# Mathematical Statistics
# Stockholm University

# Analysing the Effect of Test-and-Trace Strategy in an SIR Epidemic Model

Dongni Zhang
Tom Britton

# Research Report 2021:7

**Postal address:**
Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden


**Internet:**
http://www.math.su.se

# Analysing the Effect of Test-and-Trace Strategy in an SIR Epidemic Model

Dongni Zhang and Tom Britton

October 2021

**Abstract**

Consider a Markovian SIR epidemic model in a homogeneous community. To this model we add a rate at which individuals are tested, and once an infectious individual tests positive it is isolated and each of their contacts are traced and tested independently with some fixed probability. If such a traced individual tests positive it is isolated, and the contact tracing is iterated. This model is analysed using large population approximations, both for the early stage of the epidemic when the "to-be-traced components" of the epidemic behaves like a branching process, and for the main stage of the epidemic where the process of to-be-traced components converges to a deterministic process defined by a system of differential equations. These approximations are used to quantify the effect of testing and of contact tracing on the effective reproduction numbers (for the components as well as for the individuals), the probability of a major outbreak, and the final fraction getting infected. Using numerical illustrations when rates of infection and natural recovery are fixed, it is shown that Test-and-Trace strategy is effective in reducing the reproduction number. Surprisingly, the reproduction number for the branching process of components is not monotonically decreasing in the tracing probability, but the individual reproduction number is conjectured to be monotonic as expected. Further, in the situation where individuals also self-report for testing, the tracing probability is more influential than the screening rate (measured by the fraction infected being screened).

# 1 Introduction

An important reason for modelling the spread of infectious diseases lies in better understanding the effect of various preventive measures, such as lockdown, social distancing, contact tracing, testing, self-isolating and quarantining. During the ongoing pandemic of Covid-19 the Test-and-Trace (TT) strategy has received a lot of attention (Kendall et al., 2020; Lucas et al., 2020; Bradshaw et al., 2021). The test-part of the strategy means that testing (of suspected cases and/or randomly chosen individuals) is increased, with the hope that finding infected individuals quickly and isolating them will reduce the transmission. The trace-part of the strategy is that individuals who are tested positive are quickly questioned about their recent contacts, and such contacts are then localized, tested and isolated if testing positive.

Contact tracing and its effect has been studied both from a theoretical perspective, and during the ongoing Covid-19 pandemic also from an applied point of view, including procedures for estimation of model parameters. For Covid-19 it has been observed in several countries that contact tracing is a highly powerful intervention measure. For instance, in the UK study (Kendall et al., 2020) of the TT-program first carried out on the Isle of Wight, they concluded that the number of new confirmed Covid-19 cases decreased more sharply after the TT-intervention. In (Lucas et al., 2020), they found that it is unlikely for strict self-isolation policies to improve the effectiveness by means of contact tracing.

More theoretical studies often make simplifying assumptions in order to make more analytical progress. For example, under the assumption of a homogeneous mixing population, (Ball et al., 2011, 2015) consider the traditional SIR model with forward tracing (where either a fraction of the infectees of a parent case is tested or none of the contacts is reported) but without contact tracing backwards to infectors of tested individuals. The model in (Bradshaw et al., 2021) suggests that both backward and forward tracing would remarkably increase the effectiveness of Covid-19 epidemic control. (Müller et al., 2000) deals with a stochastic SIRS model among a homogeneous mixing population. Once infectious individuals are discovered, each of their possible infectious contacts will be traced and treated (equivalent to isolated) with some probability. It is analysed gradually in three cases, namely backward tracing, forward tracing and tracing both ways. They derive the critical tracing probability for reducing the effective reproduction number to below 1 so that a major outbreaks no longer may occur. (Ball et al., 2011) is concerned with a SIR epidemic model in a homogeneous mixing population. In contrast to the tracing scheme studied in (Müller et al., 2000), where all individuals will be eventually be traced, diagnosed individ-

uals in (Ball et al., 2011) are asked to name a fraction of their infectees, who will be isolated (i.e. removed) immediately if they have not been diagnosed earlier. Further, those traced individuals are asked to name their infectious contacts in the same way, otherwise none of their contacts will be named. The model studied in (Ball et al., 2015) extends the one in (Ball et al., 2011) by introducing the exposed period and tracing delays, where the infectious individuals who share the same infector are traced after independent delay times. It is also assumed that untraced individuals may not be asked to name their infectees, for instance when they are asymptomatic. Numerical results in (Ball et al., 2015) indicate that independent delay times have bigger effect on the spreading compared to the situation where the delay times from one infected individual being contact traced are all the same.

In the present paper we consider an SIR model with TT strategy in which not all individuals necessarily are traced, tracing is both backward and forward, but on the other hand assuming no latent periods and no delay before contact tracing happens. We study a stochastic epidemic model including Test-and-Trace prevention for a large finite population. The Test-feature is modelled by assuming that infectious individuals are tested (screened) at a constant rate, and for tracing we assume that an individual who tests positive reports each contact independently and reported contacts are traced and tested without delay. The contact tracing is iterated for contacts who test positive (also if they have recovered naturally), contrary to the assumptions in (Müller and Hösel, 2021).

More precisely, we analyse a homogeneous SIR epidemic model having four parameters: the rate of infectious contacts $\beta$, the recovery rate $\gamma$, the testing rate $\delta$ for infectious individuals, and the fraction $p$ of contacts that are reached in the contact tracing procedure. Using large population approximations, we analyse both the initial phase of the epidemic (where it behaves like a certain branching process) and the main phase of the epidemic when it can be approximated by a deterministic process. The main focus of the paper is to shed light on how much is gained from the TT-strategy. For example, how should resources optimally be distributed between testing and contact tracing, how much would the reproduction number be reduced for achievable levels in the TT-strategy, and so on.

In Section 2 we present the model details and our main results and some intuitions for how the results are obtained. In the Section 3 and 4 we give more details and proofs to the analyses of the initial stage of the epidemic and its main phase in Section 5. In Section 6 we report simulations and numerical studies of the model and study how effective the TT-strategy is. The paper ends with a conclusion where extensions and possible improvements are discussed.

3

# 2   Model and Main Results

## 2.1   The Standard SIR Model

We start with a Markovian **SIR** (**S**usceptible $\rightarrow$ **I**nfectious $\rightarrow$ **R**ecovered) epidemic spreading in a closed and homogeneous mixing population. By closed, we mean that there is no influx of new susceptibles or death. At any time point, each individual is either susceptible, infectious or recovered. We assume that the size of population is $n$, and that initially one individual is infectious individual and the rest are susceptible. Each susceptible becomes infectious once he/she makes contact with an infective. Times at which such contacts between two given individuals occurs are constructed by a homogeneous Poisson process with rate $\beta/n$. Equivalently, an infectious individual has contacts at rate $\beta$, each time with a uniformly selected individual each thus having probability $1/n$. Only contacts with susceptibles result in infection whereas other contacts have no effect. Once an individual gets infected, he/she remains infectious for a random time $T_I$. We assume that the period $T_I$ is independent, exponentially distributed with mean $\mathbb{E}[T_I] = 1/\gamma$, and the parameter $\gamma$ denotes the rate of natural recovery. So the underlying model is Markovian. Once naturally recovered, the individual plays no role in the spreading of epidemic. The epidemic stops when there is no infectives.

## 2.2   The Markovian SIR-TT model

Now we incorporate our Test-and-Trace scheme into this SIR model. It is additionally assumed that infectious individual are tested at rate $\delta$ (possibly also non-infectious individuals are tested at this rate but this has no effect and is hence not assumed). Individuals that test positive are called diagnosed and diagnosed individuals are immediately isolated thus not taking further part in disease spreading. So, infectious individuals can stop spreading disease either from natural recovery (rate $\gamma$) or from being tested and diagnosed (rate $\delta$). Individuals that are diagnosed are also contact traced. This is modelled by assuming that a diagnosed individual reports each of its infectious contacts (both the infector and infectees) independently with probability $p$. The individuals that are traced in this way are tested, and individuals that test positive (either still being infectious or by then having recovered) are then contact traced in the same way (so contact tracing is iterated among those that have been infected). To simplify modelling we assume no delay in this contact tracing and instead assume that it all happens instantaneously. This **SIR-TT** model can hence be seen as an upper bound on how effective a TT-strategy might be as the contact tracing procedure speed up. All con-

tact and reporting processes as well as infectious periods are defined mutually independent. In Table 1 we list all the model parameters.

It is possible to consider different models for how different individuals would report in relation to each other. For instance, would an infector $A$ report an infectee $B$ independent of whether the infectee $B$ would report the infector $A$? A "yes" to the answer could for example happen if what defines a "contact" is not clarified enough so what $A$ considers as a contact may not coincide what $B$ thinks, and a "no" could happen if some of the contacts are with friends/acquaintances and other contacts are between unknown people, e.g. on the bus. However, it is clear from the model description that once $A$ or $B$ are diagnosed and asked to report their contacts the reporting event in the opposite direction is useless. This is true also if the first reporting event resulted in not naming the other individual: a later contact tracing of that individual has no effect on the first individual since he/she has already been diagnosed. As a consequence, all models for how contacts report each other (independently, symmetrically or some partial dependence) will result in exactly the same stochastic model. In our description below we have chosen to use the symmetric description thus assuming that A reports B if and only if B reports A, but this is only for practical purposes.

Our model assumes that infectious individuals either recover naturally (at rate $\gamma$) or are tested and diagnosed at rate $\delta$. There is an **alternative model interpretation**, which is more detailed in the sense that testing could take place also prior to screening. In addition to those who are found by screening, some infectious individuals may test themselves, e.g. due to symptoms. This scenario also fits into the present model by simply adding one more parameter. The parameters $\gamma$ and $\delta$ are unchanged: $\gamma$ is the rate of natural recovery and $\delta$ denotes the testing rate (screening). But now we add a rate $\nu$ at which infectious individuals self-report and test themselves. Both self-reporting and screening trigger contact tracing, so all that matters for the epidemic spreading is the sum $\nu + \delta$ of these two rates. As a consequence, this new model interpretation with 5 parameters $(\beta, \gamma, \delta, \nu, p)$ is identical to the original model with the following 4 parameter values $(\beta, \gamma, \delta + \nu, p)$. Since the alternative model interpretation fits into the original model all mathematical results from the original model apply. In Section 6 we give some numerical results also for the alternative model.

## 2.3   Main Results

In Section 3, we apply coupling methods to derive the following theorem for the initial stage of the epidemic using large population approximations. First we define the limit process, which is denoted by $E(\beta, \gamma, \delta, p)$ (for the limit

**Table 1** Table with all model parameters

| Parameter | Notation |
| --- | --- |
| Infection rate | $\beta$ |
| Rate of natural recovery | $\gamma$ |
| Rate of diagnosis | $\delta$ |
| Tracing probability | $p$ |
| Size of population | $n$ |

process we use "alive" corresponding to being infectious).

We assume that there is one alive ancestor at time zero. Each alive individual gives birth at rate $\beta$, dies naturally with rate $\gamma$ and is removed from the population with rate $\delta$ (corresponding to being tested and diagnosed in the epidemic). Further, an individual who is removed also leads to that each of its offspring as well as its parent will be removed independently with probability $p$. All those that are removed in this step will in turn lead to that its parent and offspring will be removed independently with probability $p$ and so on. This limit process is identical to the SIR-TT model defined above (denoted by $E_n(\beta, \gamma, \delta, p)$) with one single exception. In the epidemic model an infectious individual infects new individuals at rate $\beta * (S_n(t)/n)$ where $S_n(t)$ denotes the number of susceptibles at $t$, since only contacts with susceptibles result in infection and this has probability $S_n(t)/n$. In the limit process alive individuals give birth at constant rate $\beta$. Nevertheless, in the beginning and assuming a large population then these two rates will be close to each other since then $S_n(t) \approx n$.

The coupling used in Section 3 is however not for the epidemic and limit processes defined above. Rather than studying *individuals* we analyse the process of to-be-reported *components* (of individuals). More precisely, a new infection/birth is immediately decided if the involved individuals would report the other (with probability $p$) or not. If it will, then the new individual belongs to the same component but if it will not, the newly infected/born will create a new to-be-reported component. The reason for studying this more complicated description of the same process is that the to-be-reported components of the limit process behave completely independent thus making it a branching process. It is hence possible to use theory for branching process to determine if the process is sub- or super-critical and derive the probability for extinction/minor outbreak. We are now ready for our first main result.

**Theorem 1.** *For any finite time interval $[0, t_0]$, the SIR-TT epidemic process*

$E_n(\beta, \gamma, \delta, p)$ converges in distribution to the limit process $E(\beta, \gamma, \delta, p)$ as $n \to \infty$.

In Section 4 we use coupling methods to show that the branching process of to-be-reported components converges to the epidemic described in terms of to-be-reported components. Having done this it remains to derive properties of such a limiting to-be-reported component. It turns out that such a to-be-reported component can be described by a jump Markov chain having births (increased by 1), deaths (decreased by one) and killing (the whole component being removed), all occurring with linear rates. Suppose that there are currently $k$ alive individuals in the component, then each such individual gives birth to a new to-be-reported individual at rate $\beta p$ and thus the total birth rate is $k\beta p$. Each individual dies naturally at rate $\gamma$ so the overall death rate equals $k\gamma$. Finally, the whole component is removed as soon as one of the $k$ alive individuals is removed, so this happens at rate $k\delta$. Until the component is removed, it generates index cases to new independent to-be-reported components at rate $k\beta(1-p)$. This describes the evolution of the to-be-reported components. Viewed as a branching process, the most interesting quantity is the distribution of the number of offspring $Z$ (= index cases of new to-be-reported components) that one to-be-reported component produces before being removed. The mean offspring distribution, corresponding to the reproduction number of the components in the epidemic setting, is then given by

$$R_*^{(c)} = \mathbb{E}[Z]. \tag{1}$$

By considering the jump Markov chain we can write the total offspring $Z$ as a sum

$$Z = \sum_{i=1}^{N_C} X_i, \tag{2}$$

where $N_C$ denotes the number of jumps the Markov process makes until it is removed, and $X_i$ denotes the number of newly generated index cases between the $(i-1)$-th and $i$-th jump. Because all three jumps the process can make (birth, death and removal) happen at linear rates, the current number of alive individuals only affects the speed of the process but not which jump it makes. As a direct consequence, the components $X_1, X_2, \ldots$ are not only independent but also identically distributed: $X_i \sim X$. It hence follows that

$$R_*^{(c)} = \mathbb{E}[N_C]\mathbb{E}[X]. \tag{3}$$

In Section 4 we show that

$$\mathbb{E}[X] = \frac{\beta(1-p)}{\beta p + \gamma + \delta}, \tag{4}$$

and

$$\mathbb{E}[N_C] = 1 + \sum_{k=1}^{\infty} \mathbf{P}(N_c > k), \tag{5}$$

where

$$\mathbf{P}(N_C > k) = \left(1 - \sum_{j=1}^{\lceil k/2 \rceil} \frac{1}{2j-1} \binom{2j-1}{j} \left(\frac{\beta p}{\gamma + \beta p}\right)^{j-1} \left(\frac{\gamma}{\gamma + \beta p}\right)^{j}\right)$$
$$\cdot \left(\frac{\beta p + \gamma}{\beta p + \gamma + \delta}\right)^{k}. \tag{6}$$

The reproduction number defined above was for the to-be-reported components (the average number of new components it produces before being removed, i.e. completely diagnosed). Even though the original limit process is not a branching process, it is possible to determine the effective reproduction number $R_*^{(ind)}$ for it. Its interpretation is easier: it equals the average number of individuals a typical infected infects during the early stage of the epidemic.

In Section 4 we derive the following relation between the two reproduction numbers.

$$R_*^{(ind)} = \frac{\mu_c - 1 + R_*^{(c)}}{\mu_c} = 1 - \frac{1}{\mu_c} + \frac{R_*^{(c)}}{\mu_c}, \tag{7}$$

with

$$\mu_c = 1 + \frac{\beta p}{\beta p + \gamma} \mathbb{E}[N_c] \tag{8}$$

the expected number of born individuals before the component is removed.

**Remark 2.** *It is easily observed that $R_*^{(ind)} < 1$ if and only if $R_*^{(c)} < 1$, and similarly for "=" and ">". The limit process is hence sub-critical (i.e. will die out with probability 1), when $R_*^{(c)} < 1$ and super-critical if $R_*^{(c)} > 1$ (so will grow beyond all limits with positive probability). The same holds true if $R_*^{(c)}$ is replaced by $R_*^{(ind)}$. This indicates the following corollary.*

**Remark 3.** *In Section 6 the two reproduction numbers are computed numerically for different parameter values. Surprisingly, the component reproduction number $R_*^{(c)}$ turns out not to be monotonically decreasing in the tracing probability $p$. However, the individual reproduction number seems to be decaying in $p$ as expected. We have failed in producing a formal proof of this result.*

**Corollary 4.** *Let $Z_n$ denote the final number, and $\bar{Z}_n = Z_n/n$ the final fraction, that get infected during the entire epidemic. If $R_*^{(ind)} \leq 1$ it then*

follows that $\bar{Z}_n \xrightarrow{p} 0$ namely there will be a minor outbreak for sure. If $R_*^{(ind)} > 1$, then $Z_n \to \infty$ with probability $1 - \pi$ where $\pi$ is the smallest solution on $[0, 1]$ of the equation

$$s = \rho_Z(s) = \rho_{N_c}(\rho_X(s)), \tag{9}$$

where $\rho_Z$, $\rho_{N_c}$ and $\rho_X$ are the probability generating functions of $Z, N_C$ and $X$, respectively.

In the last part of Section 4, we give special attention to the case where there is no natural recovery ($\gamma = 0$) which accordingly can be called the **SI-TT** model. In this situation, the expressions become simpler and are given in the following corollary.

**Corollary 5.** *In the SI-TT model having $\gamma = 0$, the component reproduction number is given by*

$$R_{*,SI-TT}^{(c)} = \frac{\beta(1 - p)}{\delta}, \tag{10}$$

*the individual reproduction number equals*

$$R_{*,SI-TT}^{(ind)} = \frac{\beta}{\beta p + \delta}, \tag{11}$$

*and the minor outbreak probability becomes*

$$\pi = \frac{1}{R_{*,SI-TT}^{(c)}} = \frac{\delta}{\beta(1 - p)}. \tag{12}$$

**Remark 6.** *Again in this case, we see from Equation (10) and (11) that $R_{*,SI-TT}^{(c)}$ is smaller than or equal to or larger than 1, if and only if $R_{*,SI-TT}^{(ind)}$ is smaller than or equal to or larger than 1, respectively.*

We now switch attention to the main phase of the epidemic rather than its beginning (the corresponding proofs are given in Section 5). In order to surpass the initial phase of the epidemic we therefore assume a small initial *fraction $\varepsilon > 0$* of infectives (instead of only one initial infective).

We start by introducing notations for the epidemic and its limiting process, where we keep track of the fraction of susceptibles as well as the fractions of infectives belonging to to-be-reported components with each given number of infectives.

For $t \geq 0$ and $\varepsilon > 0$, let $S^{(n)}(t)$ denote the number of susceptible individuals with initial value $S^{(n)}(0) = (1 - \varepsilon)n$. For $j = 1, 2, ..., n$, let $I_j^{(n)}(t)$

be the number of infectious individuals that belong to a to-be-reported component containing $j$ infectives, and $I^{(n)}(t) = \sum_{j=1}^{n} I_j^{(n)}(t)$ denotes the total number of infectious individuals at time $t$, with initial values $I^{(n)}(0) = I_1^{(n)}(0) = \varepsilon n$, and $I_2^{(n)}(0) = I_3^{(n)}(0) = \cdots = 0$. Let $R^{(n)}(t)$ denote the number of individuals who stop being infectious including both naturally recovered and diagnosed, with initial value $R^{(n)}(0) = 0$. Since it always hold that $S^{(n)}(t) + I^{(n)}(t) + R^{(n)}(t) = n$, we eliminate $R^{(n)}$ from our analysis. Further, let $E^{(n)} = \{E^{(n)}(t); t \geq 0\} = \{(S^{(n)}(t)/n, I_1^{(n)}(t)/n, I_2^{(n)}(t)/n, ..., I_n^{(n)}(t)/n\}$ be the stochastic epidemic density process which becomes infinite-dimensional, as the population size $n$ goes to infinity.

The limiting deterministic process denoted by $E^{\infty} = \{E^{\infty}(t); t \geq 0\} = \{s(t), i_1(t), i_2(t), \cdots\}$ is obtained by considering the jumps that the components make. An infection in a $j$-component moves the component to a $(j+1)$-component implying that $S$ is reduced by 1, $I_j$ reduced by $j$ and $I_{j+1}$ increased by $j+1$. A natural recovery in such a component increases $R$ by 1, decreases $I_j$ by $j$ and increases $I_{j-1}$ by $j-1$. Finally, a test-detection in such a component reduces $I_j$ by $j$ and increases $R$ by $j$.

In Section 5, we prove of the following proposition.

**Proposition 7.** *For $t \geq 0$, let $i(t) = \sum_{i=1}^{\infty} i_j(t)$, and we set*

$$s'(t) = -\beta s(t) i(t), \tag{13}$$

$$i_1'(t) = \beta(1-p)i(t)s(t) + \gamma i_2(t) - \beta p i_1(t)s(t) - (\gamma + \delta)i_1(t), \tag{14}$$

*for $j \geq 2$,*

$$i_j'(t) = \beta p j i_{j-1}(t)s(t) + \gamma j i_{j+1}(t) - \beta p j i_j(t)s(t) - (\gamma + \delta)j i_j(t), \tag{15}$$

*with the corresponding initial configuration*

$$s(0) = 1 - \varepsilon, \tag{16}$$

$$i(0) = i_1(0) = \varepsilon, \tag{17}$$

*for $j \geq 2$,*

$$i_j(0) = 0. \tag{18}$$

*Then the infinite-dimensional stochastic epidemic process $E^{(n)}$ converges to the deterministic process $E^{\infty}$ defined by Equation (13)-(18) on any finite time interval $[0, t_{end}]$.*

When proving the theorem we first truncate both systems such that there is a maximal component size $K$ making the processes finite dimensional,

for which theory of population processes gives convergence. Then we argue that the component sizes for the original processes will be exponentially small in maximal component size, thus making the truncated models good approximations of the original processes.

If the SIR-TT model was started with one initial infective the time it takes until a fraction $\varepsilon$, i.e. a number $n\varepsilon$, have been infected, tends to infinity. For this reason this initial condition does not converge to the deterministic process above. Similarly, the end of the epidemic where the final small fraction $\epsilon$ gets infected also takes longer and longer time the larger $n$ is. However, like in many similar but simpler epidemic models we expect that, when it comes to the final number getting infected, the start of the epidemic determines if there is a major outbreak or not, and end of the epidemic has negligible effect. We formulate this more precisely in the following corollary.

**Conjecture 8.** *Consider the SIR-TT epidemic starting with one initially infective. The final fraction infected $\bar{Z}_n$ converges to a two point distribution $\zeta$, where $\zeta = 0$ (minor outbreak) happens with probability $\pi$, and with probability $1 - \pi$, $\zeta = r_\infty = \lim_{\varepsilon \to 0} \lim_{t \to \infty} r(t)$ (major outbreak), where $\pi$ is defined in Corollary 4 and $r(t) = 1 - s(t) - i(t)$ in Proposition 7.*

**Remark 9.** *In Section 6, we show several simulations in support of Conjecture 8 and also indicating that the distribution of $\bar{Z}_n$ appears to satisfy a central limit theorem concentrated around the deterministic limit $r_\infty$.*

In Section 6 we perform simulations and numerical illustrations confirming our results and investigating the effect of Test-and-Trace strategy for parameter values inspired from the Covid-19 pandemic.

# 3   Early Stage Approximation of the Epidemic

In this section, we aim to approximate the early stages of the epidemic using large population approximations. We first denote the sequence of our epidemic processes with one initial infective by $\{E_n(\beta, \gamma, \delta, p), n \geq 1\}$, where we recall that $\beta$ is rate of infection, the rate of natural recovery is $\gamma$, $\delta$ denotes the testing (diagnosis) rate and the probability of a contact being reported equals $p$.

Then we describe the limiting process denoted by $E(\beta, \gamma, \delta, p)$. At time $t = 0$, there is only one initial ancestor. Each individual gives birth at rate $\beta$ during their lifetimes, dies naturally (naturally recovered) with rate $\gamma$ and is removed (diagnosed) with rate $\delta$. Once removed, each of its descendants and its parent is said to be reported and immediately removed independently

with probability $p$. Meanwhile, every parent and offspring of those who are removed, will be removed independently with probability $p$ as well and so on. In particular, if a to-be-reported individual has been already died naturally, its alive to-be-reported offspring (or parent) will also be removed. Moreover, we notice that each not to-be-reported offspring becomes a new ancestor which independently produces a process in the same pattern. Finally we show the proof of Theorem 1.

*Proof of Theorem* 1. First of all, it is worth noting that the two processes $E_n(\beta, \gamma, \delta, p)$ and $E(\beta, \gamma, \delta, p)$ behave the same way besides one slight difference. An infection occurs in the epidemic whenever a birth occurs in the branching process, whereas an infection is "effective" only if an susceptible gets infected. And in the $n$-th epidemic, the probability that an infective gives new infections to susceptibles is $S_n(t)/(n-1)$ ($\approx S_n(t)/n$ when n large), where $S_n(t)$ is the number of susceptibles at time $t$. So, an infective infects new individuals at rate $\beta S_n(t)/n$. In contrast to that, an alive individual in the limiting process give birth at rate $\beta$. However, if the size of population $n$ is large and in the beginning of epidemic we have $S_n(t) \approx n$, then we have $\beta S_n(t)/n \approx \beta$, i.e. the rate of new infection to susceptibles in $E_n(\beta, \gamma, \delta, p)$ is close to the birth rate in $E(\beta, \gamma, \delta, p)$.

As compared to the early stage approximation of standard SIR epidemic (Andersson and Britton, 2000), the number of alive individuals in this limiting process $E(\beta, \gamma, \delta, p)$ behaves not like a branching process, since it is possible that several death occur at the same time and thus the jumps of this limiting process can not only be up or down by one.

On the other hand, if the limiting process is described in terms of to-be-reported components, then it behaves like a branching process. A component starts with one newly born (infected) individual which would not report its infector and we call this individual the index case. During its life duration (infectious period) this individual gives birth to new individuals, some of which will be reported and others will not. Each of those not-to-be-reported individuals becomes a new index case of new components, whereas those who will be reported belong to the same component. Given that there are currently $k$ to-be-reported individuals in one component, then each such individual gives birth at rate $\beta$, where each newborn belongs to the same component with probability $p$ and generates a new component with probability $1-p$. Thus, each of individual in this component gives birth to new to-be-reported individuals at rate $\beta p$ and thus the total birth rate is $k\beta p$. Each alive (infectious) individual dies naturally(naturally recovered) at rate $\gamma$. The whole component is diagnosed if and only if one of those $k$ to-be-reported individuals is diagnosed, implying that the death rate of this process of components

is $k\delta$. Until all these $k$ individuals are removed, it generates new index cases at rate $k\beta(1-p)$. This describes the birth and death of the to-be-reported components.

By applying the coupling method in (Andersson and Britton, 2000), we show that the epidemic process $E_n(\beta, \gamma, \delta, p)$ described in terms of components converges in to the branching process of to-be-reported components. A contact (infection) in the epidemic corresponds to a birth in the branching process. Obviously, the branching process and the epidemic process of components are perfectly coupled with each other up until the time $T_n$ when the first "ghost" appears, where by "ghost" we mean the newly contacted individual which has been infected in the epidemic. If we label each $i-$th contact as $c_i$, then for any time $t_0 \geq 0$, the event $T_n \geq t_0$ that there has been no "ghost" occur before time $t_0$, is equivalent to the case that all the contacts $c_1, ..., c_{[t_0]}$ are distinct. Using the classic birthday-problem method, we see that

$$\mathbf{P}(T_n \geq t_0) \approx e^{-n^2/2[t_0]} \rightarrow 1,$$

as $n \rightarrow \infty$. This completes the proof of Theorem 1. $\qquad\square$

# 4 Properties of the limiting branching process

Now we explore the properties of the limiting branching process $E(\beta, \gamma, \delta, p)$ of to-be-reported components which can be used to approximate the epidemic during the early phase.

## 4.1 Process of the to-be-reported components in the full model

First, we note that our reporting process can be decided in advance, and recall that we use the same reporting or not decision in both ways between each pair of individuals since at most one direction will be used. We then focus on this Markov jump process of the components having births, deaths and sudden killing of the whole component.

We define the size $k$ of an to-be-reported component by the number of alive (infectious) individuals in the component and hence ignore the dead (recovered) individuals. A component currently having size $k$ produces new index cases of new components at rate $k\beta(1-p)$. The component itself remains with size $k$ for an exponentially distributed time with rate $k(\beta p+\gamma+\delta)$, next event would be one of the three following independent cases. The first case is that a new infection occurs at rate $k\beta p$, which corresponds to increasing the size of component by one. Secondly, we note that each of the
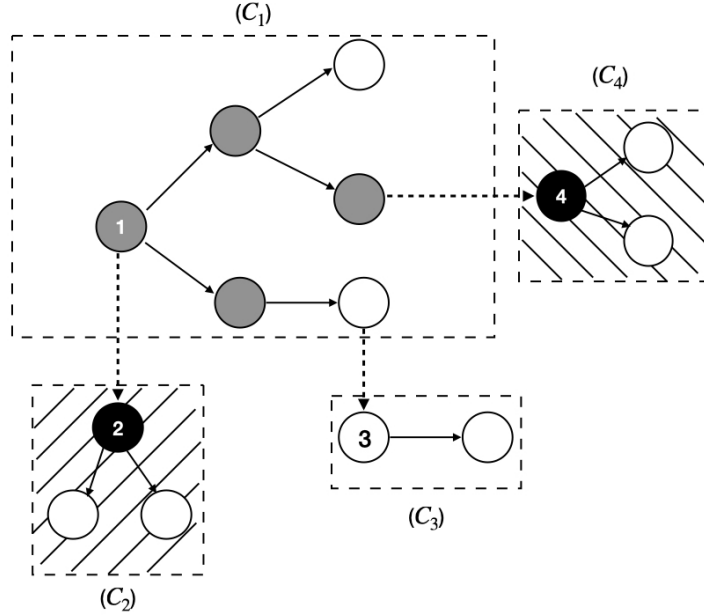
**Figure 1** Example of a "reporting tree": The white nodes stands for "infectious", the grey ones for "naturally recovered", whereas the black ones for "diagnosed". A directed edge from one node to another means that the latter one is infected by the previous one. Full edges reflect to-be-reported contacts (probability $p$) and the dashed ones for those not to be reported

individuals in the component becomes naturally recovered at rate $\gamma$. If this happens, then the size of component would decrease by one. The remaining case is that one of the to-be-reported individuals is diagnosed and so the whole component is diagnosed at rate $k\delta$, which means that upon this event, the size of component goes down to zero.

In Fig. 1, we show an example illustrating how a **"reporting branching tree"** of to-be-reported components grows: at first, we have a newly infected case, namely the node 1. An edge goes from one node to another node, meaning that the latter one is infected by the previous one. The dashed edge between two nodes means for the not-to-be-reported case, whereas the full edge stands for the to-be-reported case. After a certain period, there is a to-be-reported component, denoted by $C_1$, produced by our basic index case 1, and three newly generated index cases 2, 3 and 4, each of which produces their own to-be-reported components, denoted by $C_2$, $C_3$ and $C_4$ respectively. Furthermore, the white nodes stands for "infectious", the grey ones for "naturally recovered" and the black ones for "diagnosed". We can

see that at this stage when the index cases 2 and 4 are diagnosed, the whole components $C_2$ and $C_4$ are reported and immediately diagnosed.

Our interest is to derive the important quantity for our epidemic model, namely **the effective component reproduction number $R_*^{(c)}$**, which is defined as the expected number $R_*^{(c)} = \mathbb{E}[Z]$ of new index cases generated by one index case before its component is diagnosed. Since we aim at examining the effect of testing and tracing, so given fixed rates $\beta$ and $\gamma$, we consider $R_*^{(c)} = R_*^{(c)}(\delta, p)$ as a function of testing rate $\delta$ and tracing probability $p$. Later in Section 6, we will show how the $R_*^{(c)}$ varies with the testing fraction $\delta/(\delta + \gamma)$ and tracing probability $p$.

To find the distribution of $Z$, we first discuss the number of events before the whole component is diagnosed by computing the probability that the whole component is not diagnosed before $k$ events. We recall that at each time of event, there would be only one of three following events occurs. A birth occurs with rate $k\beta p$, whereas the death of whole component happens with rate $k\delta$ and the size of component decreases by one with rate $k\gamma$. As a consequence the probability of giving a birth, which corresponds to increasing the component size by one, is given by

$$\frac{k\beta p}{k\beta p + k\gamma + k\delta} = \frac{\beta p}{\beta p + \gamma + \delta},$$

the probability of death of the whole component equals

$$\frac{k\delta}{k\beta p + k\gamma + k\delta} = \frac{\delta}{\beta p + \gamma + \delta},$$

and the probability that the component size decreases by one, is given by

$$\frac{k\gamma}{k\beta p + k\gamma + k\delta} = \frac{\gamma}{\beta p + \gamma + \delta}.$$

A different way of describing the evolution of the component is to consider increases and decreases by one (a simple random walk!) until some time when the whole component dies simultaneously. It is worth pointing out that the random walk may reach zero by itself and hence stop before a simultaneous death. The non-symmetric simple random walk $\{S_n, n \geq 0\}$ on $\mathbb{Z}$ starts at 1 ($S_0 = 1$) and for $n = 1, 2, 3, \cdots$, each jump of the random walk is independent and identically distributed with the jump probabilities

$$\pi_{rw} = \mathbf{P}(S_n - S_{n-1} = 1) = \frac{\beta p}{\gamma + \beta p},$$

and

$$\mathbf{P}(S_n - S_{n-1} = -1) = 1 - \pi_{rw} = \frac{\gamma}{\gamma + \beta p}.$$

On top of this, each jump may result in diagnosis of the whole component (simultaneous death), and each time at which this happens with probability $\delta/(\beta p + \gamma + \delta)$. The number $N_D$ of events until the whole component is diagnosed is hence geometrically distributed with parameter $\delta/(\beta p + \gamma + \delta)$.

Next we derive the probability that the random walk does not hit zero before $k$ jumps. Let

$$N_{rw} = \inf_{n \geq 0}\{S_n = 0\}$$

denote the first hitting zero time of random walk and it is clear that only odd steps can be taken in order to hit the origin. So, for $m$ even, the probability we have $\mathbf{P}(N_{rw} = m) = 0$. Otherwise $m = 2j - 1$ for $j = 1, 2, \cdots$, we apply the Hitting Time Theorem in (Hofstad and Keane, 2008), which yields that the probability of first hitting zero at $m-$th step is given by

$$\mathbf{P}(N_{rw} = m) = \mathbf{P}(N_{rw} = 2j - 1) = \frac{1}{2j - 1}\mathbf{P}(S_{2j-1} = 0),$$

where the probability of the (unrestricted) random walk equals 0 at $m-$th step is

$$\mathbf{P}(S_{2j-1} = 0) = \binom{2j - 1}{j}\pi_{rw}^{j-1}(1 - \pi_{rw})^j,$$

since in this case, the random walk must have taken $(j - 1)$ up-jumps and $j$ down-steps. We conclude that the probability of the random walk *not* hitting zero before $k$ steps equals

$$\begin{aligned}
\mathbf{P}(N_{rw} > k) &= 1 - \sum_{m=1}^{k}\mathbf{P}_1(N_{rw} = m) \\
&= 1 - \sum_{j=1}^{\lceil k/2 \rceil}\mathbf{P}_1(N_{rw} = 2j - 1) \\
&= 1 - \sum_{j=1}^{\lceil k/2 \rceil}\frac{1}{2j - 1}\binom{2j - 1}{j}\pi_{rw}^{j-1}(1 - \pi_{rw})^j.
\end{aligned}$$

Moreover, let $N_C$ denote the number of jumps up until the whole to-be-reported component is extinct (either from simultaneous diagnosis or all individuals having recovered naturally), i.e.

$$N_C = \min\{N_{rw}, N_D\}.$$

Recalling that $N_D$ is geometric distributed, the probability that the whole

component has not gone extinct before $k = 1, 2, \cdots$, events is given by

$$\mathbf{P}(N_C > k) = \mathbf{P}(N_{rw} > k) \cdot \mathbf{P}(N_D > k)$$

$$= \left(1 - \sum_{j=1}^{\lceil k/2 \rceil} \frac{1}{2j-1} \binom{2j-1}{j} \left(\frac{\beta p}{\gamma + \beta p}\right)^{j-1} \left(\frac{\gamma}{\gamma + \beta p}\right)^{j}\right) \left(\frac{\beta p + \gamma}{\beta p + \gamma + \delta}\right)^{k}.$$

Now, it is sufficient to analyze the number $X_i$ of newly generated index cases between each $(i-1)-$th and $i-$th jump. Given a to-be-reported component of size $k$ at that time, the index cases are generated at rate $k\beta(1-p)$ for an exponential time of parameter $k(\beta p + \gamma + \delta)$. This implies that the distribution of $X_i$ is geometrically distributed with parameter

$$1 - \frac{k\beta(1-p)}{k(\beta p + \gamma + \delta) + k\beta(1-p)} = \frac{\beta p + \gamma + \delta}{\beta + \gamma + \delta}.$$

This parameter is independent of $k$ implying that the variables $X_1, X_2, \ldots$, are identically and independently geometrically distributed as $X$. So, between any two jumps, the probability of $k$ newly generated index cases is given by

$$\mathbf{P}(X = k) = \left(\frac{\beta(1-p)}{\beta + \gamma + \delta}\right)^{k} \frac{\beta p + \gamma + \delta}{\beta + \gamma + \delta}.$$

Based on the former discussion, we conclude that the total number of index cases produced by a to-be-reported component can be written as

$$Z = \sum_{i=1}^{N_C} X_i.$$

As stated in Equation (3), due to independence, it follows that

$$R_*^{(c)} = \mathbb{E}[Z] = \mathbb{E}[N_C] \cdot \mathbb{E}[X],$$

where the expectation of $X$ is given by

$$\mathbb{E}[X] = \frac{\beta(1-p)}{\beta p + \gamma + \delta},$$

and for $\mathbb{E}[N_C]$ we have

$$\mathbb{E}[N_C] = 1 + \sum_{k=1}^{\infty} \mathbf{P}(N_c > k).$$

Next we give the proof of Corollary 4.

*Proof of Corollary* 4. Intuitively, we note that the limiting process of components will become extinct when $\mathbb{E}[Z] \leq 1$. This implies that a minor outbreak will occur, if the component reproduction number $R_*^{(c)} = \mathbb{E}[Z]$ is smaller than or equal to one. Next, regarding to the situation when $R_*^{(c)} > 1$, the branching process is possible to explode, and so there will be a major outbreak in the epidemic. Now we put our focus on the probability $\pi$ of minor outbreak and the probability of major outbreak, namely $1 - \pi$.

We assume that there is one initial infective. As discussed in previous section, the probability of minor outbreak in the epidemic can be approximated by the probability of extinction in the limiting process at the early stage of outbreak. Given $k$ newly generated index cases, the conditional probability of extinction is then clearly $\pi^k$. Thus, the probability $\pi$ is the solution on $[0,1]$ of the following equation:

$$\pi = \sum_{k=1}^{\infty} \pi^k \mathbf{P}(Z = k), \tag{19}$$

where we note that the right-side of Equation (19) is exactly the probability generating function $\rho_Z(\pi)$ of $Z$. For the computation of $\rho_Z$, we have

$$\rho_Z(s) = \mathbb{E}[s^Z] = \mathbb{E}[s^{\sum_{i=1}^{N_C} X_i}] = \rho_{N_C}(\rho_X(s)),$$

where the probability generating function $\rho_X$ of $X$ is given by

$$\rho_X(s) = \mathbb{E}[s^X] = \frac{\theta}{1 - (1 - \theta)s}$$

with $\theta = (\beta p + \gamma + \delta)/(\beta + \gamma + \delta)$, and the probability generating function $\rho_{N_C}$ of $N_C$ is given by

$$\rho_{N_C}(t) = \sum_{k=1}^{\infty} t^k \mathbf{P}(N_C = k)$$

with

$$\mathbf{P}(N_C = k) = \mathbf{P}(N_{rw} = k)\mathbf{P}(N_D \geq k) + \mathbf{P}(N_{rw} > k)\mathbf{P}(N_D = k).$$

Finally we obtain the probability generating function $\rho_Z$ of $Z$:

$$
\begin{aligned}
\rho_Z(s) &= \sum_{k=1}^{\infty} \left( \frac{\theta}{1 - (1 - \theta)s} \right)^k \mathbf{P}(N_C = k) \\
&= \frac{\theta}{1 - (1 - \theta)s} \cdot (1 - \mathbf{P}(N_C > 1)) \\
&\quad + \sum_{k=2}^{\infty} \left( \frac{\theta}{1 - (1 - \theta)s} \right)^k (\mathbf{P}(N_C > k - 1) - \mathbf{P}(N_C > k)).
\end{aligned}
\tag{20}
$$

18

Solving the equation

$$s = \rho_Z(s)$$

with Equation (20) on $[0, 1]$ gives us the smallest solution $\pi$, which equals the probability of minor outbreak in the epidemic with one initial infective. In addition to that, if there are initially $m$ infectives, small outbreak occurs with probability $\pi^m$. $\qquad\square$

We then aim to derive the effective *individual* reproduction number $R_*^{(ind)} = R_*^{(ind)}(\delta, p)$, which equals the expected number of infected cases generated by a random infectious individual (before being tested or recovering).

We start with a new expression for our effective component reproduction number. Let $I_c$ be the overall number of individuals who have been infected in a to-be-reported component before it goes extinct, starting with one single infectious individual. For $k \geq 1$, let

$$p_k = \mathbf{P}(I_c = k)$$

be the probability that there have been in total $k$ individuals infected in a component, and let $r_k$ be the expected number of new index cases generated by such a component, given that $I_c = k$. It then follows that

$$R_*^{(c)} = \sum_{k=1}^{\infty} r_k p_k.$$

Further, during the early stage of epidemic, the probability $\tilde{p}_k$ that an individual belongs to a component with $I_c = k$, is given by

$$\tilde{p}_k = \frac{k p_k}{\mu_c}$$

(the size-biased distribution) with

$$\mu_c := \mathbb{E}[I_c] = \sum_{j=1}^{\infty} j p_j.$$

Given that $I_c = k$, there would be $(k-1)$ infections occurred in the component and $r_k$ average infections out of the component before it dies. In total such a component hence on average generate $k - 1 + r_k$ infections and randomly chosen individual hence infects $((k-1) + r_k)/k$ on average. This implies that our effective individual reproduction number is given by

$$R_*^{(ind)} = \sum_{k=1}^{\infty} \frac{(k-1) + r_k}{k} \tilde{p}_k = \sum_{k=1}^{\infty} ((k-1) + r_k) \frac{p_k}{\mu_c}. \qquad (21)$$

19

Simplifying the expression for $R_*^{(ind)}$ in Equation (21) gives us

$$R_*^{(ind)} = 1 - \frac{1}{\mu_c} + \frac{R_*^{(c)}}{\mu_c}.$$

The equation above shows us that $R_*^{(ind)}$ is smaller than, equal to, or larger than 1, if and only if $R_*^{(c)}$ is smaller than, equal to, or larger than 1.

Intuitively, we can explain this relation between $R_*^{(ind)}$ and $R_*^{(c)}$ as follows. On one hand, the $R_*^{(c)}$ can be considered as the average infections produced outside the component. On other hand, $\mu_c$ is the average number of individuals who have been infectious in the component. We notice that one of $\mu_c$ would be the index case, while the other $\mu_c - 1$ are internal infections in this component. Then there are in total $(R_*^{(c)} + \mu_c - 1)$ infections on average by this component. Hence, the average number of infections per individual becomes $(R_*^{(c)} + \mu_c - 1)/\mu_c$.

It remains to compute the expected number $\mu_c$ of the individuals who have been infectious in a component, since we have already derived $R_*^{(c)}$. We start by letting $J_+ = \sum_{k=1}^{\infty} I_k$ denote the number of up-jumps of the random walk $\{S_n, n \geq 0\}$ before it dies out, where $I_k$ is the indicator variable with $I_k = 1$ if the $k$-th jump of the random walk is an up-jump, i.e. for any $k \geq 1$,

$$\mathbf{P}(I_k = 1 | N_c > k - 1) = \pi_{rw}$$

and

$$\mathbf{P}(I_k = 1 | N_c \leq k - 1) = 0.$$

Then we conclude that

$$\mu_c = 1 + \mathbb{E}[J_+] = 1 + \sum_{k=1}^{\infty} \mathbf{P}(I_k = 1) = 1 + \pi_{rw} \sum_{k=1}^{\infty} \mathbf{P}(N_c > k - 1) = 1 + \frac{\beta p}{\beta p + \gamma} \mathbb{E}[N_c].$$

Together with Equation (7), we prove the final expression for $R_*^{(ind)}$ given by

$$R_*^{(ind)} = 1 - \frac{1}{1 + \frac{\beta p}{\beta p + \gamma} \mathbb{E}[N_c]} + \frac{R_*^{(c)}}{1 + \frac{\beta p}{\beta p + \gamma} \mathbb{E}[N_c]}. \tag{22}$$

## 4.2 The limiting process in the SI-TT Model

In the **SI-TT** model there is no natural recovery: $\gamma = 0$. This special case turns out to give simpler explicit expressions. The reason for the simplification is that a component can then only go extinct due to an infectious individuals being diagnosed (resulting in the whole to-be-reported component being

contact traced) whereas for the general model extinction may also happen because all individuals has recovered before a new to-be-reported infection took place.

We are interested in the number $Z_{SI-TT}$ of new index cases produced by a to-be-reported component before it dies (i.e. is diagnosed). As before, the current number of infectious individuals does not affect the probability of the next jump being a new index case or a diagnosis event. We can hence neglect the infections and conclude that there will be a geometrically distributed number of index cases before the component is diagnosed. The parameter of the geometric distribution is simply the probability of a diagnosis rather than a new index case: $\delta/(\delta + \beta(1-p))$. So for $n = 0, 1, 2, \cdots$, we have the unconditional density of $Z_{SI-TT}$ given by

$$\mathbf{P}(Z_{SI-TT} = n) = \Big(1 - \frac{\delta}{\delta + \beta(1-p)}\Big)^n \frac{\delta}{\delta + \beta(1-p)}.$$

Then in the SI-TT model, we are able to derive the **effective component reproduction number** $R_{*,SI-TT}^{(c)}$ as

$$R_{*,SI-TT}^{(c)} = \mathbf{E}[Z_{SI-TT}] = \frac{\beta(1-p)}{\delta}.$$

Again following the idea proving Corollary 4, we show Corollary 5 as follows.

*Proof of Corollary* 5. By finding the smallest solution $s$ on $[0,1]$ of equation $s = \rho(s)$ with $\rho(\cdot)$ the probability generating function of $Z_{SI-TT}$, we obtain that if $R_{*,SI-TT}^{(c)} > 1$, the probability of minor epidemic outbreak equals

$$\pi = \frac{\delta}{\beta(1-p)} = \frac{1}{R_{*,SI-TT}^{(c)}}. \tag{23}$$

In the case when $R_{*,SI-TT}^{(c)} \le 1$, the branching process will be extinct with probability $\pi = 1$, implying that a major outbreak occurs with probability 0. $\qquad\square$

Moreover, using the same idea of computing the **effective individual reproduction number** in general case, here we first have the expected number of infected cases generated by one index case before diagnosed given by

$$\mu_{c,SI-TT} = 1 + \mathbb{E}[N_c^{(SI-TT)}] = 1 + \frac{\beta p}{\delta} = \frac{\beta p + \delta}{\delta}.$$

Then as stated in Section 2, the effective individual reproduction number for this case without natural recovery has the form

$$R_{*,SI-TT}^{(ind)} = 1 - \frac{1}{\mu_{c,SI-TT}} + \frac{R_{*,SI-TT}^{(c)}}{\mu_{c,SI-TT}} = \frac{\beta}{\beta p + \delta}. \tag{24}$$

It is also easily observed from Equation (24) that the effective individual reproduction number in this SI-TT model is monotone decreasing with tracing probability $p$.

## 5 Approximation of the Main Phase

In previous section, we applied coupling methods to approximate the epidemic at its early stage. In this section, we give an approximation of the main phase, where the epidemic is initiated with positive fraction $\varepsilon$ of infectives. Here, we describe our original full model in the way of evolution of the clumps. By "clumps", we mean the to-be-reported components, and we only need to keep track of number of infectious individuals in each clump. In a population of size $n$, we assume that the number of initial infectives equals $\varepsilon n$ and the number of initial susceptibles equals $(1 - \varepsilon)n$. At time $t \geq 0$, let $S^{(n)}(t)$ be the number of susceptible individuals with initial value

$$S^{(n)}(0) = (1 - \varepsilon)n. \tag{25}$$

For $j = 1, 2, ..., n$, let $I_j^{(n)}(t)$ be the number of individuals that are infectious and belong to a to-be-reported component currently containing $j$ infectives. So, we have the total number of infectious individuals at time $t$,

$$I^{(n)}(t) = \sum_{j=1}^{n} I_j^{(n)}(t), \tag{26}$$

with initial values

$$I^{(n)}(0) = I_1^{(n)}(0) = \varepsilon n, \ I_2^{(n)}(0) = \cdots = I_n^{(n)}(0) = 0. \tag{27}$$

Let $R^{(n)}(t)$ denote the number of individuals which are recovered (counting both naturally recovered and diagnosed) with initial value

$$R^{(n)}(0) = 0. \tag{28}$$

It is then clear that for any time $t \geq 0$,

$$S^{(n)}(t) + I^{(n)}(t) + R^{(n)}(t) = n. \tag{29}$$

Next we prove Proposition 7, stating that the stochastic epidemic process denoted by

$$E^{(n)}(t) = (S^{(n)}(t)/n, I_1^{(n)}(t)/n, I_2^{(n)}(t)/n, ..., I_n^{(n)}(t)/n),$$

converges to a deterministic process.

*Proof of Proposition* 7. Below we study the corresponding truncated processes by maximizing the clump sizes to some large positive integer $K$. The corresponding processes are finite dimensional for which we apply theory for density dependent population processes. These results can then be extended to the original infinite dimensional systems (with arbitrary clump size) by observing that that the maximal clump sizes are exponentially small in $K$. As a consequence, the truncated processes can be made arbitrarily close to the original infinite dimensional processes by choosing $K$ large enough. The fact that the processes are exponentially small is a direct consequence of that the epidemic processes SIR-TT may be dominated by the SI-TT (without natural recovery), and this process will have a geometrically distributed maximal clump size with parameter $\delta/(\beta p + \delta)$. We omit the details of this argument and now should show that the truncated stochastic epidemic process converges to the truncated deterministic system.

More precisely, using Kurtz's theory of Markovian Population processes (Andersson and Britton, 2000), we show that the truncated stochastic "density" process, denoted by

$$E_K^{(n)}(t) = (S^{(n)}(t)/n, I_1^{(n)}(t)/n, I_2^{(n)}(t)/n, ..., I_K^{(n)}(t)/n)$$

converges to a $K-$dimensional deterministic process

$$E_K^\infty(t) = (s(t), i_1(t), i_2(t), ..., i_K(t)),$$

which is defined by the finite system of differential equations as below.

$$s'(t) = -\beta s(t)i(t), \tag{30}$$

$$i_1{}'(t) = \beta(1 - p)i(t)s(t) + \gamma i_2(t) - \beta p i_1(t)s(t) - (\gamma + \delta)i_1(t), \tag{31}$$

and for $j = 2, 3, ..., (K - 1)$ we have

$$i_j{}'(t) = j\beta p i_{j-1}(t)s(t) + j\gamma i_{j+1}(t) - j\beta p i_j(t)s(t) - j(\gamma + \delta)i_j(t), \tag{32}$$

whereas in the case of $j = K$,

$$i_K{}'(t) = K\beta p i_{K-1}(t)s(t) - K\beta p i_K(t)s(t) - K(\gamma + \delta)i_K(t). \tag{33}$$

23

And the corresponding initial conditions are

$$s(0) = 1 - \varepsilon, \tag{34}$$

$$i(0) = i_1(0) = \varepsilon, \tag{35}$$

and

$$i_2(0) = \ldots = i_K(0) = 0. \tag{36}$$

Essentially, we check if we are allowed to use the Theorem 5.2 stated in (Andersson and Britton, 2000) to show the convergence of truncated density process $E_K^{(n)}(t)$. First of all, we notice that there are several jumps which can affect the process. In the case of a new non-to-be-reported infection, the process changes by $(-1, 1, 0, ..., 0)$ with the corresponding jump intensity function

$$f_{(-1,1,0,\ldots,0)}(s, i_1, i_2, ..., i_K) = \beta(1 - p)s \sum_{j=1}^{K} i_j.$$

If there is a new to-be-reported infection comes to the component of size 1, then the process changes by $(-1, -1, 2, 0, ..., 0)$ with the corresponding jump intensity function

$$f_{(-1,-1,2,0,\ldots,0)}(s, i_1, i_2, i_3..., i_K) = \beta p s i_1.$$

When there is a natural recovery comes to the component of size 1 the process changes by $(0, -1, 0, ..., 0)$ with the corresponding jump intensity function

$$f_{(0,-1,0,\ldots,0)}(s, i_1, i_2, ..., i_K) = \gamma i_1,$$

whereas if the whole component of size 1 is diagnosed, the process would change by $(0, -1, 0, ..., 0)$ with the corresponding jump intensity function

$$f_{(0,-1,0,\ldots,0)}(s, i_1, i_2, ..., i_K) = \delta i_1.$$

In the case when there is a new to-be-reported infection comes to the component of size $j = 2, .., (K - 1)$, the process changes by $(-1, 0, ..., 0, -j, j + 1, 0, ..., 0)$ with the corresponding jump intensity function

$$f_{(-1,0,\ldots,0,-j,j+1,0,\ldots,0)}(s, i_1, ..., i_{j-1}, i_j, i_{j+1}, i_{j+2}, ..., i_K) = \beta p s i_j.$$

Moreover, for the component of size $K$, if there is a new to-be-reported infection occurs, then the process changes by $(-1, 0, ..., 0, -K)$ with the corresponding jump intensity function

$$f_{(-1,0,\ldots,0,-K)}(s, i_1, ..., i_{K-1}, i_K) = \beta p s i_K.$$

For $j = 2, .., K$, if there is a natural recovery comes to the component of size size $j$, then the process changes by $(0, 0, ..., 0, j-1, -j, 0, ..., 0)$ with the corresponding jump intensity function

$$f_{(0,0,...,0,j-1,-j,0,...,0)}(s, i_1, ..., i_{j-2}, i_{j-1}, i_j, i_{j+1}, ..., i_K) = \gamma i_j,$$

Further, the process changes by $(0, 0, ..., 0, 0, -j, 0, ..., 0)$ if a component of size $j$ is diagnosed, the corresponding jump intensity function is given by

$$f_{(0,0,...,0,0,-j,0,...,0)}(s, i_1, ..., i_{j-2}, i_{j-1}, i_j, i_{j+1}, ..., i_K) = \delta i_j.$$

Then we obtain the drift function $F$ defined in the Section 5.3 of (Andersson and Britton, 2000) ,which is here given by

$$F(s, i_1, i_2, ..., i_K) = \begin{pmatrix} -\beta si \\ \beta(1-p)si + \gamma i_2 - \beta psi_1 - (\gamma + \delta)i_1 \\ 2\beta psi_1 + 2\gamma i_3 - 2\beta psi_2 - 2(\gamma + \delta)i_2 \\ \vdots \\ K\beta psi_{K-1} - K\beta psi_K - K(\gamma + \delta)i_K \end{pmatrix}.$$

It can be shown that for any $x = (s, i_1, i_2, ..., i_K)$ and $y = (s', i_1', i_2', ..., i_K')$ in domain

$$C = \{x = (x_k)_k \in \mathbb{R}^{K+1} : 0 \le x_k \le 1, k = 1, .., K+1\},$$

there exists a bound $M > 0$ such that

$$|F(x) - F(y)| \le M|x - y|,$$

with the absolute norm $|\cdot|$ in $\mathbb{R}^{K+1}$. This bound $M$ can be roughly given by

$$M = \max\{2\beta + 2\beta pK^2, 2\beta + (2\beta p + 2\gamma + \delta)K^2 - \beta p\}.$$

Finally, we are now allowed to apply the Theorem 5.2 in (Andersson and Britton, 2000), which showing that the truncated "density" process

$$E_K^{(n)}(t) = (S^{(n)}(t)/n, I_1^{(n)}(t)/n, I_2^{(n)}(t)/n, ..., I_K^{(n)}(t)/n)$$

converges to the deterministic process

$$E_K^\infty(t) = (s(t), i_1(t), i_2(t), ..., i_K(t)),$$

which is defined by Equation (30)-(36).

This convergence of the truncated processes combined with the earlier sketch of why the truncated processes approximate the infinite systems well by choosing $K$ large completes the proof of Proposition 7.

$\square$

# 6 Numerical Illustrations

## 6.1 Original Model

In this section we perform simulations supporting our large population results, and also investigate the effect of the TT-strategy. We do this mainly for the following parameter values (inspired from the Covid-19 pandemic). Before the TT-strategy is applied we have the Markovian SIR epidemic model with $\beta = 0.75$ and $\gamma = 0.25$, implying an average infectious period of $1/\gamma = 4$ days and a basic reproduction number $R_0 = \beta/\gamma = 3$. When the TT-strategy is considered fix, we assume that $\delta = 0.125$ and $p = 0.5$ implying that $1/3$ of the infected individuals are tested and isolated while still infectious and that half of their contacts are reported for contact tracing ((Lucas et al., 2020) believed that the fraction $p$ of contacts that were successfully traced varies between 40% and 80%).

First we performed 10 000 simulations of the epidemic and stored the final number infected in each simulation. We did this for three different population sizes, $n = 1000$, 5000 and 10 000, each simulation starting with one initial infectious individual. We say (quite arbitrarily) that there is a minor outbreak when at most 10% get infected during the outbreak, otherwise a major outbreak occurs. We summarize the fraction of minor outbreaks and the empirical mean of total number of infected individuals among the major outbreak cases in Table 2. To these simulations we add a line for the limiting results (denoted by $n = \infty$). In this line we have derived the minor outbreak probability using Equation (9) and the mean fraction of the major outbreaks is computed numerically using Equations (30)-(36) with truncation size $K = 100$ where $r_\infty$ is approximated by $r(t)$ for $t = 100$ Days and $\varepsilon = 0.01\%$ which shows evidence that our limiting approximations work quite well already for these moderate population sizes. As shown in Fig. 2 the distribution for the major outbreaks is more peaked when the population is larger. We also note from those zoomed histograms for the major outbreaks, that they seem to follow a normal distribution with the deterministic limit as center, especially for larger $n$.

Next, we illustrate the threshold results saying that when $R_*^{(c)} \leq 1$ we expect only minor outbreaks to take place whereas when $R_*^{(c)} > 1$ also large outbreaks may occur. We first fix the parameters $(\gamma, \delta, p) = (0.25, 0.125, 0.5)$, and choose the $\beta$ to be 0.40, 0.50, 0.59 and 0.67, so that the corresponding effective component reproduction number takes values of 0.75, 1.00, 1.25 and 1.50 using Equation (3). In Fig. 3a we see that in the case of $R_*^{(c)} < 1$ there are no major outbreaks, and in Fig. 3b where $R_*^{(c)} = 1$ the distribution is still unimodal but some outbreaks are now fairly big, whereas there are
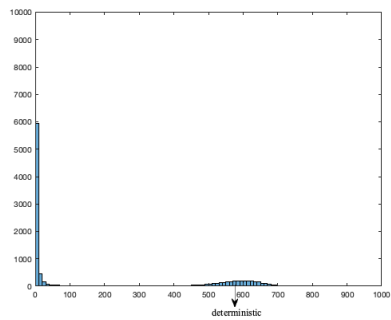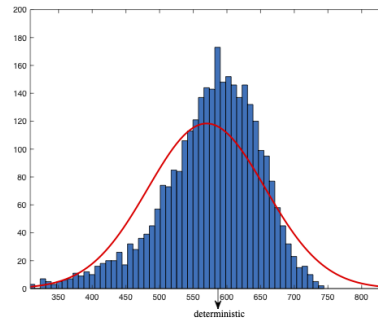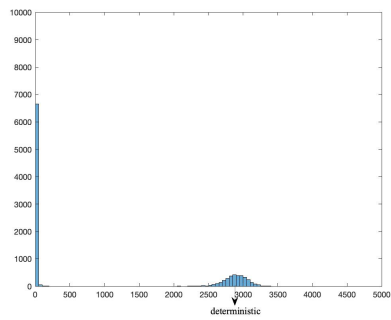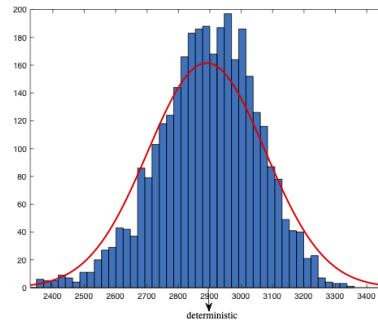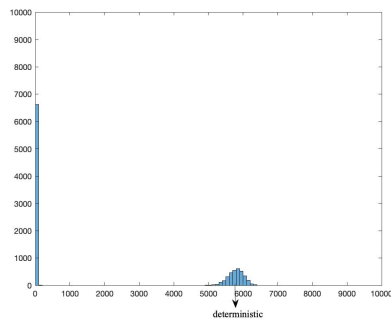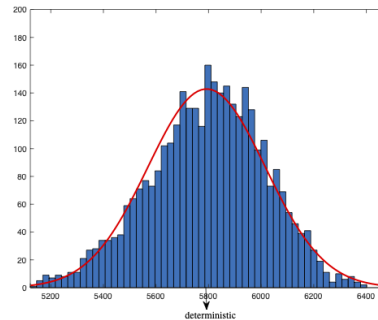
**Figure 2** Histogram of the final size in 10 000 simulations of epidemic with population size in 2a-2b $n = 1000$, in 2c-2d $n = 5000$ and in 2e-2f $n = 10\ 000$, starting with one initial infective with full histogram to the left and zoomed in on the major outbreaks to the right with normally fitted curve in red

27

**Table 2** Simulation results about final fraction infected

| size of population | mean fraction of infected among major outbreaks | standard deviation | fraction of minor outbreaks |
|---|---|---|---|
| 1000 | 0.5698 | 0.0873 | 0.6803 |
| 5000 | 0.5786 | 0.0323 | 0.6707 |
| 10000 | 0.5793 | 0.0224 | 0.6622 |
| $\infty$ | 0.5790 | 0 | 0.6667 |

major outbreaks in Fig. 3c and Fig. 3d where the distributions are clearly bimodal. As $R_*^{(c)}$ grows bigger, the size of the peak (the probability of a major outbreak) increases and the location of the peak shifts to the right (i.e. larger outbreaks).

We now study the time evolution of the epidemics showing that it becomes less random as population size $n$ increases, as stated in Proposition 7. We do this by plotting random epidemic processes $\{I_n(t)/n\}$ and comparing it with the limiting deterministic process $\{i(t)\}$. As before, we use the parameter values $(\beta, \gamma, \delta, p) = (0.75, 0.25, 0.125, 0.5)$. More specifically we plot the deterministic (in red) curves of the fraction of infectives when the population size is 1000, 5000 and 10 000 respectively. For each population size, we plot the fraction of infected for one simulation (in black), then we did 10 simulations given each size of population and plot the empirical mean of the fraction of infected (in blue). We can see from Fig. 4 that the larger the population size, the better the truncated deterministic process approximates the epidemic process. All simulations were started with 1% being infectious ($I_n(0)/n = 0.01$) and the rest susceptible. The deterministic fraction of infectives are derived by solving Equations (30)-(36) with $\varepsilon = 1\%$ and truncation size $K = 100$.

Moreover, we investigate the effect of TT strategy. We recall that $\delta$ denotes the rate of testing (either broad screening or more targeted testing) and isolate those who test positive immediately, and $p$ denotes the fraction of all contacts of infectious individuals that are successfully contact traced. In Fig. 5, we plot the effective reproduction numbers $R_*^{(c)}$ and $R_*^{(ind)}$ derived by Equations (3) and (7) respectively, as a function of the fraction of infectives being tested (before natural recovery) $\delta/(\delta + \gamma)$ in $[0, 0.5]$ and of $p$ in $[0, 1]$, keeping the other two parameters fixed at $\beta = 0.75$ and $\gamma = 0.25$. Fig. 5a shows that, surprisingly, $R_*^{(c)}$ is not monotone in $p$, whereas Fig. 5b shows that the individual reproduction number $R_*^{(ind)}$ seems to be, as expected. As seen from the contour lines where $R_*^{(ind)} = 2.5, 2, 1.5, 1$, the lines are steeper with lower $R_*^{(ind)}$. When it comes to comparing the effects of $p$ and $\delta/(\delta + \gamma)$
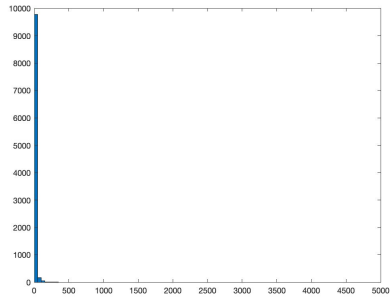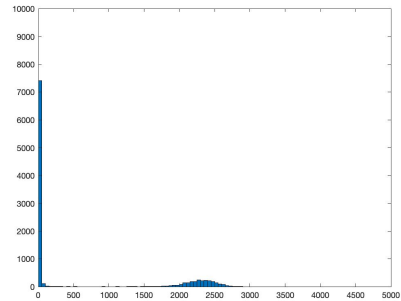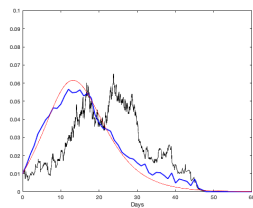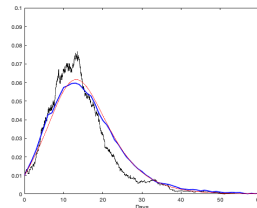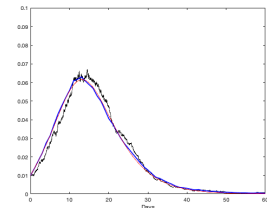
**(a)**
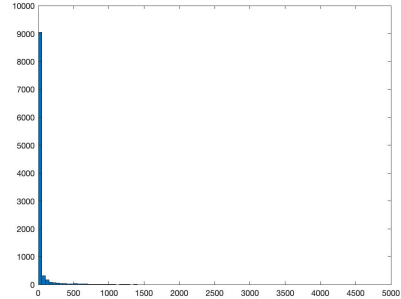
**(b)**
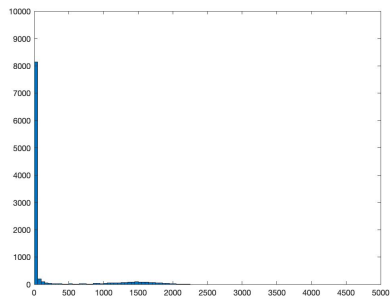
**(c)**

**(d)**

**Figure 3** Histogram of final size in 10 000 simulations of epidemic with size of population $n = 5000$ and one initial infective, where the $R_*^{(c)}$ in 3a is 0.75, in 3b is 1.00, in 3c is 1.25, and in 3d equals 1.50



**(a)**

**(b)**

**(c)**

**Figure 4** Fraction of infectives with population size of 4a 1000, 4b 5000 and 4c 10 000 with 1% initial infectives. The fraction of infectious individuals for one stochastic simulation is in black, the one for deterministic is in red, whereas the empirical mean of ten simulations is in blue

on $R_*^{(ind)}$, it seems as if $\delta/(\delta + \gamma)$ is more influential for high values (larger than 2.5 in this case) on $R_*^{(ind)}$, whereas for lower values (smaller than 2) on $R_*^{(ind)}$, tracing is more influential in preventing a major outbreak (i.e. reducing $R_*^{(ind)}$ below 1).



**Component Reproduction Number** (a)

**Individual Reproduction Number** (b)
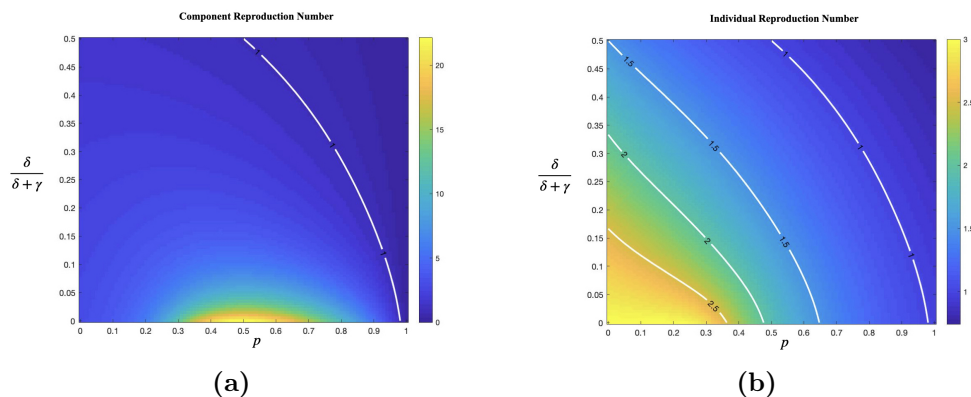
**Figure 5** Heat map of the effective reproduction number, 5a for component $R_*^{(c)}$ and 5b for individual $R_*^{(ind)}$, in a fine grid of $p = 0, 0.05, ..., 0.95, 1$ and $\delta/(\delta + \gamma) = 0, 0.005, ..., 0.495, 0.5$. with $\beta = 0.75, \gamma = 0.25$ fixed. The white lines indicate where $R_*^{(c)} = 1$ and $R_*^{(ind)} = 2.5, 2, 1.5, 1$, respectively
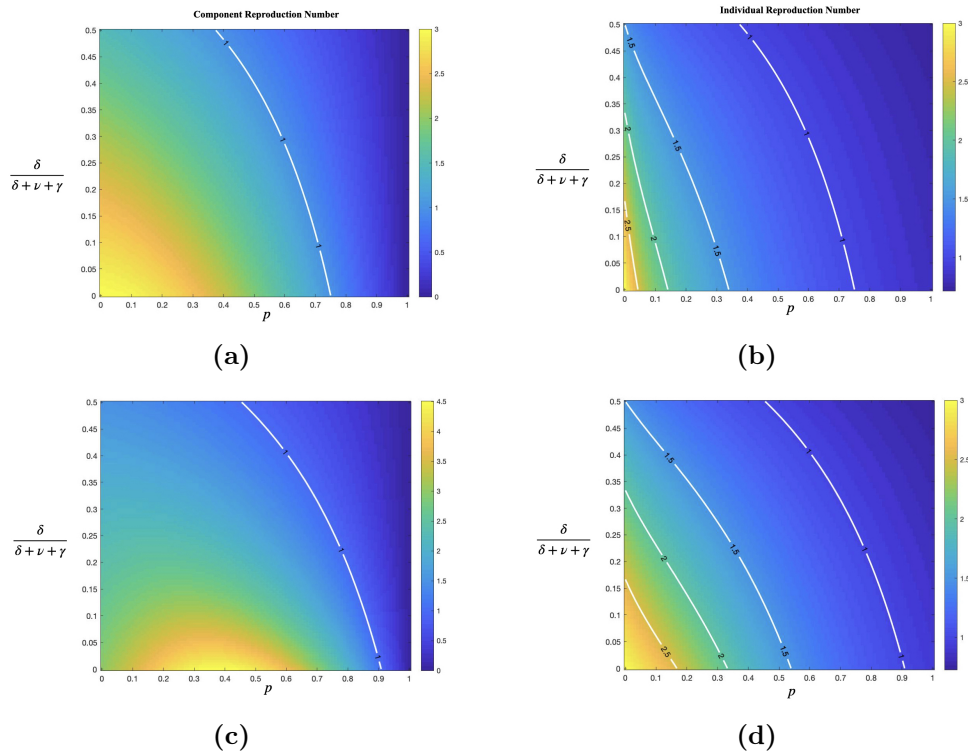
## 6.2 Alternative model interpretation

Finally, we turn to focus to the alternative model interpretation, where instead of $(\gamma, \delta)$, we have $(\gamma, \nu + \delta)$ with $\gamma$ being the rate of natural recovery, $\nu$ the rate of self-reporting and $\delta$ the rate of screening. To start we fix $\beta = 0.75$ and $\gamma = 1/12, \nu = 2\gamma = 1/6$ implying that before screening ($\delta = 0$), there would be $\nu/(\gamma + \nu) = 2/3$ of the infectious individuals get tested and self-isolated by own initiative. In Fig. 6, we plot two reproduction numbers $R_*^{(c)}$ and $R_*^{(ind)}$ as a function of $p$ in $[0, 1]$ and of the screening fraction $\delta/(\delta + \nu + \gamma)$ in $[0, 0.5]$.

In the lower panels we show the corresponding heat maps, but now for the case $\gamma = 0.2$ and $\nu = 0.05$ implying that only $1/5$ of infectives self-report, thus being closer to the original model where no individuals get tested prior to the introduction of screening. It is seen that whether the component reproduction number is monotone in $p$ or not depends on what fraction that self-report when having symptoms. Another difference as compared to the original model is that the tracing probability $p$ clearly has a bigger impact on reducing $R_*^{(ind)}$. An explanation to this would be that both self-tested individuals and those being screened will be contact traced.

Furthermore, if we assume that infectives who develop symptoms recover only due to diagnosis/self-reporting, then the fraction of asymptomatic individuals is exactly the fraction $\gamma/(\gamma + \nu)$ of infectives who do not self-report and are naturally recovered (without screening). By observing the steeper contour lines in Fig. 6b with smaller fraction of asymptomatic infectives compared with that in Fig. 6d, it implies that the tracing plays an even bigger role on reducing the individual reproduction number when there are larger fraction of individuals who are symptomatic.



(a)

(b)

(c)

(d)

**Figure 6** Heat map of the effective reproduction numbers in the alternative model interpretation. Left panels are for $R_*^{(c)}$ and right panels for $R_*^{(ind)}$, varying $p$ in $[0,1]$ and $\delta/(\delta + \nu + \gamma)$ in $[0, 0.5]$ with $\beta = 0.75$ being fixed. In 6a and 6b assuming that $\gamma = 1/12$ and $\nu = 1/6$ whereas in 6c and 6d $\gamma = 0.2$ and $\nu = 0.05$. The white lines show where $R_*^{(c)} = 1$ in 6a, 6c and $R_*^{(ind)} = 2.5, 2, 1.5, 1$ in 6b, 6d respectively

# 7    Conclusions and Discussion

In the paper we have analysed a Markovian epidemic model also incorporating the effect of testing and contact tracing (the Markovian SIR-TT-model).

By analysing the process of to-be-reported components, rather than individuals themselves, it was shown that the early stage of the epidemic could be approximated by a suitable branching process, and that if an epidemic takes off, its behaviour becomes less random as the population size $n$ increases. The reproduction numbers, both for the components as well as for the individuals, were derived. Their dependence on the amount of testing and effectiveness of contact tracing were evaluated analytically as well as numerically. It was observed that the tracing probability $p$ had a bigger impact on reducing the individual reproduction number as compared to the fraction being tested through screening, and this difference was even more pronounced in the situation when some infectives self-test also without being screened (the alternative model interpretation). Surprisingly, the reproduction number for the components was not monotonically decreasing in $p$, but the individual reproduction numbers seem to be (as expected).

There are several possible extensions to the model making it more realistic. For instance, the model assumes that there are no delays in either contact tracing or testing. The results in the present paper can hence be seen as a best possible scenario, but allowing for a delay would of course give information on how important such delays are and how much would be gained if contact tracing would be quicker. Further, the model assumes no latent period and that the infectious period follows an exponential distribution. Introducing a latent period most likely makes testing and contact tracing more effective in that individuals may get screened as well traced before even becoming infectious, but how to quantify this effect remains to be analysed. A different step towards realism would be to consider a structured community as opposed to the current assumption of a uniformly mixing community. Such structure could for example include households, spatial aspects, or some other network structures.

As the model was defined, only contacts that resulted in infection are considered for contact tracing. In reality it may of course also happen that contacts that did not result in infection are reported and traced. During the early phase of an epidemic such tracing events will rarely find new infected cases, but later in the epidemic when transmission is extensive it could (the individuals may have been infected by other individuals). To allow also for these type of contact tracing events is much harder to analyse and remains an open problem.

On the mathematical side two conjectures deserve to be proven (or disproved). The first is the statement for the final size of the epidemic in case of a major outbreak starting with one index case (see Conjecture 8). As in many similar epidemic models it seems highly plausible that this limiting final size agrees with that of the deterministic process taking $t$ to infinity

and looking at ever smaller initial starting fractions $\varepsilon$, but a proof of this is missing. Further and perhaps a lower hanging fruit, is to prove that the individual reproduction number $R_*^{(ind)}$ is monotonically decreasing both in $p$ and testing fraction $\delta/(\delta + \gamma)$.

From an applied point of view it is of course important to have parameter estimates in order to say something quantitatively useful. The model has four parameters: $(\beta, \gamma, \delta, p)$. The average infectious period $1/\gamma$ is quite often known from earlier studies, and when the basic reproduction number $R_0 = \beta/\gamma$ is known, estimates of $\beta$ would also be available. Nevertheless, the test-and-trace parameters $\delta$ and $p$ may be harder to estimate. In the case that testing comes from general broad screening it could be very well available: if for instance 1% of the community is tested each day would lead to $\delta = 0.01$ with day as time unit. If testing is targeted towards suspected cases it might be harder to know the rate $\delta$ at which infectious people are tested. Finally, estimates of the fraction $p$ of all infectious contacts that were detected by contact tracing is often hard to obtain. Perhaps a rough estimate could be obtained from studies investigating different type of contacts and how many infections they are responsible for. There are some statistical methods developed to estimate the tracing probability, e.g. a maximum-likelihood estimator in (Müller and Hösel, 2007) and an approximate Bayesian computation in (Blum and Tran, 2010). When it comes to digital contact tracing (by means of mobile tracing apps), the tracing probability $p$ is expected to be higher as compared with the traditional contact tracing (potentially due to the quicker identification and notification of infectious contacts), and would correspond to the fraction of infectious individuals that use the tracing apps.

Analyses of epidemic models incorporating various preventive measures, and statistical studies relating to them, remains a research area deserving more attention in the future.

# Acknowledgement

# Declarations

- Funding:
  See acknowledgements.

- Conflict of interest:

33

We declare that we have no conflict of interest.

- Ethics approval:
  Not applicable.

- Consent to participate:
  Not applicable.

- Consent for publication:
  Not applicable.

- Availability of data and materials:
  The paper uses no data or material.

- Code availability:
  D.Z. used Matlab for all simulations and numerical illustrations. Codes are available upon request.

- Authors' contributions:
  T.B. initiated the project, the theory was developed jointly, D.Z. performed simulations and numerical computations. D.Z. wrote most of the first draft and both revised the manuscript.

# References

H. Andersson and T. Britton. *Stochastic Epidemic Models and Their Statistical Analysis*, volume 151. 01 2000. doi: 10.1007/978-1-4612-1158-7.

F. G. Ball, E. S. Knock, and P. D. O'Neill. Threshold behaviour of emerging epidemics featuring contact tracing. *Advances in Applied Probability*, 43 (4):1048–1065, 2011. doi: 10.1239/aap/1324045698.

F. G. Ball, E. S. Knock, and P. D. O'Neill. Stochastic epidemic models featuring contact tracing with delays. *Mathematical Biosciences*, 266:23–35, Aug 2015. ISSN 0025-5564. doi: 10.1016/j.mbs.2015.05.007.

M. G. B. Blum and V. C. Tran. Hiv with contact tracing: a case study in approximate bayesian computation. *Biostatistics*, 11(4):644–660, 2010. doi: 10.1093/biostatistics/kxq022.

W. Bradshaw, E. Alley, J. Huggins, A. Lloyd, and K. Esvelt. Bidirectional contact tracing could dramatically improve covid-19 control. *Nature Communications*, 12:232, 01 2021. doi: 10.1038/s41467-020-20325-7.

R. Hofstad and M. Keane. An elementary proof of the hitting time theorem. *American Mathematical Monthly - AMER MATH MON*, 115, 10 2008. doi: 10.1080/00029890.2008.11920588.

M. Kendall, L. Milsom, L. Abeler-Dörner, C. Wymant, L. Ferretti, M. Briers, C. Holmes, D. Bonsall, J. Abeler, and C. Fraser. Epidemiological changes on the isle of wight after the launch of the nhs test and trace programme: a preliminary analysis. *The Lancet Digital Health*, 2, 10 2020. doi: 10.1016/S2589-7500(20)30241-7.

T. C. D. Lucas, E. L. Davis, D. Ayabina, A. Borlase, T. Crellen, L. Pi, G. F. Medley, L. Yardley, P. Klepac, J. Gog, and T. D. Hollingsworth. Engagement and adherence trade-offs for sars-cov-2 contact tracing. *medRxiv*, 2020. doi: 10.1101/2020.08.20.20178558.

J. Müller and V. Hösel. Estimating the tracing probability from contact history at the onset of an epidemic. *Mathematical Population Studies*, 14 (4):211–236, 2007. doi: 10.1080/08898480701612857.

J. Müller and V. Hösel. Contact tracing & super-spreaders in the branching-process model, 2021.

J. Müller, M. Kretzschmar, and K. Dietz. Contact tracing in stochastic and deterministic epidemic models. *Mathematical biosciences*, 164:39–64, 04 2000. doi: 10.1016/S0025-5564(99)00061-9.