

# Perfect Epidemics

## Applied Probability Seminar

Department of Statistics, University of Warwick

W S Kendall    S B Connor

Warwick, York

4 February 2025



Warwick  
Statistics

W S Kendall, S B Connor (Warwick, York)

Perfect Epidemics

4 February 2025

1 / 51

## 1. Introduction to Perfect Simulation

- Propp & Wilson (1996) invented exact simulation / Coupling from the Past (CFTP) / perfect simulation;
- The term “perfect simulation” (WSK, 1998) was chosen so as to make you suspicious: perfection is never achieved!
- 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).

## 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](https://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.

Warwick  
Statistics

W S Kendall, S B Connor (Warwick, York)

Perfect Epidemics

4 February 2025

2 / 51

## Perfect Epidemics

2025-02-04

### └ 1. Introduction to Perfect Simulation

Here is 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 :-). And, as with all simulation, its validity depends on correctness of code!
3. We will illustrate these key ideas by considering 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.

#### 1. Introduction to Perfect Simulation

- Propp & Wilson (1996) invented exact simulation / Coupling from the Past (CFTP) / perfect simulation;
- The term “perfect simulation” (WSK, 1998) was chosen so as to make you suspicious: perfection is never achieved!
- Key idea 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).

## Classic CFTP for a simple random walk (I)

- Consider a simple random walk on  $0:9 = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ .
  - $\mathbb{P}[+1 \text{ jump}] = p \in (0, 1)$ , while  $\mathbb{P}[-1 \text{ jump}] = 1-p$ , except that at state 9 replace the +1 jump by “staying still”, and at state 0 replace the -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

2025-02-04

Classic CFTP for simple random walk (I)

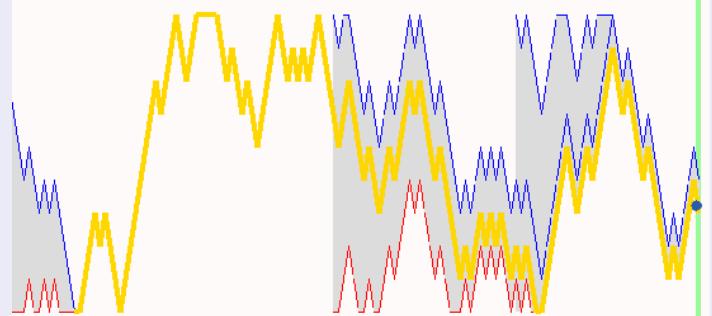
- Consider a simple random walk on  $0:9 = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ .  
 $\mathbb{P}[+1 \text{ jump}] = p \in (0, 1)$ , while  $\mathbb{P}[-1 \text{ jump}] = 1-p$ , except that at state 9 replace the +1 jump by “staying still”, and at state 0 replace the -1 jump by “staying still”.
- Conventional 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!

This random process is asymmetric simple random walk on the integers, modified by forbidding any transitions outside the specified state-space  $0:9$ .

- The process is irreducible and aperiodic, so equilibrium is a truncated Geometric distribution (hint: use reversibility to check this!).
- Of course even conventional MCMC is overkill here: simulation directly from the equilibrium distribution is simple and efficient.
- This observation was originally formalized by Aldous. Here we use “synchronous” coupling – variants are possible but not particularly useful.
- In our case coupling can only occur at 0 or 9! So here the blue and red evolutions can only couple at a boundary; thus self-evidently the value at time of coupling cannot be a draw from the equilibrium!

## Classic CFTP for a simple random walk (I)

- Start at top (9) and bottom (0) at negative time  $-T$ , run to time 0.



- If not coupled by time 0, then back-off to time  $-2T$  and repeat.  
**NB** re-use randomness!
- May need to iterate back-off doubling several times.
- When coupled, top and bottom yield a common value at time 0.
- The common value (golden thread) is an exact draw from equilibrium!

## Perfect Epidemics

2025-02-04

Classic CFTP for simple random walk (I)

- Start at top (9) and bottom (0) at negative time  $-T$ , run to time 0.
- If not coupled by time 0, then 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 (golden thread) is an exact draw from equilibrium!

- Ideally one needs to choose  $T$  neither too small nor too large. But the result is not particularly sensitive to this.
- 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**.
- 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!
- 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.
- 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.)

## Perfect Epidemics

2025-02-04

### └ A little theory about CFTP

#### 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) ideas still require a (perhaps weak) notion of partial order, and also the ability to simulate the dominating process.
2. Basic ideas: use the notion of regenerative sets (“small sets”), and (Foster-)Lyapunov arguments. Note that the resulting recipes tend not to be computationally practical: they simply suggest the *possibility* of (possibly computationally infeasible) CFTP. They are intended as a challenge: when can one find *practical* methods?
3. (Locally finite) point processes are generally not accessible via Classic CFTP; 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

$$\begin{array}{llll} \text{Infection: } & S \rightarrow S-1, & I \rightarrow I+1 & \text{at rate } \alpha S I, \\ \text{Removal: } & I \rightarrow I-1, & R \rightarrow R+1 & \text{at rate } \beta I. \end{array}$$

Both models make an *unrealistic assumption*: homogeneous mixing.

## Perfect Epidemics

2025-02-04

### └ 2. Perfect Epidemics: a challenge problem for CFTP

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:  $S \rightarrow S-1, \quad I \rightarrow I+1$  at rate  $\alpha S I$ ,  
Removal:  $I \rightarrow I-1, \quad R \rightarrow R+1$  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.



Warwick Statistics

## Perfect Epidemics

2025-02-04

└ The first question asked about a new epidemic

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.  
While the R-number is  $\gg 1$ , the R-number  $> 1$  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  $\beta$ .

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

**Question:** is there a dataset for the *Diamond Princess* outbreak? and how detailed is it?

## 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**?

## Perfect Epidemics

2025-02-04

└ Inference on the R-number

Inference on the R-number

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

- ➊ Modelling is **tough**. Either massive assumptions (homogeneous mixing) *or* very many parameters;
- ➋ Inference is **really tough**: hard to get information about infection times;
- ➌ 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);
- ➍ Markov chain Monte Carlo (MCMC) can be used (see next slide) but what about burn-in?
- ➎ 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.

Reference (for example a blog) to the “simplified example”?

## 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.

- ➊ Thus  $n, \alpha, \beta$  are known, removal times are observed, but unobserved infection times **must be inferred**.
- ➋ 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.
- ➌ Poisson point processes of *appropriate rates* yield an S-I-R epidemic.
- ➍ First step: evolve whole S-I-R trajectory in *algorithmic time* (alter potential infections and removals using immigration-death in discrete algorithmic time).
- ➎ Result: *trajectory-valued chain*, unconditioned S-I-R as equilibrium.

## From incidents to unconditioned epidemic trajectories (1/3)

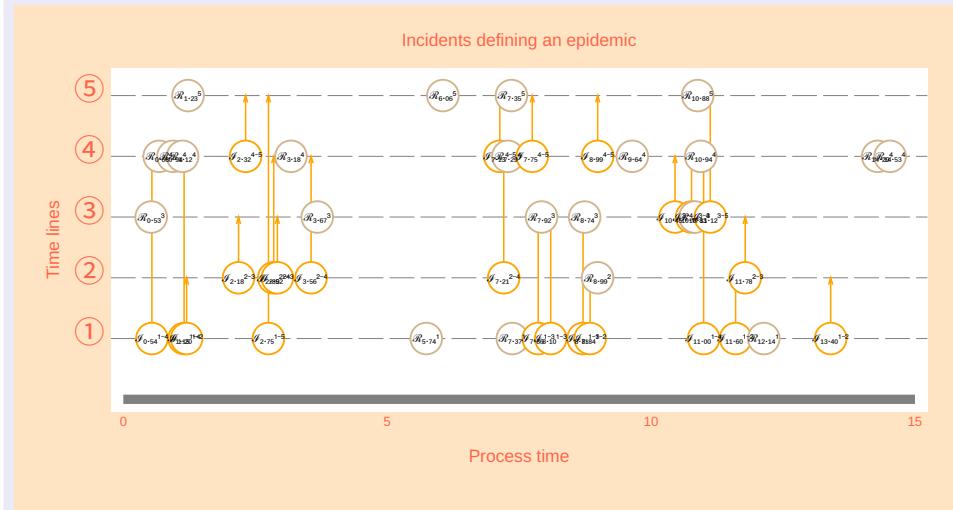


Figure 1: Light-orange circles denote potential infections (arrows point upwards to targets); light-brown circles denote potential removals.

## Perfect Epidemics

### └ An easier question

2025-02-04

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

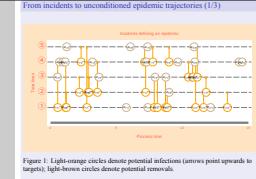
- ➊ That  $n, \alpha, \beta$  are known, removal times are observed, but unobserved infection times must be inferred
- ➋ Visualize  $n$  timelines, along which incidents are scattered
  - ➌ Incident times are the times of removals or infections
  - ➍ Potential infections, activated if timeline is infected and if designated target timeline is lowest uninfected timeline
  - ➎ Poisson point processes of appropriate rates yield an S-I-R epidemic
  - ➏ First step: evolve whole S-I-R trajectory in algorithmic time (alter potential infections and removals using immigration-death in discrete algorithmic time)
  - ➐ 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, we concentrate on the large number of nuisance parameters provided 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. This will allow us to exploit the classic connection for dynamically reversible Markov processes, between conditioning and restriction of state space.

## Perfect Epidemics

2025-02-04



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.

## From incidents to unconditioned epidemic trajectories (2/3)

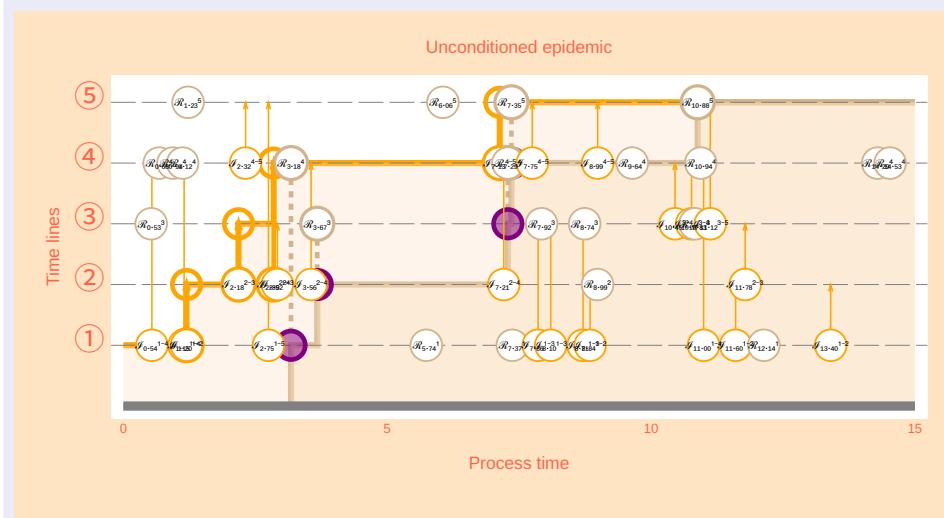
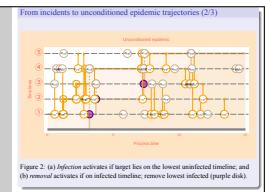


Figure 2: (a) *Infection* activates if target lies on the lowest uninfected timeline; and (b) *removal* activates if on infected timeline; remove lowest infected (purple disk).

## Perfect Epidemics

2025-02-04



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

## From incidents to unconditioned epidemic trajectories (3/3)

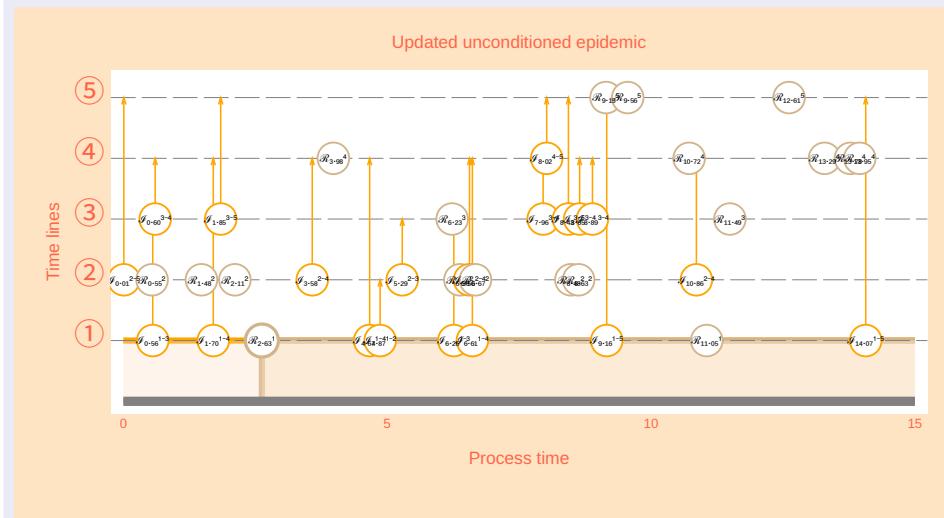
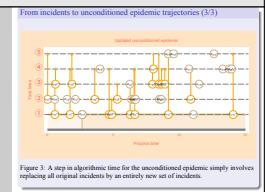


Figure 3: A step in algorithmic time for the unconditioned epidemic simply involves replacing all original incidents by an entirely new set of incidents.

## Perfect Epidemics

2025-02-04



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  $\tau$  are then (algorithmic-)*time-reversible*: so restriction to subset  $S$  of state-space (the *activated / conditioned* removals must occur precisely at specified set of process 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:
  - Replace removals ( $\mathcal{R}$ s);
  - Re-sample timelines (though not times) of  $\mathcal{R}$ s;
  - Replace infections ( $\mathcal{I}$ s).
- Re-express using continuously varying  $\tau$ . Process time runs over  $[0, T]$ .
  - For  $2nT < \tau < (2n+1)T$ , update old  $\mathcal{R}$ s with times in  $(0, \tau - 2nT)$ ;
  - For  $\tau = (2n+1)T$ , resample timelines (not times) of  $\mathcal{R}$ s;
  - For  $(2n+1)T < \tau < (2n+2)T$ , update old  $\mathcal{I}$ s 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” still holds.
- Crucially, step 2 ensures composition action is irreducible over  $S$ !  
(So equilibrium under conditioning is unique.)

## Perfect Epidemics

### Crucial technical point

2025-02-04

#### Crucial technical point

- Updates in algorithmic time  $\tau$  are then (algorithmic-)time-reversible: so restriction to subset  $S$  of state-space (the *activated / conditioned* removals must occur precisely at specified set of process 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:
  - Replace removals ( $\mathcal{R}$ s);
  - Re-sample timelines (though not times) of  $\mathcal{R}$ s;
  - Replace infections ( $\mathcal{I}$ s).
- Re-express using continuously varying  $\tau$ . Process time runs over  $[0, T]$ .
  - For  $2nT < \tau < (2n+1)T$ , update old  $\mathcal{R}$ s with times in  $(0, \tau - 2nT)$ ;
  - For  $\tau = (2n+1)T$ , resample timelines (not times) of  $\mathcal{R}$ s;
  - For  $(2n+1)T < \tau < (2n+2)T$ , update old  $\mathcal{I}$ s 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” will be equilibrated.
- 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)

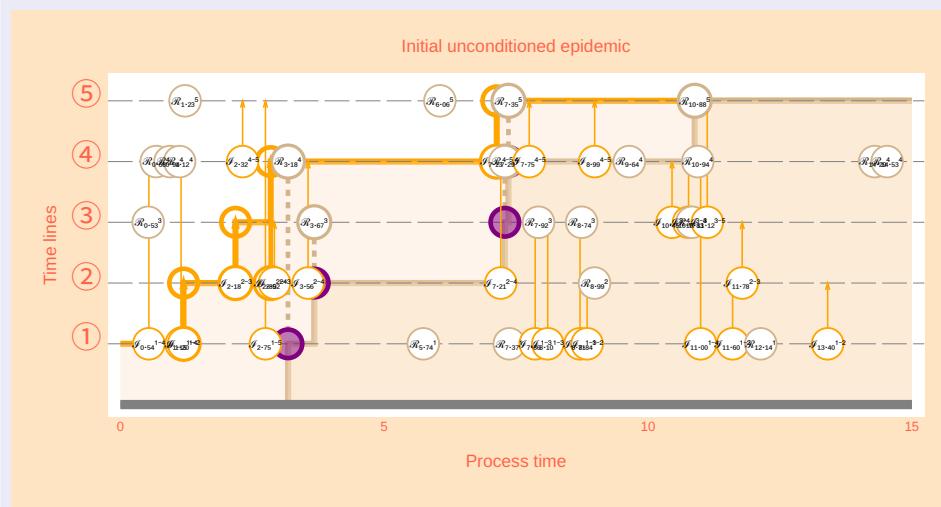
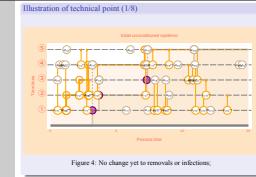


Figure 4: No change yet to removals or infections;

## Perfect Epidemics

2025-02-04



Here is the initial configuration again.

## Illustration of technical point (2/8)

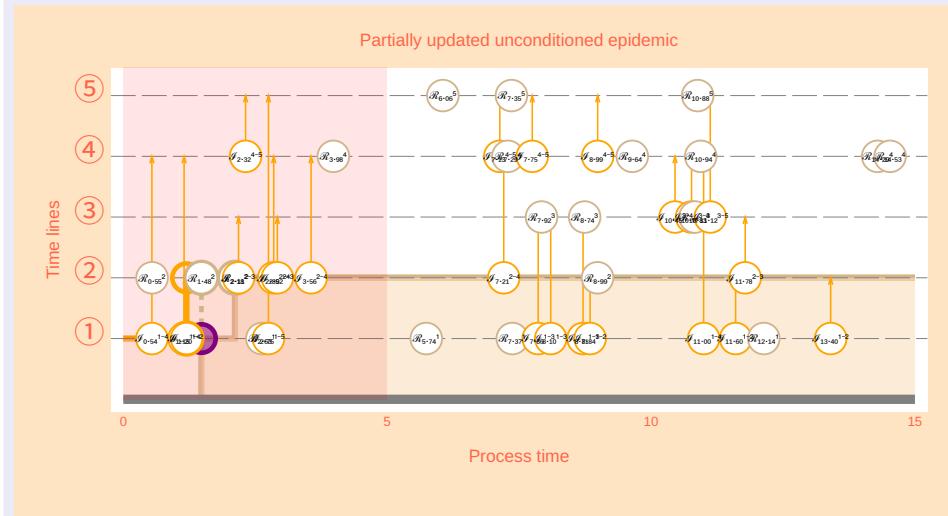
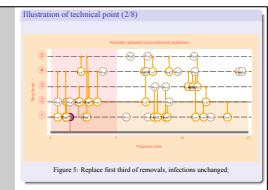


Figure 5: Replace first third of removals, infections unchanged;

## Perfect Epidemics

2025-02-04



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!

## Illustration of technical point (3/8)

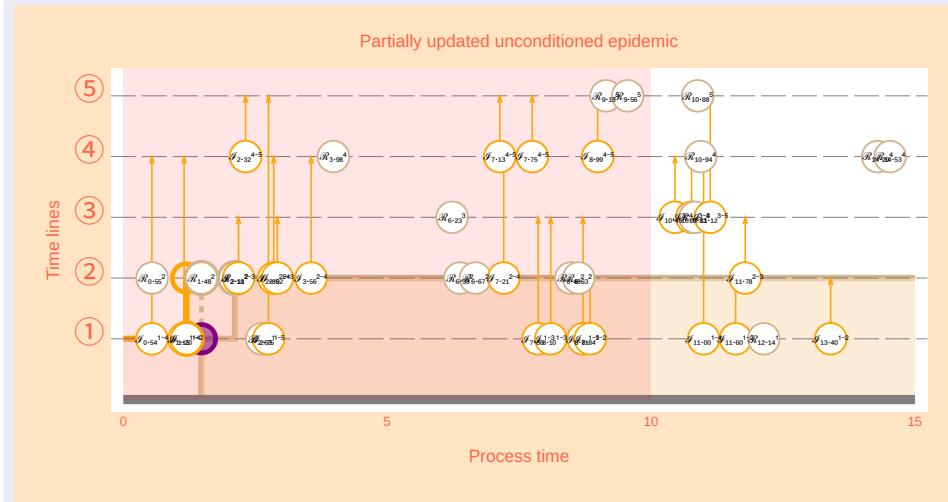
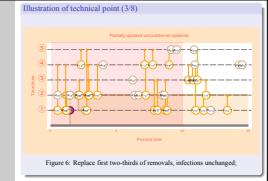


Figure 6: Replace first two-thirds of removals, infections unchanged;

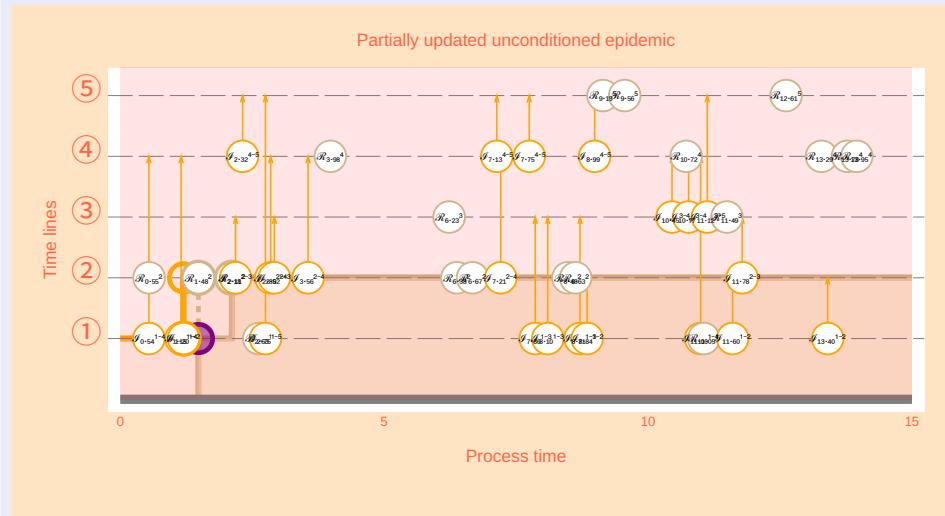
## Perfect Epidemics

2025-02-04



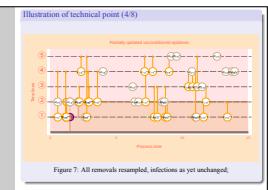
Here we are 2/3 of the way through dealing with removals.  
Flip backwards and then forwards one slide, and confirm that the light-orange nodes with vertical arrows (infections) are not changed at this stage.

## Illustration of technical point (4/8)



## Perfect Epidemics

2025-02-04



## Illustration of technical point (6/8)

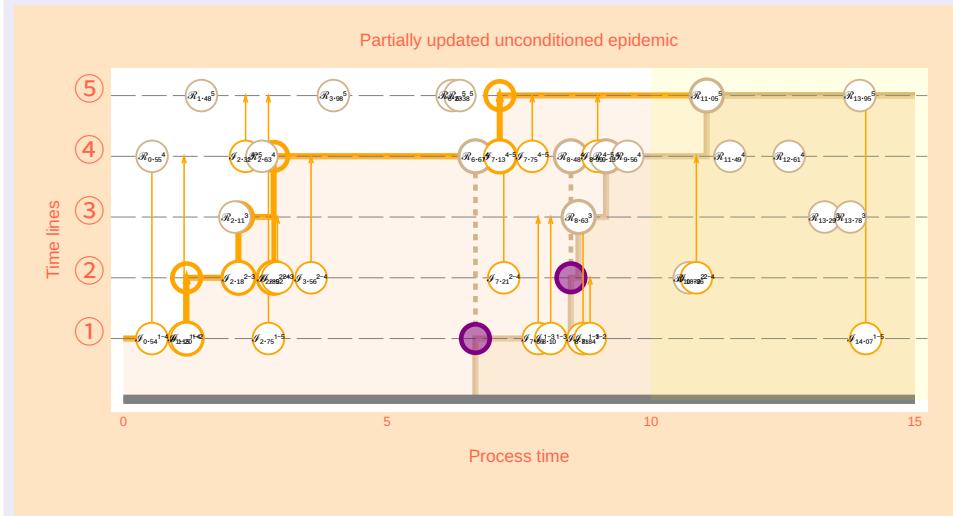
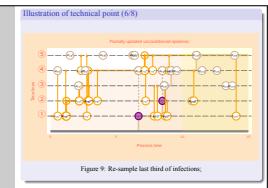


Figure 9: Re-sample last third of infections;

## Perfect Epidemics

2025-02-04



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.

## Illustration of technical point (7/8)

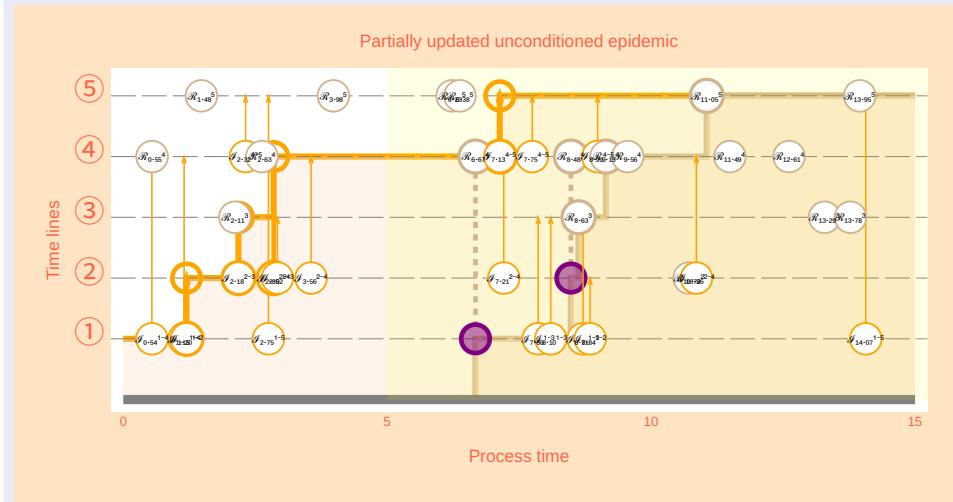
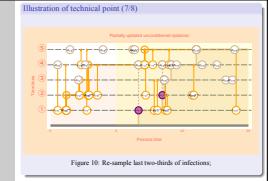


Figure 10: Re-sample last two-thirds of infections;

## Perfect Epidemics

2025-02-04



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

## Illustration of technical point (8/8)

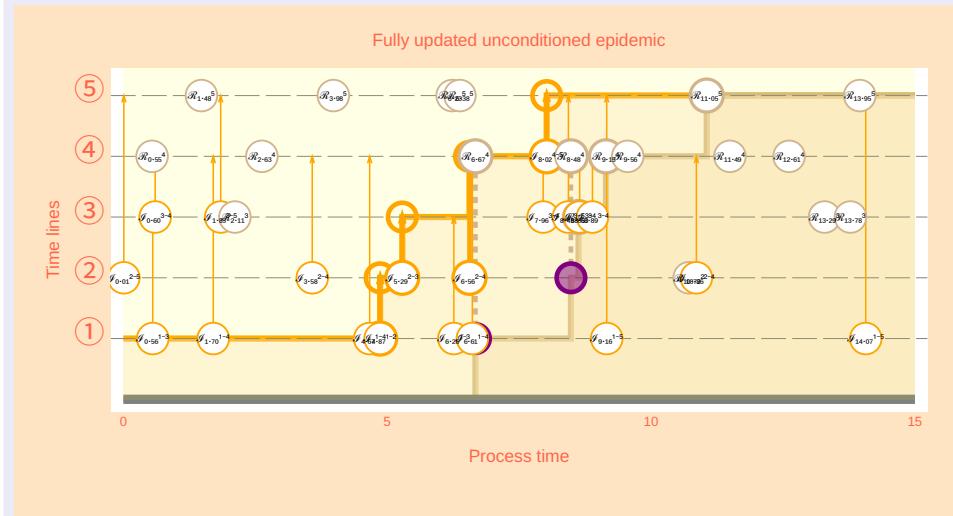


Figure 11: All infections now re-sampled.

### 3. Conditioning on observed removals

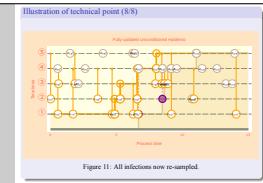
- The trajectory-valued chain is *dynamically reversible*, in *continuous algorithmic time*.
- Irreducibility is *vital* (otherwise equilibrium depends on starting point). Consequently:
  - conditioned removals must be able to change timeline (but not time of occurrence);
- Forbidding removal of observed 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** used to establish that monotonicity still works: *laziest feasible epidemic (LFE)* and *no-fly zone (NFZ)*.

## Perfect Epidemics

2025-02-04



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.

## Perfect Epidemics

2025-02-04

### └ 3. Conditioning on observed removals

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).
  - 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 used to establish that monotonicity still works: *laziest 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 often reveals the falsity of hidden assumptions!).
4. LFE and NFZ are useful vocabulary for arguing for monotonicity, but also computationally helpful.

## Initial conditioned epidemic

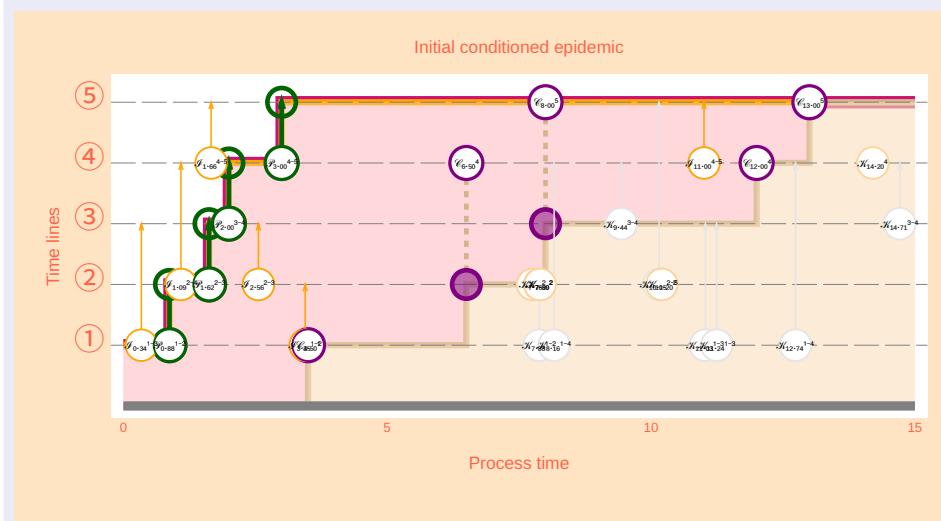
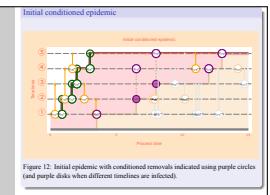


Figure 12: Initial epidemic with conditioned removals indicated using purple circles (and purple disks when different timelines are infected).

## Perfect Epidemics

2025-02-04



Here is a different initial epidemic, including *conditioned 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 conditioned removal timelines, accepting only those changes which do not de-activate the conditioned removals.
- But we need to take more care about whether or not we can take away or add in *infections*.

## Conditional epidemic update

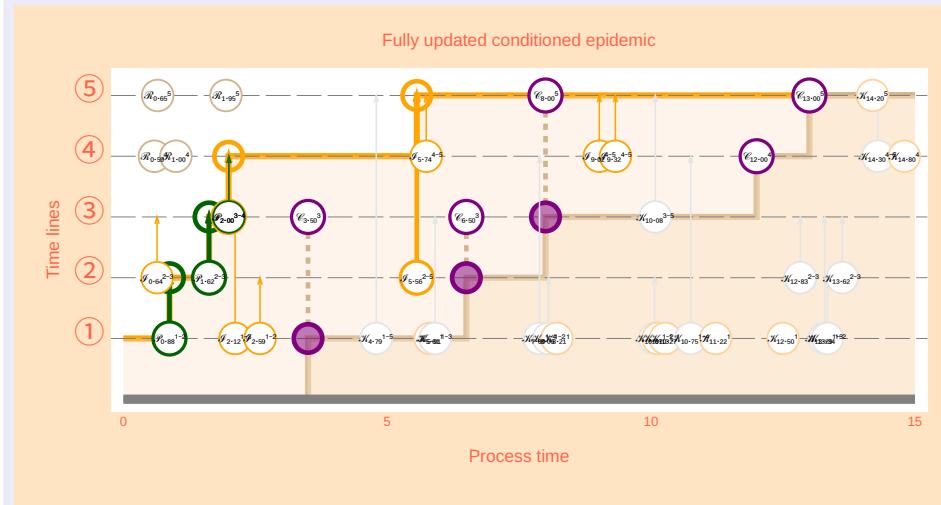
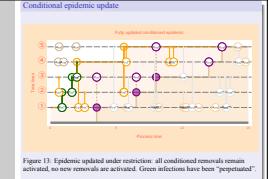


Figure 13: Epidemic updated under restriction: all conditioned removals remain activated, no new removals are activated. Green infections have been “perpetuated”.

## Perfect Epidemics

2025-02-04



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 targeting infected region is also rather simple:

- Always add them (because it won't make a difference).
- Delete them only if it is clear that this will not result in a conditioned removal falling out of the infected region.
- Implementation amounts to referring to a recursively-defined *laziest 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.
- Infections targeting the susceptible region are trickier, and we will discuss them briefly below.

## Last feasible epidemic (LFE)

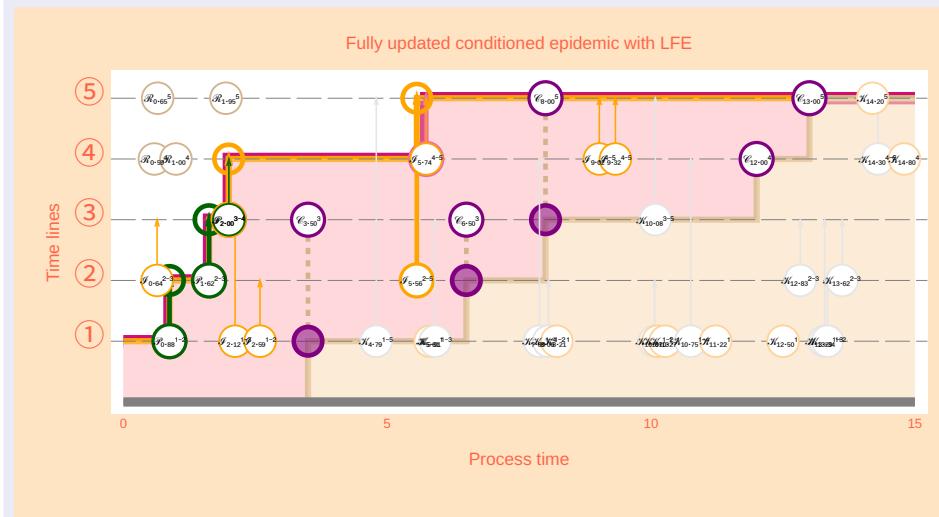
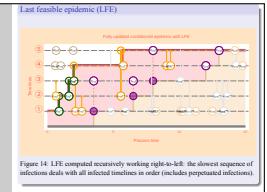


Figure 14: LFE computed recursively working right-to-left: the slowest sequence of infections deals with all infected timelines in order (includes perpetuated infections).

## Perfect Epidemics

2025-02-04



First we draw the **LFE**. As it is defined recursively working backwards in process time, it can be computed in entirety from the updated conditioned 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!

W S Kendall, S B Connor (Warwick, York)

Perfect Epidemics

4 February 2025

26 / 51

## No-fly zone (NFZ)

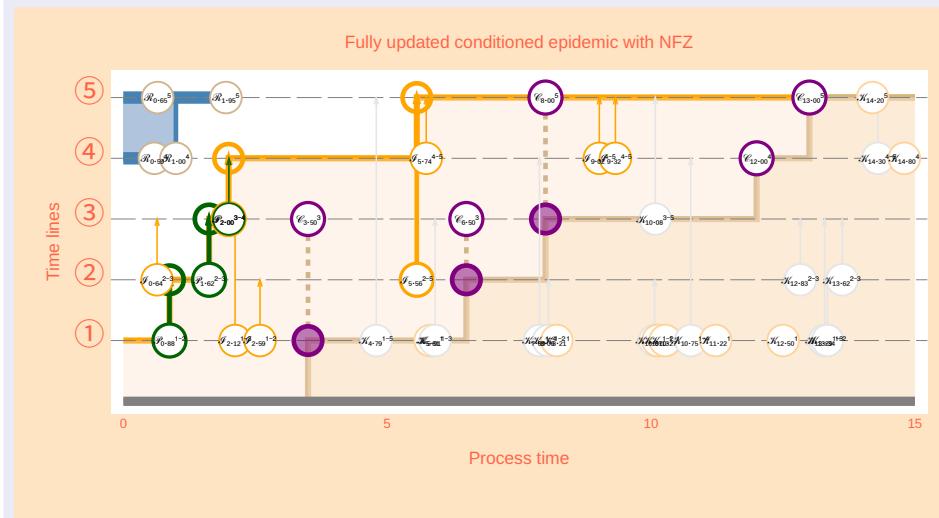
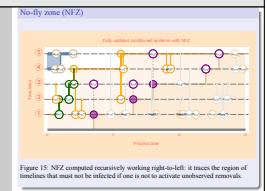


Figure 15: NFZ computed recursively working right-to-left: it traces the region of timelines that must not be infected if one is not to activate unobserved removals.

## Perfect Epidemics

2025-02-04



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!

W S Kendall, S B Connor (Warwick, York)

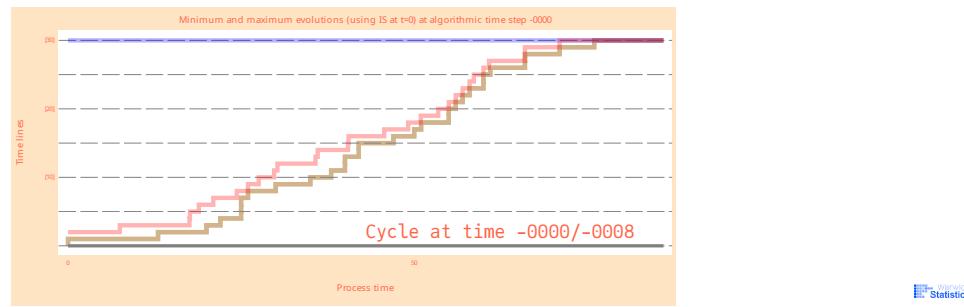
Perfect Epidemics

4 February 2025

27 / 51

## 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.



## So what?

- Why this emphasis on unobserved infections given fixed  $\alpha$  and  $\beta$ , when we need inference on R-number  $n/\beta$  for *unknown*  $\alpha$  and  $\beta$ ?
- Good question. But a re-weighting argument allows us to get (unbiased) estimates based on *different*  $\alpha$  and  $\beta$ . 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  $\alpha$  and  $\beta$ ;
  - construct steepest ascent algorithm (in effect, variant of Robbins-Monro stochastic optimization) to find *maximum a posterior* estimates of  $\alpha$  and  $\beta$ ;
  - or even, with some computational effort, compute the entire posterior joint density for  $\alpha$  and  $\beta$ !
- Finally: can we generalize to other suitable compartment models?

## Perfect Epidemics

2025-02-04

### └ 4. Example

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;
  - no further removals after last observed removal makes life easier)
- Code: [julia](#) (Bezanson *et al.*, 2017) produces a [perfect simulation](#) [GIF](#) or [MP4](#) yielding a draw from unobserved pattern of infections.

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, if the prior for the initial number of infectives is suitably monotonic.)

#### Technical note:

The language [julia](#) (Bezanson *et al.*, 2017), is chosen because it allows for rapid development by using an expressive type-based syntax (very useful for involved algorithms), and can deliver remarkably fast execution using very good “just in time” compilation techniques.

## Perfect Epidemics

2025-02-04

### └ So what?

So what?

- Why this emphasis on unobserved infections given fixed  $\alpha$  and  $\beta$ , when we need inference on R-number  $n/\beta$  for *unknown*  $\alpha$  and  $\beta$ ?
- Good question. But a re-weighting argument allows us to get (unbiased) estimates based on *different*  $\alpha$  and  $\beta$ . The perfect simulation provides exact simulation-based computation to integrate out pattern of unobserved infections.
- So next steps after SBC & WSK, 2024?
  - construct likelihood test statistic for specified  $\alpha$  and  $\beta$
  - construct steepest ascent algorithm (in effect, variant of Robbins-Monro stochastic optimization) to find *maximum a posterior* estimates of  $\alpha$  and  $\beta$ .
  - with some computational effort, compute the entire posterior joint density for  $\alpha$  and  $\beta$ !
- Finally: can we 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: a preliminary discussion is given in an appendix.

- Split population into a small number of interacting sub-populations;
- (For purposes of insight even if impractical.) Treat each individual as a sub-population on its own;
- Allow parameters to change at specific times of day or week;
- 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?**



Warwick Statistics

## Perfect Epidemics

### └ Conclusion

2025-02-04

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 have not yet found a closer connection.

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?**



## 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.
- Liggett, T.M. (1985) *Interacting particle systems*. Springer Verlag.

Warwick Statistics

## References II

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

Warwick Statistics

## References III

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
<a href="#">Terry Pratchett</a> Classic CFTP for a simple random walk	Luigi Novi <a href="#">Result of code written by WSK</a>
<a href="#">Diamond Princess</a> Epidemic CFTP images and animation	Alpsdake <a href="#">Result of code written by WSK</a>

## Previous instances of this talk

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

W S Kendall, S B Connor (Warwick, York)

Perfect Epidemics

4 February 2025

Warwick Statistics

33 / 51

## Appendix A: A “near-maximal” configuration

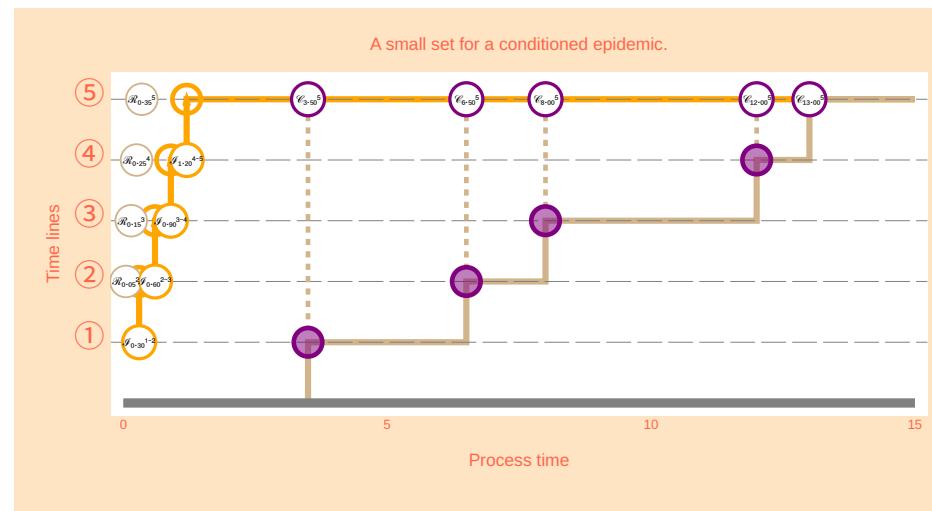


Figure 16: A conditioned epidemic in which all activated infections occur before time 3.0, also before smallest observed removal time.

W S Kendall, S B Connor (Warwick, York)

Perfect Epidemics

4 February 2025

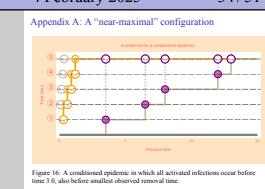
Warwick Statistics

34 / 51

## Perfect Epidemics

2025-02-04

### Appendix A: A “near-maximal” configuration



- 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  $2\delta$  is smaller than any of the conditioned removal times.
- All inactivated removals have to occur before their respective infections (in order to avoid activation).
- Variants on this construction can be used to show that CFTP for the perfect epidemic must coalesce in finite time;
- However to get a sensible stochastic bound on coalescence time would require a lot more work!

W S Kendall, S B Connor (Warwick, York)

Perfect Epidemics

4 February 2025

Warwick Statistics

35 / 51

## Appendix B: Updating a conditioned epidemic INCOMPLETE

### Perfect Epidemics

Appendix B: Updating a conditioned epidemic  
INCOMPLETE

We now work through the update of the conditioned epidemic in stages.

### Appendix B: Updating a conditioned epidemic INCOMPLETE

Seeking to convey the pixel-by-pixel nature of the update!

We now work through the update of the conditioned epidemic in stages.

### Initial conditioned epidemic (1/8)

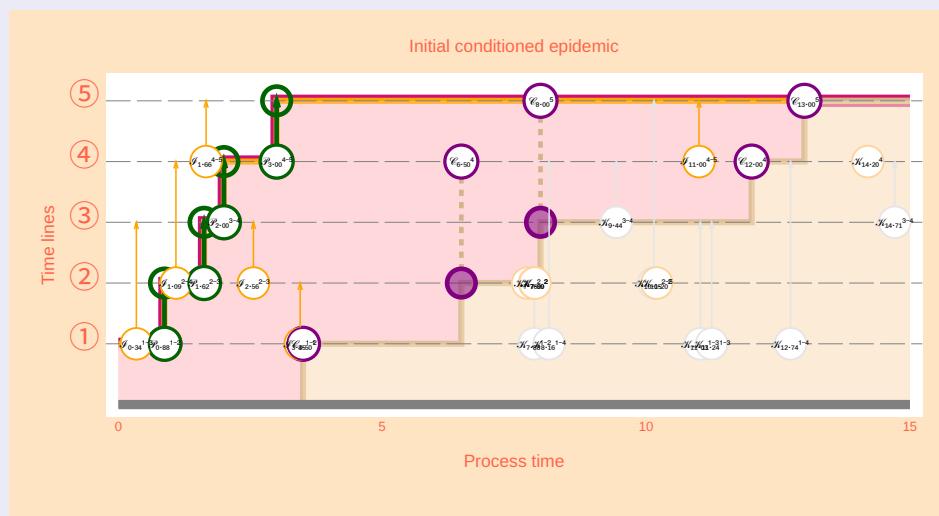
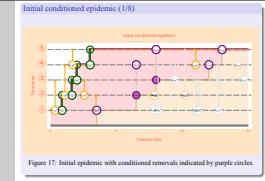


Figure 17: Initial epidemic with conditioned removals indicated by purple circles.

### Perfect Epidemics

2025-02-04



Here is the initial epidemic (actually itself produced using a conditioned update from a rather extreme initial state).

## Conditional epidemic (2/8)

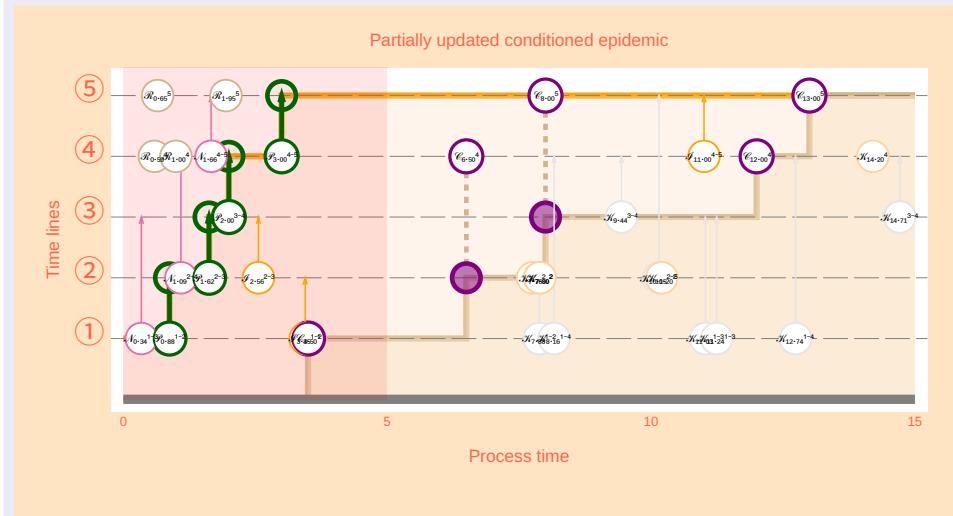


Figure 18: Replace first third of removals, infections unchanged;

## Perfect Epidemics

2025-02-04

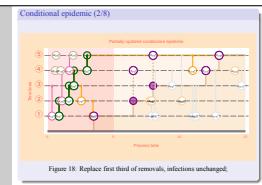


Figure 18: Replace first third of removals, infections unchanged.

Now we replace inactivated removals. A Conditioning does not affect this step. Here we are 1/3 of the way through dealing with removals.  
Not many changes arise from removal-replacement, because the initial epidemic is fast so leaves no room for inactivated removals.

## Conditional epidemic (3/8)

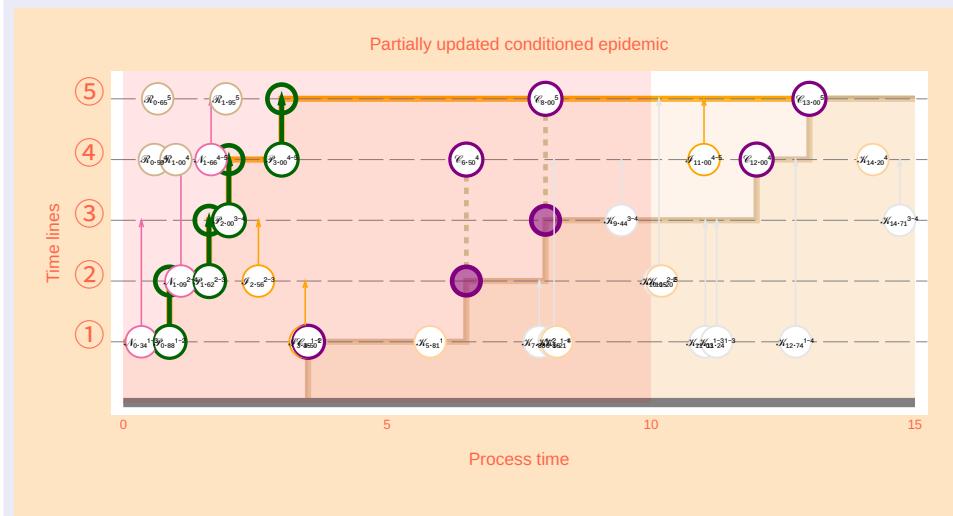


Figure 19: Replace second third of removals, infections unchanged;

## Perfect Epidemics

2025-02-04

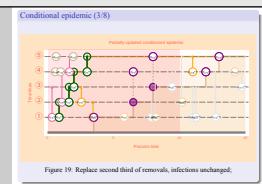


Figure 19: Replace second third of removals, infections unchanged.

Now 2/3 of the way through dealing with removals.

## Conditional epidemic (4/8)

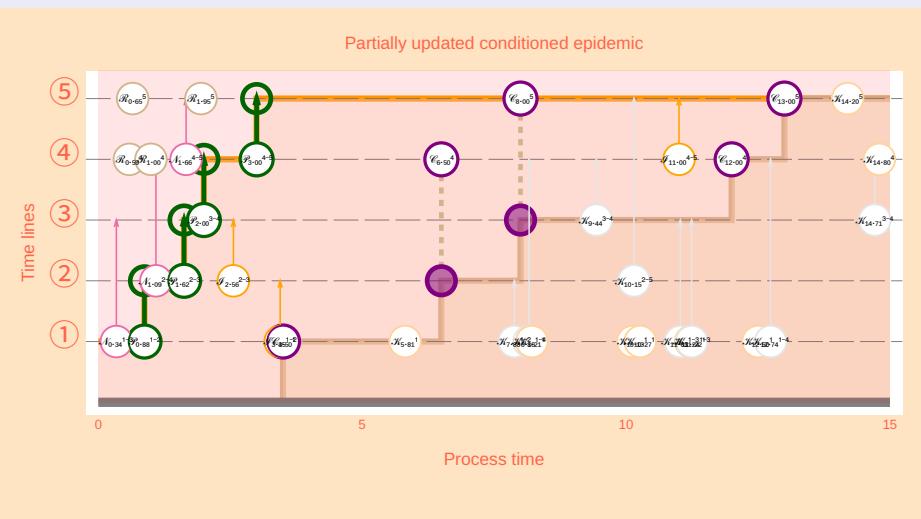
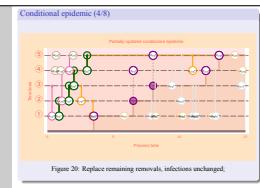


Figure 20: Replace remaining removals, infections unchanged;

## Perfect Epidemics

2025-02-04



Now all the way through dealing with removals.

## Conditional epidemic (5/8)

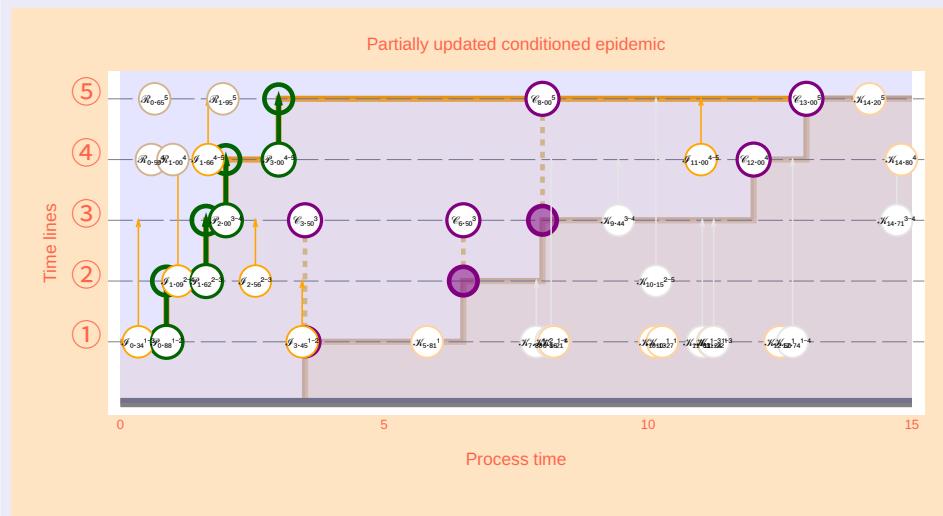
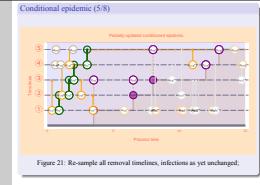


Figure 21: Re-sample all removal timelines, infections as yet unchanged;

## Perfect Epidemics

2025-02-04



At this stage we now re-sample the timelines for all removals.  
No nodes change time, but timelines for removal nodes are all re-sampled.  
**NEED TO ADD RE-SAMPLING OF INACTIVATED REMOVALS**

## Eventual conditioned epidemic after use of an innovation (6/8)

**Still to be done:** 1/3 of way through new infections, display current LFE and NFZ.

## Eventual conditioned epidemic after use of an innovation (7/8)

**Still to be done:** 2/3 of way through new infections, display current LFE and NFZ.

## Eventual conditioned epidemic after use of an innovation (8/8)

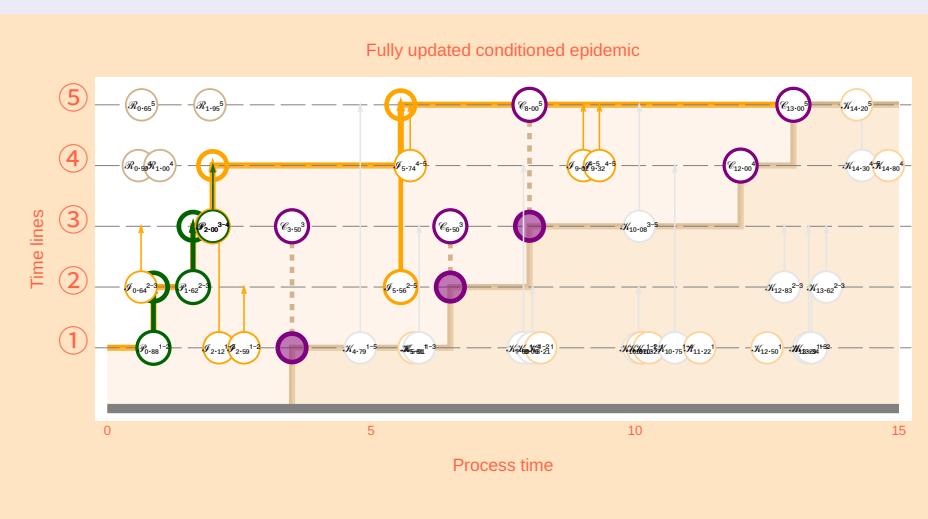
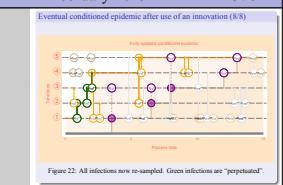


Figure 22: All infections now re-sampled. Green infections are “perpetuated”.

## Perfect Epidemics

2025-02-04



Finally the entire process time range is dealt with. In this conditioned case each addition or deletion of removal or infections is tested to see whether it would lead to a contravention of the conditioning.

**ADD NFZ AND LFE!**

## Appendix C: Naive approach to compartment models fails

- Consider a modification of this approach to the case of compartmentalized populations.
- Focus on the extreme case in which every individual  $j$  has infectivity (outgoing to individual  $k$ )  $\alpha_{j,k}$  and removal parameter  $\beta_j$ .
- Suppose the conditioning on removals is specifically about named individuals  $j$  being removed at specified times  $r_j$ ; suppose also there are no “occult” (unobserved) removals for any other individuals.
- This would apply, for example, in the case of the *Diamond Princess* if  $\alpha, \beta$  depended on age and location of cabin on the ship.

## Perfect Epidemics

2025-02-04

### Appendix C: Naive approach to compartment models fails

We expect to model  $\alpha_{j,k}$  and  $\beta_j$  in terms of a simple parametric model using explanatory variables that are observable *ab initio*.

Appendix C: Naive approach to compartment models fails

- Consider a modification of this approach to the case of compartmentalized populations.
- Focus on the extreme case in which every individual  $j$  has infectivity (outgoing to individual  $k$ )  $\alpha_{j,k}$  and removal parameter  $\beta_j$ .
- Suppose the conditioning on removals is specifically about named individuals  $j$  being removed at specified times  $r_j$ ; suppose also there are no “occult” (unobserved) removals for any other individuals.
- This would apply, for example, in the case of the *Diamond Princess* if  $\alpha, \beta$  depended on age and location of cabin on the ship.

## Timelines and incidents for the compartmental generalization

- Individuals no longer exchangeable, so S-I-R state space is unsuitable.
- Given population  $N \in \mathbb{N}$ , the ground space is a tuple of  $n$  timelines  $(i, [0, T])$ , one timeline per individual  $i$ , where  $T$  is the final time of observation of the epidemic.
- Typical element of state-space: a locally-finite point pattern of
  - infections  $\mathcal{I} = \mathcal{I}_{i,j}(t)$ , marked by timelines  $i$  at various times  $t$ , each infection marked by a target timeline  $j$  other than its mark  $k$ ;
  - (inactivated) removals  $\mathcal{R} = \mathcal{R}_i(t)$ , marked by timelines  $i$  at times  $t$ ;
  - conditioned removals  $\mathcal{C} = \mathcal{C}_i(t)$ , marked by timelines  $i$  at times  $t$ .
- Initial stipulations:
  - the  $\mathcal{I}$ s,  $\mathcal{R}$ s and  $\mathcal{C}$ s all occur at different times;
  - there is at most one  $\mathcal{C}$  per timeline;
  - for convenience, no  $\mathcal{I}$  or  $\mathcal{R}$  occur after a  $\mathcal{C}$  on the same timeline;
- Epidemic can be viewed as a union of intervals on different timelines;
  - intervals end at the  $\mathcal{C}$  in the timeline or at time  $T$ ;
  - intervals on initially infected timelines start at time 0;
  - intervals on eventually infected timelines start at the first time  $t$  an  $\mathcal{I}$  targets the timeline while marked by a timeline infected at  $t$ .

## Perfect Epidemics

2025-02-04

Timelines and incidents for the compartmental generalization

- Individuals no longer exchangeable; so S-I-R state space is unsuitable
- Given population  $N \in \mathbb{N}$ , the ground space is a tuple of  $n$  timelines  $(i, [0, T])$ , one timeline per individual  $i$ , where  $T$  is the final time of observation of the epidemic
- Typical element of state-space: a locally-finite point pattern of
  - infections  $\mathcal{I} = \mathcal{I}_{i,j}(t)$ , marked by timelines  $i$  at various times  $t$ , each infection marked by a target timeline  $j$  other than its mark  $k$ ;
  - (inactivated) removals  $\mathcal{R} = \mathcal{R}_i(t)$ , marked by timelines  $i$  at times  $t$ ;
  - conditioned removals  $\mathcal{C} = \mathcal{C}_i(t)$ , marked by timelines  $i$  at times  $t$ .
- Initial stipulations:
  - there is at most one  $\mathcal{C}$  per timeline;
  - for convenience, no  $\mathcal{I}$  or  $\mathcal{R}$  occur on the same timeline;
- Epidemic can be viewed as a union of intervals on different timelines;
  - intervals end at the  $\mathcal{C}$  in the timeline or at time  $T$ ;
  - intervals on initially infected timelines start at time 0;
  - intervals on eventually infected timelines start at the first time  $t$  an  $\mathcal{I}$  targets the timeline while marked by a timeline infected at  $t$ .

- “S-I-R state space”:  $\{(s, i, r) : s, i, r \in \mathbb{N}, s + i + j = n\}$ ;
- one timeline per individual;
- locally-finite point pattern of incidents, derived from several independent Poisson processes on  $[0, T]$  for each timeline  $i$ ;
  - intensity of  $\mathcal{I}_{i,j}$  is  $\alpha_{i,j}$ ;
  - intensity of  $\mathcal{R}_i$  is  $\beta_i$ ;
  - the  $\mathcal{C}_i$  are not random but stipulated by the conditioning.
- “different times” follows from independent Poisson process assumption. We censor  $\mathcal{I}$  or  $\mathcal{R}$  occurring after a  $\mathcal{C}$  on the same timeline, as they will play no part in the epidemic dynamics.
- the simple epidemic trajectory is replaced by a set! We can re-order individual so that (for example) conditionally removed timeline come first, and initially infected timelines take priority in the two categories of “conditionally removed” and “not conditionally removed”.

## Process dynamics

Recall that infections and removals *after* a conditioned removal have been censored out. A valid configuration must satisfy the following, derived from the process dynamics:

- ➊ initially infected timelines  $i$  possess no  $\mathcal{R}_i$ : if they possess a (single)  $\mathcal{C}_i(t)$  then they contribute  $(i, [0, t])$  to the epidemic, otherwise  $(i, [0, T])$ ;
- ➋ other timelines  $k$  only contribute if they possess  $\mathcal{I}_{k,i}(s)$  such that timeline  $k$  is infected at time  $s$ , in which case the earliest  $s$  is chosen and the contribution is  $(i, [s, t])$  when  $\mathcal{C}_i(t)$  is the conditioned removal of the timeline, otherwise  $(i, [s, T])$ ;
- ➌ every inactivated removal  $\mathcal{R}_i(t)$  satisfies  $t < s$ , where  $s$  is the time of first infection of the timeline  $i$ .
- ➍ every conditioned removal  $\mathcal{C}_i(t)$  is at the right  $t$  of an infected interval.

So each timeline is divided into a *susceptible interval* (empty if it is initially infected), an *infected interval* (empty if it is never infected), and a *removed interval* (empty if it has no conditioned removal).

## Dynamics in algorithmic time

This closely corresponds to the evolution of the S-I-R epidemic above, but does not resample the mark  $i$  for each conditioned removal  $\mathcal{C}_i$ :

- ➊ Remove all  $\mathcal{R}$ s, and re-sample inactivated removals on the susceptible intervals of each timeline. Recompute **NFZ** as a union of  $(i, [0, t_i])$  using the latest time  $t_i = t$  of re-sampled  $\mathcal{R}_i(t)$ .
- ➋ List in *time-reverse order* original infections together with sampled *new* candidate  $\tilde{\mathcal{I}}$ s in complements of the removed intervals.
- ➌ Work iteratively through this list. Would discarding original  $\mathcal{I}_{i,j}(t)$  result in failure to infect a conditioned removal? If so, **retain**  $\mathcal{I}_{i,j}(t)$  as *perpetuated infection*  $\mathcal{P}_{i,j}(t)$ , otherwise **discard**.
- ➍ Otherwise, at  $\tilde{\mathcal{I}}_{a,b}(u)$ , consider the latest update of **NFZ**.
  - ▶ If  $a$  infected at  $u$  and  $u$  is in  $b$  component of latest **NFZ**, **discard**  $\tilde{\mathcal{I}}_{a,b}(u)$ ;
  - ▶ If  $a$  is not infected at  $u$  but  $u$  lies on  $b$  component of latest **NFZ**, **update** **NFZ** by adding/replacing interval of **NFZ** at  $a$  using  $(a, [0, u])$ ;
  - ▶ Otherwise **retain**  $\tilde{\mathcal{I}}_{a,b}(u)$  as  $\mathcal{I}_{a,b}(u)$ .

As in S-I-R case, the conditioned epidemic is the unique equilibrium.

## Perfect Epidemics

2025-02-04

Process dynamics
Recall that infections and removals after a conditioned removal have been censored out. A valid configuration must satisfy the following, derived from the process dynamics:
<ul style="list-style-type: none"> <li>➊ initially infected timelines <math>i</math> possess no <math>\mathcal{R}_i</math>: if they possess a (single) <math>\mathcal{C}_i(t)</math> then they contribute <math>(i, [0, t])</math> to the epidemic, otherwise <math>(i, [0, T])</math>;</li> <li>➋ other timelines <math>k</math> only contribute if they possess <math>\mathcal{I}_{k,i}(s)</math> such that timeline <math>k</math> is infected at time <math>s</math>, in which case the earliest <math>s</math> is chosen and the contribution is <math>(i, [s, t])</math> when <math>\mathcal{C}_i(t)</math> is the conditioned removal of the timeline, otherwise <math>(i, [s, T])</math>;</li> <li>➌ every inactivated removal <math>\mathcal{R}_i(t)</math> satisfies <math>t &lt; s</math>, where <math>s</math> is the time of first infection of the timeline <math>i</math>.</li> <li>➍ every conditioned removal <math>\mathcal{C}_i(t)</math> is at the right <math>t</math> of an infected interval. So each timeline is divided into a <i>susceptible interval</i> (empty if it is initially infected), an <i>infected interval</i> (empty if it is never infected), and a <i>removed interval</i> (empty if it has no conditioned removal).</li> </ul>

1. a  $\mathcal{R}_i$  on an initially infected timeline would violate conditioning by being activated.
2. the first  $\mathcal{I}$  connecting from an infected timeline will infect the target timeline if it is not already infected.
3. inactivated removals simply occur as a Poisson point pattern on the initial uninfecte dinterval of each timeline not initially infected.

## Perfect Epidemics

2025-02-04

Dynamics in algorithmic time
This closely corresponds to the evolution of the S-I-R epidemic above, but does not resample the mark $i$ for each conditioned removal $\mathcal{C}_i$ :
<ul style="list-style-type: none"> <li>➊ Remove all <math>\mathcal{R}_i</math>s, and re-sample inactivated removals on the susceptible intervals of each timeline. Recompute <b>NFZ</b> as a union of <math>(i, [0, t_i])</math> using the latest time <math>t_i = t</math> of re-sampled <math>\mathcal{R}_i(t)</math>.</li> <li>➋ List in <i>time-reverse order</i> original infections together with sampled new candidate <math>\tilde{\mathcal{I}}</math>s in complements of the removed intervals.</li> <li>➌ Work iteratively through this list. Would discarding original <math>\mathcal{I}_{i,j}(t)</math> result in failure to infect a conditioned removal? If so, <b>retain</b> <math>\mathcal{I}_{i,j}(t)</math> as <i>perpetuated infection</i> <math>\mathcal{P}_{i,j}(t)</math>, otherwise <b>discard</b>.</li> <li>➍ Otherwise, at <math>\tilde{\mathcal{I}}_{a,b}(u)</math>, consider the latest update of <b>NFZ</b>.           <ul style="list-style-type: none"> <li>If <math>a</math> infected at <math>u</math> and <math>u</math> is in <math>b</math> component of latest <b>NFZ</b>, <b>discard</b> <math>\tilde{\mathcal{I}}_{a,b}(u)</math>;</li> <li>If <math>a</math> is not infected at <math>u</math> but <math>u</math> lies on <math>b</math> component of latest <b>NFZ</b>, <b>update</b> <b>NFZ</b> by adding/replacing interval of <b>NFZ</b> at <math>a</math> using <math>(a, [0, u])</math>;</li> <li>Otherwise <b>retain</b> <math>\tilde{\mathcal{I}}_{a,b}(u)</math> as <math>\mathcal{I}_{a,b}(u)</math>.</li> </ul> </li> </ul>
As in S-I-R case, the conditioned epidemic is the unique equilibrium.

The S-I-R epidemic resamples the marks of  $\mathcal{C}$ s to maintain irreducibility. The conditioning here means irreducibility does not fail.

1. Use Poisson process of intensity  $\beta_i$  on susceptible interval of timeline  $i$ . The **NFZ** will be iteratively updated in what follows.
2. Use a Poisson process of intensity  $\alpha_{a,b}$  to generate the candidate  $\tilde{\mathcal{I}}_{a,b}$ .
3. When computing whether removal of an infection results in uninfected conditioned removals, use current pattern of infections. The S-I-R case uses **LFE**; the necessary reformulation in set-theoretic terms means more computation is required here. (But note the transition from  $\mathcal{I}_{i,j}(t)$  to  $\mathcal{P}_{i,j}(t)$  can occur only if timeline  $b$  transits from susceptible to infected at time  $t$ , and even then only if no subsequent infection happens early enough!).
4. Note how **NFZ** grows in extent as this iteration proceeds, so as to prohibit infection of  $\mathcal{R}$ s previously encountered.

## Requirements for monotonicity

For CFTP we need to know that, for coupled iterations (using the same pattern of innovations of new  $\mathcal{I}$ s and  $\mathcal{R}$ s), if two variants are started so that the infected region of one contains the other, then this persists through development of the algorithmic time.

It would suffice to prove two technical results:

- ① If the old infected region of one contains the other, then the **NFZ** of the one is contained in the **NFZ** of the other;
- ② If the old infected region of one contains the other, and the **NFZ** of the one is contained in the **NFZ** of the other, then the new infected region of the one is contained in the other.

Then CFTP would make sense, and it would only be necessary to show that accessibility of a set of near-maximal configurations guarantees eventual coalescence.

## Perfect Epidemics

2025-02-04

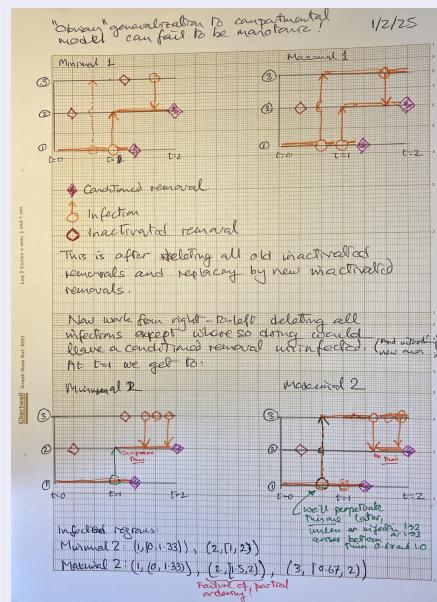
### Requirements for monotonicity

For CFTP we need to know that, for coupled iterations (using the same pattern of innovations of new  $\mathcal{I}$ s and  $\mathcal{R}$ s), if two variants are started so that the infected region of one contains the other, then this persists through development of the algorithmic time.  
It would suffice to prove two technical results:  
 ① If the old infected region of one contains the other, then the **NFZ** of the one is contained in the **NFZ** of the other;  
 ② If the old infected region of one contains the other, and the **NFZ** of the one is contained in the **NFZ** of the other, then the new infected region of the one is contained in the other.  
Then CFTP would make sense, and it would only be necessary to show that accessibility of a set of near-maximal configurations guarantees eventual coalescence.

Unfortunately, although

1. the first result can be proved used a case-by-case analysis;
2. **nevertheless the second result is not true**: see next slide for a counterexample!

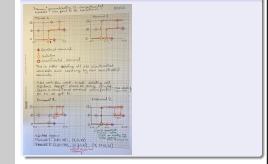
## Counterexample to monotonicity



## Perfect Epidemics

2025-02-04

### Counterexample to monotonicity



There is a counterexample for a population of size 3:

- in the minimal case only timelines 1 and 2 are infected, and on the second step it turns out that the  $1 \rightarrow 2$  infection must be perpetuated;
- however in the maximal case there is an early  $1 \rightarrow 3$  infection and a late  $3 \rightarrow 2$  infection which survives to step 2 with a minor modification of the  $3 \rightarrow 2$  infection time: the presence of this means the  $1 \rightarrow 2$  infection is *not* perpetuated, but it turns out (because there happen to be no suitable new infections) that the  $1 \rightarrow 3$  infection must be perpetuated.

We find that the Minimal 2 infection region is not a subset of the Maximal 2 infection region! (though a renumbering of the Minimal 2 population *would* be a subset – not a surprise given our previous results!).

**Question:** is there a way round this using the CFTP “crossover” trick?

## Other technical information

### Software used in computations

<i>Software</i>	<i>Version</i>	<i>Branch</i>	<i>Last commit</i>
quarto	1.6.39	—	
Running under julia	1.11.3	—	
EpidemicsCFTP	2.2.492	main	Thu Jan 23 20:50:07 2025
EpidemicsUtilities	0.1.2.156	main	Thu Jan 23 12:06:14 2025
This quarto script	2.2.621	Wilfrid-2025-01-30-compartment	Mon Feb 3 18:33:41 2025

### Revision notes

These notes were produced from `PerfectEpidemics.qmd`:

---

Version: 2.2.623 [Wilfrid-2025-01-30-compartment]  
Author: Wilfrid Kendall <[w.s.kendall@warwick.ac.uk](mailto:w.s.kendall@warwick.ac.uk)>  
Date: Tue Feb 4 18:19:36 2025 +0000  
Summary: See what this does!

---

