Perfect Epidemics

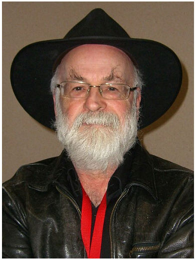
Applied Probability Seminar
Department of Statistics, University of Warwick

W S Kendall   S B Connor

Warwick, York

17 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;

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!

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*":

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;

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);

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.

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).

Classic CFTP for a simple random walk (I)

- Consider a simple random walk on $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

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**

Classic CFTP for a simple random walk (I)

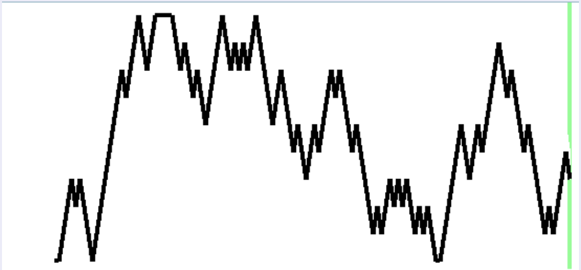
- Consider a simple random walk on $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.
  - $\mathbb{P}[+1 \text{ jump}] = p \in (0, 1)$, while $\mathbb{P}[-1 \text{ jump}] = 1-p$, **except that**
  - at state 9 replace $+1$ jump by staying still, **and**

Classic CFTP for a simple random walk (I)

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

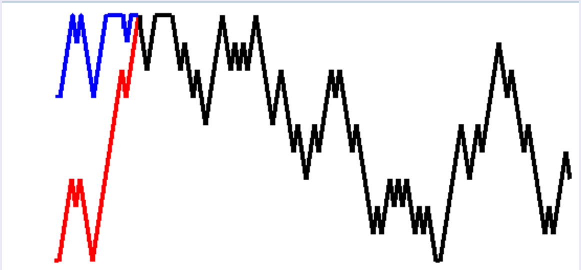
Classic CFTP for a simple random walk (I)

- Consider a simple random walk on $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.
  - $\mathbb{P}\left[+1 \text{ jump}\right] = p \in (0, 1)$, while $\mathbb{P}\left[-1 \text{ jump}\right] = 1-p$, **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.



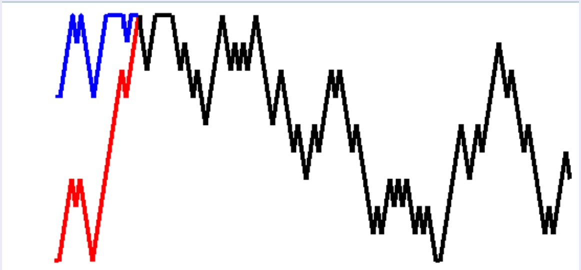
Classic CFTP for a simple random walk (I)

- Consider a simple random walk on $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.
  - $\mathbb{P}\left[+1 \text{ jump }\right] = p \in (0, 1)$, while $\mathbb{P}\left[-1 \text{ jump }\right] = 1-p$, **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;



Classic CFTP for a simple random walk (I)

- Consider a simple random walk on $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.
  - $\mathbb{P}\left[+1 \text{ jump}\right] = p \in (0, 1)$, while $\mathbb{P}\left[-1 \text{ jump}\right] = 1-p$, **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.

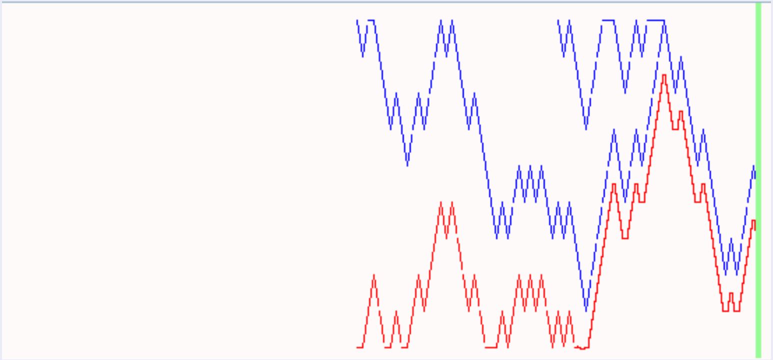


Classic CFTP for a simple random walk (I)

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

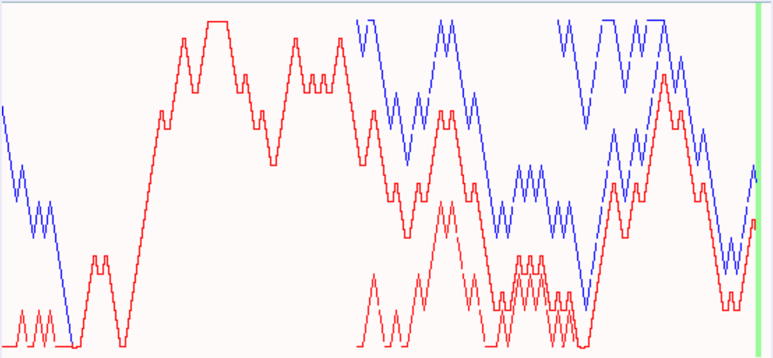
Classic CFTP for a simple random walk (I)

- 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.



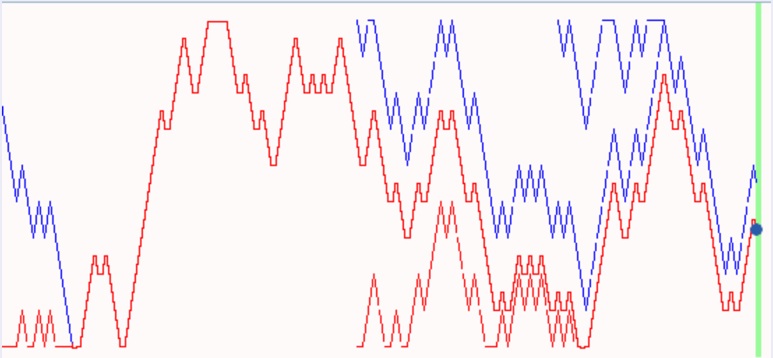
Classic CFTP for a simple random walk (I)

- 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.



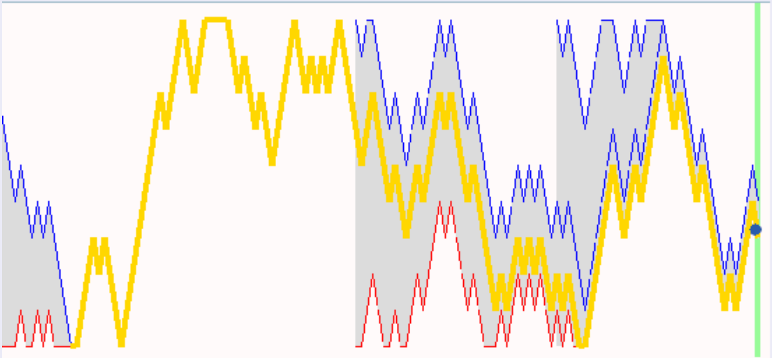
Classic CFTP for a simple random walk (I)

- 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.



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!



2: A short section on some theory about CFTP

- What about cases where monotonicity fails? or there isn't a sensible "maximal" process? Ideas (WSK, 1998):

2: A short section on some 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;

Warwick Statistics

2: A short section on some 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).

2: A short section on some 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

2: A short section on some 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);

2: A short section on some 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);

2: A short section on some 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).

2: A short section on some 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);

Warwick Statistics

2: A short section on some 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.)

Warwick Statistics

3: 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 $s + i + r = n$):

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

Warwick Statistics

3: 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	$\beta\, I\,.$

3: 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.

The first question asked about a new epidemic

"What is the R-number?"

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.

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

4: 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;

4: 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;

Warwick Statistics

4: 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);

Warwick Statistics

4: 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?

4: 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?

5: 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)

5: 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.

5: 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:

5: 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;

5: 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.

5: 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.

5: 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: 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.

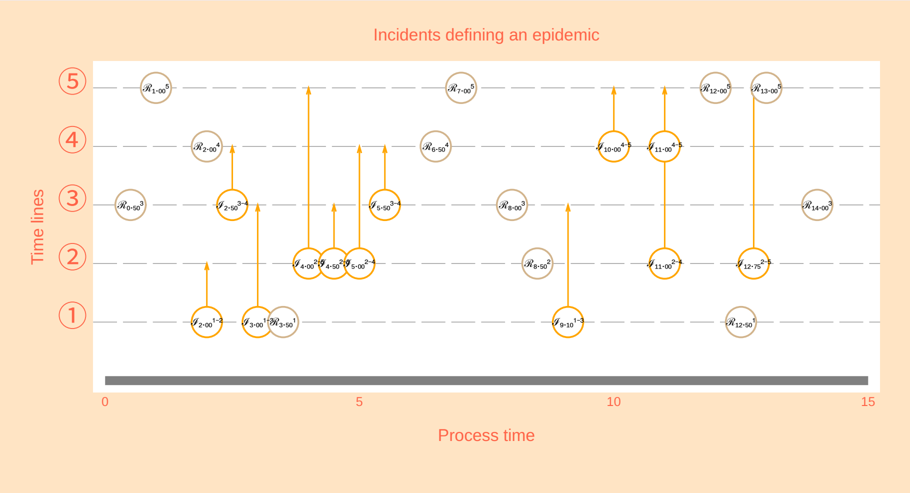


Figure 1: Light-orange circles denote potential infections (arrows point upwards to targets); light-brown circles denote potential removals.

From incidents to unconditioned epidemic trajectories (2/3)

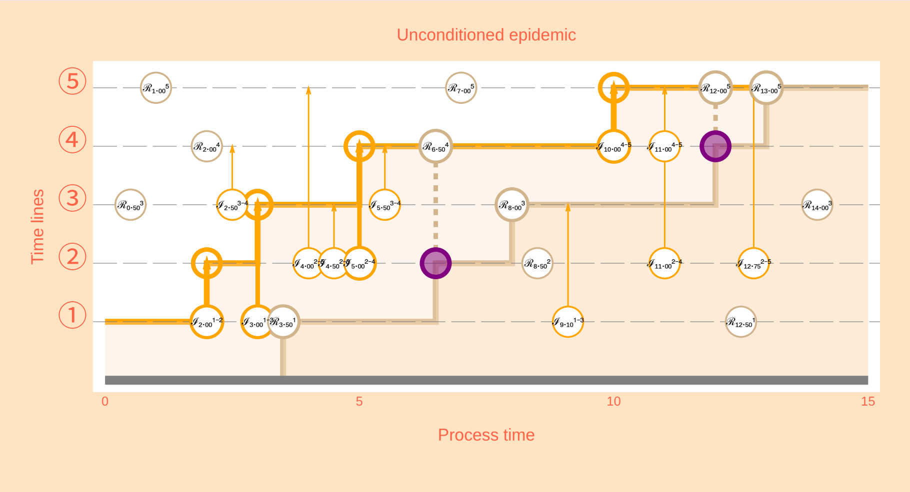


Figure 2: Activate (a) *infection* if target on lowest uninfected timeline; (b) *removal* if in infected region, then remove lowest infected (purple disk if different timeline).

From incidents to unconditioned epidemic trajectories (3/3)

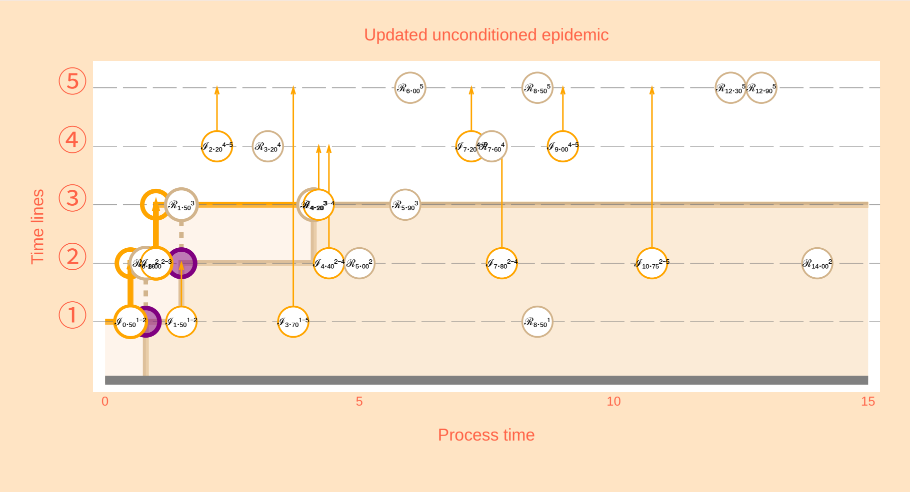


Figure 3: A step in algorithmic time for the unconditioned epidemic simply involves replacing the original incidents by a new set of incidents.

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.

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:

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);

Warwick Statistics

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;

Warwick Statistics

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).

Warwick Statistics

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 $[0, T]$.

Warwick Statistics

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 $(0, \tau - 2nT)$;

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;

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 $((2n+2)T - \tau, T)$.

Warwick Statistics

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.

Warwick Statistics

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.

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.)

Illustration of technical point (1/8)

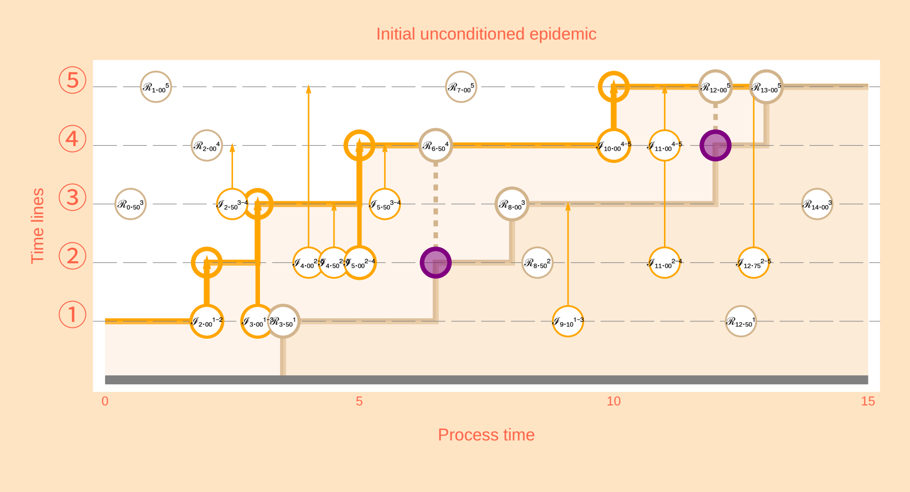


Figure 4: No change yet to removals or infections;

Illustration of technical point (2/8)

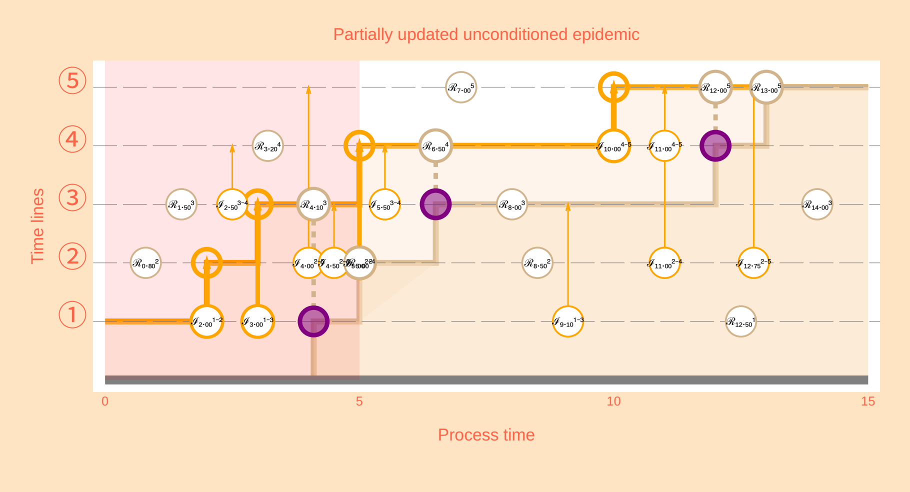


Figure 5: Replace first third of removals, infections unchanged;

Illustration of technical point (3/8)

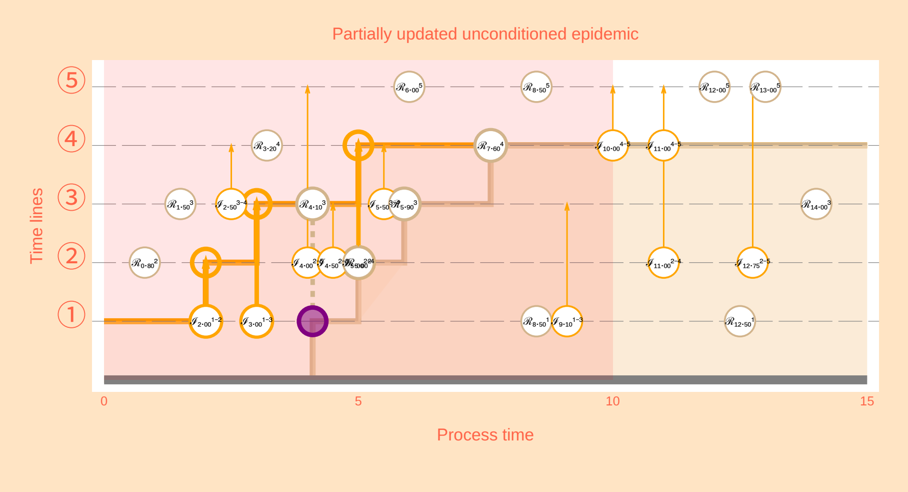


Figure 6: Replace first two-thirds of removals, infections unchanged;

Illustration of technical point (4/8)

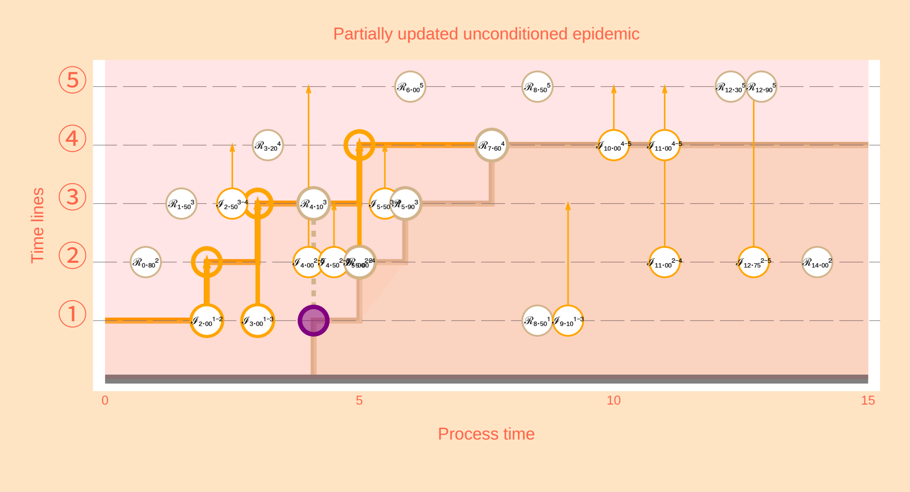


Figure 7: All removals resampled, infections as yet unchanged;

Illustration of technical point (5/8)

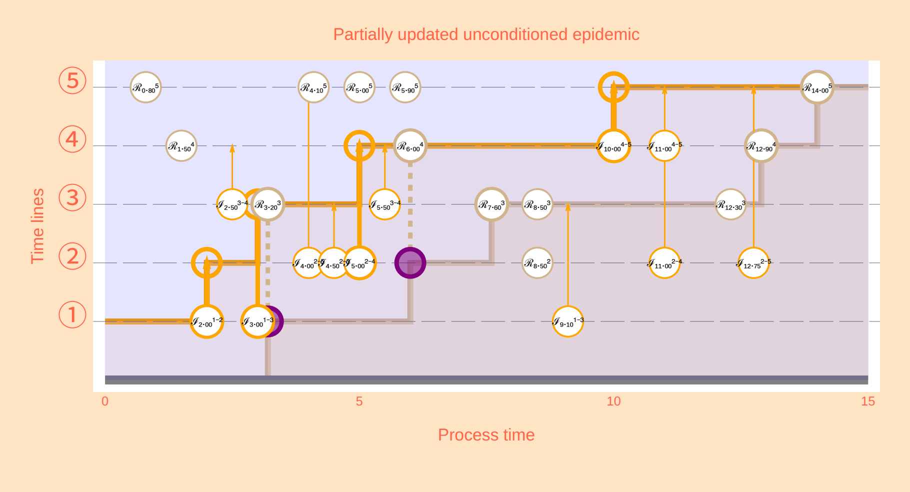


Figure 8: Re-sample all removal timelines, infections as yet unchanged;

Illustration of technical point (6/8)

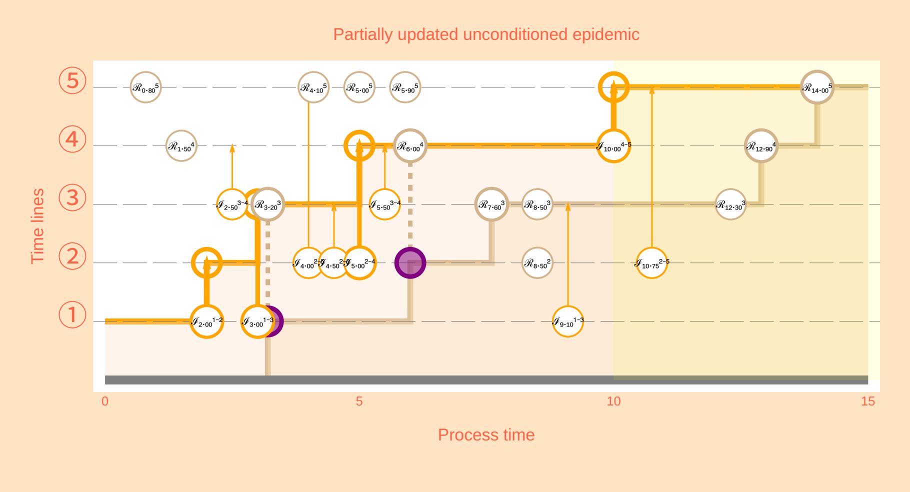


Figure 9: Re-sample last third of infections;

Illustration of technical point (7/8)

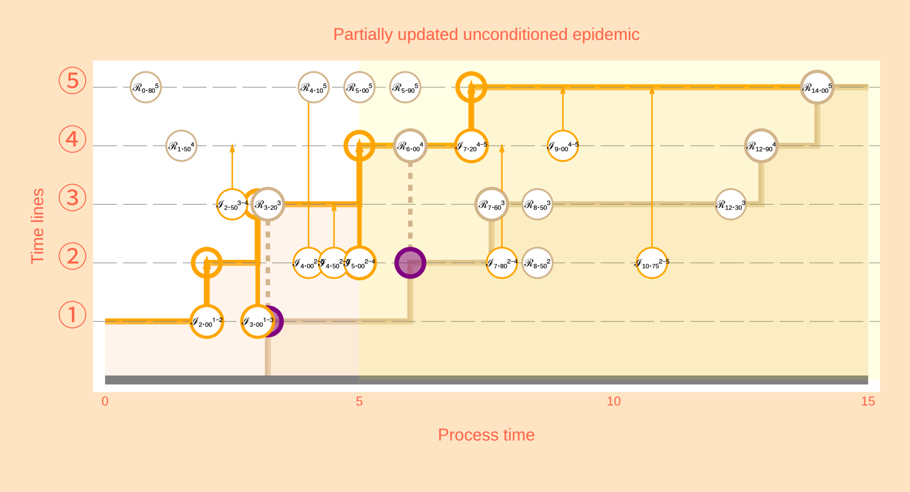


Figure 10: Re-sample last two-thirds of infections;

Illustration of technical point (8/8)

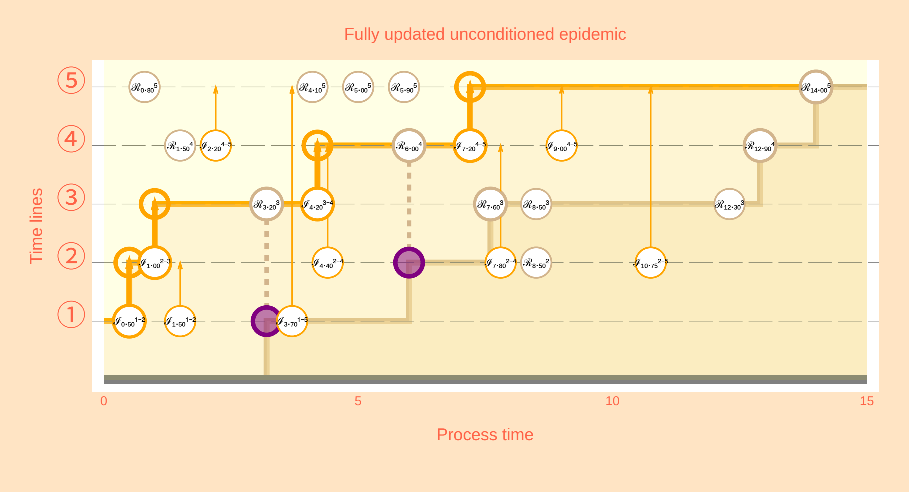


Figure 11: All infections now re-sampled.

Conditioning on observed removals

- The trajectory-valued chain is *dynamically reversible*, in *continuous algorithmic time*.

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:

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);

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:

Warwick Statistics

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;

Warwick Statistics

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;

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;

Warwick Statistics

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.

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?

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).

Warwick Statistics

Initial conditional epidemic

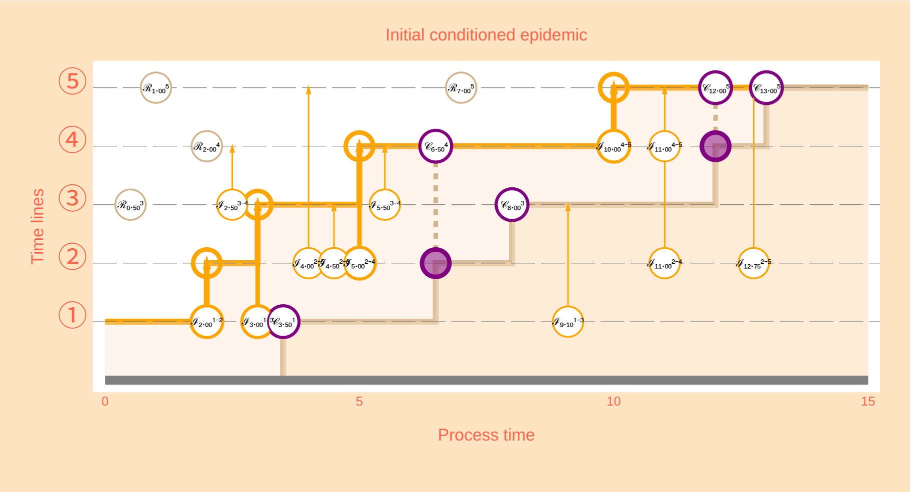


Figure 12: Initial epidemic with conditioned removals indicated using purple circles (and purple disks when different timelines are infected).

Conditional epidemic update

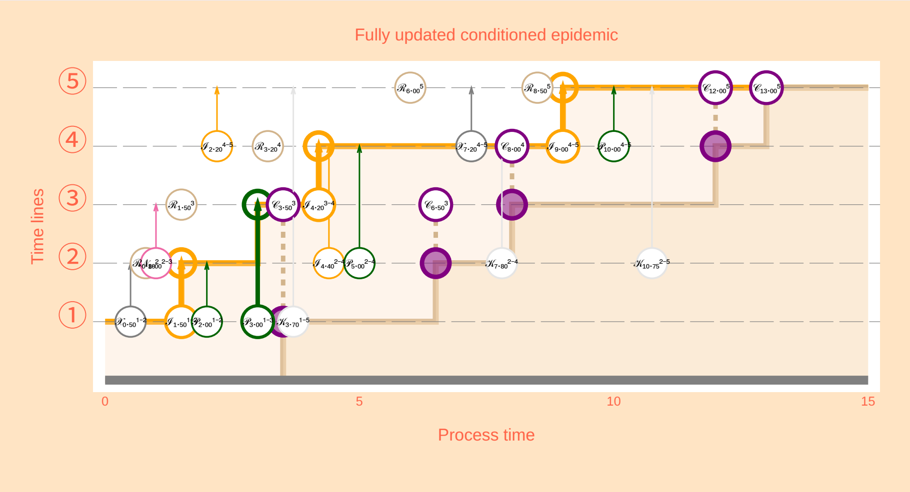


Figure 13: Epidemic updated under restriction: all conditioned removals remain activated, no new removals are activated. Green infections have been "perpetuated".

Last feasible epidemic (LFE)

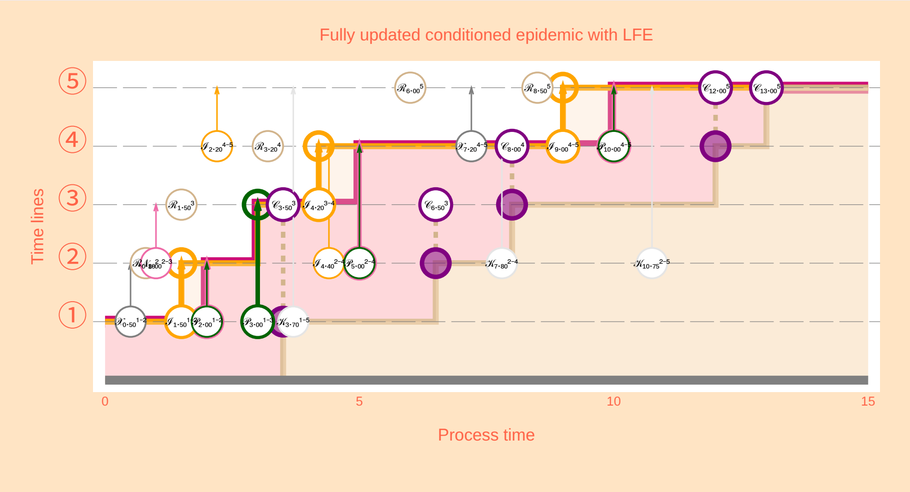


Figure 14: LFE computed recursively working right-to-left: the slowest sequence of infections deals with all infected timelines in order (includes perpetuated infections).

No-fly zone (NFZ)

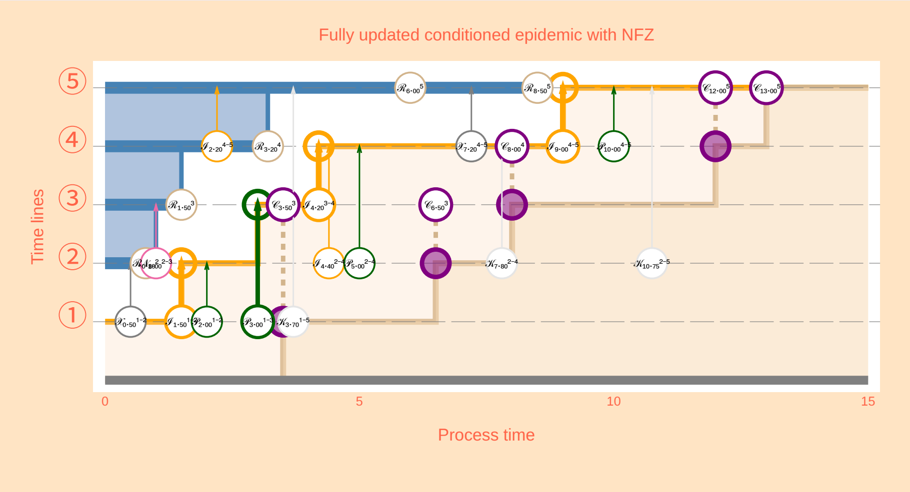


Figure 15: NFZ computed recursively working right-to-left: it traces the region of timelines that must not be infected if one is not to activate unobserved removals.

Example

- Smallpox outbreak in a closed community of 120 individuals in Abakaliki, Nigeria (much studied! see page 125 of Bailey, 1975).

Warwick Statistics

Example

- Smallpox outbreak in a closed community of 120 individuals in Abakaliki, Nigeria (much studied! see page 125 of Bailey, 1975).
- **Assume**

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;

Example

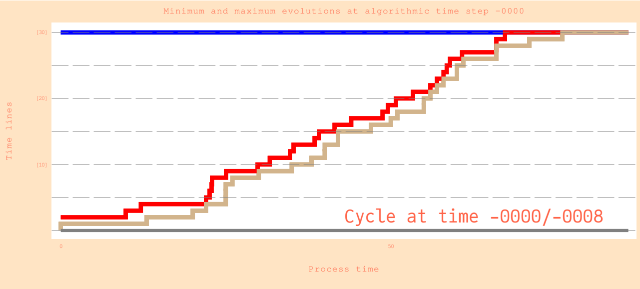
- Smallpox outbreak in a closed community of 120 individuals in Abakaliki, Nigeria (much studied! see page 125 of Bailey, 1975).
- **Assume**
  - first observed removal is also the first removal: under a plausible improper prior we can then deduce what is the distribution of infectives I_0 at time 0;
  - *all* removals are recorded;

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).

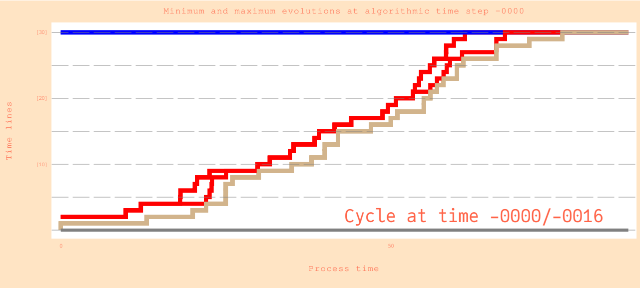
Example

- 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 can construct a perfect simulation GIF yielding a draw from unobserved pattern of infections.



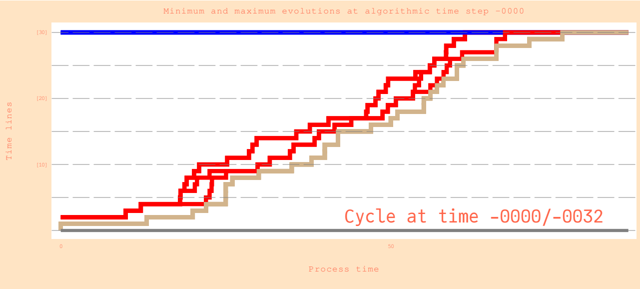
Example

- 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 can construct a perfect simulation GIF yielding a draw from unobserved pattern of infections.



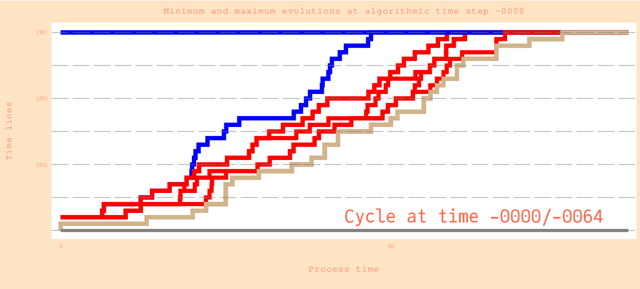
Example

- 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 can construct a perfect simulation GIF yielding a draw from unobserved pattern of infections.



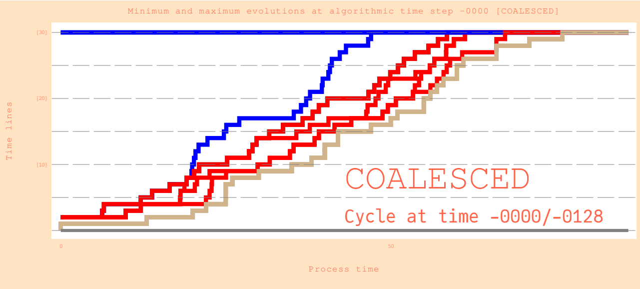
Example

- 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 can construct a perfect simulation GIF yielding a draw from unobserved pattern of infections.



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 can construct a perfect simulation GIF yielding a draw from unobserved pattern of infections.



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?

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.

Warwick Statistics

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)

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;

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;

Warwick Statistics

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!

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?

Conclusion

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

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.

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).

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.

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.

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.

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	
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

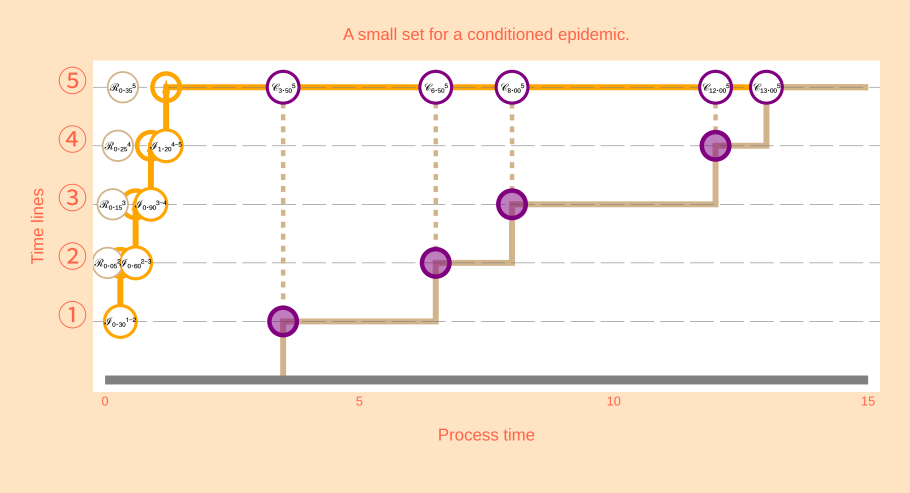


Figure 16: A conditional epidemic in which all activated infections occur before time 3.0, also before smallest observed removal time.

Other technical information

Software versions

Software used in computations:

Software	*Version*	*Branch*	*Date of last commit*
Quarto	`1.6.39`	—	
Running under `julia`	`1.11.2`	—	
Module `EpidemicsCFTP`	`2.2.487`	`main`	`Fri Jan 10 20:17:28 2025`
Module `EpidemicsUtilities`	`0.1.2.154`	`main`	`Wed Jan 15 13:23:36 2025`
This Quarto script	`2.2.589`	`master`	`Fri Jan 17 16:59:04 2025`

Revision history

These notes were produced from `PerfectEpidemics.qmd`:

Author:	Wilfrid Kendall W.S.Kendall@warwick.ac.uk
Date:	Fri Jan 17 16:59:04 2025 +0000
Summary:	This version has edits arising from the 17/1/2025 talk in Warwick. 2.2.589