

# Perfect Epidemics

## Seminar at University College Dublin

W S Kendall   S B Connor

Warwick, York

20 October 2025



# Introduction

Homage to Dublin  
(Book of Kells, 9th century)



Work on perfect simulation (**CFTP**) for epidemics, now being written up.  
WSK acknowledges the support of UK EPSRC grant EP/R022100.



Handout is on the web: use the QR-code or visit  
[wilfridskendall.github.io/talks/PerfectEpidemics](https://wilfridskendall.github.io/talks/PerfectEpidemics).

# Plan of talk

*Gregory:* Is there any other point to which you would wish to draw my attention?

*Holmes:* To the curious incident of the dog in the night-time.

*Gregory:* The dog did nothing in the night-time.

*Holmes:* That was the curious incident.

(“The Adventure of Silver Blaze”, Sir Arthur Conan Doyle, 1892)

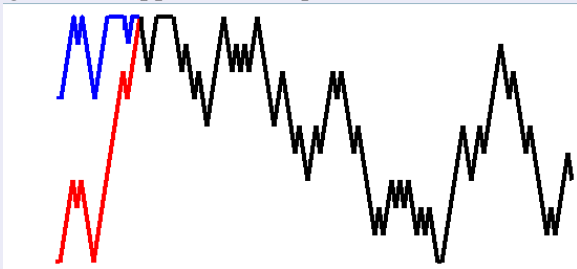
- (a) Introduction to perfect simulation:
- (b) A little theory about CFTP;
- (c) Epidemics and the  $R$ -number;
- (d) “Contact tracing”: inferring infection pattern if removals observed;
- (e) Example with real data.

# 1. A Visual Introduction to Perfect Simulation

- ➊ Propp & Wilson (1996) invented exact simulation / Coupling from the Past (CFTP) / perfect simulation;
- ➋ The term “perfect simulation” (WSK, 1998) was chosen to encourage you to be suspicious: perfection is never achieved!
- ➌ Key ideas of “*classic CFTP*”:
  - ▶ extend simulation *backwards* through time not forwards;
  - ▶ exploit monotonicity (*couple* maximal and minimal processes);
  - ▶ seek coalescence.
- ➍ 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)

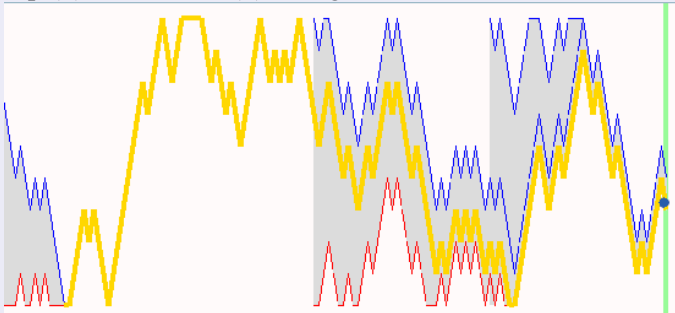
- 1 Consider a simple random walk on  $0:9 = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ .
  - ▶  $\mathbb{P}[+1 \text{ jump}] = p \in (0, 1)$ , while  $\mathbb{P}[-1 \text{ jump}] = 1-p$ , **except that**
  - ▶ at state 9 replace the  $+1$  jump by “staying still”, **and**
  - ▶ at state 0 replace the  $-1$  jump by “staying still”.
- 2 Conventional MCMC picks a starting point, then runs the simple random walk for long time till approximate equilibrium.



- 3 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;
- 4 Generally **not true** that location *at* coupling is a draw from equilibrium.

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

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



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

# Some more CFTP theory

- ① What if 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 Theoretical limits: *in principle*

- ▶ 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).

- ③ Dominated CFTP delivers perfect simulation for stable point processes (WSK & Møller, 2000);

- ④ Detailed expositions: WSK (2005), Huber (2015).  
(Want to implement CFTP in R? see WSK, 2015.)

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

S-I-R deterministic epidemic:

based on susceptibles  $s$ , infectives  $i$ , removals  $r$ :

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

Constant total population  $s + i + r = n$ .

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

In contrast, Fraser *et al* (2023) use a **UK model** with  $N=10^6$  agents!

There are *many* important inferential questions (Cori & Kucharski, 2024).



# 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 strongly 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.



# 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**?

# An easier question: “Contact Tracing”

The simplest possible variant of contact tracing:

“When did the infections occur, supposing we only observe removals?”

(Gibson & Renshaw, 1998; O'Neill & Roberts, 1999; Gibson & Renshaw, 2001)

Important first step: think about generation of an *unconditioned* epidemic.

- ① Suppose  $n$ ,  $\alpha$ ,  $\beta$  are known. Eventually 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.

# 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.

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

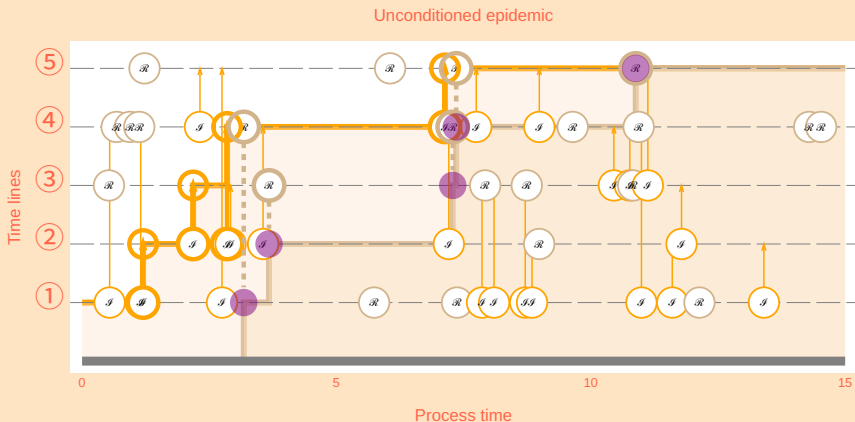


Figure 2: (a) *Infection* activates if target lies on the lowest uninfected timeline; and (b) *removal* activates if on infected timeline; remove lowest infected (purple disk).

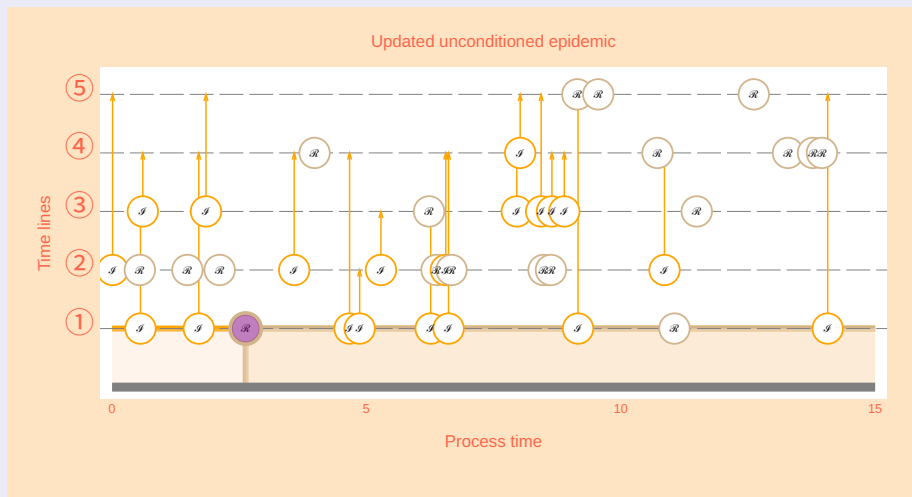
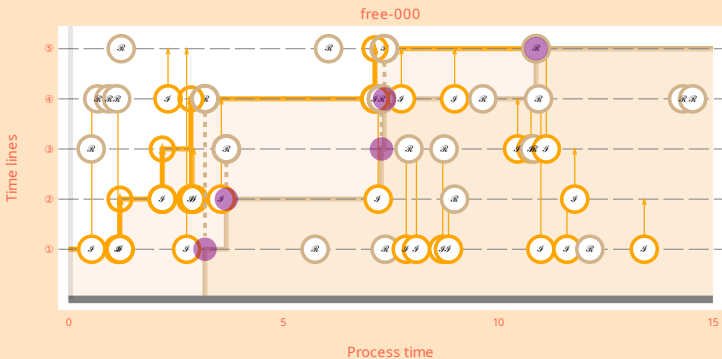


Figure 3: A step in algorithmic time for the unconditioned epidemic simply involves replacing all original incidents by an entirely new set of incidents.

## Crucial technical point

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

## Free evolution evolving in continuous algorithmic time



GIF MP4



### 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)*.

# Initial conditioned epidemic

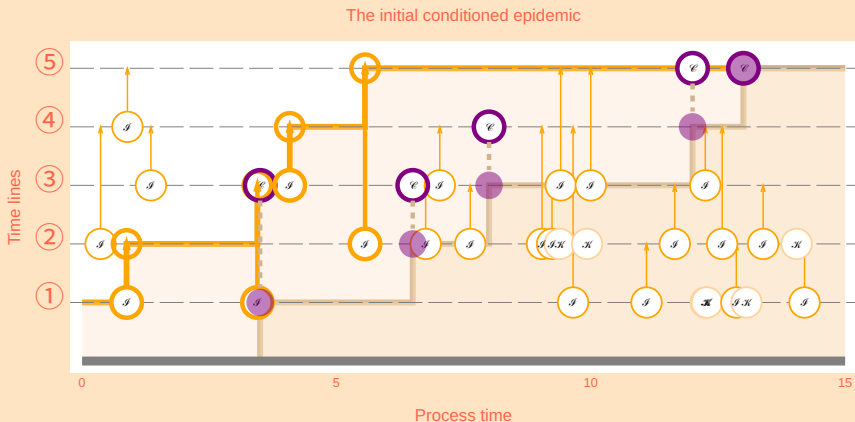


Figure 4: Initial conditioned epidemic, with conditioned removals indicated using purple circles (and purple disks when non-target timelines are infected).

# Conditional epidemic update

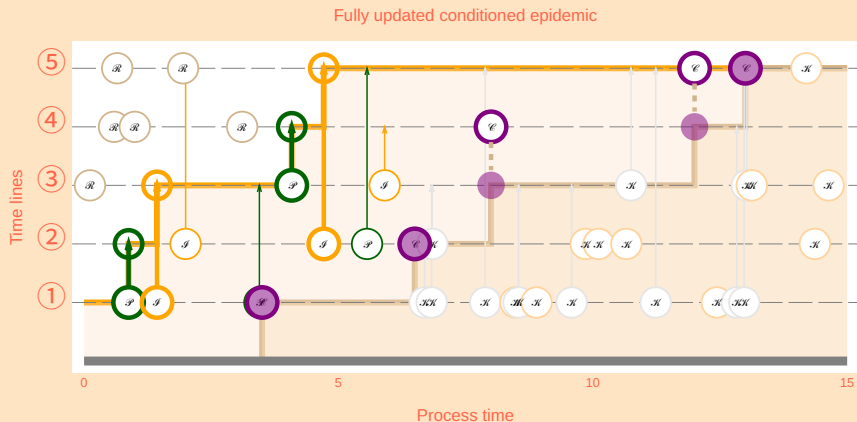


Figure 5: Epidemic updated under restriction: all conditioned removals remain activated, no new removals are activated. Green infections have been “perpetuated”.

# Laziest feasible epidemic (LFE)

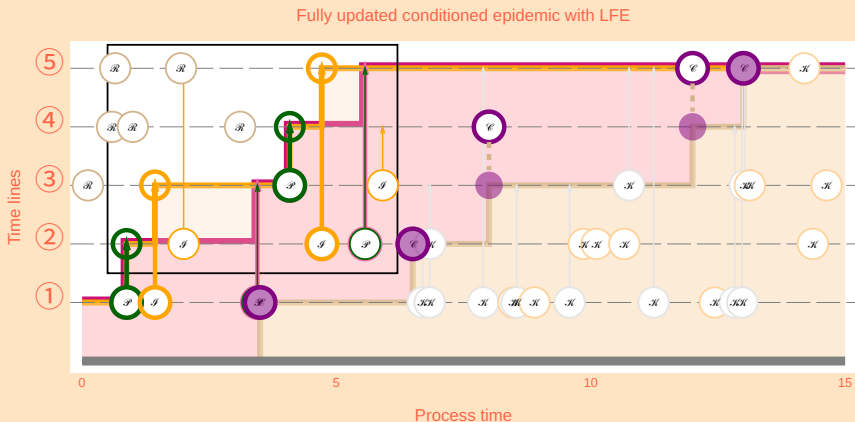
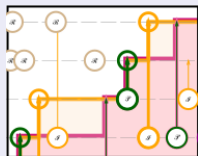


Figure 6: LFE computed recursively working right-to-left: slowest sequence of infections (and perpetuated infections) generating all conditioned removals. Can be used to identify perpetuated infections.

## LFE: construction details



- ➊ Recursive definition of **LFE**: working over  $[0, T)$ ,

$$\begin{aligned} s_N &= T \\ s_i &\leq \min \left\{ s_{i+1}, \inf \{ s : \text{there is a } \mathcal{C}_s^i \} \right\} . \end{aligned}$$

- ② Intrinsic definition of LFE:

Slowest/lowest epidemic activating all re-marked  $\mathcal{C}$ s, formed from *subset* of

- ▶ new potential infections,
- ▶ *previous* epidemic history (supplies  $\mathcal{P}$ s).

- 3 Comparisons based on intrinsic definition show monotonic dependence of LFE on previous epidemic history.

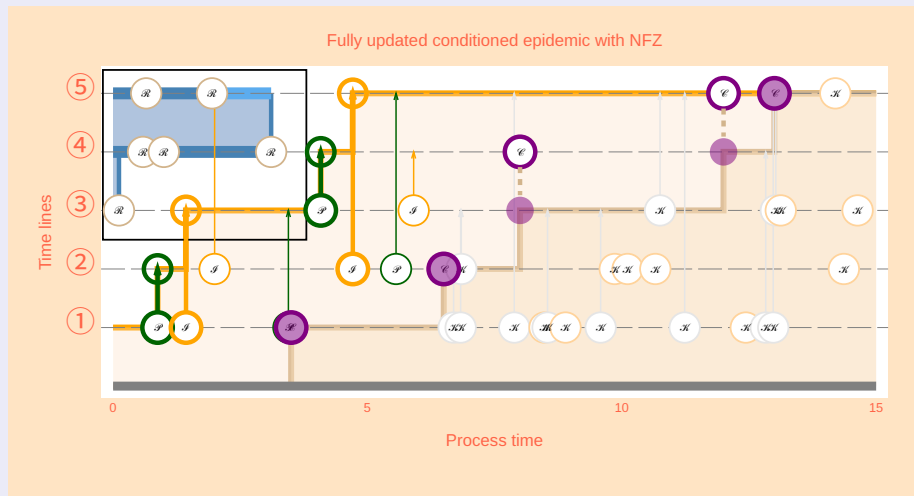
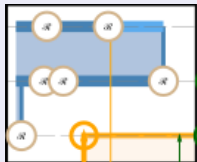


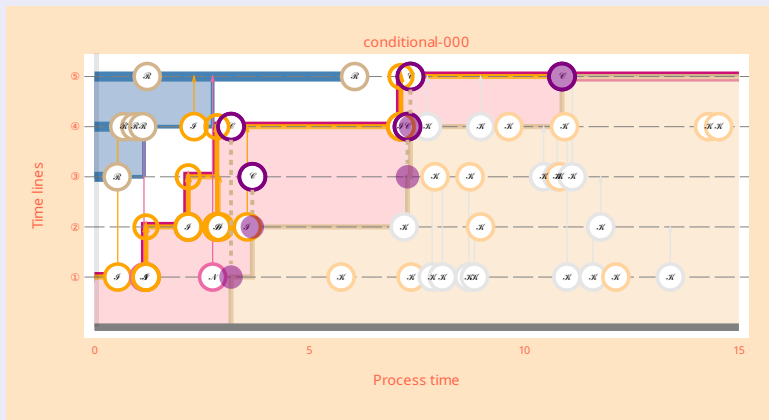
Figure 7: NFZ computed recursively working right-to-left: it traces a region of timelines such that unobserved removals are not activated if region not infected.

## NFZ: construction details



- 1 NFZ is union of timeline intervals  $[0, u_i] \times \{i\}$ , expressible as union of regions (for each new *inactivated*  $\mathcal{R}_t^i$ ) each with right-boundary
  - ▶  $(r, t)$  for  $r = i, \dots, N$ ,
  - ▶ working backwards in time from  $\mathcal{R}_t^i$  and following potential infections down by one step per infection that does *not* have a target in the infected or removed regions.
- 2 Can then show monotonic dependence of NFZ on previous epidemic history.

# Conditioned evolution evolving in continuous algorithmic time



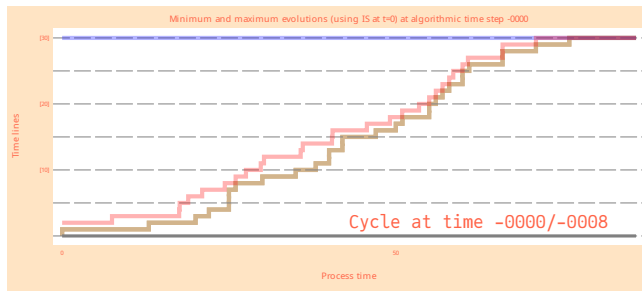
GIF MP4

If a new  $\mathcal{J}_t^{i < j}$  has  $i, j$  in infectious zone then **LFE** is relevant;  
 if  $i$  in infectious zone and  $j$  in susceptible zone then **NFZ** is relevant.  
*New epidemic is monotonic in LFE, NFZ!*



## 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](#)), animates (GIF or MP4) a perfect simulation of a draw from unobserved pattern of infections.



# So what?

- What about accept-reject methods?
- Why this emphasis on unobserved infections given fixed  $\alpha$  and  $\beta$ , when we need inference on R-number  $\alpha n / \beta$  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, 2025**)
  - ▶ estimate likelihood test statistic for specified  $\alpha$  and  $\beta$ ;
  - ▶ Rao-Blackwell-ize: re-sample infection times given  $I$  at removals;
  - ▶ construct steepest ascent algorithm (in effect, variant of Robbins-Monro stochastic optimization) to find *maximum a posteriori* estimates of  $\alpha$  and  $\beta$ ;
  - ▶ or even, with some computational effort, compute an approximation to the entire posterior joint density for  $\alpha$  and  $\beta$ !
- **Finally:** generalize to other suitable 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).
- Experiments suggest **CFTP** out-competes non-naïve accept-reject.
- Still to be done: seek faster **CFTP**; 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.
- Bensoussane, H. (2025) Bayesian Individual-level Epidemic Models : Accounting for Missing Data and Utilising Covariate Information (PhD No. January).
- 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.
- Diaconis, P. (2009) The Markov Chain Monte Carlo Revolution. *Bulletin of the American Mathematical Society*, **46**, 179–205.
- 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.

# References II

- 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.
- 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.
- 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.
- Rocklöv, J., Sjödin, H., & Wilder-Smith, A. (2020) COVID-19 outbreak on the Diamond Princess cruise ship: estimating the epidemic potential and effectiveness of public health countermeasures. *Journal of Travel Medicine*, **27**, 7 pp.
- SBC (2020) Omnithermal Perfect Simulation for Multi-server Queues. *ACM Transactions on Modeling and Computer Simulation*, **30**, 1–15.
- SBC & WSK (2007a) Perfect simulation for a class of positive recurrent Markov chains. *Annals of Applied Probability*, **17**, 781–808.
- SBC & WSK (2007b) Perfect simulation for a class of positive recurrent Markov chains (corrigendum). *Annals of Applied Probability*, **17**, 1808–1810.
- SBC & WSK (2015) Perfect simulation of M/G/c queues. *Advances in Applied Probability*, **47**, 1039–1063.

# References III

SBC & WSK (2025) Perfect Epidemics.

Sigman, K. (2011) Exact simulation of the stationary distribution of the FIFO M/G/c queue. *Journal of Applied Probability*, **48**, 209–213.

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.

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

<i>Image</i>	<i>Attribution</i>	
<i>Book of Kells</i> Classic CFTP for a simple random walk	Huber Gerhard Result of code written by WSK	<i>CC BY 4.0</i>
<i>Diamond Princess</i> Epidemic CFTP images and animation	Alpsdake Result of code written by WSK	<i>CC BY-SA 4.0</i>

## Previous instances of this talk

<i>Date</i>	<i>Title</i>		<i>Location</i>
19/04/24	Perfect Epidemics	Short Research Talk	12mn Warwick
15/05/24	McMC+Perfect Simulation	Graduate Seminar, Aristotle Univ.	50mn Thessaloniki
17/01/25	Perfect Epidemics	Applied Probability Seminar	50mn Warwick
27/06/25	Perfect Epidemics	UK Research Network Stochastics	45mn Liverpool
20/10/25	Perfect Epidemics	Seminar	Dublin

# 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	1.12.0	—	
EpidemicsCFTP	2.2.532	develop	Tue Jul 8 17:13:42 2025 +0100
EpidemicsUtilities	0.1.2.177	main	Fri Sep 26 15:35:26 2025 +0100
This quarto script	0.2.2.725	2025-10-09-Dublin-preparation	Tue Oct 14 18:01:39 2025 +0100

## Project information

<b>Version:</b>	0.2.2.725 (2025-10-09-Dublin-preparation)
<b>Author:</b>	Wilfrid Kendall <W.S.Kendall@warwick.ac.uk>
<b>Date:</b>	Tue Oct 14 18:01:39 2025 +0100

### Comment:

Near-final preparation for Dublin October 2025 talk. Added material on [LFE](#) and [NFZ](#) including some sketches of monotonicity arguments. Added note on Rao-Blackwell-ization.