very bad news statistically speaking. A lot of research has been done on how to deal with more realistic models! (For example, the UK model using 1000000 agents mentioned here!)

However in this work we deal only with perfect simulation for homogeneous mixing. Before running, one must learn to walk!
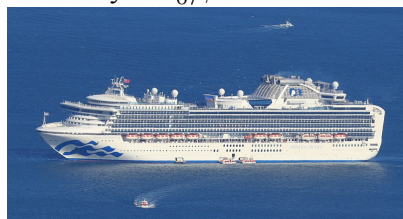
The first question asked about a new epidemic

"What is the R-number?"

The R-number is $\alpha\, s_0/\beta$: mean number of new infectives produced per infective at *start* of epidemic with initially s_0 susceptibles.

Whittle (1955)'s threshold theorem: R-number $\gg 1$ means positive chance of epidemic infecting significant proportion of the population.

Wikipedia: "The British-registered *Diamond Princess* was the first cruise ship to have a major [COVID-19] outbreak on board, with the ship quarantined at Yokohama from 4 February 2020 for about a month. Of 3711 passengers and crew, around 700 people became infected and 9 people died."

Evidently $\alpha\, s_0/\beta \gg 1$ – as was sadly later confirmed, a sorrow for us all.
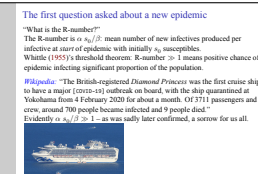
I am skipping over very influential early work on this incident, which particularly focussed on the incubation period (a time when the subject is infectious, perhaps at time-varying rate, before symptoms appear). This is related to drawing inferences about β.

————————————————

See Mizumoto *et al* (2020) for very early analysis of the *Diamond Princess* outbreak.

Inference on the R-number

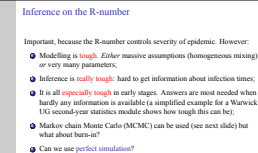Important, because the R-number controls severity of epidemic. However:

1. Modelling is tough. *Either* massive assumptions (homogeneous mixing) *or* very many parameters;

2. Inference is really tough: hard to get information about infection times;

3. It is all especially tough in early stages. Answers are most needed when hardly any information is available (a simplified example for a Warwick UG second-year statistics module shows how tough this can be);

4. Markov chain Monte Carlo (MCMC) can be used (see next slide) but what about burn-in?

5. Can we use perfect simulation?

The "R-number" was the subject of continuing interest throughout the COVID-19 pandemic!

Notice that the "R-number" is as much a social construct as a biological one, and in real situations will interact in subtle and complicated ways with behaviour.

————————————————

We will seek to make headway in using perfect simulation to aid inference.

An easier question

An absurdly simple variant of contact tracing:

"When did the infections occur, supposing we only observe removals?"
(Gibson & Renshaw, 1998; O'Neill & Roberts, 1999; Gibson & Renshaw, 2001)

An important step on the way: generating an *unconditioned* epidemic.

1. Thus n, α, β are known, removal times are observed,
   but unobserved infection times must be inferred.

2. Visualize n timelines, along which incidents are scattered:
   - potential removals, activated if timeline is infected;
   - potential infections, activated if timeline is infected *and* if designated target timeline is lowest uninfected timeline.

3. Poisson point processes *of appropriate rates* yield an S-I-R epidemic.

4. First step: evolve whole S-I-R trajectory in *algorithmic time* (alter potential infections and removals using immigration-death in discrete algorithmic time).

5. Result: *trajectory-valued chain*, unconditioned S-I-R as equilibrium.

An important step on the way is to focus on just one aspect of inference:

1. Instead of trying to tackle inference concerning the parameters, let's concentrate on the large number of nuisance parameters provide by (typically unobserved) infection times.
2. This is a compromise: instead of tracking individuals as in the stochastic epidemic model, we effectively track the numbers of individuals in various categories.
3. Poisson point processes generate incidents, which then collectively influence the S-I-R trajectory.
4. I learned the phrase "algorithmic time" from Andrew Stuart.
   We move from Poisson points spread out along timelines to discrete immigration-death processes producing patterns that evolve in algorithmic time.
5. But we shall quickly move to *continuous* algorithmic time.

From incidents to unconditioned epidemic trajectories (1/3)
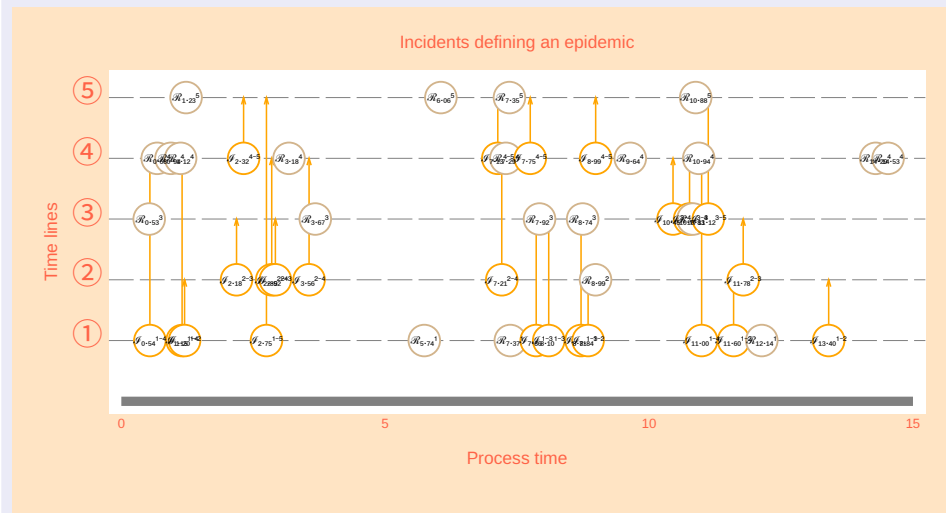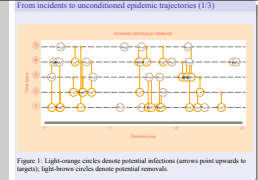

Figure 1: Light-orange circles denote potential infections (arrows point upwards to targets); light-brown circles denote potential removals.

As implied above, incidents comprising a single innovation are produced by Poisson point processes on each of the timelines.

Colour-coding distinguishes infections from deaths. Note that infections need to specify which timeline is going to be infected.

- The rates of the Poisson processes generating the (proposed) incidents are chosen to ensure that the numbers of actual susceptibles, infectives and removals form a Markov chain with rates as required for a S-I-R epidemic.
- For an infection incident to be activated, it must sit on an infected portion of its timeline and must point to the least of the uninfected timelines (calculating at the process time of the incident).
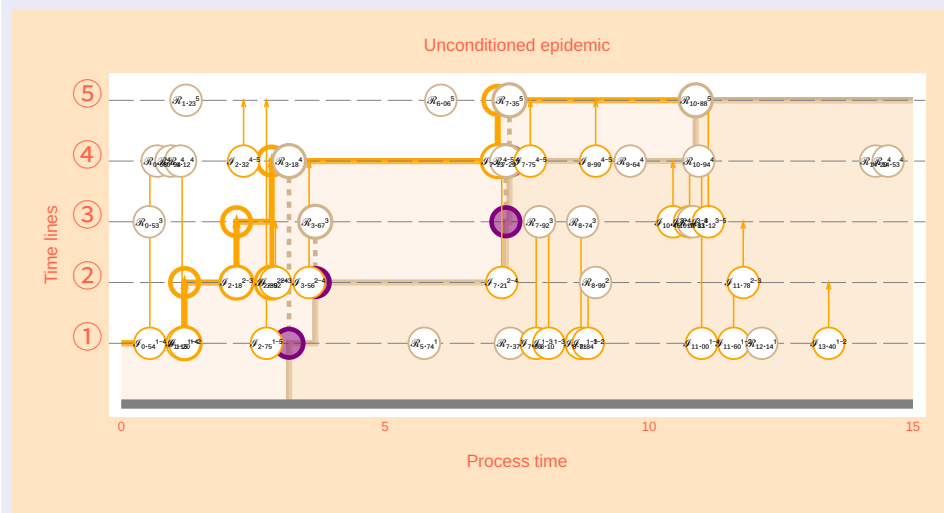- For a removal incident to be activated, it must sit on an infected portion of its timeline.
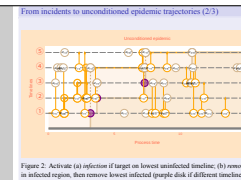
Figure 2: Activate (a) *infection* if target on lowest uninfected timeline; (b) *removal* if in infected region, then remove lowest infected (purple disk if different timeline).

Perfect Epidemics

2025-01-24

Here is the resulting evolution of an epidemic. Note however that we need to specify the initial number of infectives (here we specify $I_0 = 1$).
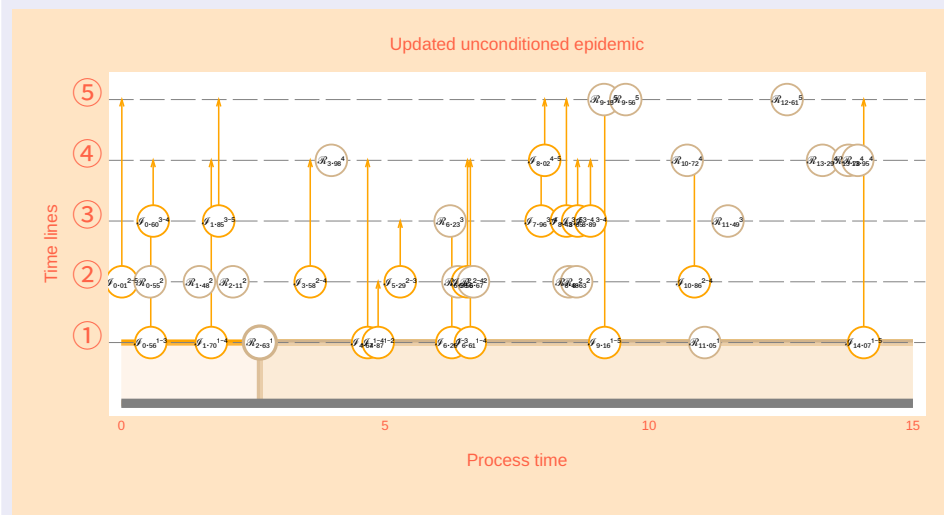

Figure 3: A step in algorithmic time for the unconditioned epidemic simply involves replacing the original incidents by a new set of incidents.

Perfect Epidemics

2025-01-24

An update in algorithmic time, if there is no conditioning, simply involves replacement of one set of incidents by another. (Here there is a radical change, as epidemic dies out at an early stage.)

But we need to refine this if we are to take account of conditioning: a simple accept/reject procedure will almost always reject proposals of entire innovations involving all of both infection and removal incidents.
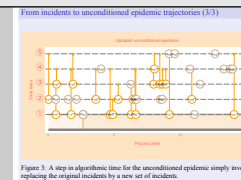
Crucial technical point

- Updates in algorithmic time τ are then (algorithmic-)*time-reversible*: so restriction to subset S of state-space (of *activated* removals occurring precisely at specified set of times) implies a new equilibrium which is the old equilibrium conditioned to lie in S.
- For later purposes it is convenient to stage the replacement as follows:
  1. Replace removals (Rs);
  2. Re-sample timelines (though not times) of Rs;
  3. Replace infections (Is).
- Re-express using continuously varying τ. Process time runs over $[0, T]$.
  1. For $2nT < \tau < (2n+1)T$, update old Rs with times in $(0, \tau - 2nT)$;
  2. For $\tau = (2n+1)T$, resample timelines (not times) of Rs;
  3. For $(2n+1)T < \tau < (2n+2)T$, update old Is in $((2n+2)T - \tau, T)$.
- Thus the original update is expressible as a (continuous) composition of updates, each of which satisfies detailed balance in equilibrium.
- The connection "restriction=conditioning" is thereby preserved.
- Crucially, step 2 ensures composition action is irreducible over S! (So equilibrium under conditioning is unique.)

In the case of discrete algorithmic time, it seems that arranging for conditioning would involve horrendous computation. Instead we "unroll'' into continuous algorithmic time.

Intuitively, we analyze first removals and then infections "pixel-by-pixel" along the process time axis.

It turns out to be convenient to work forwards in time for removals, then backwards in time for infections – and to re-assign random timelines (but not times) for removals in between these two phases.

Illustration of technical point (1/8)
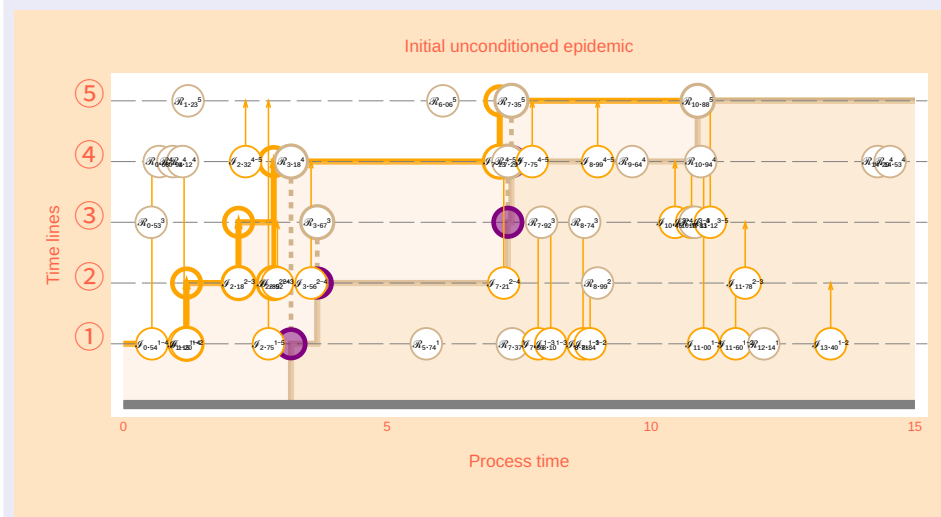
Figure 4: No change yet to removals or infections;
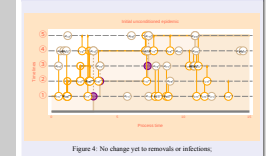
Here is the initial configuration again.

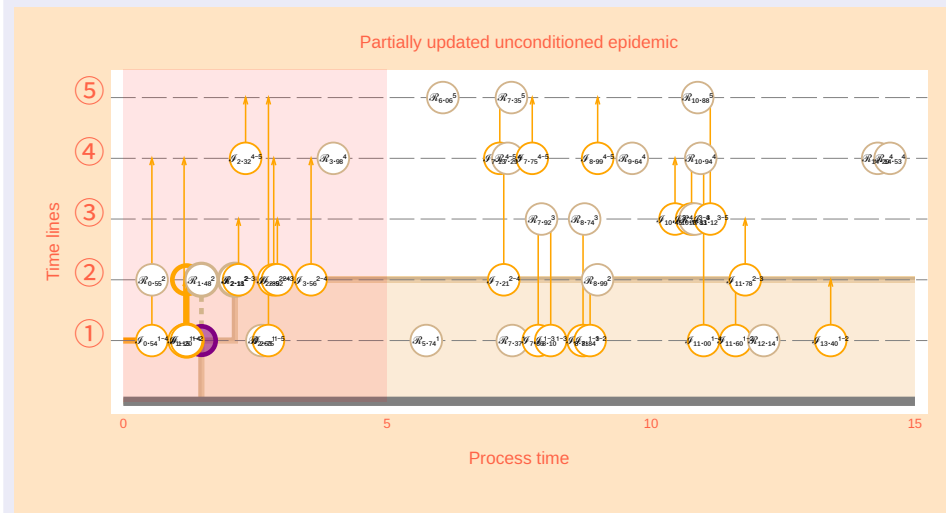Figure 5: Replace first third of removals, infections unchanged;

Perfect Epidemics

2025-01-24

Figure 5: Replace first third of removals, infections unchanged;

Now we use the immigration-death process of removals to work forwards along the process-time axis removing old removals and adding new ones. At present we are not applying conditioning. Here we are 1/3 of the way through dealing with removals. Note that the removals have now rendered the epidemic extinct!

Figure 6: Replace first two-thirds of removals, infections unchanged;

Perfect Epidemics

2025-01-24

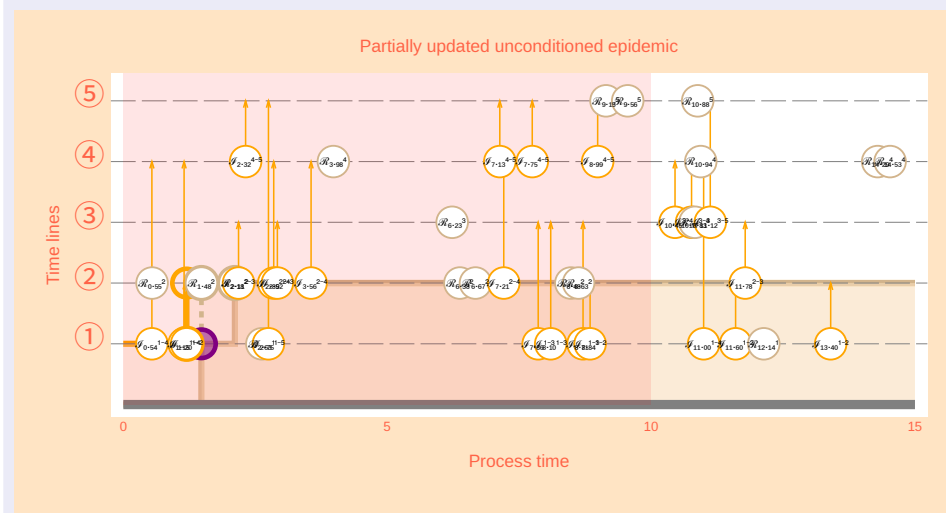Figure 6: Replace first two-thirds of removals, infections unchanged;

Here we are 2/3 of the way through dealing with removals.
Flip backwards and then forwards one slide, and confirm that the light-orange nodes with vertical arrows (infections) are not changed at this stage.
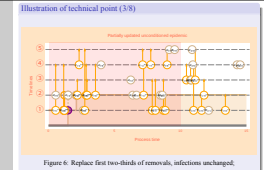
Partially updated unconditioned epidemic

Figure 7: All removals resampled, infections as yet unchanged;

Perfect Epidemics

2025-01-24
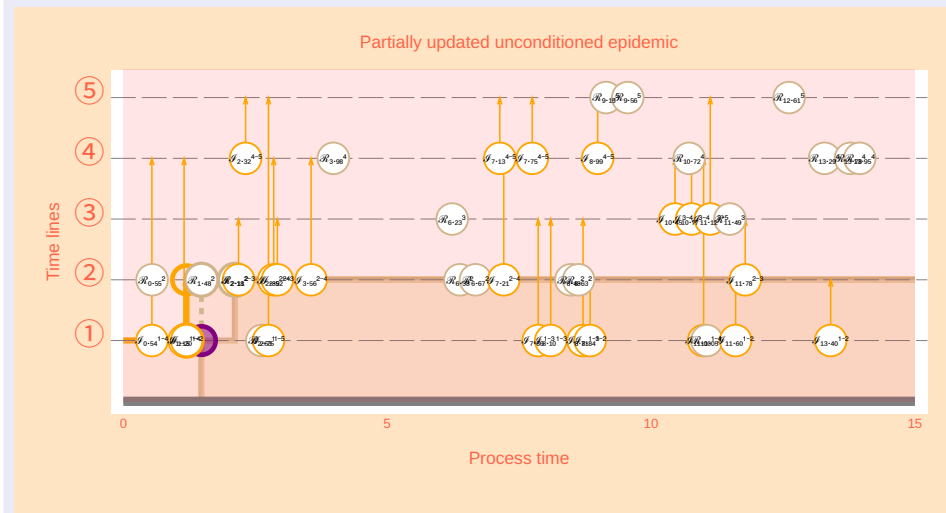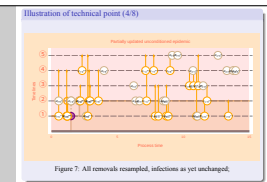
Figure 7: All removals resampled, infections as yet unchanged;

All removals have now been processed. Because we are not yet conditioning, all old removals have now been removed, and all new removals have been added.

_____

Notice that nothing was done for this simulation in the last 2/3 of removal work, because the new epidemic had by then run its course.

Partially updated unconditioned epidemic

Figure 8: Re-sample all removal timelines, infections as yet unchanged;

Perfect Epidemics

2025-01-24

Figure 8: Re-sample all removal timelines, infections as yet unchanged;

At this stage we now re-sample the timelines for all removals. Since we are not conditioning, this results in a big change.
No nodes change time, but timelines for removal nodes are all re-sampled.

1. When we proceed to consider conditioning on observing certain removals, the observed ("conditioned' ') removals will not have been moved so far (their removal will have been rejected!) but at this stage the change in timeline will be accepted if the dynamics do not then lead to a change in conditioned removals.
2. On the other hand, timeline-resampling for an inactivated removal will be rejected if it leads to that removal being activated, but statistically the effect will be the same as if it wasn't re-sampled.
3. Because of this re-sampling, we **won't** get the same final result as in the preliminary run.

Partially updated unconditioned epidemic
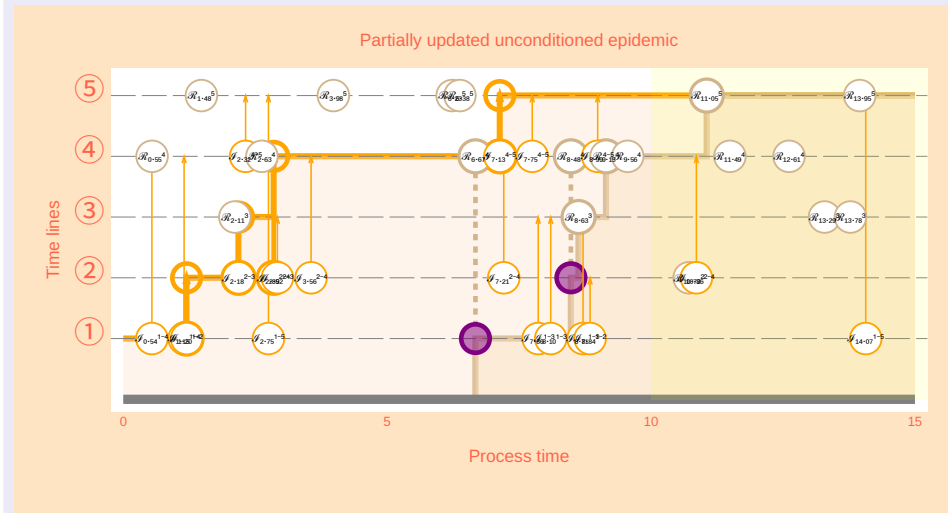
Figure 9: Re-sample last third of infections;

Perfect Epidemics

2025-01-24

Figure 9: Re-sample last third of infections;

Now we resample infections, *but this time working backwards through process time.* Here is how matters stand after 1/3 of the work is done. Again, there is little scope for anything to be done yet: one infection has been dropped.
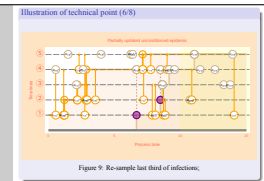
Partially updated unconditioned epidemic

Figure 10: Re-sample last two-thirds of infections;

Perfect Epidemics

2025-01-24
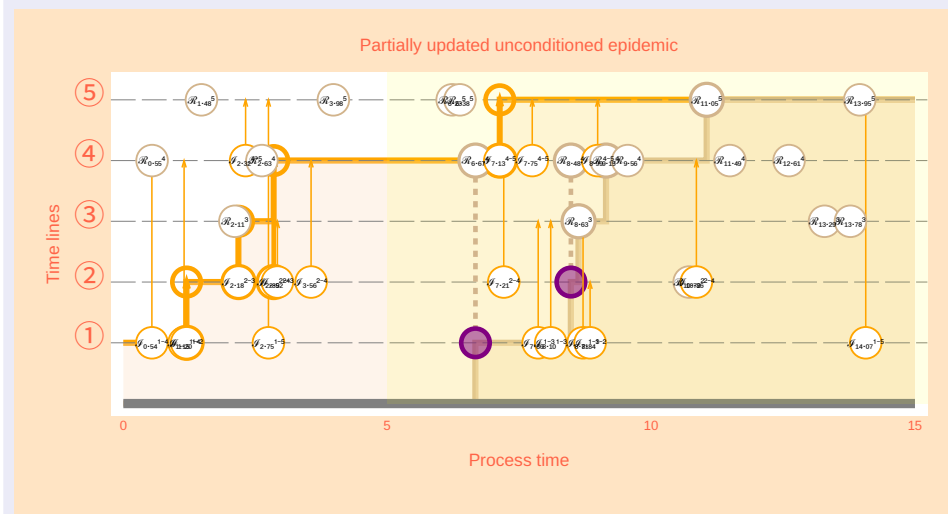

Figure 10: Re-sample last two-thirds of infections;

And here is how matters stand after 2/3 of the work is done. Still little change at this stage.

Figure 11: All infections now re-sampled.
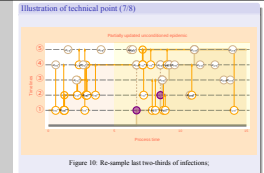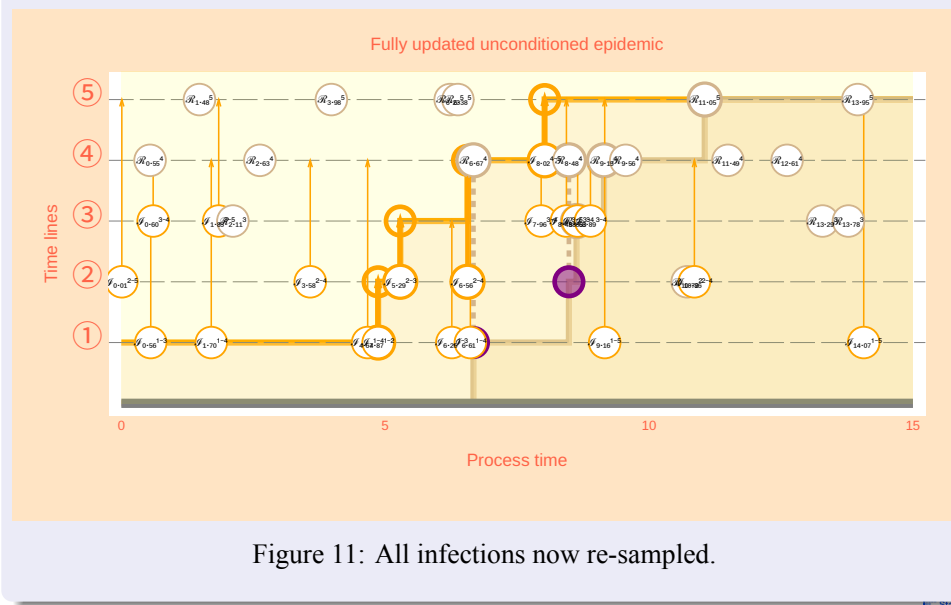
Finally the entire process time range is dealt with. In this unconditioned case all removals and infections are removed and replaced by new ones. Under conditioning, however, each individual decision (to remove, resample, or introduce) has to be tested to see if it affects the pattern of conditioned removals. We now need to consider how to do this efficiently, but also in a manner that allows us to reason about the crucial monotonicity.

---
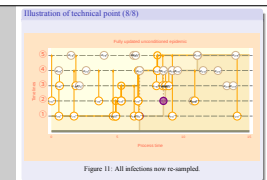
3. Conditioning on observed removals

- The trajectory-valued chain is *dynamically reversible*, in *continuous algorithmic time*.
- Irreducibility is *vital* (otherwise equilibrium depends on starting point). Consequently:
  - conditioned removals must be able to change timeline (but not time of occurrence);
- Forbidding removal of observed removals, and forbidding creation of new activated removals, yields a modified chain whose invariant probability measure conditions on observed pattern of removals.

  Implications:
  - a removal can be introduced only if it doesn't activate;
  - a conditioned removal timeline can be altered only if it doesn't de-activate;
  - an infection cannot be removed if that action loses a conditioned removal;
  - an infection can be introduced only if no new observed removals result.
- Does this produce a *feasible* and suitably monotonic algorithm?
- Housekeeping details required to establish that monotonicity still works.

  Key notions: *last feasible epidemic* (LFE) and *no-fly zone* (NFZ).

Monotonicity is a key concept here. (Note however that we can still sometimes generate perfect simulation algorithms even when monotonicity does not obtain: we don't need to do this here but should bear it in mind for more general problems in the future.)

1. We can also deal with "unbounded" sample space using the notion of Dominated CFTP.
2. The heart of the matter lies in establishing (some variant of) monotonicity.
3. Additionally. for a complicated algorithm such as this one, it is important to test the underlying logic by implementing the algorithm in computer code (which has a way of revealing false hidden assumptions!).
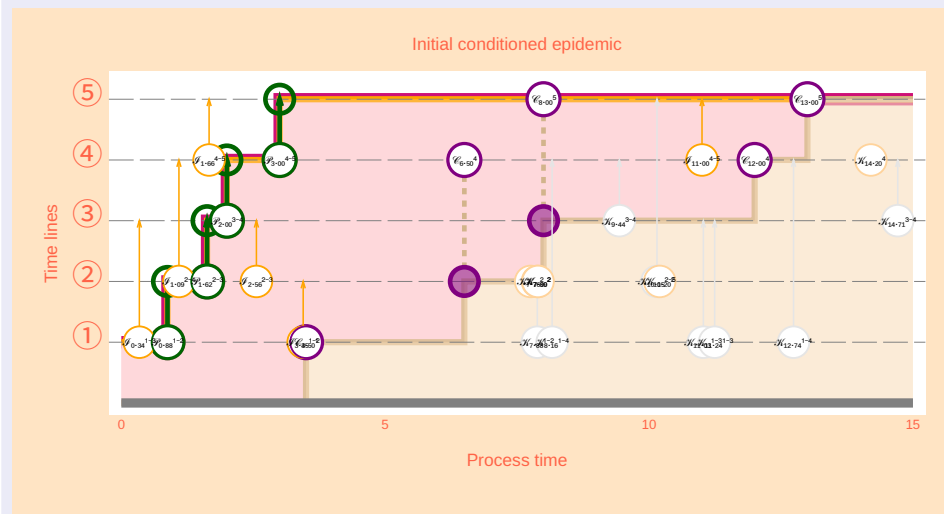
Figure 12: Initial epidemic with conditioned removals indicated using purple circles (and purple disks when different timelines are infected).

Perfect Epidemics

2025-01-24

Here is a different initial epidemic, including *conditional removals*, indicated by purple circles.

- Essentially we change it in the same way as in the unconditioned case (working "pixel-by-pixel"), but forbidding any change if that breaks conditioning.
- Dealing with removals is easy: it amounts to removing all inactivated removals and replacing them with a new bunch, and then resampling the conditional removal timelines, accepting only those changes which do not de-activate the conditional removals.
- But we need to take more care about whether or not we can take away or add in *infections*.

Figure 13: Epidemic updated under restriction: all conditioned removals remain activated, no new removals are activated. Green infections have been "perpetuated".

Perfect Epidemics

2025-01-24

Here is the result of the update, which we now explain general terms. We have noted removals are easy. Re-sampling the conditioned removal timelines is a simple accept-reject step for each conditioned removal.

The key to dealing with infections that target the infected region is also rather simple:

1. Always add them (because it won't make a difference).
2. Delete them only if it is clear that this will not result in a conditioned removal falling out of the infected region.
3. Implementation amounts to referring to a recursively-defined *last feasible epidemic* (LFE). The LFE coincides with the actual epidemic trajectory exactly when deletion of the infection in question will lead to a conditioned removal falling out of the infected region; therefore this deletion is rejected and the retained infection is designated as **perpetuated** (coloured green!). Otherwise the infection can be deleted.
4. Infections targeting the susceptible region are trickier, and we will discuss them briefly below.

Figure 14: LFE computed recursively working right-to-left: the slowest sequence of infections deals with all infected timelines in order (includes perpetuated infections).

Perfect Epidemics

2025-01-24

First we draw the LFE. Because it is defined recursively working backwards in process time, it can be computed in entirety from the updated conditional epidemic.

- Of course, when computing the update, we have to compute it set-by-step when working backwards in process time.
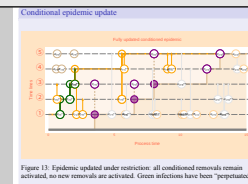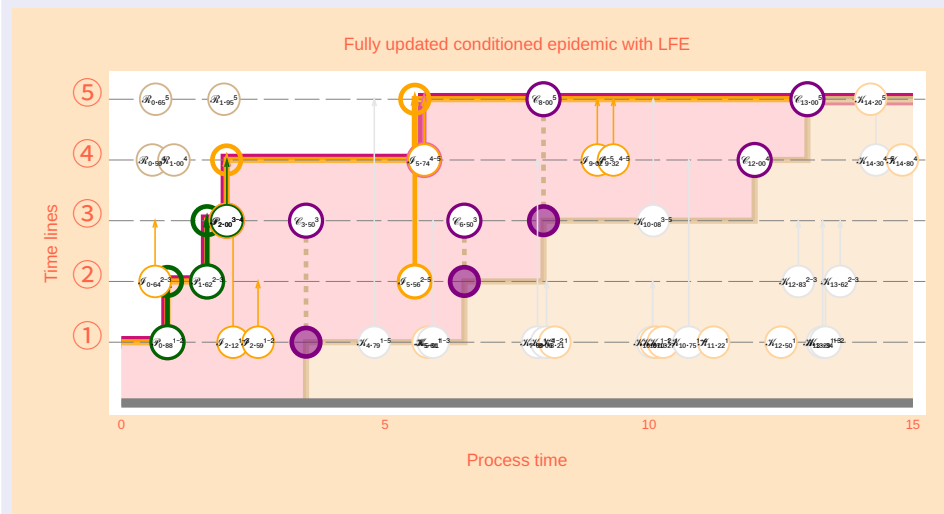- Here there is a single discrepancy between LFE and epidemic trajectory, over the time interval $[5.56, 5.74]$.
- At times 0.88, 1.62, 2.00 there are perpetuations (an old infection has to be retained to ensure conditioning is not broken); at the first two times (0.88, 1.62) the perpetuation ends up forming part of the epidemic trajectory.

A relatively simple induction argument shows that the LFE depends monotonically on the epidemic trajectory!

Figure 15: NFZ computed recursively working right-to-left: it traces the region of timelines that must not be infected if one is not to activate unobserved removals.

Perfect Epidemics

2025-01-24

The NFZ is used to decide whether to accept proposals of infections targeting the susceptible region. More complicated than LFE; two phenomena interact.

1. An infection rooted in the infection zone must be assessed to see whether its introduction would lead to activating a removal, because it connects the infectious zone to a timeline which would lead to such an activation (a "no-fly" portion of the timeline);
2. An infection rooted in the susceptible zone may be such that it makes a portion of its timeline *before* its time into "no-fly".

In this particular case neither of these two mechanisms play a part.
The NFZ is evaluated recursively, working backwards in process time. A more involved induction argument shows that the NFZ depends monotonically on the epidemic trajectory. This implies monotonicity of the algorithmic-time evolution of the conditioned epidemic: hence CFTP is feasible!

4. Example

- Smallpox outbreak in a closed community of 120 individuals in Abakaliki, Nigeria (much studied! see page 125 of Bailey, 1975).
- **Assume**
  - ▶ first observed removal is also the first removal: under a plausible improper prior we can then deduce what is the distribution of infectives I_0 at time 0;
  - ▶ *all* removals are recorded;
  - ▶ no further removals after last observed removal (makes life easier).
- Coding in *julia* (Bezanson *et al.*, 2017), we obtain a perfect simulation GIF or MP4 yielding a draw from unobserved pattern of infections.

Minimum and maximum evolutions (using I5 at t=0) at algorithmic time step -0000

Cycle at time -0000/-0008

Process time

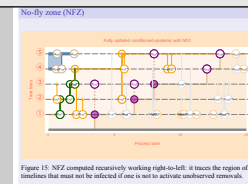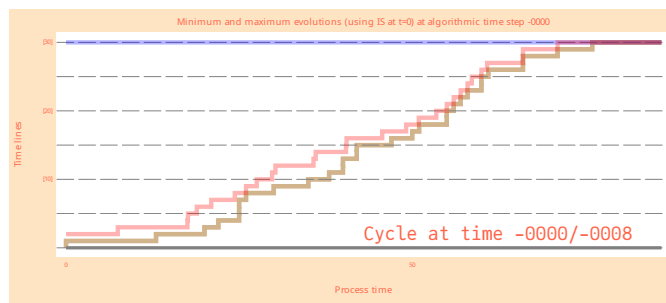We could vary the assumption concerning the initial and final numbers of infectives while still using essentially the same perfect simulation algorithm! This amounts to presuming some Bayesian prior knowledge.

- How do we allow for the (unobserved) number of initial infectives being unobserved?
- Use an independence sampler: draw from the (unconditioned) number of initial infectives, accept if this leads to observed removals exactly at the list of conditioned removals.
- (One can alternatively employ a random-walk Metropolis sampler, so long as the prior for the initial number of infectives is suitably monotonic.)

Technical note:
The simulation implementation uses `julia` (Bezanson *et al.*, 2017), which combines rapid development, an expressive syntax (very useful for involved algorithms), and remarkably fast execution using very good "just in time" compilation techniques.
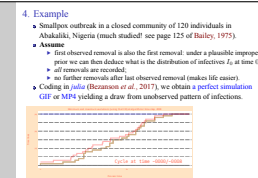
So what?

- Why this emphasis on unobserved infections given fixed α and β, when we need inference on R-number $\alpha\, n/\beta$ for *unknown* α and β?

- Good question. But a re-weighting argument allows us to get (unbiased) estimates based on *different* α and β. The perfect simulation provides exact simulation-based computation to integrate out pattern of unobserved infections.

- So (**next steps after** SBC & WSK, 2024)
  - ▶ estimate likelihood test statistic for specified α and β;
  - ▶ construct steepest ascent algorithm (in effect, variant of Robbins-Monro stochastic optimization) to find *maximum a posterior* estimates of α and β;
  - ▶ or even, with some computational effort, compute the entire posterior joint density for α and β!

- Finally: generalize to other suitable compartment models?

In effect, perfect simulation supplies a well-behaved stochastic integration mechanism, reducing an MCMC algorithm to use a technique equivalent to a much more amenable Monte Carlo calculation. Other compartment models presenting varying degrees of challenge:

1. Split population into a small number of interacting sub-populations;

2. (For purposes of insight even if impractical.) Treat each individual as a sub-population on its own;

3. Allow parameters to change at specific times of day or week;

4. Allow for some activated removals to be unobserved!
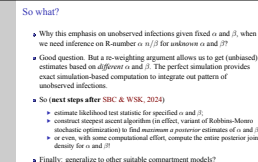
Conclusion

- If MCMC burn-in is a concern, try to build a perfect simulation!

- CFTP works even for significantly complex and relevant models of real-life phenomena.

- *Of course* detailed models resist perfect simulation (but it will be helpful to compare with a simpler model using fewer parameters).

- Still to be done: statistical estimation of parameters, generalization to other compartment models.

- Thank you for your attention! **QUESTIONS?**

Lots still to do!

At the Warwick talk in January 2025, Paul Jenkins asked whether this might be related to the "look-down" argument in population genetics (Donnelly & Kurtz, 1996)? My provisional answer, after recalling Donnelly & Kurtz (1996) later in the day, is:
as far as I can tell, the "look-down" argument is related in the sense that it also involves a particle model as discussed in Liggett (1985). However I do not yet see a closer connection.

References I

Bailey, N.T.J. (1975) *The mathematical theory of infectious diseases and its applications*, 2nd Ed. ed. Griffin.

Bezanson, J., Edelman, A., Karpinski, S., & Shah, V.B. (2017) Julia: A Fresh Approach to Numerical Computing. *SIAM Review*, **59**, 65–98.

Cori, A. & Kucharski, A. (2024) Inference of epidemic dynamics in the COVID-19 era and beyond. *Epidemics*, **48**, 100784.

Donnelly, P. & Kurtz, T.G. (1996) A countable representation of the Fleming-Viot measure-valued diffusion. *The Annals of Probability*, **24**, 698–742.

Foss, S.G. & Tweedie, R.L. (1998) Perfect simulation and backward coupling. *Stochastic Models*, **14**, 187–203.

Fraser, C. & Others (2023) OpenABM-Covid19: Agent-based model for modelling the Covid-19 and Contact-Tracing.

Gibson, G.J. & Renshaw, E. (1998) Estimating parameters in stochastic compartmental models using Markov chain methods. *Mathematical and Medical Biology*, **15**, 19–40.

Gibson, G.J. & Renshaw, E. (2001) Likelihood estimation for stochastic compartmental models using Markov chain methods. *Statistics and Computing*, **11**, 347–358.

Huber, M.L. (2015) *Perfect Simulation*. Boca Raton: Chapman; Hall/CRC.

References II

Liggett, T.M. (1985) *Interacting Particle Systems*, Grundlehren der mathematischen wissenschaften. Berlin, Heidelberg: Springer Berlin.

O'Neill, P.D. & Roberts, G.O. (1999) Bayesian Inference for Partially Observed Stochastic Epidemics. *Journal of the Royal Statistical Society Series A: Statistics in Society*, **162**, 121–129.

Pratchett, T. & Baxter, S. (2012) *The Long Earth*. Doubleday.

Propp, J.G. & Wilson, D.B. (1996) Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms*, **9**, 223–252.

SBC (2020) Omnithermal Perfect Simulation for Multi-server Queues. *ACM Transactions on Modeling and Computer Simulation*, **30**, 1–15.

SBC & WSK (2007b) Perfect simulation for a class of positive recurrent Markov chains (corrigendum). *Annals of Applied Probability*, **17**, 1808–1810.

SBC & WSK (2007a) Perfect simulation for a class of positive recurrent Markov chains. *Annals of Applied Probability*, **17**, 781–808.

SBC & WSK (2015) Perfect simulation of M/G/c queues. *Advances in Applied Probability*, **47**, 1039–1063.

SBC & WSK (2024) Perfect Epidemics.

Whittle, P. (1955) The outcome of a stochastic epidemic—a note on Bailey's paper. *Biometrika*, **42**, 116–122.

References III

WSK (1998) Perfect Simulation for the Area-Interaction Point Process. *Probability towards 2000* (Accardi, L. & Heyde, C.C. eds). Springer-Verlag, pp. 218–234.

WSK (2004) Geometric ergodicity and perfect simulation. *Electronic Communications in Probability*, **9**, 140–151.

WSK (2005) Notes on Perfect Simulation. Singapore: World Scientific, pp. 93–146.

WSK (2015) Introduction to CFTP using R. *Stochastic geometry, spatial statistics and random fields, Lecture notes in mathematics*. Springer, pp. 405–439.

WSK & Møller, J. (2000) Perfect simulation using dominating processes on ordered spaces, with application to locally stable point processes. *Advances in Applied Probability*, **32**, 844–865.

Image information

Image	Attribution	
Terry Pratchett	Luigi Novi	CC BY 3.0
Classic CFTP for a simple random walk	Result of code written by WSK	
Diamond Princess	Alpsdake	CC BY-SA 4.0
Epidemic CFTP images and animation	Result of code written by WSK	

Previous instances of this talk

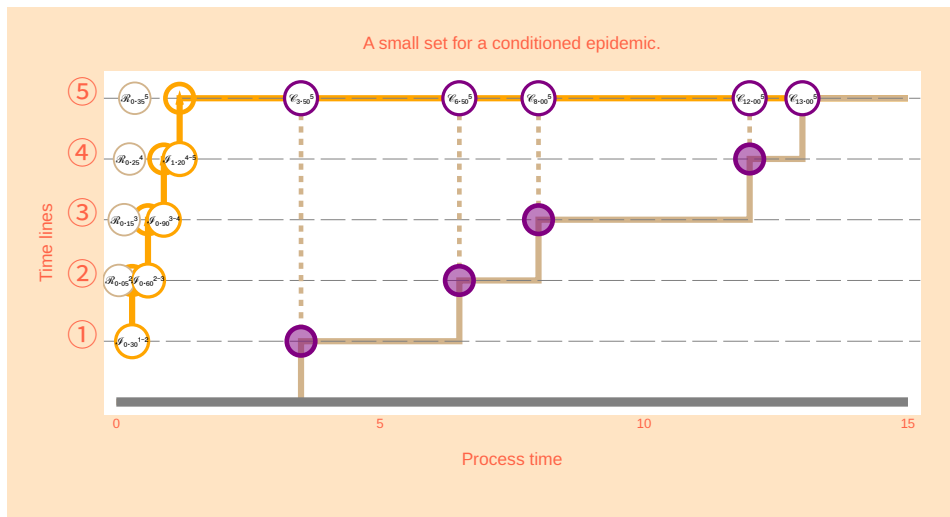| Date	Title		Location
19/04/24	Perfect Epidemics	Short Research Talk (12min)	Warwick
15/05/24	McMC and Perfect Simulation	Graduate Seminar, Aristotle Univ. (50min)	Thessaloniki
17/01/25	Perfect Epidemics	Applied Probability Seminar (50min)	Warwick

---

A "near-maximal" configuration

Figure 16: A conditional epidemic in which all activated infections occur before time 3.0, also before smallest observed removal time.

Perfect Epidemics

2025-01-24

└─ A "near-maximal" configuration

1. For a "near-maximal configuration", we set relevant infections to occur in order on or before a fixed small time $\delta = 1.5 > 0$, where 2δ is smaller than any of the conditioned removal times.
2. All inactivated removals have to occur before their respective infections (in order to avoid activation).
3. Variants on this construction can be used to show that CFTP for the perfect epidemic must coalesce in finite time;
4. However to get a sensible stochastic bound on coalescence time would require a lot more work!

Other technical information

Software versions

Software used in computations:

Software	Version	Branch	Date of last commit
Quarto	1.6.39	—	
Running under `julia`	1.11.3	—	
Module `EpidemicsCFTP`	2.2.492	main	Thu Jan 23 20:50:07 2025
Module `EpidemicsUtilities`	0.1.2.156	main	Thu Jan 23 12:06:14 2025
This Quarto script	2.2.600	Wilfrid	Fri Jan 24 11:43:26 2025

Revision notes

These notes were produced from `PerfectEpidemics.qmd`:

Author:	Wilfrid Kendall W.S.Kendall@warwick.ac.uk
Date:	Fri Jan 24 15:14:27 2025 +0000
Summary:	Cosmetics.
	2.2.603