

# Note for $\mathcal{R}_{\text{eff}}$ question

## 1. Instantaneous $\mathcal{R}_{\text{eff}}$ in homogeneous model

For homogeneous SIR model:

$$\begin{cases} \dot{S}(t) = -\beta S(t)I(t) \\ \dot{I}(t) = +\beta S(t)I(t) - \gamma I(t) \\ \dot{R}(t) = +\gamma I(t) \end{cases}$$

One can define the effective reproductive number  $\mathcal{R}_{\text{eff}}(t)$  as the (expected/average) instantaneous incidence caused by any randomly chosen infected individual in  $I(t)$  at the moment  $t$  before they recover.

Consider the expected recovery time from the exponential distribution is given by  $\frac{1}{\gamma}$ , this gives us:

$$\mathcal{R}_{\text{eff}} = -\frac{\dot{S}(t)}{I(t)} \times \frac{1}{\gamma} = \frac{\beta}{\gamma} S(t)$$

for homogeneous model.

If we take the general definition without specifying the incidence term  $\dot{S}(t)$ :

$$\mathcal{R}_{\text{eff}} = -\frac{\dot{S}(t)}{I(t)} \times \frac{1}{\gamma}$$

it still make sense for the definition.

This definition also indicates that the instantaneous incidence  $\dot{S}(t)$  is proportional to  $I(t)$

Moreover,

$$\mathcal{R}_{\text{eff}} = -\frac{\dot{S}(t)}{I(t)} \times \frac{1}{\gamma} = 1 \Leftrightarrow \dot{I}(t) = 0$$

which fits the intuitive understanding for the effective reproductive number.

## 2. Migration to MSV network model

Now consider the network framework developed by [J.C. Miller, A.C. Slim & E.M. Volz\(2011\)](#) (referred as **MSV**), such that the system is now a combination of compartment of nodes:

$$\begin{cases} S(t) = G_p(\theta(t)) \\ I(t) = 1 - S(t) - R(t) \\ \dot{R}(t) = \gamma I(t) \end{cases}$$

and ODE for edges:

$$\dot{\theta} = -\beta\phi_I = -\beta(\theta - \phi_S - \phi_R) = -\beta\theta + \beta\frac{G'_p(\theta)}{\delta} + \gamma(1 - \theta)$$

The equation for  $I(t)$  still leads to:

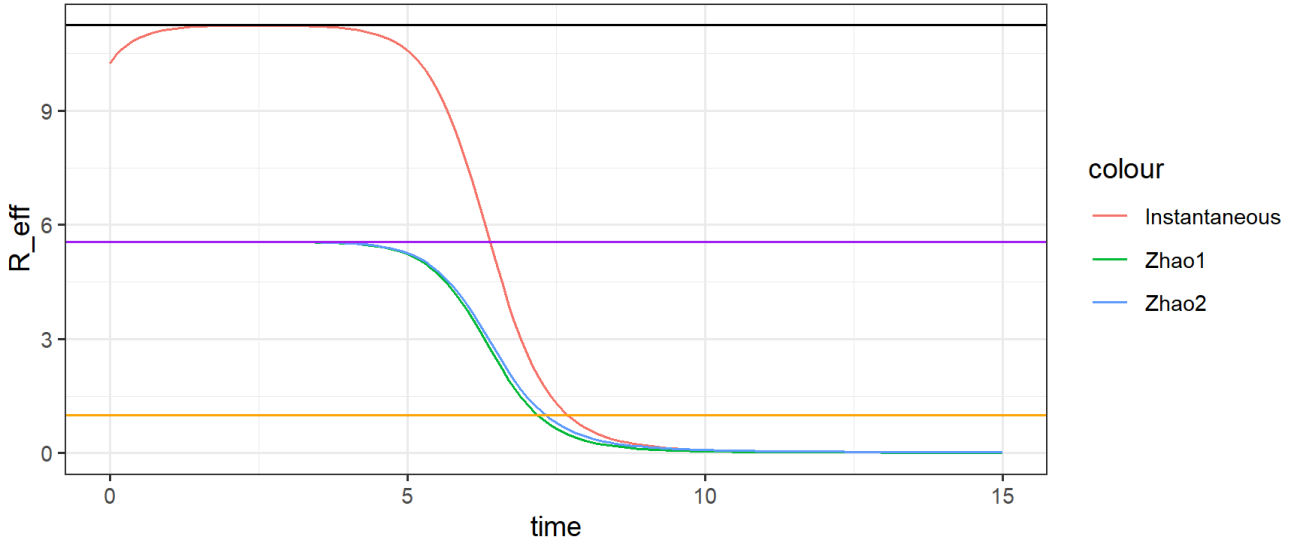
$$\dot{I}(t) = -\dot{S}(t) - \gamma I(t)$$

like the homogeneous model.

So it seems that we can migrate the  $\mathcal{R}_{\text{eff}}$  directly here.

### 3. Difference in Simulation

However, here is the simulation result for an outbreak with  $\beta = 0.25$ ,  $\gamma = 0.2$  on a network with Poisson degree distribution and mean degree  $\delta = 10$ :



The red curve is the instantaneous  $\mathcal{R}_{\text{eff}}$  as defined in section 1.

The green and blue curve is  $\mathcal{R}_{\text{eff}}^*$  corresponding to Zhao1 and Zhao2 result.

The black horizontal line is estimated peak value of  $\mathcal{R}_{\text{eff}}$  as described in section 8.

The purple horizontal line is the  $\mathcal{R}_0^*$  value given by MSV and other literatures for network.

The orange horizontal line is just reference of 1.

It is clear that instantaneous  $\mathcal{R}_{\text{eff}}$  defined by

- (Average/expected) number of incidence caused by **a randomly chosen** infected individual in  $I(t)$  at the moment  $t$  before they recover.  
does not converge to  $\mathcal{R}_0^*$  and the  $\mathcal{R}_{\text{eff}}^*$ , which is defined by
- Average/expected number of incidence caused by a random individual **newly** infected at time  $t$  **during the whole outbreak**.

If we try to estimate the peak value (represent by the black in the simulation figure), it is possible for it exceed the maximum degree of the network in some special degree distribution with some feasible value of disease parameters.

- e.g.  $\beta = 0.25$ ,  $\gamma = 0.1$ , with a network with all vertices has degree  $\delta = 10$ , this peak value for  $\mathcal{R}_{\text{eff}}$  is  $22.5 > 10$  as the maximum number of neighbours that any nodes have.

So we are interested in why and how for such difference.

## 4. Some thoughts for reason of the difference

For homogeneous model, any individual in have same infectivity to infect susceptible nodes, because of the fully-mixed mass-action assumption for contact.

In such case, every infected individual has the same infectivity like any individual in  $I(t)$ , so the  $\mathcal{R}_{\text{eff}}$  is proportion to instantaneous incident  $\dot{S}(t)$  averaged on  $I(t)$ , i.e. the incident is governed by  $\mathcal{R}_{\text{eff}} \times I(t)$ .

This definition might be less justifiable on MSV network frame with heterogeneity in contact.

Consider transmission on MSV type configuration network, it is possible to have an vertex in the  $I(t)$  compartment while no longer being able to transmit infection to any of its neighbor at the moment  $t$ , i.e. all of its neighbor are not susceptible at and after time  $t$ .

An obvious example would be infected vertex with degree one, whose only neighbor will be its infector and thus not being able to transmit the infection to any other vertices.

Unlike homogeneous case, such vertices are counted in  $I(t)$  but contribute nothing to new infections, while also flow to  $R$  compartment with same rate as those infected vertices with transmission potentials.

Similar arguments about contact heterogeneity applies to other vertices in  $I$ , where their heterogeneity in degree affect the their infection potential, but cannot be represented by just the proportion  $I(t)$ , as it is easily to see that degree distribution of infected nodes is clearly with time.

In heterogeneous network model, the instantaneous  $\mathcal{R}_{\text{eff}}$  as defined for homogenous model now have difference with retrospective definition, which are counting the average/expected number of all new cases caused by a random individual **newly** infected at time  $t$  **during the whole outbreak**.

In earlier paper of [Volz\(2008\)](#), Volz state that in their network model, the number of new infections/incidence in a small time interval is proportional to the proportion of  $S - I$  pairs at the moment (which is equivalent to number of potentially newly infected vertex) in the network.

This is in contrast to (homogeneous) compartment models in which the number of new infections/incidence is proportional the current number of infectious.

## 5. Derivation of $\mathcal{R}_0^*$

Based on this idea, [J.C. Miller, A.C. Slim & E.M. Volz\(2011\)](#) and many literature has defined the basic reproductive number  $\mathcal{R}_0^*$  from network approach as the expected number of infections a randomly chosen **newly** infected vertex causes.

This gives the expression:

$$\mathcal{R}_0^* = \frac{\beta}{\beta + \gamma} \times \frac{G_p''(1)}{\delta} = \frac{\beta}{\beta + \gamma} \times \frac{G_p''(1)}{G_p'(1)}$$

where  $G_p'(x)$  is the probability generating function (PGF) of the network degree distribution  $p$  and  $G_p'(1) = \delta$  is the mean degree of the network.

They also provide an equivalent derivations for the threshold effect of  $\mathcal{R}_0^*$  based on the ODE of  $\theta$ .

Consider  $\theta(0) = 1 + \epsilon$  ( $\epsilon < 0$ ) which is very close to 1:

$$\dot{\theta} = \dot{\epsilon} = -\beta(1 + \epsilon) + \beta \frac{G_p'(1 + \epsilon)}{\delta} + \gamma(1 - (1 + \epsilon))$$

Consider the first order:

$$\begin{aligned} G_p'(1 + \epsilon) &= \sum_d p_d d (1 + \epsilon)^{d-1} \\ &= \sum_d p_d d (1 + (d-1)\epsilon + o(\epsilon)) \\ &\approx \sum_d p_d d + \sum_d p_d d (d-1)\epsilon \\ &= \delta + \epsilon \sum_d p_d d (d-1) \\ &= \delta + \epsilon G_p''(1) \end{aligned}$$

Then for the first order we have:

$$\dot{\theta} = \dot{\epsilon} \approx -\beta(1 + \epsilon) + \beta \frac{\delta + \epsilon G_p''(1)}{\delta} - \gamma\epsilon = [-(\beta + \gamma) + \beta \frac{G_p''(1)}{\delta}] \times \epsilon$$

For the outbreak,  $\mathcal{R}_0^* = \frac{\beta}{\beta + \gamma} \frac{G_p''(1)}{\delta} > 1 \Leftrightarrow \frac{\dot{\epsilon}}{\epsilon} = [-(\beta + \gamma) + \beta \frac{G_p''(1)}{\delta}] > 0$  so  $\theta$  will decrease from 1 and cause the outbreak.

## 6. Zhao2 Result: Derivation of $\mathcal{R}_{\text{eff}}^*$

JD and RZ try to derive the corresponding  $\mathcal{R}_{\text{eff}}^*(t)$  from the stochastic process of transmission on network.

Follow [RomanescuEtAL\(2023\)](#): Define  $\mathcal{R}_{\text{eff}}^*(t)$  to be the expected number of secondary infections for a **newly** infected individual  $X_t$  at time  $t$ , s.t.  $\mathcal{R}_{\text{eff}}^*(t) = \mathbb{E}[X_t]$ .

To derive  $X_t$

- $K_S(t)$  is the random variable of degree of a random susceptible vertex at time  $t$ 
  - Proportion of susceptible vertices with degree  $k$  at time  $t$  in the entire population  
 $p_k^S(t) = p_k \theta^k$
  - Total proportion of susceptible nodes at time  $t$  is  $S(t) = G_p(\theta(t))$
  - Corresponding PGF of  $K_S(t)$  is  $G_S(x) = \frac{G_p(x\theta)}{G_p(\theta)} = \frac{G_p(x\theta)}{S}$
- Observe the process by which a susceptible individual becomes infected.

- Consider a random edge that has the potential to transmit infection at time  $t$ , which is an  $S - I$  pair: The uninfected individual at the end of this edge is chosen from the susceptible set, but not at random: an individual's chance of being selected is proportional to their degree, in the absence of higher-order features.
- Thus, the relative frequency of an individual of degree  $k$  becoming infected at the next time step is proportional to  $kp_k^S(t) = kp_k\theta^k$ .
- $K_I(t)$  is the random variable of degree of a random infective vertex at time  $t$ 
  - Corresponding PGF of  $K_I(t)$  is

$$\frac{\sum_k kp_k\theta^k x^k}{\sum_k kp_k\theta^k} = \frac{xG'_S(x)}{G'_S(1)} = G_I(x)$$

- Expectation of  $K_I(t)$  is:

$$\mathbb{E}(K_I(t)) = G'_I(1) = \frac{\theta G''_p(\theta)}{G'_p(\theta)}$$

- Since neighbors are assumed to be independent, given the a **newly infected** vertex have degree  $K_I(t)$ , the number  $X_t$  of new infective vertices infected by this vertex is distributed as a Binomial( $n = K_I(t) - 1, \mu$ ).
  - $n = K_I(t) - 1$  is because we know for any infective vertex, it can no longer infect its infector.
  - $\mu = \frac{\beta}{\beta+\gamma} \times \frac{\phi_S}{\theta}$  if we considering that the vertex is newly infected, where  $\frac{\phi_S}{\theta}$  is the probability that an edge is connected to a susceptible node given that it has not transmitted infection.
- Therefore, with law of total expectation, for a randomly newly infected node we know the expectation number  $X_t$  of infected vertices it can generate would be:

$$\begin{aligned} \mathcal{R}_{\text{eff}}^* &= \mathbb{E}[X_t] = \mathbb{E}_{K_I}[\mathbb{E}[X_t|K_I]] \\ &= \mathbb{E}_{K_I}[\mu(K_I - 1)] \\ &= \mu \mathbb{E}[K_I - 1] \\ &= \mu(\mathbb{E}[K_I] - 1) \\ &= \mu\theta \frac{G''_p(\theta)}{G'_p(\theta)} \end{aligned}$$

Take into the idea that

$$\mu = \frac{\beta}{\beta+\gamma} \times \frac{\phi_S}{\theta} = \frac{\beta}{\beta+\gamma} \times \frac{G'_p(\theta)}{\theta G'_p(1)}$$

, we have:

$$\mathcal{R}_{\text{eff}}^* = \frac{\beta}{\beta+\gamma} \frac{G''_p(\theta)}{\delta}$$

As  $t \rightarrow 0 \Leftrightarrow \theta \rightarrow 1$ ,  $\mathcal{R}_{\text{eff}}^*$  converge to

$$\mathcal{R}_0^* = \frac{\beta}{\beta + \gamma} \frac{G_p''(1)}{\delta}$$

Similar as how MSV verifying their  $\mathcal{R}_0^*$  with dynamic,  $\mathcal{R}_{\text{eff}}^* = 1$  is where  $\dot{\phi}_I = 0$ , which validate the result with the system.

- $\phi_I$  is defined as the proportion that a neighbor  $b$  is infected but the infection has not yet transmitted to the randomly chosen focal vertex  $a$ .

$$\dot{\phi}_I = [-(\beta + \gamma) + \beta \frac{G_p''(1)}{\delta}] \times \phi_I$$

## 6.1 How $\mathcal{R}_{\text{eff}}^*$ affect the incidence $-\dot{S}(t)$ ?

Consider rate of change for the incidence term

$$\begin{aligned} \frac{d}{dt}(-\dot{S}(t)) &= -\ddot{S}(t) \\ &= \beta\delta[\phi_S\dot{\phi}_I + \dot{\phi}_S\phi_I] \\ &= \beta\delta[\phi_S\phi_I(\beta + \gamma)(\mathcal{R}_{\text{eff}}^* - 1) - \beta \frac{G_p''(\theta)}{\delta} \phi_I^2] \\ &= \beta\delta\phi_I[\phi_S(\beta + \gamma)(\mathcal{R}_{\text{eff}}^* - 1) - \beta \frac{G_p''(\theta)}{\delta} \phi_I] \end{aligned}$$

If and only if  $(\mathcal{R}_{\text{eff}}^* - 1) > 0$ , the rate of change  $-\ddot{S}(t)$  have positive term (increasing force).

## 7. Attempt to Derive $\mathcal{R}_{\text{eff}}$ by Similar approach

Also, consider the Bayesian Formula and a randomly chose edge/stub  $u$ :

$$\begin{aligned} \mathbb{P}(u \in \theta \Leftrightarrow u \in \phi_I | u \text{ connect to a vertex } \in I) &= \frac{\mathbb{P}(u \in \phi_I | u \in \theta) \mathbb{P}(u \in \theta)}{\mathbb{P}(u \text{ connect to a vertex } \in I)} \\ &= \frac{\frac{\phi_I}{\theta} \times \theta}{\frac{NI \times \mathbb{E}(K_I)}{N\delta}} \\ &= \frac{\phi_I \delta}{I(t) \mathbb{E}[K_I]} \end{aligned}$$

Following previous idea for **newly** infected vertex, we could slightly modify this probability argument by replacing  $K_I$  with  $K_I - 1$  as we are sure for each infected (other than the initial patient-zero) vertex, there is one and only one edge comes from its infector, thus can no longer transmit the infection.

We name this new probability  $\eta = \frac{\phi_I \delta}{I(t) \mathbb{E}[K_I - 1]}$  which is the probability that a randomly chosen edge  $u$  is not yet transmit the infection, given it is connected to an infected focal vertex and not the edge infected the focal node.

Following the previous binomial distribution process, we might be able to construct the

expected number of new infection caused by a **RANDOM** infected vertex instead of a **NEWLY** infected vertex.

For a random edge  $u$  connected infected focal vertex at time  $t$  while not connected to its infector, the probability  $u$  can transmit the infection is

$$\eta \times \phi_S \times \alpha$$

- $\eta$  as defined earlier is probability  $u$  still not yet transmit disease.
- $\eta \times \phi_S$  is the probability that such  $u$  is connected to a susceptible node. Note we no longer need to conditional on  $\theta$  like in  $\mu$  as it is already considered in  $\eta$ .
- $\alpha$  is the probability such edge could transmit the infection before the infected focal vertex recovering.

Then following the binomial distribution, the expected number of new cases that a randomly infected vertex will causes at time  $t$  is given by:

$$\mathcal{R}_{\text{random}} = \alpha \phi_S \eta (\mathbb{E}[K_I] - 1) = \alpha \phi_S \frac{\phi_I \delta}{I(t) \mathbb{E}[K_I - 1]} (\mathbb{E}[K_I] - 1) = \alpha \times \frac{\delta \phi_S \phi_I}{I(t)}$$

Now we observe the traditional expression for  $\mathcal{R}_{\text{eff}}$ :

$$\mathcal{R}_{\text{eff}} = -\frac{\frac{dS(t)}{dt}}{I(t)} \times \frac{1}{\gamma} = -\frac{-\beta \delta \phi_S \phi_I}{I(t)} \times \frac{1}{\gamma} = \frac{\beta}{\gamma} \times \frac{\delta \phi_S \phi_I}{I(t)}$$

So they merge if  $\alpha = \frac{\beta}{\gamma}$ !

But my problem is for a newly chosen  $S - I$  edge in the whole network, the probability it transmit the infection before  $I$  recovered or  $S$  is infected by others is  $\frac{\beta}{\beta + \gamma}$ .

There is some inconsistency here and we need to better understand  $\alpha$ .

By def,  $\alpha$  should be a probability while  $\frac{\beta}{\gamma}$  could be larger than 1 while still feasible.

## 8. Estimation(?) of Peak Value of $\mathcal{R}_{\text{eff}}$

At the peak point, we must have  $\dot{\mathcal{R}}_{\text{eff}} = 0$ , which leads to

$$0 = \dot{\mathcal{R}}_{\text{eff}} = \frac{1}{\gamma} \times \frac{\ddot{S}I - \dot{S}\dot{I}}{I^2}$$

For non-zero  $I(t)$ , this just requires the numerator:

$$0 = \ddot{S}I - \dot{S}\dot{I} = \ddot{S}I - \dot{S}(-\dot{S} - \gamma I) \Leftrightarrow I_{\text{max}} = -\frac{\dot{S}^2}{\ddot{S} + \gamma \dot{S}}$$

As we could represent  $S$  and its derivatives with  $\theta$  and PGFs but have no explicit expression for  $I$ , we could take this relationship at peak back into  $\mathcal{R}_{\text{eff}}$ :

$$\max(\mathcal{R}_{\text{eff}}) = -\frac{\dot{S}}{I_{\text{max}}} \times \frac{1}{\gamma} = -\frac{\dot{S}}{-\frac{\dot{S}^2}{\ddot{S} + \gamma\dot{S}}} \times \frac{1}{\gamma} = \frac{\ddot{S} + \gamma\dot{S}}{\gamma\dot{S}}$$

Take into the previous relationships for  $\dot{S}$  and  $\ddot{S}$  we have:

$$\begin{aligned}\max(\mathcal{R}_{\text{eff}}) &= \frac{\ddot{S} + \gamma\dot{S}}{\gamma\dot{S}} \\ &= \frac{\beta}{\gamma} \left( \frac{G_p''(\theta)}{\delta} \times \frac{\phi_S - \phi_I}{\phi_S} - 1 \right) \\ &= \frac{\beta}{\gamma} \left[ \frac{G_p''(\theta)}{\delta} \left( 1 - \frac{\phi_I}{\phi_S} \right) - 1 \right] \\ &= \frac{\beta}{\gamma} \left[ \frac{G_p''(\theta)}{\delta} \left( 1 - \frac{\theta - \frac{\gamma}{\beta}(1 - \theta) - \frac{G_p'(\theta)}{\delta}}{\frac{G_p'(\theta)}{\delta}} \right) - 1 \right] \\ &= \frac{\beta}{\gamma} \left[ \frac{G_p''(\theta)}{\delta} \left( 2 - \delta \times \frac{\beta\theta - \gamma(1 - \theta)}{\beta G_p'(\theta)} \right) - 1 \right]\end{aligned}$$

If  $\theta = 1$ , this function equals to:

$$\frac{\ddot{S} + \gamma\dot{S}}{\gamma\dot{S}}|_{\theta=1} = \frac{\beta}{\gamma} \left[ \frac{G_p''(1)}{\delta} \left( 2 - \delta \times \frac{\beta - \gamma(1 - 1)}{\beta G_p'(1)} \right) - 1 \right] = \frac{\beta}{\gamma} \left[ \frac{G_p''(1)}{\delta} - 1 \right]$$

This amount equals to 1 iff  $\mathcal{R}_0^* = \frac{\beta}{\beta + \gamma} \frac{G_p''(1)}{\delta} = 1$ .

## Problem:

Unlike the  $\mathcal{R}_0^*$  or  $\mathcal{R}_{\text{eff}}^*$ ,  $\max(\mathcal{R}_{\text{eff}})|_{\theta=1}$  this amount could easily be larger than the maximum degree as there is no bond for the ratio  $\frac{\beta}{\gamma}$ .

- E.g. consider a network with every nodes has degree  $k = 5$ , then  $\delta = 5$  and  $G_p''(1) = k^2 - k = 20$ , assume  $\beta = 0.2$  and  $\gamma = 0.1$ , then

$$\max(\mathcal{R}_{\text{eff}})|_{\theta=1} = \frac{\beta}{\gamma} \left[ \frac{G_p''(1)}{\delta} - 1 \right] = \frac{0.2}{0.1} \left[ \frac{20}{5} - 1 \right] = 6 > k = 5$$

- We need a better definition for this!