



**HAL**  
open science

# A mathematical assessment of the efficiency of quarantining and contact tracing in curbing the COVID-19 epidemic

Amaury Lambert

► **To cite this version:**

Amaury Lambert. A mathematical assessment of the efficiency of quarantining and contact tracing in curbing the COVID-19 epidemic. *Mathematical Modelling of Natural Phenomena*, 2021, 16, pp.53. 10.1051/mmnp/2021042 . hal-03366943

**HAL Id: hal-03366943**

**<https://hal.science/hal-03366943v1>**

Submitted on 5 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## A MATHEMATICAL ASSESSMENT OF THE EFFICIENCY OF QUARANTINING AND CONTACT TRACING IN CURBING THE COVID-19 EPIDEMIC

AMAURY LAMBERT<sup>1,2,\*</sup> 

**Abstract.** In our model of the COVID-19 epidemic, infected individuals can be of four types, according whether they are asymptomatic ( $A$ ) or symptomatic ( $I$ ), and use a contact tracing mobile phone application ( $Y$ ) or not ( $N$ ). We denote by  $R_0$  the average number of secondary infections from a random infected individual. We investigate the effect of non-digital interventions (voluntary isolation upon symptom onset, quarantining private contacts) and of digital interventions (contact tracing thanks to the app), depending on the willingness to quarantine, parameterized by four cooperating probabilities. For a given ‘effective’  $R_0$  obtained with non-digital interventions, we use non-negative matrix theory and stopping line techniques to characterize mathematically the minimal fraction  $y_0$  of app users needed to curb the epidemic, *i.e.*, for the epidemic to die out with probability 1. We show that under a wide range of scenarios, the threshold  $y_0$  as a function of  $R_0$  rises steeply from 0 at  $R_0 = 1$  to prohibitively large values (of the order of 60–70% up) whenever  $R_0$  is above 1.3. Our results show that moderate rates of adoption of a contact tracing app can reduce  $R_0$  but are by no means sufficient to reduce it below 1 unless it is already very close to 1 thanks to non-digital interventions.

**Mathematics Subject Classification.** 92D30, 60J80, 60J85, 91A12, 91D30.

Received September 7, 2020. Accepted July 24, 2021.

### 1. INTRODUCTION

In this paper, we model the SARS-Cov-2 epidemic by a multitype branching process where infected individuals can be asymptomatic or symptomatic, use or not a contact tracing mobile phone app, be cooperators or defectors.

We let  $f$  denote the natural fraction of asymptomatics in the population,  $R_a$  denote the mean number of secondary infections from an asymptomatic individual and  $R_i$  the mean number of secondary infections from a symptomatic individual, so that the mean number of secondary infections in the population is

$$R_0 = fR_a + (1 - f)R_i. \quad (1.1)$$

---

*Keywords and phrases:* Multitype branching process - leading eigenvalue - stopping line - epidemiology - personal data.

<sup>1</sup> Laboratoire de Probabilités, Statistique & Modélisation (LPSM), Sorbonne Université, CNRS UMR8001, Université de Paris, Paris, France.

<sup>2</sup> Center for Interdisciplinary Research in Biology (CIRB), Collège de France, CNRS UMR7241, INSERM U1050, PSL Research University, Paris, France.

\* Corresponding author: [amaury.lambert@upmc.fr](mailto:amaury.lambert@upmc.fr)

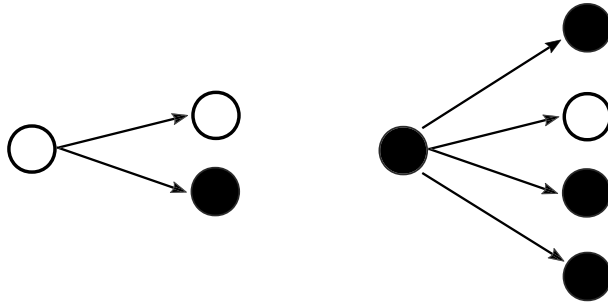


FIGURE 1. Asymptomatics ( $A$ ) infect on average  $R_a = 2$  susceptibles and symptomatics ( $I$ ) infect on average  $R_i = 4$  susceptibles. The overall fraction of  $A$ 's is  $f = 1/3$ . Legend: white =  $\mathbf{A}$  (asymptomatic), black =  $\mathbf{I}$  (symptomatic).

To fix ideas, we will exemplify our results by giving typical numerical values to these and other parameters. Let us specify that these numerical values are only here as means of illustration. Our results are general and depend symbolically on all the parameters, so that the reader can freely tune all parameter values to apply our results to any specific situation.

We will take as default value  $f = 1/3$ . Estimates of  $f$  range between 20% (data from the Guangdong province [1] and the Diamond Princess cruise ship [10]) and 40% (data from Japanese repatriation flights [12] and from the municipality of Vo in Italy [7]).

In the absence of mitigation measure, we take as default values  $R_a = 2$  and  $R_i = 4$ , resulting in a global ‘natural’  $R_0 = 3.33$ , which is the geometric growth rate of the infected population. This figure is slightly above the point estimate in France before lockdown (2.90 in [16]) and is in line with estimates from most countries, which range between 2.2 and 3.9 in the absence of mitigation measure [8]. See Figure 1 for a cartoon representing transmissions from  $A$ 's (asymptomatics) and from  $I$ 's (symptomatics).

We explore a range of scenarios susceptible to curb the epidemic, that is, reduce  $R_0$  below 1. These scenarios are:

- Case isolation upon the appearance of symptoms;
- Additionally quarantining private contacts of symptomatics;
- Additionally quarantining physical contacts of symptomatic individuals, by means of a contact tracing app (forward tracing);
- Additionally quarantining physical contacts of physical contacts of symptomatic individuals (recursive tracing).

Let us make some preliminary observations.

First note that, except maybe in the occurrence of testings sufficiently massive to reach asymptomatics, quarantines and alerts only concern symptomatic individuals. Therefore, in the absence of mass testing, a crucial quantity is  $fR_a$ , which is the growth rate of the infected population restricted to  $A$ -to- $A$  transmissions ( $A$  denoting asymptomatics). A necessary condition for curbing the epidemic is then

$$fR_a < 1, \tag{1.2}$$

which can occur either naturally (as under our assumed default values) or by the effect of social distancing – see Section 3.

In the same vein, once  $R_a$  and  $R_i$  have been optimally reduced by non-digital interventions (social distancing, case isolation), a crucial quantity is  $(1 - y)R_0$ , where  $y$  is the fraction of the population using the contact tracing app, because  $(1 - y)R_0$  is the growth rate of the infected population restricted to  $N$ -to- $N$  transmission ( $N$  denoting individuals *not* using the app). A necessary (but certainly not sufficient, as we will see) condition for

curbing the epidemic is then

$$(1 - y)R_0 < 1, \quad (1.3)$$

or equivalently

$$y > 1 - \frac{1}{R_0},$$

identical to the condition for vaccination to curb the epidemic, where  $y$  is the fraction of vaccinated people. This means for example that if non-digital interventions have lowered  $R_0$  from its natural value to an effective value of 1.5, say, then the fraction of app users must be at least  $1/3$  for contact tracing to have a chance to work. The purpose of this work (see Sects. 4 and 5) is to precisely determine the minimal value  $y_0$  of  $y$  allowing contact tracing to get  $R_0$  below 1, *i.e.* to see the epidemic die out with probability 1, a condition we will consistently denote by  $(\star)$ .

We will see that condition (1.3), which is necessary and sufficient in the case where  $y$  is the fraction of vaccinated people, is far from being sufficient in the case where  $y$  is the fraction of app users.

In the following statement and in the caption of Figure 2,  $q_0$  and  $q_1$  are probabilities of cooperation of app users:  $q_0$  is the probability of informing the app upon first symptoms and  $q_1$  is the probability of effectively isolating if alerted by the app (see Section 2).

**Theorem 1.1.** *Fix the effective average numbers of secondary infections ( $R_a$ ,  $R_i$ ,  $R_0$ ) obtained in the presence of non-digital interventions (social distancing, isolation upon symptom onset, quarantining private contacts of symptomatics).*

*Recall that  $(\star)$  denotes the condition that the epidemic dies out with probability 1, here thanks to the marginal effect of contact tracing (*i.e.*, the multitype branching process modeling the outbreak in the presence of contact tracing is subcritical).*

*In the case of forward tracing,*

$$(\star) \iff (1 - R_0(1 - y))yq_0q_1(1 - f)R_i - R_0 + 1 \geq 0, \quad (1.4)$$

*or equivalently*

$$(\star) \iff y \geq y_0 = \frac{R_0 - 1 + \sqrt{(R_0 - 1)(R_0(1 + 4/(q_0q_1(1 - f)R_i)) - 1)}}{2R_0}, \quad (1.5)$$

*where  $y_0$  is the minimal fraction of app users to curb the epidemic.*

Figure 2 shows how the threshold  $y_0$  increases as a function of  $R_0$ . The previous statement is proved in Section 4. In Section 5, we address the case of recursive contact tracing (see Thm. 5.1).

**Remark 1.2.** We recover from equation (1.4) the fact that  $y$  must be at least  $1 - 1/R_0$  in order to control the spread. The actual threshold  $y_0$  given in equation (1.5) is actually much higher.

**Remark 1.3.** In both the forward contact tracing and the recursive contact tracing cases, the outbreak is modeled by an irreducible multitype branching process, where the number of types is 4, respectively 6. Such processes are subcritical if and only if the leading eigenvalue of their mean offspring matrix is smaller than 1 [11]. In order to get the most explicit criterion on the model parameters, and notably on  $y$ , for this to hold, we reduce the dimension of the problem to 2, by introducing two new states, either by lumping existing states (forward tracing case) or by defining them through genealogical properties (recursive tracing case). For the dimension reduction to work, we need the states to be regenerative (*i.e.*, the descendance of a vertex in the transmission

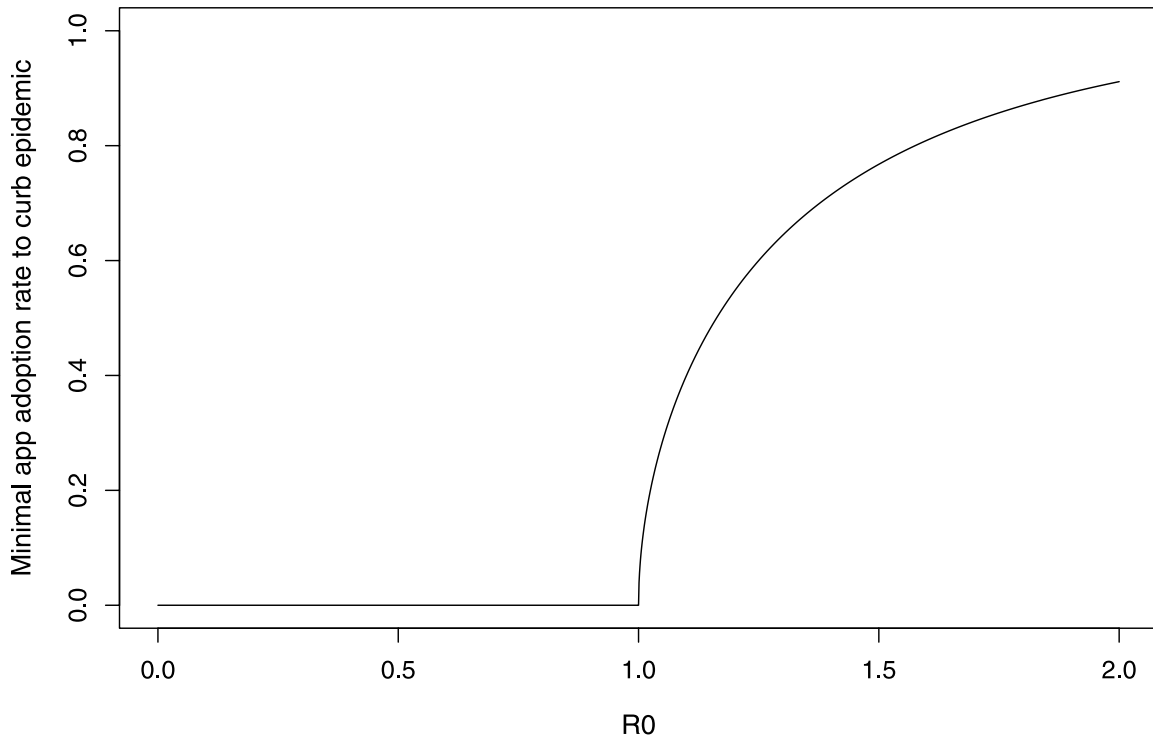


FIGURE 2. Minimal app adoption rate  $y_0$  to curb epidemic as a function of effective  $R_0$  (*i.e.*, obtained as a result of social distancing, case isolation and quarantining of private contacts, regardless of contact tracing), under forward contact tracing (as given by Eq. (1.5)). Cooperating probabilities are set equal to  $q_0 = q_1 = 1$  (best case scenario). Other parameters are set to their default values:  $f = 1/3$ ,  $b = 3/5$ ,  $kr = 1/6$ ,  $R_a = 2$ ,  $R_i = 4$ . Note that for these parameter values,  $R_0 = 2c$ , and that  $c$  only is tuned to let  $R_0$  vary. This graph is the same as the bottom curve in Figure 5.

tree is independent of the rest of the tree conditional on its state) and form stopping lines (a stopping line of a tree  $T$  is a set of vertices of  $T$  whose intersection with each path from the root to a leaf of  $T$  is a singleton).

## 2. BASIC MODELING ASSUMPTIONS

We call a given infected individual a case or index case, and secondary infections the individuals she infects. We will also speak of mother/daughter (implicitly: in the transmission tree).

Now a typical index case can, independently:

- be Asymptomatic (**A**) with probability  $f$ , or Symptomatic (**I**) with probability  $1 - f$ . We will take  $f = 1/3$  as the default value (see Introduction). We denote the mean number of secondary infections from an  $A$ -individual by  $R_a$  and the mean number of secondary infections from an  $I$ -individual by  $R_i$ . The mean number  $R_0$  of secondary infections in the population is then given by equation (1.1).
- use a contact tracing app (**Y** for yes) with probability  $y$ , or not (**N**) with probability  $1 - y$ .
- cooperate or defect. Here, cooperating/defecting can mean different things depending on the state of the index case:
  - A symptomatic individual can cooperate by self-isolating upon symptom onset, and thus ceasing to infect other people after that time (probability of cooperating  $p$ ).

- A private contact (work, family...) of a symptomatic individual can also cooperate by self-quarantining after being alerted by plain talk/phone/email, and thus ceasing to infect other people after that time (probability of cooperating  $r$ ).
- A symptomatic individual who uses a contact tracing mobile phone app can cooperate by entering the information in her app after she feels the first symptoms (probability of cooperating  $q_0$ ).
- An individual who uses the app and is alerted by her app can cooperate by self-quarantining upon receiving the digital alert. We will distinguish whether the receiver of the alert has degree 1 or 2 with the original index case in the contact network (probabilities of cooperating  $q_1$  and  $q_2$ , respectively) – see Sections 4 and 5 for details.

**Important assumptions.** We will make the following two assumptions.

- Branching assumption. We assume independence of infection events (branching assumption). This means in particular that 1) susceptibles are always in excess and that 2) the contact network is tree-like, neglecting the existence of shared contacts. This biases our predictions in two ways, because 1) we neglect the possible reduction of the effective  $R_0$  thanks to the accumulation of recovered, immune individuals; 2) we underrate the efficiency of alerting in case when transmission has occurred in clusters, but we also underrate the speed of propagation by ignoring these clusters.
- Multiple alerts. Note that an individual can be confronted to the decision of cooperating or defecting several times and in particular be alerted by several different sources (private vs public, sources of degree 1 vs degree 2). In contrast to standard models where each individual sticks to a single strategy of always cooperating or always defecting (and thus can be labelled cooperator/defector), we assume here that each time an individual can cooperate or defect, she decides to do so with the aforementioned probabilities, independently and independently of her past decisions to cooperate or defect. In addition, for defection to actually occur, an individual receiving several alerts must defect independently to each of these alerts, modeling the multiplicative effect of multiple alerts.

**Non-digital mitigation measures.** The values of  $R_a$ ,  $R_i$  and  $R_0$  may vary (but Eq. (1.1) always holds) according to four scenarios of non-digital interventions (*i.e.*, independent of contact tracing):

- No intervention. In this case, we use the notation  $R_a^\emptyset$ ,  $R_i^\emptyset$  and  $R_0^\emptyset$ . We will take as default values  $R_a^\emptyset = 2$  and  $R_i^\emptyset = 4$ , so that  $R_0^\emptyset = 3.33$ .
- Social distancing. When social distancing is in force, we use the notation  $R_a^c$ ,  $R_i^c$ ,  $R_0^c$ .
- Additionally self-isolating upon the appearance of symptoms. When these measures are in force in addition to social distancing, we use the notation  $R_a^{cb}$ ,  $R_i^{cb}$  and  $R_0^{cb}$ .
- Additionally quarantining private contacts of symptomatics. When these measures are in force in addition to the previous ones, we use the notation  $R_a^{cbk}$ ,  $R_i^{cbk}$  and  $R_0^{cbk}$ .

### 3. NON-DIGITAL INTERVENTIONS

#### 3.1. Social distancing

Social distancing scales indistinctively  $R_a^\emptyset$  and  $R_i^\emptyset$  by a factor  $c$ , so that  $R_a^c = cR_a^\emptyset$ ,  $R_i^c = cR_i^\emptyset$  and  $R_0^c = cR_0^\emptyset$ .

#### 3.2. Case isolation

Assume a certain fraction  $\mathbf{p}$  of symptomatic individuals cooperate by self-isolating and let  $m$  be the average fraction of the total number of *potential* secondary infections made before her isolation from a symptomatic, cooperating individual. In other words,  $1 - m$  is the fraction of secondary infections avoided thanks to case isolation. More specifically, a symptomatic, cooperating individual self-isolates upon day  $D + T$ , where  $D$  is the day of onset of symptoms and  $T$  is the waiting time before taking action (self-isolation) in cooperating

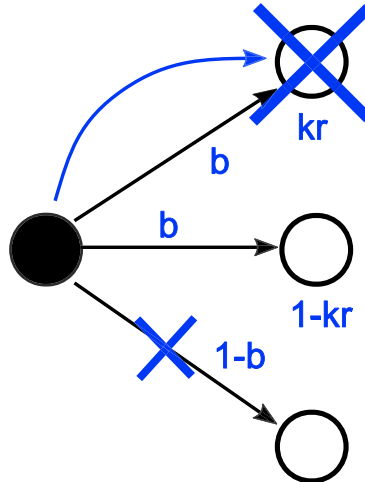


FIGURE 3. Upon symptom onset, an  $I$ -individual decides with probability  $p$  to self-isolate, resulting in removing an average fraction  $1 - b = p(1 - m)$  of her daughters from the epidemic (small, blue cross). A fraction  $k$  of the ‘surviving’ daughters is assumed to be private contacts, who are alerted (blue arrow) and then do self-quarantine with probability  $r$  (large, blue cross). Legend: white =  $\mathbf{A}$  (asymptomatic), black =  $\mathbf{I}$  (symptomatic).

individuals, *e.g.*,  $T$  is distributed between 0 and 2 days. Then  $m$  is the average fraction of the total number of secondary infections, that are already made by day  $D + T$  from a typical symptomatic individual.

To select a default value for  $m$ , we can rely on empirical estimates of the average fraction  $m_0$  of secondary infections made before symptoms ( $m_0 \leq m$ ). In [9], the authors find  $m_0 \approx 0.4$ , but since case isolation must have somewhat been in force in this study, this estimate actually is an upper bound for  $m_0$ . We will take  $m = 0.5$  as default value.

Now set  $b$  the average fraction of secondary infections made before isolation ( $1 - b$  is the fraction of secondary infections avoided in a regime of partial case isolation)

$$b := 1 - p + pm,$$

so that  $R_i^{cb} = bR_i^c$  is the average number of secondary infections from an  $I$ -individual in this partial case isolation regime, whereas  $R_a^{cb} = R_a^c$  remains unchanged. Note that  $m_0 \leq m \leq b \leq 1$ .

In this scenario, the condition  $(\star)$  that the epidemic dies out with probability 1 holds iff

$$R_0^{cb} < 1,$$

where  $R_0^{cb} = cfR_a^\varnothing + c(1 - f)bR_i^\varnothing$ .

**Application.** Here we use the default values  $f = 1/3$ ,  $R_a^\varnothing = 2$  and  $R_i^\varnothing = 4$ .

If we assume that 100% of symptomatic individuals self-isolate upon symptoms ( $p = 1$ ), then  $b = 1/2$ , and we need  $c < 1/2$  to control the epidemic, which means that social distancing would have to cut down transmissions by at least 50%.

If we assume that only 50% of symptomatic individuals self-isolate upon symptoms ( $p = 1/2$ ), then  $b = 3/4$ , and we need  $c < 3/8$ , which means that social distancing would have to cut down transmissions by at least 62.5%.

The possibility that mere social distancing does not come anywhere near these figures cannot at all be discarded. In such a situation, case isolation would not be sufficient in itself to curb the epidemic. We will take

as default value  $b = 3/5$ , which corresponds to the optimistic estimate of a fraction  $p = 4/5$  of cooperators and yields  $R_0^b = fR_a^\varnothing + (1-f)bR_i^\varnothing = 2.27$ . Then the average number of secondary infections in the presence of both social distancing and case isolation is  $R_0^{cb} = cR_0^b$ , which remains larger than 1 whenever  $c > 44\%$ .

For example, with  $c = 3/4$ ,  $R_0^{cb} = 1.7$ , with  $c = 2/3$ ,  $R_0^{cb} = 1.51$  and with  $c = 1/2$ ,  $R_0^{cb} = 1.13$ .

We will now investigate (assuming  $c > 0.44$ ) the effect of quarantining private contacts (which does not require a contact tracing app).

### 3.3. Case isolation and quarantining private contacts

Here we consider the possibility that symptomatic individuals alert, by plain talk/phone/email their private contacts (work, family). We denote by  $k$  (for ‘known’) the fraction of secondary infections that are private contacts and by  $r$  the fraction of private contacts who are alerted and do quarantine. We assume that all self-quarantining daughters are removed from the epidemic because they self-quarantine before being infectious.

In this scenario, the condition ( $\star$ ) that the epidemic dies out with probability 1 holds iff

$$R_0^{cbk} < 1,$$

where  $R_0^{cbk} = cR_0^{bk}$  and  $R_0^{bk} = fR_a^\varnothing + (1-f)b(1-kr)R_i^\varnothing$ .

**Application.** In [14], out of a total of 7324 well documented cases in 120 Chinese towns in January–February 2020, only 1,245 could be clustered into mini outbreaks involving 3 or more people in the same household, transport, restaurant, mall... This implies that all other cases (83% of the data) had untraceable secondary infections (or were clusters of size 2), so that a large part of secondary infections were due to socially non-connected physical contacts.

Additionally, in the national survey ALCoV2 conducted among the French population in May 2020, we (the author and his research group) have collected data on 6000 households with at least one symptomatic member, representing approximately 20,000 people. In these data, households of 2 or 3 people had on average 60% symptomatic members, and larger households had less than 50% symptomatic members (pers. comm.). This confirms that a large proportion of transmissions contributing to the value of  $R_i$  are due to transmissions outside home.

Altogether, these figures suggest that private contacts play a minor role in the epidemic so that in reality  $k$  is quite low. If we take  $kr = 1/6$  and stick to the default values given earlier of  $f = 1/3$ ,  $R_a^\varnothing = 2$ ,  $R_i^\varnothing = 4$  and  $b = 3/5$ , we get  $R_i^{bk} := b(1-kr)R_i^\varnothing = 2$ , so that  $R_0^{bk} = 2$ , and we need  $c < 1/2$  to curb the epidemic.

We will call  $R_0^{cbk}$  the effective  $R_0$ , compared to the natural  $R_0^\varnothing$ . We will now assume that  $c > 1/2$  and investigate for a given effective  $R_0$ , whether a contact tracing app can manage to control the spread. Note that with the default parameter values,  $R_a^{cbk} = R_i^{cbk} = R_0^{cbk} = 2c$ , which can be tuned by merely varying  $c$ .

## 4. FORWARD CONTACT TRACING

Now we assume that a proportion  $y$  of the population uses a contact tracing mobile phone app. Such individuals are denoted  $\mathbf{Y}$  (‘yes to the app’), and the others  $\mathbf{N}$  (‘no to the app’).

In this section we consider that an alert is always of degree 1, that is, originates from an individual of type  $YI$  (using the app, symptomatic) and is only transmitted to her close physical contacts. We reserve for the next section the case of alerts of degree 2, that is, which originate from the index case  $YI$  but are conveyed through an intermediate physical contact of the case, to a contact of this contact.

Recall the probabilities of cooperation  $q_0$  and  $q_1$  defined as follows. A  $YI$ -individual informs the app of her symptoms with probability  $q_0$  and a  $Y$ -individual alerted by a cooperating  $YI$  self-quarantines with probability  $q_1$ . See Figure 4 for a cartoon depicting the possible effects of alerts of degree 1.

We assume that:



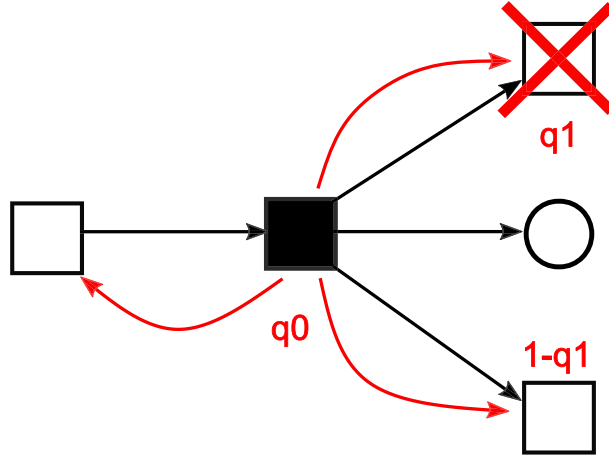


FIGURE 4. A  $YI$ -individual (symptomatic, using the app) decides with probability  $q_0$  to inform the app of her symptoms, resulting in alerting (red arrows) all her physical contacts of type  $Y$  (among which her mother and daughters in the epidemic). Each alerted daughter does self-quarantine with probability  $q_1$  and is then removed from the epidemic (red cross). Legend: white =  $\mathbf{A}$ , black =  $\mathbf{I}$ , square =  $\mathbf{Y}$ , circle =  $\mathbf{N}$ .

- decisions to cooperate or defect that do not require the app (isolating and alerting private contacts) are independent of using the app or not.
- decisions of the same individual to cooperate or defect in different situations are independent.

Let us compute the average number of secondary infections in each class,  $N$  or  $Y$ , depending on the class of the index case,  $YI$ ,  $YA$ ,  $NI$ ,  $NA$ .

Let us start with an index case  $YI$  (using the app, symptomatic) who feels her first symptoms. An individual  $Y$  infected by this index case:

- receives a private injunction to quarantine with probability  $k$  and if this is the case, cooperates with probability  $r$ ;
- receives independently a digital injunction with probability  $q_0$  and if this the case, cooperates with probability  $q_1$ .

Then the probability for a  $Y$ -individual of not being removed when infected from a  $YI$ -individual is

$$(1 - q_0)(1 - kr) + q_0(1 - kr)(1 - q_1) = (1 - kr)(1 - q_0q_1).$$

In cases when the mother is symptomatic ( $I$ ), but either she ( $NI$ ) or her daughter ( $N$ ) is of type  $N$ , this probability is always  $1 - kr$ . In cases when the mother is asymptomatic ( $YA$  or  $NA$ ), this probability is always 1, regardless of the daughter's type ( $Y$  or  $N$ ).

Now let  $M$  be the matrix with entries  $m_{y,y}$ ,  $m_{y,n}$ ,  $m_{n,y}$ ,  $m_{n,n}$ , where  $m_{y,y}$  denotes the average number of secondary infections of type  $Y$  made by a random  $Y$ -individual and so on and so forth. We easily get

$$m_{y,y} = fcR_a^\emptyset y + (1 - f)bc(1 - kr)(1 - q_0q_1)R_i^\emptyset y = fR_a^{cbk} y + (1 - f)R_i^{cbk} y(1 - q_0q_1)$$

$$m_{y,n} = m_{n,n} = fcR_a^\emptyset (1 - y) + (1 - f)bc(1 - kr)R_i^\emptyset (1 - y) = fR_a^{cbk} (1 - y) + (1 - f)R_i^{cbk} (1 - y)$$

and finally

$$m_{n,y} = fcR_a^\varnothing y + (1-f)bc(1-kr)R_i^\varnothing y = fR_a^{cbk}y + (1-f)R_i^{cbk}y.$$

**Remark 4.1.** We see from the previous equations that whenever the nature and number of non-digital interventions enforced is fixed and known, the mean numbers of secondary infections from  $Y/N$  to  $Y/N$  only depend on these interventions through  $R_a^{cbk}$  and  $R_i^{cbk}$ . Then from now on, we will drop superscripts and use the generic notation  $R_a$  and  $R_i$ , defined as effective average numbers of secondary infections. The values of these two parameters will depend upon the nature and number of non-digital interventions enforced, or more precisely on the values of  $c$ ,  $b$ ,  $k$  and  $r$ . Using the same default values as in the previous section yields  $R_a = R_i = 2c$ .

We now prove Theorem 1.1, applying the heuristic introduced in Remark 1.3. Here,  $Y$  and  $N$  are two regenerative states, because the probability of being  $A/I$  conditional on being  $Y/N$  is independent of past transmissions. Therefore, the epidemic dies out with probability 1 iff the leading eigenvalue of  $M$  is smaller than 1.

Using the notation  $R_0 = R_0^{cbk}$ ,  $R_i = R_i^{cbk}$  (see previous Remark) and  $s_i = q_0q_1(1-f)R_i$ , we get

$$M := \begin{pmatrix} m_{y,y} & m_{y,n} \\ m_{n,y} & m_{n,n} \end{pmatrix} = \begin{pmatrix} yR_0 - ys_i & (1-y)R_0 \\ yR_0 & (1-y)R_0 \end{pmatrix}$$

Now let  $Q$  be the characteristic polynomial of  $M$ :

$$Q(X) = X^2 - (R_0 - ys_i)X - y(1-y)s_iR_0.$$

Since  $Q(0) \leq 0$ , the leading eigenvalue of  $M$  is the unique positive root of  $Q$ . Also, this root is smaller than 1 if and only if  $Q(1) \geq 0$ , which yields

$$(\star) \iff y(1 - R_0(1 - y)) - h(R_0 - 1) \geq 0,$$

where

$$h = \frac{1}{q_0q_1(1-f)R_i},$$

which is equation (1.4) of Theorem 1.1.

We define  $y_0$  the minimal fraction of users of the app necessary to control the spread, or minimal app adoption rate, by

$$(\star) \iff y \geq y_0.$$

Elementary calculus yields

$$y_0 := \frac{R_0 - 1 + \sqrt{(R_0 - 1)(R_0(1 + 4h) - 1)}}{2R_0},$$

as in equation (1.5) of Theorem 1.1.

**Remark 4.2.** Note that for digital interventions to be able to curb the epidemic, we need  $y_0$  to be actually smaller than 1, that is,

$$q_0q_1 \geq \frac{R_0 - 1}{(1-f)R_i},$$

which can from the start only hold if  $(R_0 - 1)/(1 - f)R_i < 1$ . This inequality is equivalent to  $fR_a < 1$ , which we have already seen as equation (1.2). Indeed, if  $fR_a > 1$ , the epidemic restricted to  $A$ -to- $A$  transmissions would be growing exponentially, with no control possible by neither digital nor non-digital interventions (in the absence of mass testing).

**Application.** The actual values of  $R_0$  and  $R_i$  depend on the nature and strength of non-digital interventions. Let us assume that social distancing is in force, parameterized by an unknown scaling factor  $c$ , in addition to case isolation and quarantining of private contacts (respectively parameterized by  $b$  and  $kr$ ). If we stick to the default values given earlier ( $f = 1/3$ ,  $R_a^\varnothing = 2$ ,  $R_i^\varnothing = 4$ ,  $b = 3/5$ ,  $kr = 1/6$ ), which yield  $R_0 = R_i = 2c$ , we can study how  $y_0$  varies as a function of the effective  $R_0$ .

If social distancing cuts down infections by  $1/4$ , *i.e.*, if  $c = 3/4$ , then the effective  $R_0 = 3/2$  and

$$y_0 = \frac{1}{6} \left( 1 + \sqrt{1 + \frac{12}{q_0 q_1}} \right).$$

Note that  $q_0 q_1$  is the probability that the index case using the app does enter the information about her symptoms into the app and that the physical contact receiving the alert does self-quarantine. For the threshold  $y_0$  to be smaller than 1, we need that  $q_0 q_1 \geq 1/2$ . If app users are 100% reliable, that is  $q_0 q_1 = 1$ , we get that the minimal adoption rate of the app is

$$y_0 = \frac{1 + \sqrt{13}}{6} \approx 0.77.$$

If social distancing cuts down infections by  $1/3$ , *i.e.*, if  $c = 2/3$ , then the effective  $R_0 = 4/3$  and

$$y_0 = \frac{1}{8} \left( 1 + \sqrt{1 + \frac{18}{q_0 q_1}} \right).$$

For the threshold  $y_0$  to be smaller than 1, we need that  $q_0 q_1 \geq 3/8$ . If app users are 100% reliable, that is  $q_0 q_1 = 1$ , we get that the minimal adoption rate of the app is

$$y_0 = \frac{1 + \sqrt{19}}{8} \approx 0.67.$$

Let us summarize this part:

- to get from effective  $R_0 = 3/2$  to  $R_0 = 1$ , the minimal adoption rate of the app assuming perfect cooperative behavior is  $y_0 = 77\%$ .
- to get from effective  $R_0 = 4/3$  to  $R_0 = 1$ , the minimal adoption rate of the app assuming perfect cooperative behavior is  $y_0 = 67\%$ .

Figure 5 shows more generally how  $y_0$  varies as a function of  $R_0$  for 4 different values of  $q_1$ .

It can be seen in Figure 5 that the minimal adoption rate of the app to reduce  $R_0$  below 1 increases very steeply from  $R_0 = 1$ , requiring the fraction  $y$  of app users to be very large, even in the best scenarios when the fraction  $f$  of asymptomatics is small (not shown here, see Figure B.1 in the appendix) and the probability  $q_1$  of cooperation is large (bottom curve). Current rates of adoption of the kind of mobile phone app considered in this paper are in most countries much lower than  $y_0$ , unless non-digital measures (here, reduced to the effect of social distancing via parameter  $c$ ) already curb  $R_0$  to values very close to 1.

We will now study the case of recursive contact tracing.

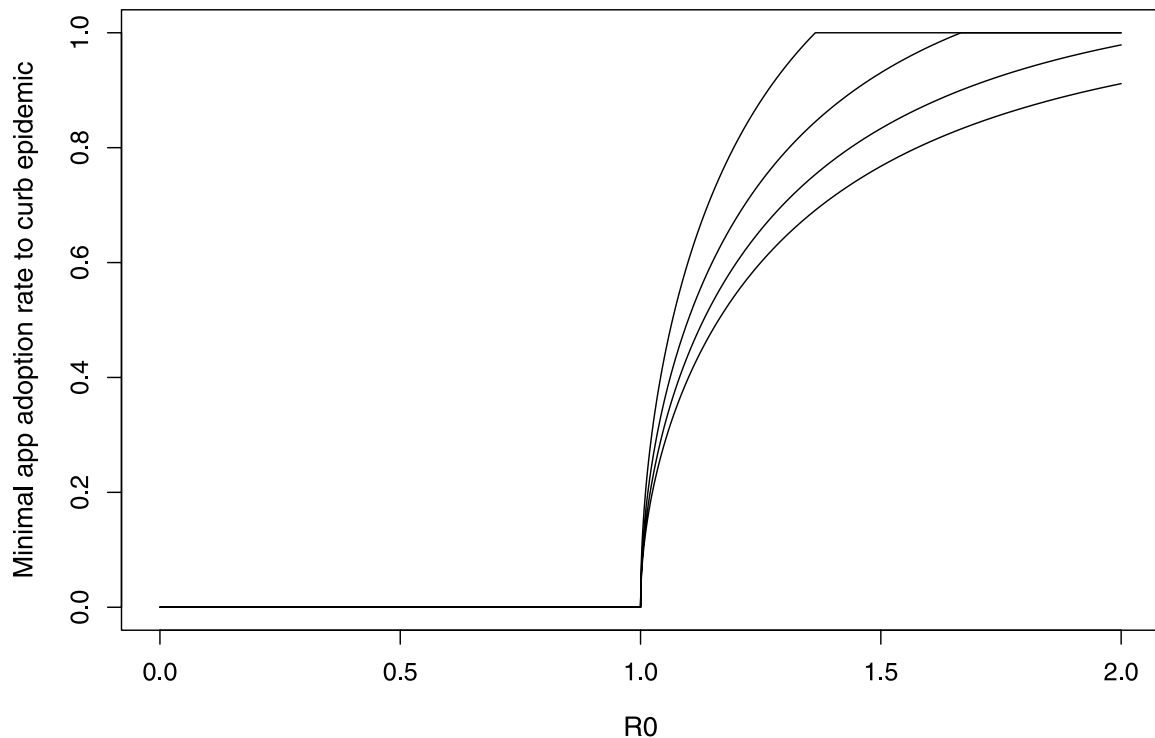


FIGURE 5. Minimal app adoption rate  $y_0$  to curb epidemic as a function of effective  $R_0$  under forward contact tracing (as given by Eq. (1.5)). Curves differ according to the value of the probability  $q_1$  of cooperating (*i.e.*, of quarantining upon app alert): from bottom to top,  $q_1 = 1$  (same curve as in Fig. 2), 0.8, 0.6 and 0.4. When contact tracing is unable to curb epidemic,  $y_0$  is set equal to 1 by convention (top two curves). Other parameter values as in Figure 2.

## 5. RECURSIVE CONTACT TRACING

### 5.1. Preliminary observations

In this section, we assume that alerts can be of degree 2 in the graph of contacts. Since the transmission tree is a subgraph of the contact network (assumed to also be tree-like), a physical contact of degree 2 is either a sibling or a grand-daughter in the transmission tree.

If a contact of degree 1 or 2 related to an index case decides to cooperate and quarantine upon being alerted, we will assume that this quarantine:

- removes her from the epidemic if she is a daughter of the index case (contact of degree 1, alerted by her mother), as previously;
- has no effect on her if she is the mother of the index case (contact of degree 1, alerted by her daughter), as previously;
- removes her from the epidemic if she is a grand-daughter of the index case (contact of degree 2, alerted by her mother);
- removes her from the epidemic with probability  $\ell$  if she is a sibling of the index case (contact of degree 2, alerted by her mother).

The effect of alerting a sibling implicitly depends on the temporality of infections. Indeed an individual receiving an alert coming from an index case via their common mother has better chance of being removed if she is infected later than the index case. Let us be more specific by considering a sibling of the index case,

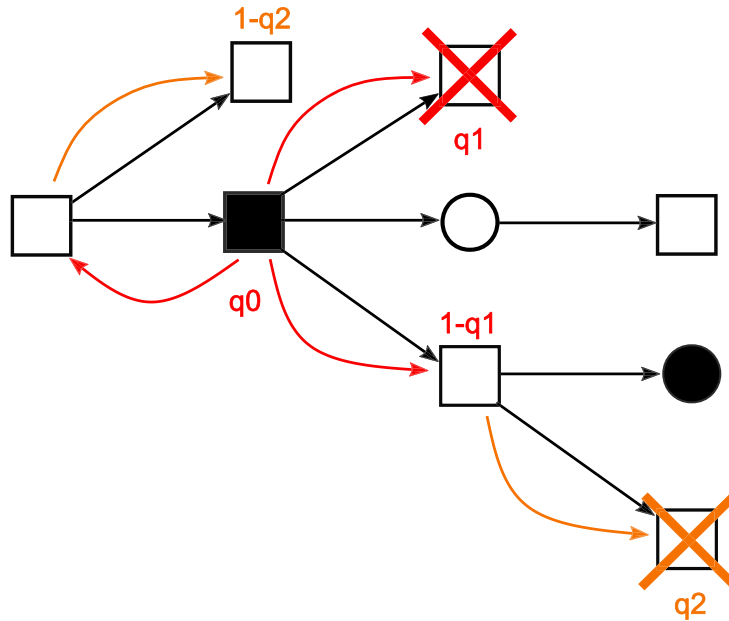


FIGURE 6. A  $YI$ -individual (symptomatic, using the app) decides with probability  $q_0$  to inform the app of her symptoms. This results in alerting all her physical contacts (degree 1, red arrows) of type  $Y$  (among which her mother and daughters in the epidemic) and in alerting all the contacts (degree 2, orange arrows) of type  $Y$  of these contacts (among which her siblings and grand-daughters in the epidemic). Each alerted daughter (degree 1) does quarantine with probability  $q_1$  and is then removed from the epidemic (red cross). Each alerted grand-daughter (degree 2) does quarantine with probability  $q_2$  and is then removed from the epidemic (orange cross). Each alerted sibling (degree 2) does quarantine with probability  $q_2$  and is then removed from the epidemic with probability  $\ell$  (not shown). Legend: white =  $\mathbf{A}$ , black =  $\mathbf{I}$ , square =  $\mathbf{Y}$ , circle =  $\mathbf{N}$ .

assumed to be infected  $x$  time units after the index case, and assumed to be alerted (alert of degree 2) and willing to quarantine. Let  $s(x)$  be the probability of being removed from the epidemic for this quarantined sibling of the index case. Also let  $N \geq 2$  denote the number of secondary infections coming from the common mother of the index case and this sibling, occurring at times  $\tau_1 < \dots < \tau_N$ . Then considering that the index case and the sibling have equal chances of being any of these  $N$  infected individuals,

$$\ell = E \left( \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j \neq i}^N s(\tau_j - \tau_i) \right).$$

In the particular case when  $s(x) = s_0$  if  $x > 0$  and  $s(x) = 0$  if  $x < 0$ , we get  $\ell = s_0/2$ . We will take  $\ell = 1/2$  as a default value, corresponding to  $s_0 = 1$ .

See Figure 6 for a cartoon depicting the possible effects of alerts of degrees 1 and 2.

In the next subsection, we quantify the effect of sibling's alerts.

## 5.2. Quantifying the effect of sibling's alerts

In this section, we take into account the fact that a  $Y$ -individual may receive alerts (of degree 1) from her daughters of type  $YI$  and forward them (in the form of a degree 2 alert) to her other daughters of type  $Y$ .

Start with an individual of type  $YZ$ ,  $Z = A$  or  $I$ . We denote by  $N_z$  the number of her daughters (in particular  $E(N_z) = R_z$ ) and by  $\tilde{R}_z$  the mean number of her daughters of type  $Y$  who survive alerts of degree 2 given by their siblings via their common mother, either by choosing to defect to all these alerts or by failing, whatever their cooperation choice, to be removed (interpreted as: ‘infected too early compared to their alerting sibling’), the latter event occurring with probability  $1 - \ell$ . Then

$$\tilde{R}_z = yE\left((1 - \ell)N_z + \ell N_z(1 - q_2)^{K_z}\right),$$

where, conditional on  $N_z = n \geq 1$ ,  $K_z$  is a binomial random variable with parameters  $n - 1$  and the probability  $q := q_0y(1 - f)$  of being a cooperative individual of type  $YI$ . In particular,

$$\begin{aligned} E\left((1 - q_2)^{K_z} \mid N_z = n\right) &= (1 - q + q(1 - q_2))^{n-1} \\ &= (1 - q_0y(1 - f) + q_0y(1 - f)(1 - q_2))^{n-1} = (1 - q_0q_2y(1 - f))^{n-1}, \end{aligned}$$

so we get

$$\tilde{R}_z = yE\left((1 - \ell)N_z + \ell N_z(1 - q_0q_2y(1 - f))^{N_z-1}\right) = y\left[(1 - \ell)R_z + \ell E\left(N_z(1 - q_0q_2y(1 - f))^{N_z-1}\right)\right].$$

This also reads

$$\tilde{R}_z = yx_zR_z,$$

with

$$x_z := 1 - \ell + \ell t_z$$

and

$$t_z := \frac{E\left(N_z(1 - q_0q_2y(1 - f))^{N_z-1}\right)}{R_z}$$

Notice that because  $f \neq 0$  and  $R_z \neq 0$ ,  $t_z \neq 0$  even if  $q_0q_2y = 1$ , so that  $x_z \neq 0$ .

In applications we will assume that  $N_z$  follows the Poisson distribution with parameter  $R_z$ . Since  $t_z = F'_z(1 - q_0q_2y(1 - f))/R_z$ , where  $F_z$  is the probability generating function of  $N_z$ , and  $F_z(s) = e^{R_z(s-1)}$  in the Poisson case, we easily get that  $F'_z(s)/R_z = F_z(s)$ , so that

$$t_a = e^{-q_0q_2y(1-f)R_a} \quad \text{and} \quad t_i = e^{-q_0q_2y(1-f)R_i}. \quad (5.1)$$

As soon as  $q_0$ ,  $q_2$  or  $y$  is zero, we find as expected that  $t_z = 1$  and so  $x_z = 1$ .

### 5.3. Main result

We can now state the main result of Section 5, which gives the condition on the rate  $y$  of app users required for the epidemic to die out with probability 1 in the case of recursive tracing.

**Theorem 5.1.** *As in Theorem 1.1, fix the effective average numbers of secondary infections  $(R_a, R_i, R_0)$  obtained in the presence of non-digital interventions (social distancing, isolation upon symptom onset, quarantining private contacts of symptomatics).*

Recall that  $(\star)$  denotes the condition that the epidemic dies out with probability 1, here thanks to the marginal effect of contact tracing (i.e., the multitype branching process modeling the outbreak in the presence of contact tracing is subcritical).

In the case of recursive tracing,  $(\star) \implies q_0 y(1-f)(1-q_1)(1-q_2)x_i R_i < 1$ , and if this last inequality holds,

$$\begin{aligned} (\star) \iff & -q_0 q_2 (1-q_1) y^2 (1-y)(1-f)x_i R_i R_0^2 \\ & + (1-R_0(1-y))y[f(1-x_a)R_a + (1-f)(1-x_i)R_i] \\ & + (1-R_0(1-y))yq_0(q_1+q_2-q_1q_2)(1-f)x_i R_i - R_0 + 1 \geq 0. \end{aligned} \quad (5.2)$$

For comparison, we recall the condition obtained in the case of forward tracing, namely equation (1.4):

$$(\star) \iff (1-R_0(1-y))yq_0q_1(1-f)R_i - R_0 + 1 \geq 0.$$

**Remark 5.2.** When  $q_2 = 0$ , there is no additional effect of recursive tracing compared to forward tracing, and indeed, because then  $x_i = x_a = 1$ , equation (5.2) then becomes equation (1.4) as expected.

**Remark 5.3.** When  $q_1 = 1$ , the only additional effect of recursive tracing compared to forward tracing is the alert of siblings and indeed equation (5.2) becomes in this case

$$(\star) \iff (1-R_0(1-y))y[f(1-x_a)R_a + (1-f)(1-x_i)R_i] + (1-R_0(1-y))yq_0(1-f)x_i R_i - R_0 + 1 \geq 0,$$

which boils down to equation (1.4) whenever  $x_a$  and  $x_i$  (embodying the effect of siblings' alerts) are set to 1 (for example by taking  $\ell = 0$ ).

Theorem 5.1 is proved in an appendix, Appendix A.

## 6. DISCUSSION

### 6.1. Robustness of results

Contact tracing supposes that case isolation is possible and is in force. Since contact tracing has no effect (at least in democracies?) on the efficiency of case isolation, and case isolation can be enforced independently of contact tracing, a measure of the effect of contact tracing should not include the effect of case isolation. This explains why we have chosen to express the minimal rate  $y_0$  (of adoption of a contact tracing app to curb the epidemic) as a function of the effective  $R_0$ , that is, the  $R_0$  obtained by non-digital interventions, notably case isolation.

By measuring only the net effect of contact tracing, our results are effectively insensitive to assumptions on the natural value of  $R_0$ , as well as on crucial parameters like  $f$  (fraction of asymptomatics) and  $b$  (one minus the fraction of infections avoided thanks to case isolation). We also found empirically that our results hold also for a wide range of values of the cooperating probabilities, as testified by the striking similarity of the four curves of Figures 5 and 7.

### 6.2. Interpretation of parameters

Let us discuss briefly the interpretation of some parameters in terms of the natural history of the virus and of the nature of healthcare policies.

- $1 - b$  is the fraction of secondary infections barred thanks to case isolation. Its value depends both on the time  $T$  taken to actually isolate after day  $D$  of symptom onset and on the natural history of the virus, via the fraction  $m$  of secondary infections made before  $D + T$ .

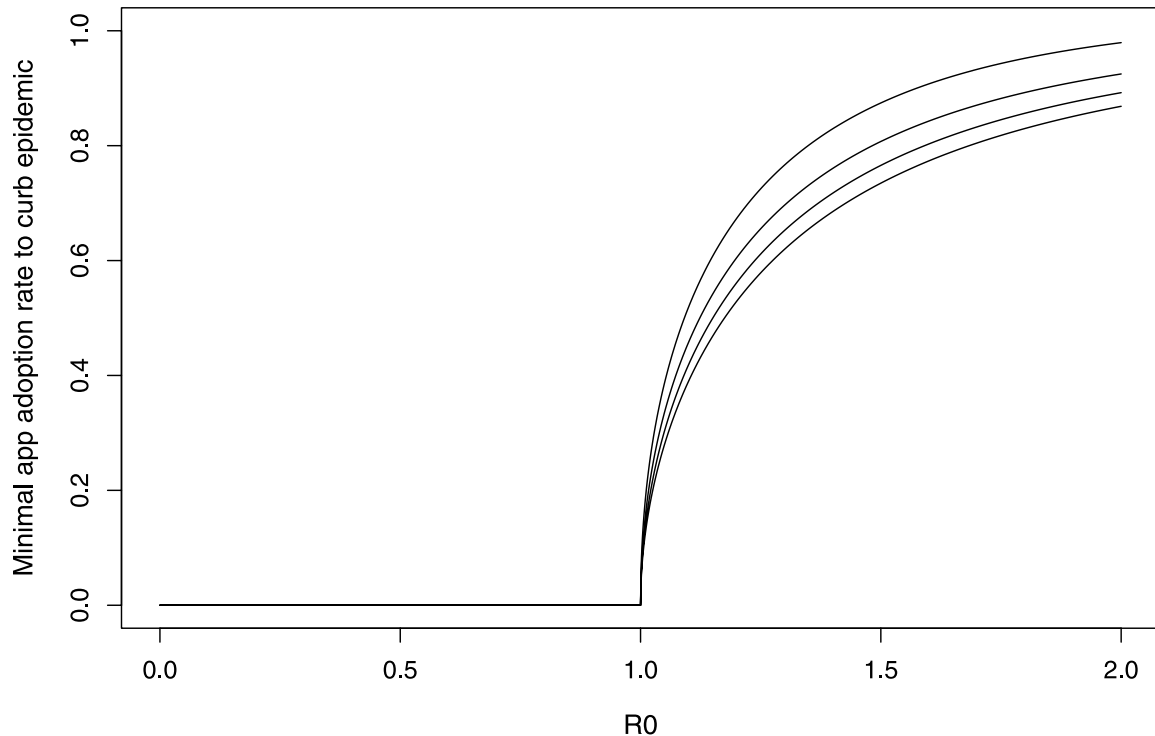


FIGURE 7. Minimal app adoption rate  $y_0$  to curb epidemic as a function of effective  $R_0$  under recursive contact tracing (as given by Eq. (5.2)). Curves differ according to the values of the probabilities  $q_1$  and  $q_2$  of cooperating (*i.e.*, of quarantining upon app alert as a contact of degree 1/of degree 2): from bottom to top,  $q_1 = q_2 = 1, 0.8, 0.6$  and  $0.4$ . Other parameter values as in Figure 2.

We have parameterized  $b$  as  $b = 1 - p + pm$ , where  $p$  is the probability of actually self-isolating. Parameter  $b$  can be tuned optimally by minimizing  $T$  so as to minimize  $m$  and by testing more systematically so as to maximize  $p$ , *e.g.*, because a symptomatic individual tested positive may feel more inclined to self-isolate.

- $q_0$  is the probability that a symptomatic app user informs the app of her symptoms when they first appear. In applications, we have assumed throughout the manuscript that  $q_0 = 1$ , but we cannot discard the existence of individuals who will download the app only to be aware of whether they have been in contact with sick individuals (‘leecher’ vs ‘seeder’ strategy).
- $q_1$  (resp.  $q_2$ ) is defined here as the probability of self-quarantining upon being alerted by app from a contact of degree 1 (resp. of degree 2). In this sense, we have  $q_1 = q_2$  whenever app users are not aware of whether they are contacts of degree 1 or 2 of the alerting index case; if they are, it is more reasonable to assume  $q_2 < q_1$ . Again, testing more systematically can help increase  $q_1$  and  $q_2$ .

Alternatively, the quantity  $q_1$  (resp.  $q_2$ ) can be interpreted as the likelihood of actual removal from the epidemic as a daughter of index case (resp. as a grand-daughter/sibling of index case). In this interpretation, one has on the contrary  $q_1 < q_2$ , bearing in mind that some daughters are infected too early to be actually removed, as opposed to grand-daughters.

### 6.3. Comparing forward and recursive contact tracing

The effect of recursive tracing is two-fold: to secure removal of grand-daughters of index cases when removal of daughters has failed ( $q_1 < 1$ , see Rem. 5.2) and to remove siblings of index cases by alerting their mother



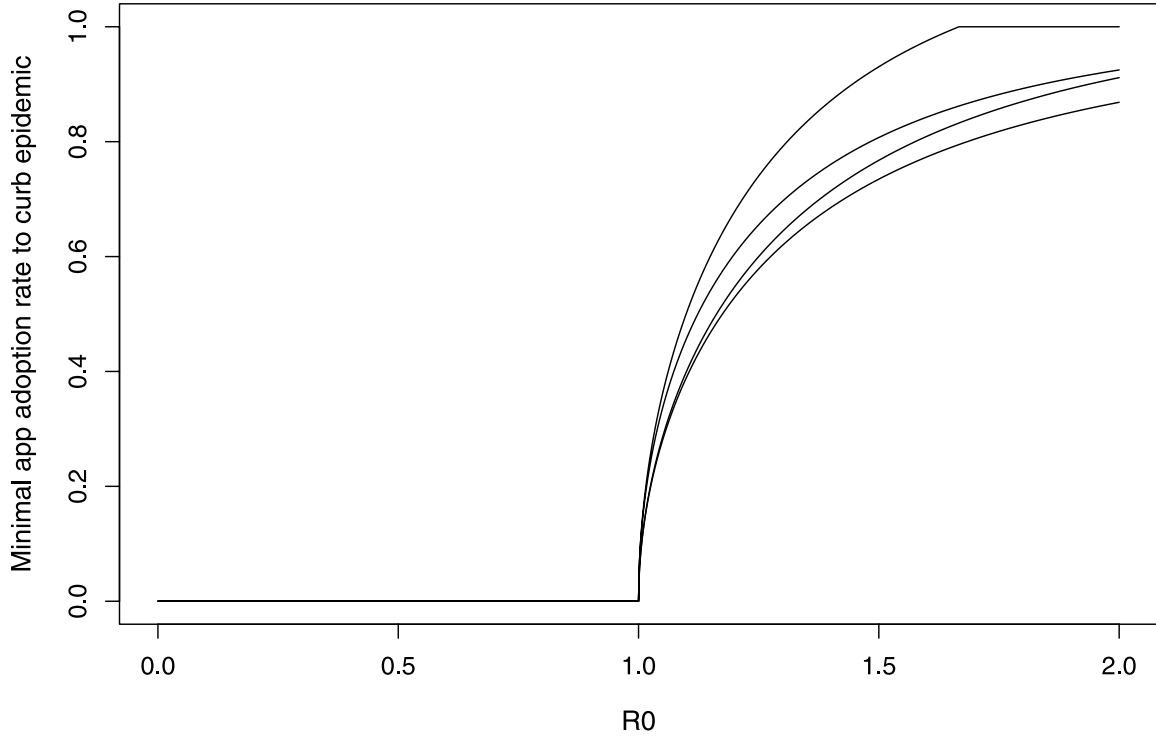


FIGURE 8. Minimal app adoption rate  $y_0$  to curb epidemic as a function of effective  $R_0$ , comparing forward vs recursive contact tracing. The top two curves differ from the bottom two curves according to the values of the probabilities  $q_1$  and  $q_2$ : bottom two curves:  $q_1 = q_2 = 1$ ; top two curves:  $q_1 = q_2 = 0.6$ . For each pair of (top vs bottom) curves, the bottom curve corresponds to recursive tracing and the top curve to forward tracing. When contact tracing is unable to curb epidemic,  $y_0$  is set equal to 1 by convention (top curve). Other parameter values as in Figure 2.

(see Rem. 5.3). Note that degree 2-contacts of an index case are a combination of two populations: an average number  $R_0^2$  of grand-daughters in the transmission tree and an unknown, larger number of non-infected physical contacts (of physical contacts). Due to the psychological and economical costs of quarantining all these secondary contacts, it is important to evaluate the marginal benefit of recursive tracing compared to forward tracing.

Figure 8 compares the minimal app adoption rate required to curb the epidemic with forward vs recursive tracing, when  $q_1 = q_2 = 0.6$  (top two curves) and when  $q_1 = q_2 = 1$  (bottom two curves). When  $q_1 = q_2 = 0.6$ , alerts of degree 2 (recursive tracing) can rescue the failure of forward tracing to curb the epidemic when the effective  $R_0$  is high, provided  $y$  is accordingly high. When  $q_1 = q_2 = 1$ , the only benefit of recursive tracing is through alerting siblings, and we see by comparing the two bottom curves that this benefit is hardly detectable (see also Fig. B.1). In addition, whatever the value of  $q_1$ , both strategies have basically the same effect for small values of  $R_0$  and  $y$ .

We conclude that the marginal benefits of recursive tracing are negligible compared to its costs, so that in particular, the explicit result given in equation (1.5) can be used for all practical purposes.

#### 6.4. Relation to previous work

The model (but not the approach) that we use here is similar to the one used in [6] and in two other works specifically interested in the current epidemic [3, 5]. See also [4, 13] for seminal works on the topic

of quantifying the effect of non-pharmaceutical interventions on epidemics and [15, 17] for works on contact networks and contact tracing.

We now explain why our predictions seem somewhat less optimistic than those given in [3, 5]. In these works, the  $R_0$  given corresponds to what we have termed  $R_0^\circ$  or  $R_0^c$ , possibly taking into account social distancing (in view of the values considered) but not case isolation. Recall that case isolation is parameterized in our notation by the fraction  $b$  of infections made before isolation. Also recall that with our notation, symptoms appear  $D$  days after infection, case isolation occurs  $T$  days after symptoms and  $m$  is the fraction of infections made before  $D + T$ , so that  $b = 1 - p + pm$ , where  $p$  is the probability of actually self-isolating.

In [5], the baseline scenario has  $f = 0$  (or 0.1),  $T = 3.4$  (‘short delay’) and the fraction of infections made before  $D$  is 0.15, which corresponds to  $b \approx 0.7$  (see Fig. 2 in this paper). In addition, the only scenario for which contact tracing works has  $R_0^\circ = 1.5$ . This corresponds to an effective  $R_0$  equal to  $bR_0^\circ = 1.05$  when  $f = 0$  (and equal to  $fR_0^\circ + b(1 - f)R_0^\circ = 1.095$  when  $f = 0.1$ ), in agreement with our findings that moderate adoption rates of the contact tracing app are sufficient only when the effective  $R_0$  is very close to 1.

In [3],  $R_0^\circ = 2$  and  $fR_a = 0.1$ . The main results can be seen on Figure 3 in this paper. Each panel corresponds to a different value of  $T$ , decreasing from left to right. The rightmost panel (best case scenario) has  $T = 0$ , which implies that  $m$  is the fraction of infections made before symptoms by symptomatics and is approximately 0.5 (see Fig. 2 in this paper). The panel shows the region of parameter space  $(X, Y)$  for which the epidemic dies out, where  $X$  is the ‘success rate of instant isolation of symptomatic cases’ and  $Y$  is the ‘success rate of instant contact tracing’. In our notation,  $X = p$  and  $Y = q_0q_1y^2$ . The effective  $R_0$  is roughly  $fR_a + (1 - p + pm)(R_0^\circ - fR_a) = 2 - 0.95p$ , ranging from 1.05 to 2 when  $p$  ranges between 0 and 1. Taking  $q_0 = q_1 = 1$  and referring to top right panel of our Figure 5, our prediction is that  $y_0$  ranges between 0.2 and 0.9 as  $p$  goes from 1 to 0, *i.e.*  $Y(p) = q_0q_1y_0^2$  ranges between 0.04 to 0.8, which is actually slightly more optimistic for large values of  $p$  than what shows Figure 3 in [3]. Referring to calculations made page 10 and taking  $p = 1/2$  so that the effective  $R_0 \approx 3/2$ , our prediction is  $y_0 \approx 0.77$ . This yields  $Y(p) = q_0q_1y_0^2 = 0.6$ , which is visually the same prediction as on Figure 3 in [3].

In conclusion, we see that our predictions are actually in line with those given in [5] and [3]. The most prominent differences come from the facts that 1) we measure the net effect of contact tracing by comparison with an effective  $R_0$  that takes into account the effect of case isolation (as a rule of thumb there can be a factor 2 between the effective  $R_0$  and the natural  $R_0$  other studies refer to) and 2) we measure this effect in terms of a minimal rate of app users rather than in terms of a minimal efficiency of contact tracing (as a rule of thumb the latter is the square of the former). The bottomline is that all three studies agree that the minimal rate of contact tracing app users must be larger than 60-70% to curb the epidemic unless the effective  $R_0$ , taking case isolation and social distancing into account, is already very close to 1.

## APPENDIX A. PROOF OF THEOREM 5.1

We prove Theorem 5.1 using a stopping line technique (see heuristic mentioned in Rem. 1.3) that allows us to display the semi-explicit criterion (5.2) for the leading eigenvalue of the underlying multitype branching process to be smaller than 1.

Because the behavior of an individual who is alerted possibly influences the removal of her daughters, we have to distinguish whether a  $Y$ -individual has received an alert (degree 1) or not. A  $Y$ -individual who has been alerted by an alert of degree 1 coming from her mother will be said ‘in excited state’ or simply ‘alerted’ and the corresponding type denoted with a star. Here are the following kinds of types to consider:  $YI^*$ ,  $YA^*$ ,  $YI$ ,  $YA$ ,  $NA$ ,  $NI$ . An individual of the four latter types will be said ‘in ground state’.

An individual in ground state who is in state  $YA$  or  $YI$  (resp.  $NA$  or  $NI$ ) with probabilities  $f$  and  $1 - f$  will merely be denoted  $Y$  (resp.  $N$ ) and called a regenerative state. In the genealogical tree of transmissions starting from a single individual, we follow all lines of descent descending from her and stop them at the first regenerative state encountered. The set of regenerative states forms what is called a stopping line [2] in the transmission tree. We will call seed-tree the tree obtained by pruning from the initial transmission tree all vertices downstream

of the stopping line. The leaves of a seed tree are all in a regenerative state, either  $Y$  or  $N$ . We call them the  $Y$ -regenerative leaves and  $N$ -regenerative leaves of the seed tree, respectively. See Figure A.1.

We can then define a Galton–Watson branching process with two types  $Y$  and  $N$  by saying that the offspring of type  $Z$  ( $Z = Y$  or  $N$ ) of a  $X$ -individual ( $X = Y$  or  $N$ ) are the  $Z$ -regenerative leaves of a seed-tree seeded by  $X$ . This process has no interest in itself except that the epidemic dies out iff it is subcritical.

We let  $m_{ya,y}$  denote the average number of  $Y$ -regenerative leaves of a seed tree seeded by a  $YA$ -individual. We define similarly  $m_{yi,y}$ ,  $m_{ya,n}$ ,  $m_{yi,n}$ ,  $m_{na,y}$ ,  $m_{ni,y}$ ,  $m_{na,n}$ ,  $m_{ni,n}$ . We then define

$$m_{y,y} := fm_{ya,y} + (1-f)m_{yi,y} \quad \text{and} \quad m_{y,n} := fm_{ya,n} + (1-f)m_{yi,n}.$$

Similarly, we define

$$m_{n,y} := fm_{na,y} + (1-f)m_{ni,y} \quad \text{and} \quad m_{n,n} := fm_{na,n} + (1-f)m_{ni,n}.$$

We still define  $M$  as the matrix with entries  $m_{y,y}$ ,  $m_{y,n}$ ,  $m_{n,y}$  and  $m_{n,n}$ , despite the fact that these quantities have a different meaning from theirs in the previous section. However, since  $M$  is the mean matrix of the two-type branching process defined previously, we still have that the epidemic dies out with probability 1 iff the leading eigenvalue of  $M$  is smaller than 1.

We will now compute the expected number of individuals of each type at generation  $k$  of a seed-tree, for example denoted  $[YI^*]_k$  for individuals of type  $YI^*$ . Let us make some preliminary observations:

- A seed-tree seeded by a  $NA$ , a  $NI$  or a  $YA$ -individual stops at generation 1, because all her daughters are in ground state ( $Y$  or  $N$ ).
- When the seed-tree is seeded by a  $YI$ -individual, there are two possibilities:
  - if the seed cooperates (*i.e.*, informs the app), then the daughters of the seed can be of type  $YA^*$ ,  $YI^*$  or  $N$ ;
  - if the seed does not cooperate, the seed-tree stops at generation 1 as previously.

Now we consider a seed-tree starting from a  $YI$ -individual. Recall that for the infection of a multiply alerted individual to succeed, this individual must defect independently to all alerts she has received.

**Infections from an individual of type  $YI$  of generation 0 (seed).** With probability  $1 - q_0$ , the daughters of the seed are

- an expected number  $yx_i R_i$  of type  $Y$ ,
- an expected number  $(1-y)R_i$  of type  $N$ .

Recall that  $R_i = R_i^{cbk}$ , where the values of  $c$ ,  $b$ ,  $k$  and  $r$  can be tuned once for all. With probability  $q_0$ , the daughters of the seed are

- an expected number  $yf(1-q_1)x_i R_i$  of type  $YA^*$ ,
- an expected number  $y(1-f)(1-q_1)x_i R_i$  of type  $YI^*$ ,
- an expected number  $(1-y)R_i$  of type  $N$ .

**Infections from an individual of type  $YA^*$  of generation  $k \geq 1$ .** Daughters of an individual of type  $YA^*$  of generation  $k \geq 1$  are

- an expected number  $y(1-q_2)x_a R_a$  of type  $Y$ ,
- an expected number  $(1-y)R_a$  of type  $N$ .

**Infections from an individual of type  $YI^*$  of generation  $k \geq 1$ .** With probability  $q_0$ , daughters of an individual of type  $YI^*$  of generation  $k \geq 1$  are

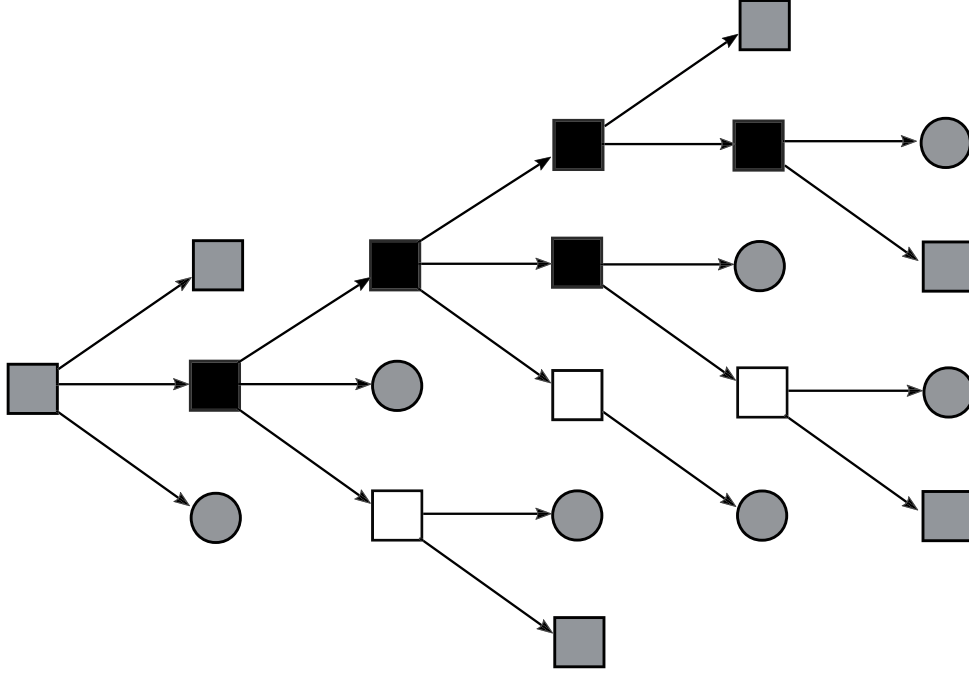


FIGURE A.1. A seed-tree starting from a  $Y$ -individual. A seed-tree is a maximal subtree of the transmission tree such that all internal transmission edges are doubled with an alerting arrow (not shown). In particular, a leaf of a seed-tree is by definition an individual who receives no alert of degree 1 from her mother (also called regenerative), either because she is of type  $N$  or because her mother is a  $YI$  who does not cooperate, or because her mother is a  $YA$  merely forwarding an alert (degree 1) to which she does not comply (degree 2). Legend: white =  $\mathbf{A}$ , black =  $\mathbf{I}$ , square =  $\mathbf{Y}$ , circle =  $\mathbf{N}$ , gray = regenerative state.

- an expected number  $yf(1 - q_1)(1 - q_2)x_iR_i$  of type  $YA^*$ ,
- an expected number  $y(1 - f)(1 - q_1)(1 - q_2)x_iR_i$  of type  $YI^*$ ,
- an expected number  $(1 - y)R_i$  of type  $N$ .

With probability  $1 - q_0$ , her daughters are

- an expected number  $y(1 - q_2)x_iR_i$  of type  $Y$ ,
- an expected number  $(1 - y)R_i$  of type  $N$ .

Then we obtain the following equations

$$[YA^*]_1 = q_0yf(1 - q_1)x_iR_i \quad (\text{A.1})$$

$$[YI^*]_1 = q_0y(1 - f)(1 - q_1)x_iR_i \quad (\text{A.2})$$

$$[Y]_1 = (1 - q_0)yx_iR_i \quad (\text{A.3})$$

$$[N]_1 = (1 - y)R_i, \quad (\text{A.4})$$

and for any  $k \geq 1$ ,

$$[YA^*]_{k+1} = q_0yf(1 - q_1)(1 - q_2)x_iR_i[YI^*]_k \quad (\text{A.5})$$

$$[YI^*]_{k+1} = q_0y(1 - f)(1 - q_1)(1 - q_2)x_iR_i[YI^*]_k \quad (\text{A.6})$$

$$[Y]_{k+1} = y(1 - q_2)x_a R_a [Y A^*]_k + (1 - q_0)y(1 - q_2)x_i R_i [Y I^*]_k \quad (\text{A.7})$$

$$[N]_{k+1} = (1 - y)R_a [Y A^*]_k + (1 - y)R_i [Y I^*]_k. \quad (\text{A.8})$$

Now define  $T_y$  (resp.  $T_n$ ) the total expected number of  $Y$ -regenerative (resp.  $N$ -regenerative) leaves of the seed-tree seeded by a  $YI$ -individual:

$$T_y := \sum_{k \geq 0} [Y]_{k+1} \quad \text{and} \quad T_n := \sum_{k \geq 0} [N]_{k+1}.$$

First observe that thanks to (A.2) and (A.6), we get

$$[Y I^*]_k = q_0 y (1 - f) (1 - q_1) x_i R_i \rho_i^{k-1} \quad k \geq 1,$$

with

$$\rho_i := q_0 y (1 - f) (1 - q_1) (1 - q_2) x_i R_i.$$

As a consequence, thanks to (A.1) and (A.5),

$$\begin{aligned} [Y A^*]_{k+1} &= \{q_0 y f (1 - q_1) (1 - q_2) x_i R_i\} \{q_0 y (1 - f) (1 - q_1) x_i R_i\} \rho_i^{k-1} \\ &= q_0 y f (1 - q_1) x_i R_i \rho_i^k \quad k \geq 0. \end{aligned}$$

Next, thanks to (A.7), we get

$$\begin{aligned} [Y]_{k+1} &= y(1 - q_2) \{x_a R_a [Y A^*]_k + (1 - q_0) x_i R_i [Y I^*]_k\} \\ &= q_0 y^2 (1 - q_1) (1 - q_2) \{f x_a R_a + (1 - q_0) (1 - f) x_i R_i\} x_i R_i \rho_i^{k-1} \quad k \geq 1. \end{aligned}$$

Finally, thanks to (A.8),

$$\begin{aligned} [N]_{k+1} &= (1 - y) \{R_a [Y A^*]_k + R_i [Y I^*]_k\} \\ &= q_0 y (1 - y) (1 - q_1) \{f R_a + (1 - f) R_i\} x_i R_i \rho_i^{k-1} \quad k \geq 1. \end{aligned}$$

Using (A.3), we have

$$T_y = [Y]_1 + \sum_{k \geq 1} [Y]_{k+1} = (1 - q_0) y x_i R_i + \frac{q_0 y^2 (1 - q_1) (1 - q_2) \{f x_a R_a + (1 - q_0) (1 - f) x_i R_i\} x_i R_i}{1 - \rho_i}$$

whenever  $\rho_i < 1$ . On the other hand, if  $\rho_i = q_0 y (1 - f) (1 - q_1) (1 - q_2) x_i R_i \geq 1$ , then because  $q_0 y^2 (1 - f) (1 - q_1) (1 - q_2) x_i R_i \neq 0$  and  $f x_a R_a \neq 0$ , we get  $T_y = \infty$ .

Similarly, using (A.4), we have

$$T_n = [N]_1 + \sum_{k \geq 1} [N]_{k+1} = (1 - y) R_i + \frac{q_0 y (1 - y) (1 - q_1) \{f R_a + (1 - f) R_i\} x_i R_i}{1 - \rho_i}$$

whenever  $\rho_i < 1$ . On the other hand, if  $\rho_i \geq 1$ , then because  $q_0 y (1 - q_1) x_i R_i \neq 0$  and  $x_i R_i \neq 0$ , we get  $T_n = \infty$  whenever  $y \neq 1$ . Now we use the fact that

$$m_{y,y} = y f x_a R_a + (1 - f) T_y$$

$$m_{y,n} = (1 - y)fR_a + (1 - f)T_n$$

while

$$m_{n,y} = R_0y \quad \text{and} \quad m_{n,n} = R_0(1 - y).$$

If  $\rho_i \geq 1$ , we have  $m_{y,y} = \infty$ , so the leading eigenvalue of  $M$  cannot be smaller than 1 (translated as the first implication given in the statement of the theorem). Then assume that  $\rho_i < 1$ . Elementary algebra yields

$$m_{y,y} = (fx_aR_a + (1 - q_0)(1 - f)x_iR_i) \frac{y}{1 - \rho_i} \quad (\text{A.9})$$

and

$$m_{y,n} = (1 + q_0q_2y(1 - f)(1 - q_1)x_iR_i) \frac{R_0(1 - y)}{1 - \rho_i}. \quad (\text{A.10})$$

The determinant of  $M$  is

$$\det M = m_{y,y}m_{n,n} - m_{y,n}m_{n,y},$$

which after calculation is

$$\det M = -\frac{R_0y(1 - y)}{1 - \rho_i} \{f(1 - x_a)R_a + (1 - f)(1 - x_i)R_i + q_0(1 - f)x_iR_i [1 + q_2(1 - q_1)yR_0]\}$$

Also recall that the trace of  $M$  is  $\text{Tr } M = m_{y,y} + m_{n,n}$ . Now as in the case of forward tracing, we denote by  $Q$  the characteristic polynomial of  $M$ , *i.e.*,

$$Q(X) = X^2 - (\text{Tr } M)X + \det M.$$

Again  $Q(0) \leq 0$  so the leading eigenvalue of  $M$  is the unique positive root of  $Q$ . Also, this root is smaller than 1 if and only if  $Q(1) \geq 0$ , which yields

$$(\star) \iff 1 - \text{Tr } M + \det M \geq 0.$$

After some algebra, we get

$$\begin{aligned} (\star) \iff & -q_0q_2(1 - q_1)y^2(1 - y)(1 - f)x_iR_iR_0^2 \\ & + (1 - R_0(1 - y))y[f(1 - x_a)R_a + (1 - f)(1 - x_i)R_i] \\ & + (1 - R_0(1 - y))yq_0(q_1 + q_2 - q_1q_2)(1 - f)x_iR_i - R_0 + 1 \geq 0, \end{aligned}$$

which is exactly equation (5.2) and so ends the proof.

## APPENDIX B. EFFECT OF $f$

We display an additional figure showing that the effect of contact tracing as well as the rescuing effect of recursive vs forward tracing are not much improved if the rate  $f$  of asymptomatics is reduced compared to its default value of  $1/3$ .

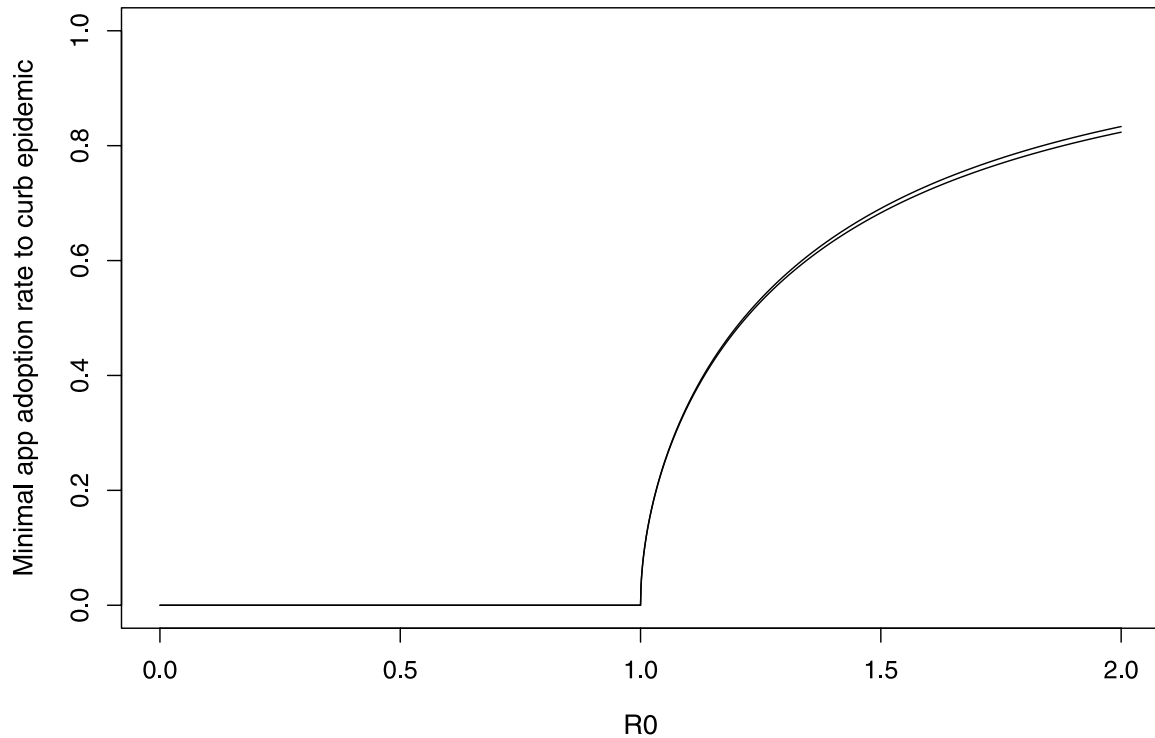


FIGURE B.1. Minimal app adoption rate  $y_0$  to curb epidemic as a function of effective  $R_0$ , comparing forward (top curve) vs recursive (bottom curve) contact tracing. Best case scenario:  $q_0 = q_1 = q_2 = 1$ . Other parameter values as in other figures except  $f = 1/10$  instead of  $f = 1/3$ .

*Acknowledgements.* The author thanks the *Center for Interdisciplinary Research in Biology* (CIRB) for funding and his entire research group SMILE (Stochastic Models for the Inference of Life Evolution) for numerous fruitful (online) discussions around the COVID-19 pandemic. He wishes to thank two reviewers for their thorough reading, insightful comments and extremely relevant, helpful remarks. He also thanks the task force MODCOV19 of the INSMI (CNRS) for their technical/scientific support during the epidemic.

## REFERENCES

- [1] Q. Bi, Y. Wu, S. Mei, C. Ye, X. Zou, Z. Zhang, X. Liu, L. Wei, S.A. Truelove, T. Zhang *et al.*, Epidemiology and transmission of COVID-19 in Shenzhen China: analysis of 391 cases and 1,286 of their close contacts. *MedRxiv* (2020).
- [2] B. Chauvin, Product martingales and stopping lines for branching Brownian motion. *Ann. Probab.* **19** (1991) 1195–1205.
- [3] L. Ferretti, C. Wymant, M. Kendall, L. Zhao, A. Nurtay, L. Abeler-Dörner, M. Parker, D. Bonsall and C. Fraser, Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* (2020).
- [4] C. Fraser, S. Riley, R.M. Anderson and N.M. Ferguson, Factors that make an infectious disease outbreak controllable. *Proc. Natl. Acad. Sci.* **101** (2004) 6146–6151.
- [5] J. Hellewell, S. Abbott, A. Gimma, N.I. Bosse, C.I. Jarvis, T.W. Russell, J.D. Munday, A.J. Kucharski, W.J. Edmunds, F. Sun *et al.*, Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *Lancet Global Health* (2020).
- [6] D. Klinkenberg, C. Fraser and H. Heesterbeek, The effectiveness of contact tracing in emerging epidemics. *PloS one* **1** (2006).
- [7] E. Lavezzo, E. Franchin, C. Ciavarella, G. Cuomo-Dannenburg, L. Barzon, C. Del Vecchio, L. Rossi, R. Manganelli, A. Lorigian, N. Navarin *et al.*, Suppression of a SARS-CoV-2 outbreak in the Italian municipality of vo. *Nature* **584** (2020) 425–429.
- [8] M. Lv, X. Luo, J. Estill, Y. Liu, M. Ren, J. Wang, Q. Wang, S. Zhao, X. Wang, S. Yang *et al.*, Coronavirus disease (COVID-19): a scoping review. *Eurosurveillance* **25** (2020) 2000125.
- [9] S. Ma, J. Zhang, M. Zeng, Q. Yun, W. Guo, Y. Zheng, S. Zhao, M.H. Wang and Z. Yang, Epidemiological parameters of coronavirus disease 2019: a pooled analysis of publicly reported individual data of 1155 cases from seven countries. *medRxiv* (2020).

- [10] K. Mizumoto, K. Kagaya, A. Zarebski and G. Chowell, Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020. *Eurosurveillance* **25** (2020) 2000180.
- [11] C.J. Mode, Vol. 34 of *Multitype branching processes: theory and applications*. American Elsevier Pub. Co. (1971).
- [12] H. Nishiura, T. Kobayashi, T. Miyama, A. Suzuki, S.-m. Jung, K. Hayashi, R. Kinoshita, Y. Yang, B. Yuan, A.R. Akhmetzhanov *et al.*, Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). *Int. J. Infectious Diseases* **94** (2020) 154.
- [13] C.M. Peak, L.M. Childs, Y.H. Grad and C.O. Buckee, Comparing nonpharmaceutical interventions for containing emerging epidemics. *Proc. Natl. Acad. Sci.* **114** (2017) 4023–4028.
- [14] H. Qian, T. Miao, L. Liu, X. Zheng, D. Luo and Y. Li, Indoor transmission of SARS-CoV-2. *Indoor Air* **31** (2021) 639–645.
- [15] M. Salathé, M. Kazandjieva, J.W. Lee, P. Levis, M.W. Feldman and J.H. Jones, A high-resolution human contact network for infectious disease transmission. *Proc. Natl. Acad. Sci.* **107** (2010) 22020–22025.
- [16] H. Salje, C.T. Kiem, N. Lefrancq, N. Courtejoie, P. Bosetti, J. Paireau, A. Andronico, N. Hozé, J. Richet, C.-L. Dubost *et al.*, Estimating the burden of SARS-CoV-2 in France. *Science* **369** (2020) 208–211.
- [17] E. Yoneki and J. Crowcroft, Epimap: Towards quantifying contact networks for understanding epidemiology in developing countries. *Ad Hoc Netw.* **13** (2014) 83–93.

## Subscribe to Open (S2O)

A fair and sustainable open access model



This journal is currently published in open access under a Subscribe-to-Open model (S2O). S2O is a transformative model that aims to move subscription journals to open access. Open access is the free, immediate, online availability of research articles combined with the rights to use these articles fully in the digital environment. We are thankful to our subscribers and sponsors for making it possible to publish this journal in open access, free of charge for authors.

**Please help to maintain this journal in open access!**

Check that your library subscribes to the journal, or make a personal donation to the S2O programme, by contacting [subscribers@edpsciences.org](mailto:subscribers@edpsciences.org)

More information, including a list of sponsors and a financial transparency report, available at: <https://www.edpsciences.org/en/math-s2o-programme>