Supplementary Material: Fast likelihood calculations for emerging epidemics

Frank Ball and Peter Neal (University of Nottingham)*

January 9, 2025

In the supplementary material, the section and equation counters continue from the main paper with the first section being Section 7, the first equation being (21), the first figure being Figure 5 and the first table being Table 3. References to sections, equations, figures and tables with lower numbers refer to the main paper.

In Section 7, we derive expressions for the moments of $\tilde{R}_k$ and its Gaussian approximation in Section 3.3. In Section 8 we present a simulation study which explores the effectiveness of the approximate birth-death likelihood in a range of scenarios. Finally, in Section 9, we provide supporting material, via a simulation study, for the analysis of the emergence of Covid-19 in Europe presented in Section 5.

7 Derivation of moments of for approximate likelihood

In Section 7.1, we derive (14) and (15), the mean and variance of $\tilde{R}_k$ in terms of the moment generating function of $\tilde{R}_{k-1}$ and its first three derivatives. In Section 7.2, we derive the recursive relationships for (η_k, σ_k^2) in terms of $(\eta_{k-1}, \sigma_{k-1}^2)$ given in (10) and (11) .

*School of Mathematical Sciences, University Park, Nottingham, NG7 2RD, United Kingdom. email: peter.neal@nottingham.ac.uk

7.1 Derivation of moments of $\tilde{R}_k$

Throughout the derivation of the results, we assume that the distribution of $\tilde{R}_k$ depends on $\tilde{\mathbf{T}}_{2:k} = \tilde{\mathbf{t}}_{2:k}$, but suppress explicit mention of the dependence to simplify presentation of the results.

Remember that for $k = 2, 3, \ldots$ and $j = 2, 3, \ldots, k$, $\tilde{B}_{k,j}$ denotes $\mathbb{P}(\tilde{R}_k = j|\tilde{\mathbf{\Theta}}_{2:k})$. Thus $\tilde{\mathbf{B}}_2 = (1)$ and for $k = 3, 4, \ldots$, $\tilde{\mathbf{B}}_k = (\tilde{B}_{k,2}, \tilde{B}_{k,3}, \ldots, \tilde{B}_{k,k})$, is given by

$$\tilde{\mathbf{B}}_k = \left\{\tilde{\mathbf{B}}_{k-1}\tilde{\mathbf{M}}_{k-1}\left(\tilde{t}_k\right) \cdot \mathbf{1}_{k-1}^{\top}\right\}^{-1} \tilde{\mathbf{B}}_{k-1}\tilde{\mathbf{M}}_{k-1}\left(\tilde{t}_k\right), \tag{21}$$

where $\mathbf{1}_{k-1}^{\top}$ denotes a column vector of 1s of length $k-1$ and for $\tau \geq 0$, $\tilde{\mathbf{M}}_{k-1}(\tau)$ is the $(k-2) \times (k-1)$ matrix with $(i,j)^{th}$ element

$$\left[\tilde{\mathbf{M}}_{k-1}(\tau)\right]_{i,j} = \begin{cases} (i+1)\binom{i}{j-1} h_k(\tau)^{j-1}[1-h_k(\tau)]^{i+1-j}\tilde{r}_k(\tau)^i & \text{for } j \leq i+1, \\ 0 & \text{otherwise,} \end{cases} \tag{22}$$

where $h_k(\tau) = [1 - \tilde{\pi}_k - \tilde{\psi}_k(\tau)]/[(1-\tilde{\pi}_k)(1-\tilde{\psi}_k(\tau))]$.

From the dependence of the distribution of $\tilde{R}_k$ on the distribution of $\tilde{R}_{k-1}$ given in (21) and noting that $[\tilde{M}_{k-1}(\tilde{t}_k) \cdot \mathbf{1}_{k-1}^{\top}]_i = (i+1)\tilde{r}_k(\tilde{t}_k)^i$, it follows that

$$\tilde{\mathbf{B}}_{k-1}\tilde{\mathbf{M}}_{k-1}\left(\tilde{t}_k\right) \cdot \mathbf{1}_{k-1}^{\top} = \sum_{l=2}^{k-1} \tilde{B}_{k-1,l} l \tilde{r}_k(\tilde{t}_k)^{l-1} = \mathbb{E}[\tilde{R}_{k-1}\tilde{r}_k(\tilde{t}_k)^{\tilde{R}_{k-1}-1}]. \tag{23}$$

Therefore from (22) and (23), for $n = 1, 2$, we have that

$$\begin{aligned} \mathbb{E}\left[\tilde{R}_k^n\right] &= \sum_{j=2}^{k} j^n \tilde{B}_{k,j} \\ &= \frac{\sum_{j=2}^{k} j^n \sum_{l=j-1}^{k-1} l\binom{l-1}{j-2} h_k(\tilde{t}_k)^{j-2}[1-h_k(\tilde{t}_k)]^{l+1-j}\tilde{r}_k(\tilde{t}_k)^{l-1}\tilde{B}_{k-1,l}}{\mathbb{E}[\tilde{R}_{k-1}\tilde{r}_k(\tilde{t}_k)^{\tilde{R}_{k-1}-1}]}. \end{aligned} \tag{24}$$

The numerator on the right-hand side of (24) is

$$\begin{aligned} &\sum_{j=2}^{k} j^n \sum_{l=j-1}^{k-1} l\binom{l-1}{j-2} h_k(\tilde{t}_k)^{j-2}[1-h_k(\tilde{t}_k)]^{l+1-j}\tilde{r}_k(\tilde{t}_k)^{l-1}\tilde{B}_{k-1,l} \\ &\qquad = \sum_{l=1}^{k-1} l\tilde{r}_k(\tilde{t}_k)^{l-1}\tilde{B}_{k-1,l}\left\{\sum_{j=2}^{l+1} j^n \binom{l-1}{j-2} h_k(\tilde{t}_k)^{j-2}[1-h_k(\tilde{t}_k)]^{l+1-j}\right\}. \end{aligned} \tag{25}$$

The inner summation on the right-hand side of (25) is $\mathbb{E}[(I_{k,l}+2)^n]$, where $I_{k,l} \sim \text{Bin}(l-1, h_k(\tilde{t}_k))$ $(k = 2, 3, \ldots; l = 1, 2, \ldots, k-1)$. Hence,

$$\begin{aligned}
\mathbb{E}\left[\tilde{R}_k\right] &= \frac{\sum_{l=1}^{k-1} l\tilde{r}_k(\tilde{t}_k)^{l-1}\tilde{B}_{k-1,l}[2+(l-1)h_k(t_k)]}{\mathbb{E}\left[\tilde{R}_{k-1}\tilde{r}_k(\tilde{t}_k)^{\tilde{R}_{k-1}-1}\right]} \\
&= \frac{2\mathbb{E}\left[\tilde{R}_{k-1}\tilde{r}_k(\tilde{t}_k)^{\tilde{R}_{k-1}-1}\right] + h_k(\tilde{t}_k)\mathbb{E}\left[\tilde{R}_{k-1}(\tilde{R}_{k-1}-1)\tilde{r}_k(\tilde{t}_k)^{\tilde{R}_{k-1}-1}\right]}{\mathbb{E}\left[\tilde{R}_{k-1}\tilde{r}_k(\tilde{t}_k)^{\tilde{R}_{k-1}-1}\right]} \\
&= \frac{2\mathbb{E}\left[\tilde{R}_{k-1}\tilde{r}_k(\tilde{t}_k)^{\tilde{R}_{k-1}}\right] + h_k(\tilde{t}_k)\mathbb{E}\left[\tilde{R}_{k-1}(\tilde{R}_{k-1}-1)\tilde{r}_k(\tilde{t}_k)^{\tilde{R}_{k-1}}\right]}{\mathbb{E}\left[\tilde{R}_{k-1}\tilde{r}_k(\tilde{t}_k)^{\tilde{R}_{k-1}}\right]}. \qquad (26)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathbb{E}\left[\tilde{R}_k^2\right] &= \frac{\sum_{l=1}^{k-1} l\tilde{r}_k(\tilde{t}_k)^{l-1}\tilde{B}_{k-1,l}[4+5(l-1)h_k(\tilde{t}_k)+(l-1)(l-2)h_k(\tilde{t}_k)^2]}{\mathbb{E}\left[\tilde{R}_{k-1}\tilde{r}_k(\tilde{t}_k)^{\tilde{R}_{k-1}-1}\right]} \\
&= \frac{1}{\mathbb{E}\left[\tilde{R}_{k-1}\tilde{r}_k(\tilde{t}_k)^{\tilde{R}_{k-1}}\right]}\left\{4\mathbb{E}\left[\tilde{R}_{k-1}\tilde{r}_k(\tilde{t}_k)^{\tilde{R}_{k-1}}\right] + 5h_k(\tilde{t}_k)\mathbb{E}\left[\tilde{R}_{k-1}(\tilde{R}_{k-1}-1)\tilde{r}_k(\tilde{t}_k)^{\tilde{R}_{k-1}}\right]\right. \\
&\quad \left. + h_k(\tilde{t}_k)^2\mathbb{E}\left[\tilde{R}_{k-1}(\tilde{R}_{k-1}-1)(\tilde{R}_{k-1}-2)\tilde{r}_k(\tilde{t}_k)^{\tilde{R}_{k-1}}\right]\right\}. \qquad (27)
\end{aligned}$$

We can then express the first two moments of $\tilde{R}_k$ in terms of the moment generating function of $\tilde{R}_{k-1}$ and its first three derivatives as follows. For $k = 2, 3, \ldots$, let $\varphi_k = \log \tilde{r}_k(\tilde{t}_k)$, and hence, for $n = 0, 1, 2, 3$, let

$$\chi_k^n = \mathbb{E}\left[\tilde{R}_{k-1}^n \tilde{r}_k(\tilde{t}_k)^{\tilde{R}_{k-1}}\right] = \mathbb{E}\left[\tilde{R}_{k-1}^n \exp\left(\varphi_k \tilde{R}_{k-1}\right)\right].$$

Therefore we can rewrite (26) and (27), respectively, as

$$\mathbb{E}\left[\tilde{R}_k\right] = \frac{2\chi_k^1 + h_k(\tilde{t}_k)[\chi_k^2 - \chi_k^1]}{\chi_k^1} = 2 + h_k(\tilde{t}_k)\left[\frac{\chi_k^2}{\chi_k^1} - 1\right] = \tilde{\eta}_k, \quad \text{say}, \qquad (28)$$

which corresponds to (14), and

$$\begin{aligned}
\mathbb{E}\left[\tilde{R}_k^2\right] &= \frac{4\chi_k^1 + 5h_k(\tilde{t}_k)[\chi_k^2 - \chi_k^1] + h_k(\tilde{t}_k)^2[\chi_k^3 - 3\chi_k^2 + 2\chi_k^1]}{\chi_k^1} \\
&= 4 + 5h_k(\tilde{t}_k)\left(\frac{\chi_k^2}{\chi_k^1} - 1\right) + h_k(\tilde{t}_k)^2\left(\frac{\chi_k^3}{\chi_k^1} - 3\frac{\chi_k^2}{\chi_k^1} + 2\right). \qquad (29)
\end{aligned}$$

Finally, (15) follows from (28) and (29).

7.2 Recursive relationship for (η_k, σ_k^2)

Let $\eta_2 = \mathbb{E}\left[\tilde{R}_2\right] = 2$ and $\sigma_2^2 = \text{var}(\tilde{R}_2) = 0$. For $k = 3, 4, \ldots$, we set $W_k \sim N(\eta_k, \sigma_k^2)$ where (η_k, σ_k^2) are constructed iteratively as follows. For $k = 2, 3, \ldots$ and $\rho \in \mathbb{R}$, let

$$\xi_k(\rho) = \exp\left(\rho\eta_{k-1} + \frac{\rho^2}{2}\sigma_{k-1}^2\right),$$

the moment generating function of $W_{k-1} \sim N(\eta_{k-1}, \sigma_{k-1}^2)$. For $n = 1, 2, 3$, let $\xi_k^{(n)}(\rho) = \mathbb{E}[W_{k-1}^n \exp(\rho W_{k-1})]$, the n^{th} derivative of $\xi_k(\rho)$ with respect to ρ. Therefore, if W_{k-1} is an approximation for $\tilde{R}_{k-1}$, we have that $\xi_k^{(n)}(\varphi_k) \approx \chi_k^n$. Hence, we define

$$\eta_k = 2 + h_k(\tilde{t}_k)\left[\frac{\xi_k^{(2)}(\varphi_k)}{\xi_k^{(1)}(\varphi_k)} - 1\right] \tag{30}$$

and

$$\sigma_k^2 = h_k(\tilde{t}_k)(1 - h_k(\tilde{t}_k))\left[\frac{\xi_k^{(2)}(\varphi_k)}{\xi_k^{(1)}(\varphi_k)} - 1\right] + h_k(\tilde{t}_k)^2\left[\frac{\xi_k^{(3)}(\varphi_k)}{\xi_k^{(1)}(\varphi_k)} - \left(\frac{\xi_k^{(2)}(\varphi_k)}{\xi_k^{(1)}(\varphi_k)}\right)^2\right]. \tag{31}$$

The first three derivatives of $\xi_k(\rho)$ are

$$\begin{aligned}
\xi_k^{(1)}(\rho) &= [\eta_{k-1} + \rho\sigma_{k-1}^2]\xi_k(\rho) \\
\xi_k^{(2)}(\rho) &= [\sigma_{k-1}^2 + (\eta_{k-1} + \rho\sigma_{k-1}^2)^2]\xi_k(\rho) \\
\xi_k^{(3)}(\rho) &= [\eta_{k-1} + \rho\sigma_{k-1}^2][3\sigma_{k-1}^2 + (\eta_{k-1} + \rho\sigma_{k-1}^2)^2]\xi_k(\rho).
\end{aligned}$$

Thus

$$\frac{\xi_k^{(2)}(\varphi_k)}{\xi_k^{(1)}(\varphi_k)} = \eta_{k-1} + \varphi_k\sigma_{k-1}^2 + \frac{\sigma_{k-1}^2}{\eta_{k-1} + \varphi_k\sigma_{k-1}^2} \tag{32}$$

$$\frac{\xi_k^{(3)}(\varphi_k)}{\xi_k^{(1)}(\varphi_k)} = 3\sigma_{k-1}^2 + (\eta_{k-1} + \varphi_k\sigma_{k-1}^2)^2. \tag{33}$$

Finally using (32) and (33), (10) and (11) follow from (30) and (31), respectively.

7.3 Gaussian approximation of the likelihood

In this section we provide additional support to that given in Section 3.3 for using $\hat{L}(\tilde{\mathbf{t}}_T; \tilde{\mathbf{\Theta}}_{K(T)+1})$ as an approximation of $L(\tilde{\mathbf{t}}_T; \tilde{\mathbf{\Theta}}_{K(T)+1})$.

We begin by considering the cases where $k = 4, 5$ to illustrate that the Gaussian approximation is effective even for small k. To enable explicit results, we use the time-homogeneous birth-death process with birth rate α, death rate μ and all deaths detected, $d = 1$. The notation simplifies with $q = \mu/(\alpha+\mu)$ ($p = \alpha/(\alpha+\mu)$) being the probability that an event is a death (birth). We then have for all $k = 1, 2, \ldots$, $\tilde{q}_k = q$ and $\tilde{p}_k = p$ with $u_k = 1$, $\lambda_k = q$, $\nu_k = 0$ and $\tilde{\pi}_k = q$. For $\tau \geq 0$, let $\phi(\tau) = \exp(-(\alpha+\mu)\tau)$, $\psi(\tau) = p(1-\phi(\tau))$, $r(t) = q + p\phi(t)$ and $h(t) = \phi(t)/\{q + p\phi(t)\}$. Then, for all $k = 1, 2, \ldots$ and $\tau \geq 0$, $\tilde{\phi}_k(\tau) = \phi(\tau)$, $\tilde{\psi}_k(\tau) = \psi(\tau)$, $\tilde{r}_k(\tau) = r(\tau)$ and $h_k(\tau) = h(\tau)$.

Firstly, we note that if all deaths are detected in the time-homogeneous birth-death process, $\tilde{r}_k(t) \geq q$ is bounded away from 0, so $\varphi_k(\leq 0)$ is bounded away from minus infinity. We have that $\xi_k^{(1)}(\varphi_k) = (\eta_{k-1} + \varphi_k\sigma_{k-1}^2)\exp(\varphi_k\eta_{k-1} + \varphi_k^2\sigma_{k-1}^2/2)$ can go negative if $\varphi_k\sigma_{k-1}^2$ is sufficiently large, in absolute value, relative to η_{k-1}. The lower bound on φ_k helps control against $\xi_k^{(1)}(\varphi_k)$ being negative, as does σ_{k-1}^2 being small relative to η_{k-1}. For example, we note in general that $\sigma_3^2 = h_3(\tilde{t}_3)[1 - h_3(\tilde{t}_3)]$ and $\eta_3 = 2 + h_3(\tilde{t}_3)$, and provided $\tilde{r}_4(\tilde{t}_4) > 5.023 \times 10^{-5}$, we have that $\eta_3 + \varphi_4\sigma_3^2 > 0$.

In Table 3, we give bounds on the maximum and minimum values $\log(\chi_k^1/\xi_k^{(1)}(\varphi_k))$ $k = 4, 5$ and $\log(\chi_4^1\chi_5^1/[\xi_4^{(1)}(\varphi_4)\xi_5^{(1)}(\varphi_5)])$, in terms of q, for the time-homogeneous birth-death process. We note that the size of the potential error increases as q decreases. For $q = 0.2$ corresponding to a basic reproduction number of $R_0 = p/q = 4$, we have that the maximum possible difference, on the log-scale, is 0.0554 corresponding to an error of less than 6%. However, the difference will often be much smaller; in a simulation of 100,000 birth-death processes with $q = 0.2$ and at least 5 deaths, the absolute value of $\log(\chi_4^1\chi_5^1/[\xi_4^{(1)}(\varphi_4)\xi_5^{(1)}(\varphi_5)])$ was less than 0.001 in 57.3% of cases and less than 0.01 in 98.4% of cases. We also note that the maximum difference for $\log(\chi_4^1\chi_5^1/[\xi_4^{(1)}(\varphi_4)\xi_5^{(1)}(\varphi_5)])$ can be smaller than that for $\log(\chi_5^1/\xi_5^{(1)}(\varphi_5))$, owing to the interplay between $h_4(\tilde{t}_4)$ and $\tilde{r}_4(\tilde{t}_4)$.

We now turn to the case k is large and $\eta_{k-1} \gg 1$. We prove Lemma 3, restated below for convenience.

Lemma 3 *Suppose that $\tilde{\eta}_{k-1} \gg 1$ with $\tilde{t}_k = O(\tilde{\eta}_{k-1}^{-1})$, $\tilde{\sigma}_{k-1}^2 = O(\tilde{\eta}_{k-1})$, $\mathbb{E}[(\tilde{R}_{k-1} - \tilde{\eta}_{k-1})^3] = O(\tilde{\eta}_{k-1})$ and $\mathbb{E}[(\tilde{R}_{k-1} - \tilde{\eta}_{k-1})^4] = O(\tilde{\eta}_{k-1}^2)$. Then for $m = 1, 2, 3$,*

$$\log(\chi_k^m/\tilde{\xi}_k^{(m)}(\varphi_k)) = O(\tilde{\eta}_{k-1}^{-2}), \tag{34}$$

q	0.4	0.3	0.2	0.1
$\max\{\log(\chi_4^1/\xi_4^{(1)}(\varphi_4))\}$	0.0067	0.0067	0.0197	0.0923
$\min\{\log(\chi_4^1/\xi_4^{(1)}(\varphi_4))\}$	−0.0032	−0.0032	−0.0070	−0.0651
$\max\{\log(\chi_5^1/\xi_5^{(1)}(\varphi_5))\}$	0.0117	0.0140	0.0554	0.2276
$\min\{\log(\chi_5^1/\xi_5^{(1)}(\varphi_5))\}$	−0.0046	−0.0047	−0.0163	−0.1203
$\max\{\log(\chi_4^1\chi_5^1/[\xi_4^{(1)}(\varphi_4)\xi_5^{(1)}(\varphi_5)])\}$	0.0156	0.0156	0.0554	0.2259
$\min\{\log(\chi_4^1\chi_5^1/[\xi_4^{(1)}(\varphi_4)\xi_5^{(1)}(\varphi_5)])\}$	−0.0053	−0.0060	−0.0143	−0.1400

Table 3: The maximum and minimum possible values $\log(\chi_k^1/\xi_k^{(1)}(\varphi_k))$ $k = 4, 5$ and $\log(\chi_4^1\chi_5^1/[\xi_4^{(1)}(\varphi_4)\xi_5^{(1)}(\varphi_5)])$ for $q = 0.1, 0.2, 0.3, 0.4$.

where $\tilde{\xi}_k(\rho) = \exp(\rho\tilde{\eta}_{k-1} + \rho^2\tilde{\sigma}_{k-1}^2/2)$*, the moment generating function of* $\tilde{W}_{k-1} \sim N(\tilde{\eta}_{k-1}, \tilde{\sigma}_{k-1}^2)$.

Proof of Lemma 3. We begin by proving (34) in the case $m = 1$. For $\theta, x \in \mathbb{R}$, let $f_\theta(x) = x\exp(\theta x)$, and for $n = 0, 1, \ldots$, let $f_\theta^{(n)}(x)$ denote the n^{th} derivative, with respect to x, of $f_\theta(x)$. Then, for all $n = 0, 1, \ldots$,

$$f_\theta^{(n)}(x) = \theta^{n-1}(n + \theta x)\exp(\theta x).$$

By Taylor's Theorem, for any $x, \eta \in \mathbb{R}$, we have, there exists c lying between x and η, such that

$$f_\theta(x) = \sum_{n=0}^{3} \frac{\theta^{n-1}}{n!}(n + \theta\eta)e^{\theta\eta}(x - \eta)^n + \frac{\theta^3}{4!}(4 + \theta c)e^{\theta c}(x - \eta)^n.$$

Therefore

$$\chi_k^1 = \mathbb{E}\left[\tilde{R}_{k-1}\exp(\varphi_k\tilde{R}_{k-1})\right] = \sum_{n=0}^{3} \frac{\varphi_k^{n-1}}{n!}(n + \varphi_k\tilde{\eta}_{k-1})\exp(\varphi_k\tilde{\eta}_{k-1})\mathbb{E}\left[(\tilde{R}_{k-1} - \tilde{\eta}_{k-1})^n\right] + \frac{\varphi_k^3}{4!}(4 + \varphi_k\tilde{\eta}_{k-1})\mathbb{E}\left[\exp(\varphi_k c(\tilde{R}_{k-1}, \tilde{\eta}_{k-1}))(\tilde{R}_{k-1} - \tilde{\eta}_{k-1})^4\right], \quad (35)$$

where $c(\tilde{R}_{k-1}, \tilde{\eta}_{k-1})$ lies between $\tilde{\eta}_{k-1}$ and $\tilde{R}_{k-1}$. Since $\varphi_k \leq 0$ and $\tilde{R}_{k-1} \geq 2$, we have that $\exp(\varphi_k c(\tilde{R}_{k-1}, \tilde{\eta}_{k-1})) \leq \exp(-\varphi_k\tilde{\eta}_{k-1})$. Given that $\varphi_k = O(\tilde{t}_k) = O(\tilde{\eta}_{k-1}^{-1})$, the final term on the right-hand side of (35) is bounded above by

$$\frac{\varphi_k^3}{4!}(4 + \varphi_k\tilde{\eta}_{k-1})\exp(-\varphi_k\tilde{\eta}_{k-1})\mathbb{E}\left[(\tilde{R}_{k-1} - \tilde{\eta}_{k-1})^4\right] = \frac{1}{24}(4 + \varphi_k\tilde{\eta}_{k-1})\exp(-\varphi_k\tilde{\eta}_{k-1}) \times \varphi_k^3\mathbb{E}\left[(\tilde{R}_{k-1} - \tilde{\eta}_{k-1})^4\right] = O(1) \times O(\tilde{\eta}_{k-1}^{-1}) = O(\tilde{\eta}_{k-1}^{-1}).$$

It follows that

$$\mathbb{E}\left[\tilde{R}_{k-1}\exp(\varphi_k\tilde{R}_{k-1})\right] = \exp(\varphi_k\tilde{\eta}_{k-1})\left\{\tilde{\eta}_{k-1} + \varphi_k\left(1+\frac{\varphi_k\tilde{\eta}_{k-1}}{2}\right)\tilde{\sigma}^2_{k-1}\right\} + O(\tilde{\eta}^{-1}_{k-1}). \tag{36}$$

Also, since $\varphi_k^2\tilde{\sigma}^2_{k-1} = O(\eta^{-1}_{k-1})$, we have that

$$\begin{aligned}\tilde{\xi}^{(1)}_k(\varphi_k) = \mathbb{E}\left[\tilde{W}_{k-1}\exp(\varphi_k\tilde{W}_{k-1})\right] &= \left(\tilde{\eta}_{k-1}+\varphi_k\tilde{\sigma}^2_{k-1}\right)\exp\left(\varphi_k\tilde{\eta}_{k-1}+\frac{\varphi_k^2}{2}\tilde{\sigma}^2_{k-1}\right)\\ &= \left(\tilde{\eta}_{k-1}+\varphi_k\tilde{\sigma}^2_{k-1}\right)\exp\left(\varphi_k\tilde{\eta}_{k-1}\right)\left[1+\frac{\varphi_k^2}{2}\tilde{\sigma}^2_{k-1}+O(\eta^{-2}_{k-1})\right]\\ &= \exp\left(\varphi_k\tilde{\eta}_{k-1}\right)\left\{\tilde{\eta}_{k-1}+\varphi_k\left(1+\frac{\varphi_k\tilde{\eta}_{k-1}}{2}\right)\tilde{\sigma}^2_{k-1}\right\}+O(\tilde{\eta}^{-1}_{k-1}).\end{aligned} \tag{37}$$

From (36) and (37), the leading order terms of χ^1_k and $\tilde{\xi}^{(1)}_k(\varphi_k)$ are the same, giving

$$\frac{\chi^1_k}{\tilde{\xi}^{(1)}_k(\varphi_k)} = \frac{\exp\left(\varphi_k\tilde{\eta}_{k-1}\right)\left\{\tilde{\eta}_{k-1}+\varphi_k\left(1+\varphi_k\tilde{\eta}_{k-1}/2\right)\tilde{\sigma}^2_{k-1}\right\}+O(\tilde{\eta}^{-1}_{k-1})}{\exp\left(\varphi_k\tilde{\eta}_{k-1}\right)\left\{\tilde{\eta}_{k-1}+\varphi_k(1+\varphi_k\tilde{\eta}_{k-1}/2)\tilde{\sigma}^2_{k-1}\right\}+O(\tilde{\eta}^{-1}_{k-1})} = 1+O(\tilde{\eta}^{-2}_{k-1}). \tag{38}$$

By using similar Taylor series expansions of $x^m\exp(\theta x)$ $(m=2,3)$ about $\eta(=\tilde{\eta}_{k-1})$, we can show, under the conditions of the lemma, that

$$\begin{aligned}\chi^2_k &= \exp(\varphi_k\tilde{\eta}_{k-1})\left\{\tilde{\eta}^2_{k-1}+\left(\frac{\varphi_k^2}{2}\tilde{\eta}^2_{k-1}+2\varphi_k\tilde{\eta}_{k-1}+1\right)\tilde{\sigma}^2_{k-1}\right\}+O(1)\\ &= \tilde{\xi}^{(2)}_k(\varphi_k)+O(1)\end{aligned} \tag{39}$$

and

$$\begin{aligned}\chi^3_k &= \exp(\varphi_k\tilde{\eta}_{k-1})\left\{\tilde{\eta}^3_{k-1}+\left(\varphi_k^2\tilde{\eta}^2_{k-1}+6\varphi_k\tilde{\eta}_{k-1}+6\right)\frac{\tilde{\eta}_{k-1}}{2}\tilde{\sigma}^2_{k-1}\right\}+O(\tilde{\eta}_{k-1})\\ &= \tilde{\xi}^{(3)}_k(\varphi_k)+O(\tilde{\eta}_{k-1}).\end{aligned} \tag{40}$$

It follows from (38), (39) and (40), for $m=1,2,3$, that

$$\log\left[\frac{\chi^m_k}{\tilde{\xi}^{(m)}_k(\varphi_k)}\right] = \log[1+O\left(\tilde{\eta}^{-2}_{k-1}\right)] = O\left(\tilde{\eta}^{-2}_{k-1}\right)$$

and the lemma is proved. □

Finally, for $\eta_{k-1}\gg 1$ and $\tilde{t}_k\approx 0$, $h_k(\tilde{t}_k) = 1+h'_k(0)\tilde{t}_k+O(\tilde{t}^2_k)$. Therefore, using a Maclaurin series

expansion of $h_k(\cdot)$ and (11), we have that

$$\begin{aligned}\sigma_k^2 &= \left(-h_k'(0)\tilde{t}_k + O(\tilde{t}_k^2)\right)\left[\eta_{k-1} + \varphi_k\sigma_{k-1}^2 + \frac{\sigma_{k-1}^2}{\eta_{k-1}+\varphi_k\sigma_{k-1}^2} - 1\right] \\ &\quad + \left(1 + 2h_k'(0)\tilde{t}_k + O(\tilde{t}_k^2)\right)\sigma_{k-1}^2\left[1 - \frac{\sigma_{k-1}^2}{[\eta_{k-1}+\varphi_k\sigma_{k-1}^2]^2}\right] \\ &= \sigma_{k-1}^2\left[1 - \frac{\sigma_{k-1}^2}{[\eta_{k-1}+\varphi_k\sigma_{k-1}^2]^2}\right] - h_k'(0)\tilde{t}_k\left[\eta_{k-1} - 2\sigma_{k-1}^2\right] + o(1). \qquad (41)\end{aligned}$$

Since $h_k'(0) < 0$, the right-hand side of (41) is less than σ_{k-1}^2 if $\eta_{k-1} < 2\sigma_{k-1}^2$. Thus for large k, we expect $\sigma_{k-1}^2 < \eta_{k-1}/2$ and we find this to be the case in the simulation scenarios considered in Section 8.

8 Simulation Study

In this section we complement the comparison of exact and approximate birth-death process likelihoods presented in Sections 3.3 and 7.3 with a numerical simulation study. Remember we assume that the parameters (birth rate, death rate and detection probability) are constant between detected deaths. Throughout this section we simulate birth-death processes until the $N = 200^{th}$ detected death, and set $H = 50$, when using the hybrid likelihood given in (17), so the exact likelihood is used for the first 50 detected deaths.

We start with a time-homogeneous birth-death process with birth and death parameters $\alpha = 0.2$ and $\mu = 0.1$, and all deaths detected, $d = 1$. Denote the observed data by $\tilde{\mathbf{t}}_{2:200}^*$, say. Given that the detection probability is fixed at $d = 1$, we have only two parameters α and μ to estimate. Using the exact, approximate and hybrid likelihoods, we compute the likelihood surface on a 201×191 grid, for (α, μ) taking values $(0.1 + 0.001i, 0.01 + 0.001j)$, $(i = 0, 1, \ldots, 200; j = 0, 1, \ldots, 190)$. In Figure 5, we present contour plot estimates of the likelihood surfaces computed using each of the three likelihoods along with a plot measuring the discrepancy between the exact and approximate likelihoods. We note that the three likelihood surfaces are virtually indistinguishable with very similar plots obtained for other choices of H for the hybrid likelihood. The approximate likelihood gives an excellent approximation to the exact likelihood close to the maximum likelihood estimate (MLE) of (α, μ) with the estimation of the likelihood deteriorating the further away the parameters (α, μ) are from the MLE. The approximation is

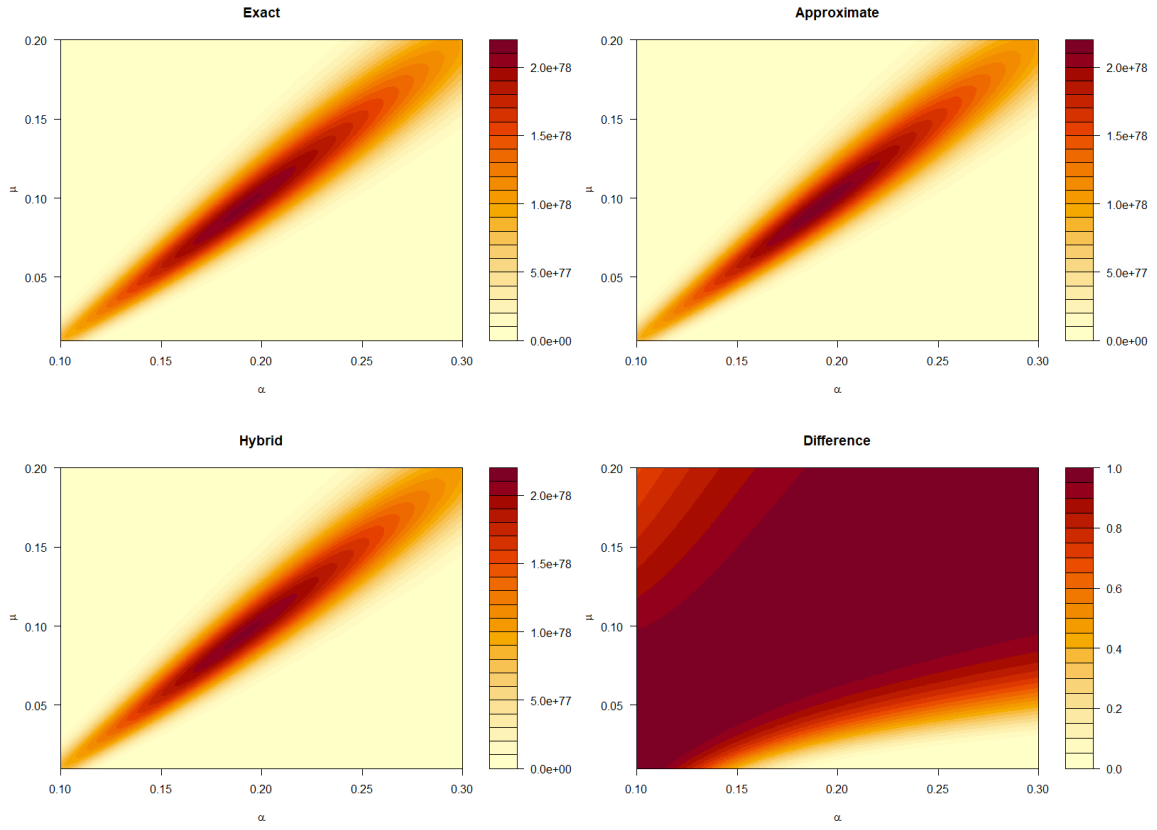

Figure 5: Contour plots of the likelihood for data $\tilde{\mathbf{t}}^*_{2:200}$ using the exact (top left), approximate (top right) and hybrid (bottom left) likelihoods. In the bottom right plot is a contour plot of $\min\{L(\tilde{t}^*_{2:200};(\alpha,\mu)),\tilde{L}(\tilde{t}^*_{2:200};(\alpha,\mu))\}/\max\{L(\tilde{t}^*_{2:200};(\alpha,\mu)),\tilde{L}(\tilde{t}^*_{2:200};(\alpha,\mu))\}$.

worst when α/μ is large corresponding to $q=\mu/(\alpha+\mu)$ being close to 0 in line with analysis presented in Section 7.3. A general observation made across the simulation examples below, and the Abakiliki data in Ball and Neal (2023), Section 4, is that the approximate likelihood tends to perform best close to modes of the likelihood, which makes it particularly useful in practice, for example, in an MCMC algorithm. Using the exact likelihood it took 22,985 seconds to compute the 38,391 likelihood values needed for the likelihood surface, whereas using the approximate and hybrid likelihoods, were orders of magnitude faster, with the calculations taking 40 and 493 seconds, respectively.

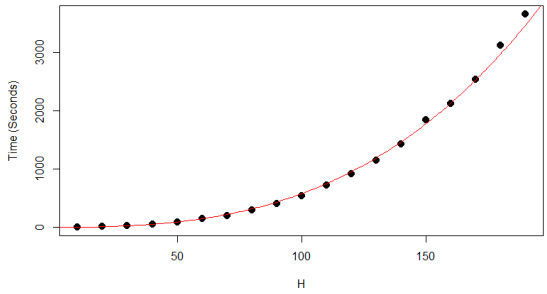

Figure 6: Plots of time taken (in seconds) to compute $10,000$ likelihood values for data $\tilde{\mathbf{t}}^*_{2:200}$ using the hybrid algorithm with $H = 10, 20, \ldots, 190$. The red line satisfies $\text{Time}^{1/3} = 0.7763 + 0.0755H$ and is obtained as the least-squares line of best fit for $\text{Time}^{1/3}$ against H.

In Figure 6 the time taken (in seconds) to calculate 10,000 likelihood values for the time-homogeneous birth-death process for different choices of H are recorded. The likelihood is computed on a 100×100 grid, for (α, μ) taking values $(0.100 + 0.002i, 0.010 + 0.002j)$, $(i = 1, 2, \ldots, 100; j = 1, 2, \ldots, 100)$. We observe that the time taken to compute the likelihood values is increasing approximately cubically in H. This is consistant with $H - 1$ matrix multiplications required for the hybrid algorithm with the k^{th} matrix multiplication (for $\mathbf{B}_k$) being $O(k^2)$.

We consider the use of the approximate and hybrid likelihoods for time-inhomogeneous birth-death processes. Below we provide details of four scenarios with the time-varying birth and death rates chosen to mimic potential changing infection (birth) and removal (death) rates in an epidemic process. The number of parameters in the three scenarios range from 4 to 7 and it is no longer feasible to compute the likelihood over a grid of values as in the time-homogeneous example above.

The four scenarios are:

1. Time-homogeneous birth-death process with a change-point in the birth rate after the L^{th} observed death (removal) and parameters $\boldsymbol{\theta} = (\beta_1, \beta_2, \gamma, \delta, L) = (0.3, 0.05, 0.1, 0.4, 50)$. Let $\tilde{\alpha}_j = \beta_1$ $(j = 1, 2, \ldots, L)$, $\tilde{\alpha}_j = \beta_2$ $(j = L+1, L+2, \ldots, 200)$, $\tilde{\mu}_j = \gamma$ and $\tilde{d}_j = \delta$ $(j = 1, 2, \ldots, 200)$.

2. Time-inhomogeneous birth-death process with the birth rate decreasing with each detected death.

The parameters are $\boldsymbol{\theta} = (\beta, \epsilon, \gamma, \delta) = (0.25, 0.2, 0.1, 0.25)$ with $\tilde{\alpha}_j = \beta\epsilon^{(j-1)/200}$, $\tilde{\mu}_j = \gamma$ and $\tilde{d}_j = \delta$ $(j = 1, 2, \ldots, 200)$. That is, β represents the initial birth rate and ϵ represents the relative drop in the birth rate after $N = 200$ detected deaths. We note that $\epsilon^{1/N} = 0.2^{1/200} \approx 0.992$ is the relative drop in the birth rate with every detected death.

3. Time-homogeneous process with a change-point in all the parameters after the L^{th} observed death (removal) with $\boldsymbol{\theta} = (\beta_1, \gamma_1, \delta_1, \beta_2, \gamma_2, \delta_2, L) = (0.2, 0.1, 0.2, 0.15, 0.15, 0.5, 100)$. Let $(\tilde{\alpha}_j, \tilde{\mu}_j, \tilde{d}_j) = (\beta_1, \gamma_1, \delta_1)$ $(j = 1, 2, \ldots, L)$ and $(\tilde{\alpha}_j, \tilde{\mu}_j, \tilde{d}_j) = (\beta_2, \gamma_2, \delta_2)$ $(j = L+1, L+2, \ldots, 200)$.

4. Time-inhomogeneous birth-death process with the birth rate decreasing, death rate increasing and detection probability increasing with each detected death. The parameters are $\boldsymbol{\theta} = (\beta, \epsilon_\beta, \gamma, \epsilon_\gamma, \delta, \epsilon_\delta) = (0.4, 0.25, 0.1, 0.5, 0.8, 0.25)$ with $\tilde{\alpha}_j = \beta\epsilon_\beta^{(j-1)/200}$, $\tilde{\mu}_j = \gamma\epsilon_\gamma^{(201-j)/200}$ and $\tilde{d}_j = \delta\epsilon_\delta^{(201-j)/200}$ $(j = 1, 2, \ldots, 200)$. Thus $(\beta, \gamma\epsilon_\gamma, \delta\epsilon_\delta)$ represents the initial birth rate, death rate and detection probability and $(\beta\epsilon_\beta, \gamma, \delta)$ represents the corresponding quantities after $N = 200$ detected deaths.

For each of the four scenarios a RWM algorithm is constructed along the same lines as the MCMC algorithm presented in Ball and Neal (2023), Section 4.2, with details provided there. The RWM algorithm is run with each of the three likelihoods with $m_B = 3$, $M_B = 5,000$ and $M = 20,000$, a total of 35,000 iterations per RWM algorithm. The (independent) priors chosen for the parametes are; discrete uniform on $\{1, 2, \ldots, 200\}$ for L, Gamma$(5, 5/\omega)$ for a rate parameter (β or γ) with true (simulated) value ω and Beta$(5\varsigma, 5(1-\varsigma))$ for a probability parameter (δ or ϵ) with true (simulated) value ς. The initial parameter values for the MCMC are chosen to be distinct from the values used in the simulation. The calculation times for the likelihoods are similar to those reported for the time-homogeneous example given above. The estimates of the posterior means and standard deviations of the parameters from the MCMC runs are given in Tables 4 - 7 for Scenarios 1 - 4, respectively. We observe that using each of the three likelihoods, within the MCMC algorithm, gives similar estimates of the posterior means and standard deviations of the parameters. The similarities between the MCMC algorithms with the different likelihoods extend to other measures of comparison, such as, trace plots and effective sample sizes of the MCMC outputs. The effective sample sizes of the parameters are in the ranges (915-1214), (702-1069),

Parameter	Simulation Value	Mean			Standard Deviation		
		Exact	Approximate	Hybrid	Exact	Approximate	Hybrid
β_1	0.3000	0.3056	0.3052	0.3051	0.0405	0.0394	0.0409
β_2	0.0500	0.0421	0.0418	0.0410	0.0176	0.0178	0.0178
γ	0.1000	0.1112	0.1110	0.1098	0.0195	0.0200	0.0195
δ	0.4000	0.3992	0.4017	0.4124	0.1442	0.1460	0.1503
L	50.00	57.14	56.87	56.59	10.22	10.00	10.13

Table 4: Posterior estimates of parameter means (standard deviations) for the parameters $(\beta_1, \beta_2, \gamma, \delta, L)$ for Scenario 1 (time-homogeneous model with a change in the birth rate) using the RWM algorithm.

Parameter	Simulation Value	Mean			Standard Deviation		
		Exact	Approximate	Hybrid	Exact	Approximate	Hybrid
β	0.2500	0.2165	0.2164	0.2154	0.0376	0.0359	0.0354
ϵ	0.2000	0.2160	0.2184	0.2168	0.1088	0.1168	0.1129
γ	0.1000	0.0741	0.0738	0.0734	0.0310	0.0306	0.0303
δ	0.2500	0.2210	0.2296	0.2235	0.1612	0.1725	0.1619

Table 5: Posterior estimates of parameter means (standard deviations) for the parameters $(\beta, \epsilon, \gamma, \delta)$ for Scenario 2 (time-inhomogeneous model with decaying infection rate) using the RWM algorithm.

(221-447) and (368-643) for scenarios 1-4, respectively. The relatively low effective sample size in scenario 3 is owing to the low variability in L with $L = 101$ having 41.7% of the posterior probability. Fixing $L = 100$, and rerunning the MCMC led to effective sample sizes in the range (245-550) for the parameters. Throughout, the estimated posterior distribution of δ closely resembles its prior demonstrating limited power in identifying the detection rate. However, in Scenario 3 there is positive correlation of 0.486 in the posterior samples of δ_1 and δ_2 providing evidence that a change in the relative rate of detection of deaths can be detected to some extent.

9 Covid-19 Analysis

In this section we provide supporting material, via a simulation study, for the analysis of the emergence of Covid-19 in Europe presented in Section 5. In Section 9.1, we show that it is difficult to identify the effect of the probability of detecting a removal from data supporting findings given in Section 8. In

Parameter	Simulation Value	Mean			Standard Deviation		
		Exact	Approximate	Hybrid	Exact	Approximate	Hybrid
β_1	0.2000	0.1887	0.1803	0.1863	0.0381	0.0343	0.0393
γ_1	0.1000	0.0931	0.0875	0.0914	0.0351	0.0323	0.0357
δ_1	0.2000	0.2196	0.2263	0.2199	0.1232	0.1222	0.1335
β_2	0.1500	0.1552	0.1532	0.1577	0.0592	0.0603	0.0590
γ_2	0.1500	0.1299	0.1349	0.1288	0.0523	0.0547	0.0508
δ_2	0.5000	0.5185	0.5005	0.4960	0.2021	0.1853	0.1850
L	100.00	101.01	100.96	100.83	2.84	2.46	2.44

Table 6: Posterior estimates of parameter means (standard deviations) for the parameters $(\beta_1, \gamma_1, \delta_1, \beta_2, \gamma_2, \delta_2, L)$ for Scenario 3 (time-homogeneous model with a change in the birth rate) using the RWM algorithm.

Parameter	Simulation Value	Mean			Standard Deviation		
		Exact	Approximate	Hybrid	Exact	Approximate	Hybrid
β	0.4000	0.3858	0.3828	0.3844	0.0602	0.0611	0.0578
ϵ_β	0.2500	0.2810	0.2702	0.2693	0.1721	0.1727	0.1680
γ	0.1000	0.0977	0.0943	0.0945	0.0458	0.0394	0.0394
ϵ_γ	0.5000	0.5012	0.5007	0.5127	0.1837	0.1829	0.1770
δ	0.8000	0.8048	0.7973	0.7872	0.1613	0.1604	0.1667
ϵ_δ	0.2500	0.2943	0.2897	0.2849	0.1350	0.1453	0.1368

Table 7: Posterior estimates of parameter means (standard deviations) for the parameters $(\beta, \epsilon_\beta, \gamma, \epsilon_\gamma, \delta, \epsilon_\delta)$ for Scenario 4 (time-homogeneous model with a change in the birth rate) using the RWM algorithm.

Section 9.2, we demonstrate that the epidemic model introduced in Section 5.2 is identifiable by using a simulation study for 5 countries and 5 non-pharmaceutical interventions (NPIs) with similar regimes for introducing NPIs as observed in Europe in March 2020.

9.1 Covid-19 Model appraisal

As noted at the end of Section 5.2, the effect of the probability of detecting a removal (a case results in death) is much harder to identify because there is very little information from the observed data. For a time-homogeneous birth-death process with constant birth rate α, death rate μ and detection probability d, we observe similar behaviour in the number of observed cases per day for all parameter sets with a given growth rate $\alpha - \mu$. This is because the expected number of individuals alive at time t in a time-homogeneous birth-death process satisfies $\mathbb{E}[B(t) \mid B(0)] = B(0)\exp(\{\alpha - \mu\}t)$, depending only on the parameters through the growth rate, and d does not affect the ratio of the expected number of deaths from one day to the next. To illustrate this, in Figure 7 we show the number of observed deaths per day for six birth-death processes with a changepoint in the birth rate. For $i = 1, 2, \ldots 6$, the i^{th} birth-death process has death rate $\mu = 0.05(i + 1)$. All birth-death processes have growth rate 0.10 prior to the changepoint and growth rate -0.05 after the changepoint. For $i = 1, 2, \ldots, 6$, the detection probability of a death, d_i, is drawn independently from $U(0.1, 0.5)$. The changepoint occurs immediately after the first day on which more than 500 deaths are observed. After the change in the growth rate has occurred, 20 further days of data are observed. In Figure 7 the number of detected deaths per day for the final 50 days are plotted, and hence the changepoint occurs at day 30 in the plot. We observe that the plotted data for the number of detected deaths per day for the six birth-death processes are virtually indistinguishable.

9.2 Covid-19 Simulation

The simulation study consists of the outbreak of a disease across five countries with five NPIs implemented to control the disease using the model outlined in Section 5.2. The data consist of the number of detected cases (corresponding to deaths in Covid-19) per day, and hence, data augmentation is used to impute the arrival (inter-arrival) times of detected cases. We assume that there is no delay between detection of

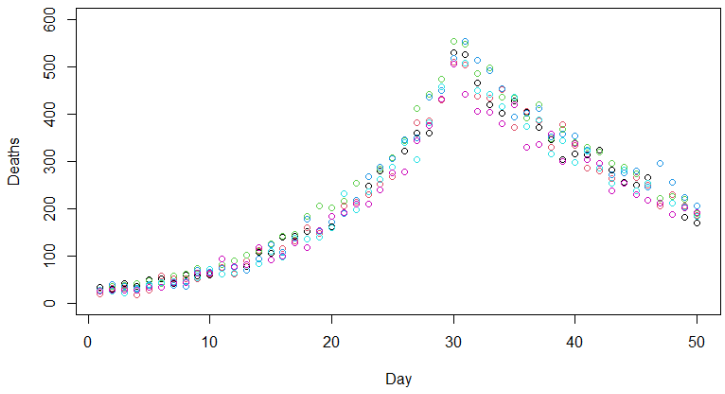

Figure 7: The number of detected deaths per day for the final 50 days of birth-death processes with changepoints. The growth rate is 0.10 before the changepoint and -0.05 after the changepoint with the changepoint occurring at day 30 in the plot corresponding to the first day on which the number of detected deaths exceeds 500. The death rates (probabilities of detection) μ (d) are 0.10 (0.1508) - black, 0.15 (0.3250) - red, 0.20 (0.3635) - green, 0.25 (0.2898) - blue, 0.30 (0.4255) - cyan, 0.35 (0.1625) - purple.

	Control 1	Control 2	Control 3	Control 4	Control 5
Country A	15	12	20	25	25
Country B	18	22	22	22	27
Country C	12	15	22	20	20
Country D	16	20	20	14	24
Country E	15	12	18	22	20

Table 8: Day control measures are implemented in simulation study.

cases and reporting but we would obtain the same data with a time shift if we assumed a constant delay between detection of cases and reporting, as we assume for the Covid-19 data. The daily case counts are available for between 45 and 55 days after the first detected case in the five countries with control measures implemented between the twelth and twenty seventh day after the first detected case. The days after the first detected cases on which control measures are implemented are given in Table 8. The length of time between the first and last control measure being implemented in a country ranges from nine to thirteen days. In all countries, except Country E, there are at least two control measures implemented on the same day.

The simulation study parameters were; $\boldsymbol{\alpha} = (0.4, 0.35, 0.45, 0.38, 0.42)$, $\mathbf{d} = (0.10, 0.08, 0.09, 0.11, 0.12)$, $\mu = 0.1$, $\boldsymbol{\zeta} = (0.9, 0.8, 0.8, 0.75, 0.35)$ and $\boldsymbol{\xi} = (-0.15, -0.01, 0.07, -0.24, 0.17)$. Therefore the mean infectious period is 10 days with between 8% and 12% of cases detected. The initial values of R_0 range from 3.5 to 4.5. The control measures are increasingly effective with $\prod_{j=1}^{5} \zeta_j = 0.1512$ corresponding to control measures reducing transmission rates by approximately 85% with the effect of $\boldsymbol{\xi}$ leading to the reduction in transmission rate between 82% and 89%. The number of detected cases in the five countries were 1740, 1067, 6785, 345 and 4123, respectively, and exhibit similar levels of variability to the 11 European countries in the Covid-19 analysis.

The MCMC algorithm was initiated with $\alpha_i = 0.4$, $\xi_i = 0$ $(i = 1, 2, \ldots, 5)$, $\zeta_j = 0.7$ $(j = 1, 2, \ldots, 5)$ $d = 0.1$ and $\mu = 0.1$, corresponding to an initial $R_0 = 4$ in each country, no country level variation in the effect of the implementation of the final NPI and each NPI reducing infectivity by 30%. The priors on α_i $(i = 1, 2, \ldots, 5)$ were Gamma$(10, 25)$ and μ was Gamma$(10, 100)$ with prior means of 0.4 and 0.1, respectively. The priors on ξ_i $(i = 1, 2, \ldots, 5)$ were $N(0, 0.1^2)$. The prior on d was Beta$(10, 90)$ corresponding to a prior mean of 0.1. Finally, Beta$(1, 1)$ (uniform) priors were chosen for ζ_j $(j = 1, 2, \ldots, 5)$. Note that all priors are mutually independent. Exploratory analysis led to choosing an initial $\Sigma = 10^{-5}\mathbf{I}$ for the random walk of the parameters. The implementation of the MCMC algorithm was as described in Section 5.3, with 10% of detection times proposed to be updated at each iteration.

The estimated probability density function of the marginal posterior of d closely resembles the Beta$(10, 90)$ prior, confirming that the data contain little information about the detection probabilities. Similar results are observed if each country has its own probability d_i of a removal resulting in death. The posterior mean for μ is 0.0997, which is close to the true value (and prior mean) of $\mu = 0.1$. Note that the posterior standard deviation of μ, 0.0131, is considerably smaller than the prior standard deviation of 0.1000. The growth rates for the five countries pre-interventions and post-interventions (implementation of NPIs) are presented in Table 9. The results show good agreement between the values chosen for the simulations and the posterior means, both pre- and post-intervention. The growth rates for post-intervention are approximately 10 times smaller in magnitude than the pre-intervention growth rates and this is reflected in both the posterior means and standard deviations of the estimates. Comparisons of the reproduction

Growth rate	Country A	Country B	Country C	Country D	Country E
Pre-intervention (Simulation)	0.3000	0.2500	0.3500	0.2800	0.3200
Pre-intervention (Posterior mean)	0.3087	0.2251	0.3421	0.2505	0.3591
Pre-intervention (Posterior sd)	0.0204	0.0133	0.0174	0.0275	0.0191
Post-intervention (Simulation)	-0.0479	-0.0476	-0.0270	-0.0548	-0.0247
Post-intervention (Posterior)	-0.0446	-0.0541	-0.0281	-0.0505	-0.0290
Post-intervention (Posterior sd)	0.0030	0.0042	0.0019	0.0056	0.0020

Table 9: Comparison of posterior means and standard deviations of growth rates of the disease pre-interventions and post-interventions (implementation of all NPIs) with simulation values

R_t	Country A	Country B	Country C	Country D	Country E
Pre-intervention (Simulation)	4.0000	3.5000	4.5000	3.8000	4.2000
Pre-intervention (Posterior mean)	4.1525	3.2959	4.4924	3.5565	4.6684
Pre-intervention (Posterior sd)	0.4742	0.3206	0.4976	0.4434	0.5446
Post-intervention (Simulation)	0.5206	0.5239	0.7297	0.4520	0.7527
Post-intervention (Posterior)	0.5463	0.4504	0.7134	0.4872	0.7041
Post-intervention (Posterior sd)	0.0584	0.0609	0.0443	0.0665	0.0443

Table 10: Comparison of posterior means and standard deviations of R_t of the disease pre-interventions and post-interventions (implementation of all NPIs) with simulation values

number, R_t, before and after intervenions, for each country give similarly good results between the posterior means and the values used for the simulations, see Table 10.

Declarations

The simulated data used in Sections 8 and 9.2 and the R code used in the analysis are available at: https://github.com/peteneal77/BDLikelihood