

## **Epidemics in Competition**

Partial Cross-Immunity

Andreasen, Viggo

*Published in:*  
Bulletin of Mathematical Biology

*DOI:*  
[10.1007/s11538-018-0495-2](https://doi.org/10.1007/s11538-018-0495-2)

*Publication date:*  
2018

*Document Version*  
Other version

*Citation for published version (APA):*  
Andreasen, V. (2018). Epidemics in Competition: Partial Cross-Immunity. *Bulletin of Mathematical Biology*, 80(11), 2957-2977. <https://doi.org/10.1007/s11538-018-0495-2>

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

### **Take down policy**

If you believe that this document breaches copyright please contact [rucforsk@kb.dk](mailto:rucforsk@kb.dk) providing details, and we will remove access to the work immediately and investigate your claim.

# Epidemics in competition - partial cross-immunity

Viggo Andreasen

Received: date / Accepted: date

**Abstract** The competition between two pathogen strains during the course of an epidemic represents a fundamental step in the early evolution of emerging diseases as well as in the antigenic drift process of influenza. The outcome of the competition however, depends not only on the epidemic properties of the two strains but also on the timing and size of the introduction, characteristics that are poorly captured by deterministic mean-field epidemic models. We describe those aspects of the competition that can be determined from the mean-field models giving the range of possible final sizes of susceptible hosts and cumulated attack-rates that could be observed after an epidemic with two cross-reacting strains. In the limit where the size of the initial infection goes to zero the possible outcomes lie on a (one dimensional) curve in the outcome space.

**Keywords** Epidemic model · Final size · Competition · Cross-immunity

## 1 Introduction

The outcome of a single epidemic involving two competing types of a pathogen is key to our theoretical understanding of several important epidemiological questions. For an outbreak of an emerging disease, e.g. an influenza pandemic, the outcome of such a mixed epidemic characterizes the early stages of pathogen adaptation to the new host population (Antia et al, 2003; Holmes, 2009; Alexander and Day, 2010; Kubiak et al, 2010). In influenza drift a similar situation occurs in each season as the epidemic is seeded by a genetically and antigenically heterogeneous pathogen population or mutants arise during the early stages of the epidemic (Koelle et al, 2006;

---

Supported by the Independent Research Fund Denmark (grant 8020-00284)

Viggo Andreasen  
Department of Science, Roskilde University, DK-4000 Roskilde, Denmark  
Phone.: +45-46742542  
Fax: +45-46743020  
E-mail: viggo@ruc.dk

Boni et al, 2006; Lipsitch et al, 2007; Omori et al, 2010; Ho and Chao, 2017). For plant disease evolution the scenario describes the simultaneous spread of a wild-type and a mutant in an uninfected field (Ohtsuki and Sasaki, 2006) and co-existence of multiple strains of (bacculo) virus in annual insect populations is determined by similar epidemic interactions (Fleming-Davies et al, 2015). Furthermore the modeling structure is analogous to that describing spontaneous behavioral changes in response to a severe epidemic (Poletti et al, 2009; Funk et al, 2009; Kiss et al, 2010).

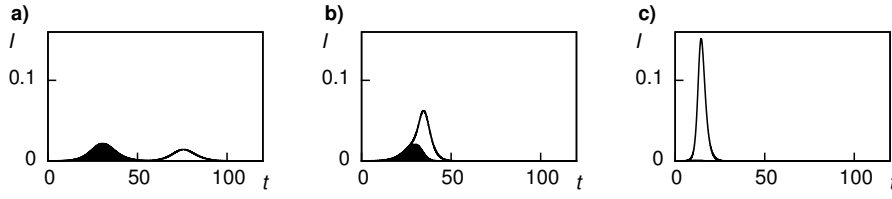
Pathogen competition has been studied extensively in the theoretical literature for the case where the host-pathogen interaction occurs at an endemic equilibrium i.e. in a situation where demographic turn-over and disease transmission have established an epidemiological equilibrium. Long term coexistence at the endemic equilibrium is not possible (Bremermann and Thieme, 1989) unless additional factors are included see for example Gjini et al (2016) and Lipsitch et al (2009). The present study differs in that we are concerned with competition during a single epidemic i.e. we study competition occurring in a host population who's immunity structure is dynamically changing in response to disease transmission and loss of immune hosts.

A previous study on competition during an epidemic by Saunders (1981) and Kendall and Saunders (1983) has focused on establishing the basic model structure and on characterizing the possible outcome of the epidemic in the case of full cross-immunity. Multi-strain epidemics have also been studied in the context of epidemics on networks (Funk and Jansen, 2010; Marceau et al, 2011; Karrer and Newman, 2011; Miller, 2013; Hebert-Dufresne et al, 2013) but their results should be compared to the stochastic formulation by Svensson and Scalia Tomba (2001).

The present paper will rely heavily on the methods of Kendall and Saunders and the expression "epidemics in competition" is a reference to the title of their original work. Specifically we extend their results to the case where the two strains confer only partial cross-immunity. Partial cross-immunity is particularly important for influenza drift, though similar issues arise when modeling the introduction of dengue and malaria to new areas. In addition, partial immunity allows us to model early stages of pathogen speciation as our formulation will span the full range from two identical pathogens to two immunologically separate species. The existence of cross-reactions among strains is well documented. However, the exact effect is often unclear and may vary from system to system (Ackerman et al, 1990; Gomes and Medley, 2001; Gog and Swinton, 2002). In this paper we focus on the case where infection confers complete immunity to the infecting strain while cross-immunity acts by reducing infectivity during a subsequent infection with the other strain. Our main motivation for this choice is mathematical, as it is well known that this formulation is particularly amenable for analytic investigation (Gupta et al, 1998; Ferguson and Andreasen, 2001).

In its simplest form – and omitting for now the possibility of partial cross-immunity – the problem may be described as follows: Imagine that a pathogen is introduced into a susceptible host population and that the pathogen occurs in two variants which differ in their transmission properties. What will be the outcome of such an epidemic in terms of the size of the epidemic and its distribution among the two types?

Clearly the answer depends on the exact way in which the two pathogens are introduced. Figure 1 shows the range of possible outcomes and in particular Figure 1a



**Fig. 1** Three characteristic epidemics that can arise for a system of two competing viral strains conferring complete cross protection in a homogeneously mixing host population. White area indicates prevalence of the most transmissible type ( $\mathcal{R}_0 = 2$ ); black area prevalence of least transmissible type ( $\mathcal{R}_0 = 1.25$ ). The three epidemic curves differ only in initial conditions.

and c show the extreme cases where one pathogen is introduced sufficiently early that it has caused and completed its epidemic prior to the arrival of the other type, while Figure 1b shows an intermediate situation where the two epidemics overlap. The extreme cases are easy to characterize analytically as they consist of a pair of *sequential epidemics*, each involving only a single pathogen type (Andreasen and Sasaki, 2006; Newman, 2005). However, for the overlapping epidemics on Figure 1b the outcome depends on the absolute size and composition of the initial introductions, suggesting that it would be necessary to include the stochastic effects of small pathogen population size during the early phases of the epidemic to obtain a complete characterization (Svensson and Scalia Tomba, 2001).

Here we will take another approach focusing on those aspects of the epidemic interaction that may be determined deterministically. This implies that we will not aim at providing a unique answer but rather provide a range of possible outcomes.

Our starting point is the two-strain epidemic with full cross-protection as this case leads to the least complicated model while still allowing us to present the basic model structure in Section 2, and to develop the necessary mathematical methods in Section 3. In Section 4 we extend the analysis to the case where partial cross-immunity acts by reducing infectivity for infected hosts who are already immune to the other strain and in Section 5 we show how the strength of cross-immunity can span the full range of epidemics in competition.

## 2 Two pathogen types and full cross-immunity

For a homogeneously mixing host population, the state of the epidemic may be captured by specifying the fraction of the host population that is susceptible  $S$  and the fractions that are currently infected by each pathogen type  $I, Y$ . Neglecting the possibility of co-infection and assuming an exponential infectious period with recovery rates  $\nu_K$  and transmission coefficients  $\beta_K, K = I, Y$ , the transmission dynamics are described by

$$\dot{S} = -\beta_I SI - \beta_Y SY \quad (1)$$

$$\dot{I} = \beta_I SI - \nu_I I \quad (2)$$

$$\dot{Y} = \beta_Y SY - \nu_Y Y. \quad (3)$$

We will assume that the disease is non-lethal such that the population size remains constant and omit explicit reference to the dynamics of the recovered class as its size is redundant.

In the next section we will show that for any non-negative initial condition  $(s, i, y)$  there exists a number  $s_\infty$  such that  $S(t) \rightarrow s_\infty$  and that  $I(t), Y(t) \rightarrow 0$  for  $t \rightarrow \infty$ . It is convenient to express the outcome in terms of the fraction of the susceptible population that remains uninfected at the end of the epidemic  $\sigma = s_\infty/s$  and the fraction of hosts that became infected by either strain during the entire course of the epidemic  $A_I = v_I \int_0^\infty I dt$  with a similar definition of  $A_Y$ .

We notice that the quantities

$$\dot{S}/S = -\beta_I I - \beta_Y Y \quad (4)$$

$$\dot{S} + \dot{I} + \dot{Y} = -v_I I - v_Y Y \quad (5)$$

are linear expressions in  $I$  and  $Y$ . We will refer to quantities with this property as (generalized) *first integrals*.

First integrals are useful because we can obtain information about  $\sigma$ ,  $A_I$ , and  $A_Y$  by integration of (4–5) over the entire epidemic period:

$$\begin{aligned} \int_0^\infty \dot{S}/S dt &= -\beta_I \int_0^\infty I dt - \beta_Y \int_0^\infty Y dt \\ \int_0^\infty \dot{S} + \dot{I} + \dot{Y} dt &= -v_I \int_0^\infty I dt - v_Y \int_0^\infty Y dt, \end{aligned}$$

which yields two conditions in the three unknowns  $(\sigma, A_I, A_Y)$

$$-\log \sigma = \mathfrak{R}_0^I A_I + \mathfrak{R}_0^Y A_Y \quad (6)$$

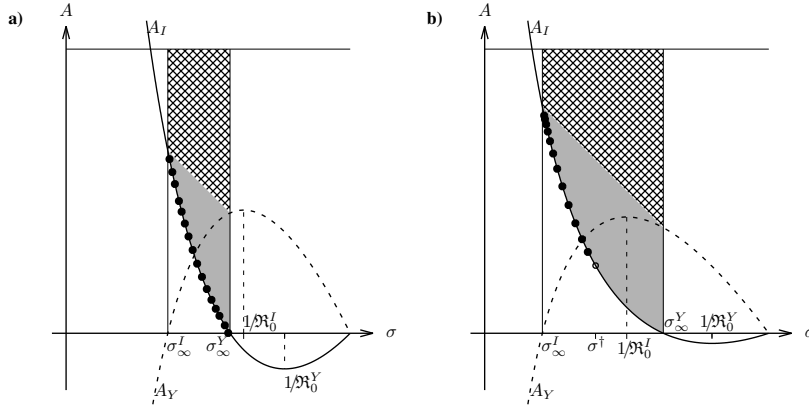
$$s(1 - \sigma) + i + y = A_I + A_Y. \quad (7)$$

It is remarkable that the range of outcomes from the competing epidemics depends only on the reproduction numbers  $\mathfrak{R}_0^K = \frac{\beta_K}{v_K}$  and not explicitly on the growth rates as is the case for epidemics on networks (Karrer and Newman, 2011).

The condition (7) may be seen as a simple balance equation dividing the total population between those still susceptible and those who were infected with either pathogen type at some point during the epidemic. The first condition (6) generalizes the final size equation to a mixed epidemic.

To determine the three quantities  $(\sigma, A_I, A_Y)$  one would need one additional condition. However this third condition will depend on the size and composition of the pathogen introduction and we will not explore this issue here. We first discuss what information can be obtained from the two conditions (6–7). Due to the invariance of the first octant we observe that every trajectory of the differential equation (1–3) with initial conditions of the form  $s, i, y \geq 0$  and  $s + i + y \leq 1$  will satisfy (6–7). It is however not clear if every positive point  $(\sigma, A_I, A_Y)$  satisfying (6–7) corresponds to a solution of the differential equation.

Our primary interest is the condition during a pandemic arising from a newly introduced pathogen so for now we focus on the situation where  $s \approx 1$  and  $0 < i, y \ll 1$ . Without loss of generality we may also assume that  $\mathfrak{R}_0^I > \mathfrak{R}_0^Y$ .



**Fig. 2** Graphical solution of the final size problem for an epidemic with two types of the disease. The fraction of hosts that are still susceptible at the end of the epidemic is denoted  $\sigma$ . The value of  $\sigma$  depends on initial conditions and thus each value of  $\sigma$  corresponds to a possible outcome of the epidemic. A vertical line divides the unit interval into three segments representing (from bottom to top) the fraction of the population infected with type  $I$ , those infected with type  $Y$  and the fraction of the host population that remains uninfected during the entire epidemic. Solid points represent the outcomes of numerical solutions to the model (1–3) for initial conditions of the form  $I(0) = 0.00001/N, Y(0) = 0.99998/N$  with varying  $N$ . Panel a) shows the case where sequential epidemics are not possible since  $\sigma_\infty^Y < \mathfrak{R}_0^I$ . Parameter values are  $\mathfrak{R}_0^I = 1.6, \mathfrak{R}_0^Y = 1.3$ , and  $v_I = v_Y = 1$ . Panel b) shows the case where sequential epidemics may occur. The open circle represents the sequential epidemics where the inferior type has completed its epidemic prior to the arrival of the superior competitor. Sequential epidemics can arise only if  $1/\mathfrak{R}_0^I < \sigma_\infty^Y$ . Parameter values are  $\mathfrak{R}_0^I = 2.0, \mathfrak{R}_0^Y = 1.25$ , and  $v_I = v_Y = 1$ .

Considering the fraction of the population that remains susceptible  $\sigma$  as a parameter we can determine the fraction that become infected with each of the pathogen types as

$$A_I(\sigma) = \frac{1}{\mathfrak{R}_0^I - \mathfrak{R}_0^Y} (\mathfrak{R}_0^Y(1 - \sigma) + \log \sigma) \quad (8)$$

$$A_Y(\sigma) = \frac{-1}{\mathfrak{R}_0^I - \mathfrak{R}_0^Y} (\mathfrak{R}_0^I(1 - \sigma) + \log \sigma) \quad (9)$$

The two curves are shown on Figure 2. Remarkably the magnitudes of  $A_I$  and  $A_Y$  are proportional to the final size expression  $f_L : \sigma \mapsto \mathfrak{R}_0^L(1 - \sigma) + \log \sigma$  for the other type  $L = Y, I$ . The graph for  $f_I$  may be seen from the curve marked  $A_Y$  at Figure 2 and we have  $f_K(0^+) = -\infty, f_K(1) = 0$ , and  $f_K'(1) = 1 - \mathfrak{R}_0^K$  so that for  $\mathfrak{R}_0^K > 1$ ,  $f_K$  has a unique root  $\sigma_\infty^K$  in the interval  $(0, 1)$  giving the fraction of hosts that are susceptible at the end of an epidemic consisting entirely of type  $K$  pathogens. We observe that  $A_I$  is positive for  $\sigma < \sigma_\infty^Y$ . Similarly  $A_Y$  is positive for  $\sigma > \sigma_\infty^I$  and hence feasible solutions occur only in the interval  $\sigma_\infty^I \leq \sigma \leq \sigma_\infty^Y$ . Figure 2b suggest that not all feasible solutions may arise as outcome of the epidemic. We shall discuss this issue in the next section.

**Remark.** In this section we have characterized the epidemic in terms of the cumulated attack-rate  $A_I$  as this allows for the simple interpretation of Figure 2. For the

more complex situation in Section 4, the cumulated force of infection  $\alpha = \int \Lambda dt = \int \beta I dt$  – or alternatively the probability of remaining uninfected  $\pi = \exp(-\int \Lambda dt)$  gives a more compact description.

### 3 Elementary properties of the model

From a mathematical view point model (1–3) is somewhat awkward as it has a continuum of equilibria where every point of the form  $(s, 0, 0)$  is a stationary point. Thus the model is not "generic" and results from dynamical systems theory are of limited use. For the classical model of an epidemic with a single infectious agent, the model is typically solved by explicitly determining the solution curves using  $S$  as the independent variable. A similar approach is not available for our model, so instead we will in this section use elementary methods to study the map  $\Phi$  that connects to every possible initial condition the corresponding final outcome of the epidemic. The results are essentially reformulations of the statements in (Kendall and Saunders, 1983, Theorem 1).

Although our main interest is the limiting case where the initial infectious seed is small  $0 < I(0), Y(0) \approx 0$  we first expand our problem allowing for any initial condition where no hosts are immune at the onset of the epidemic. This allows us to separate limiting processes in the time domain from limiting processes in the initial conditions.

If the entire population is either susceptible or infected at the onset of the epidemic, the initial condition  $(S(0), I(0), Y(0)) = (s, i, y)$  satisfies  $s + i + y = 1$  so that the set of initial conditions may be described by the 2-D simplex

$$\Gamma = \{ (i, y) \mid i, y \geq 0, 1 \geq i + y \}.$$

For the methods we develop in this section it is natural to express the final size in terms of  $s_\infty$  – the actual size of the susceptible population – rather than to measure the final size relative to the size of the susceptible population at epidemic onset  $\sigma = s_\infty/s$  as we do in Sections 2 and 4, but the conversion between  $\sigma$  and  $s_\infty$  is straight forward.

We first prove (the expected fact) that every trajectory goes to a limit point as time goes to infinity.

**Theorem 1** *For every initial condition in  $\Gamma$  there exists a number  $s_\infty$  so that the corresponding solution  $(S(t), I(t), Y(t))$  to model (1–3) satisfies*

$$\lim_{t \rightarrow \infty} S(t) = s_\infty \quad \lim_{t \rightarrow \infty} I(t) = \lim_{t \rightarrow \infty} Y(t) = 0.$$

The theorem and its proof are generalizations of results by Brauer (2008). The proof builds on the following elementary

**Lemma 1 (Brauer, 2008)** *If  $f(t)$  is a nonnegative monotone non-increasing continuously differentiable function, then there exists  $f_\infty \geq 0$  such that as  $t \rightarrow \infty$ ,  $f(t) \rightarrow f_\infty$ , and in addition  $f'(t) \rightarrow 0$ .*

To see that Theorem 1 holds, observe that by the invariance of the planes  $I = 0$  and  $Y = 0$ , we have that  $I, Y \geq 0$  and hence that  $S$  is a non-increasing function. Since  $(S + I + Y)' = -(v_I I + v_Y Y)$ , we conclude that if  $I(0) + Y(0) > 0$  then  $S + I + Y$  is decreasing and hence there exists  $s_\infty \geq 0$  such that  $(S(t) + I(t) + Y(t)) \rightarrow s_\infty$  as  $t \rightarrow \infty$ . By Brauer's Lemma we know that  $v_I I + v_Y Y \rightarrow 0$  for  $t \rightarrow \infty$ . Hence by the positivity of  $I$  and  $Y$  we have that  $I(t), Y(t) \rightarrow 0$  for  $t \rightarrow \infty$  and thus  $S(t) \rightarrow s_\infty$  for  $t \rightarrow \infty$ .

Theorem 1 shows that we can define a map from the initial conditions to the corresponding limit point of the solution to (1–3)

$$\Phi : (i, y) \mapsto s_\infty$$

as a map of the form

$$\Phi : \Gamma \rightarrow [0, 1]$$

Keeping in mind that by assumption  $\mathfrak{R}_0^I > \mathfrak{R}_0^Y$ , we can now state our main result about this map.

**Theorem 2 (Kendall and Saunders, 1983)** *The map  $\Phi : \Gamma \mapsto [0, 1]$  has the following properties*

1. *If  $\Phi(i, y) < 1/\mathfrak{R}_0^I$  then  $\Phi$  is continuous at  $(i, y)$ .*
2. *If  $\Phi(i, y) > 1/\mathfrak{R}_0^I$  then  $i = 0$ .*
3. *Assume that  $\Phi(0, y_0) > 1/\mathfrak{R}_0^I$  and let  $\Psi$  denote the unique solution to the equation*

$$\log \Psi + \mathfrak{R}_0^I \Phi(0, y_0)(1 - \Psi) = 0, \quad (10)$$

*then*

$$\Phi(i, y) \rightarrow \Psi \text{ for } (i, y) \rightarrow (0, y_0) \text{ and } (i, y) \in \Gamma_+,$$

*where  $\Gamma_+$  denotes the interior of  $\Gamma$ .*

**Remark.** Properties 2 and 3 characterize sequential epidemics. Property 2 states that after the conclusion of an epidemic, the population can only support an additional epidemic if the first epidemic did not involve the superior competitor and Property 3 then states that if in addition the inferior competitor ( $Y$ ) caused an epidemic sufficiently small such that the effective reproduction number of the superior competitor  $\Phi(0, y)\mathfrak{R}_0^I$  exceeds unity, then the superior strain ( $I$ ) will produce a sequential epidemic if it is present ( $i > 0$ ).

Property 3 also shows that in this situation, the map  $\Phi$  is discontinuous at  $(0, y_0)$ . The limit  $\Psi$  defined by (10) gives the final size - in terms of the fraction still susceptible after a double epidemic where the inferior strain ( $Y$ ) has completed its epidemic before the superior strain ( $I$ ) gives rise to an epidemic among those hosts that were still susceptible.

*Proof of theorem 2:*

The basic idea is to break  $\Phi$  into a composite map where the first map takes  $(i, y)$  into the value of the solution at a (suitably chosen) fixed time  $T$ . This map is continuous in general (Coddington and Levinson, 1955). The second map takes  $(S(T), I(T), Y(T))$  into  $(S(\infty), I(\infty), Y(\infty))$ . For the selected  $T$  we show that this map



is bounded by suitable exponential decays toward the limit point and use this observation to establish continuity. The detailed proofs may be found in Kendall and Saunders (1983), we will only sketch how to prove claim 1.

*Claim 1.* Consider a point  $(i_0, y_0) \in \Gamma$  with  $\Phi(i_0, y_0) < 1/\mathfrak{R}_0^I$  and let  $X_0(t) = (S_0(t), I_0(t), Y_0(t))$  denote the solution to (1–3) with initial conditions  $(i_0, y_0)$ . Let  $X_0^* = (\Phi(i_0, y_0), 0, 0)$  be the limit of  $X_0(t)$  as  $t \rightarrow \infty$ . Set  $\kappa = 1/\mathfrak{R}_0^I - \Phi(i_0, y_0) > 0$ .

For a given  $1 > \varepsilon > 0$  set  $\varepsilon_1 = \varepsilon\kappa/4$  and observe that there exists a  $T$  such that  $\|X_0(t) - X_0^*\| < \varepsilon_1$  for all  $t \geq T$  where  $\|\cdot\|$  denotes the 2 or 3 dimensional Euclidean norm as appropriate. Since the solution of (1–3) at the finite time  $T$  depends continuously on initial conditions (Coddington and Levinson, 1955) we can find  $\delta > 0$  such that

$$\|(i, y) - (i_0, y_0)\| < \delta \Rightarrow \|X(T) - X_0(T)\| < \varepsilon_1$$

for the solution  $X(t)$  with initial conditions  $(i, y)$ . We need to show that  $\|X^* - X_0^*\| < \varepsilon$ .

We first observe that by construction,  $S(T) \leq \Phi(i_0, y_0) + \kappa/2$ , such that  $1/\mathfrak{R}_0^I - S(T) > 1/\mathfrak{R}_0^I - (\Phi(i_0, y_0) + \kappa/2) > \kappa/2$ . Since  $S(t)$  is decreasing, the inequality holds for all  $t \geq T$ . This result provides an estimate for  $I(t)$  since  $I' = \beta_I(S - 1/\mathfrak{R}_0^I)I \leq -\frac{1}{2}\beta_I\kappa I$ , for  $t \geq T$ . We deduce that for  $t \geq T$  we have

$$I(t) \leq I(T) \exp(-\frac{1}{2}\beta_I\kappa t),$$

with a similar expression for  $Y(t)$ . This provide us with suitable bound on how fast  $I$  and  $Y$  decay and thus a bound for the decay of  $S(t) - S(\infty)$ .

We can now study  $S(T) - S(\infty)$  and first observe that

$$\begin{aligned} \log S(\infty)/S(T) &= - \int_T^\infty \beta_I I(\tau) + \beta_Y Y(\tau) d\tau \\ &\geq - \int_T^\infty \beta_I I(T) e^{-\frac{1}{2}\beta_I\kappa\tau} + \beta_Y Y(T) e^{-\frac{1}{2}\beta_Y\kappa\tau} d\tau \\ &= -\frac{2}{\kappa}(I(T) + Y(T)). \end{aligned}$$

A simple rearrangement yields

$$\begin{aligned} 0 < S(T) - S(\infty) &\leq \left(1 - \exp(-\frac{2}{\kappa}((I(T) + Y(T))))\right) S(T) \\ &\leq \frac{2}{\kappa}((I(T) + Y(T))S(T) \leq \frac{4}{\kappa}\varepsilon_1 \leq \varepsilon. \end{aligned}$$

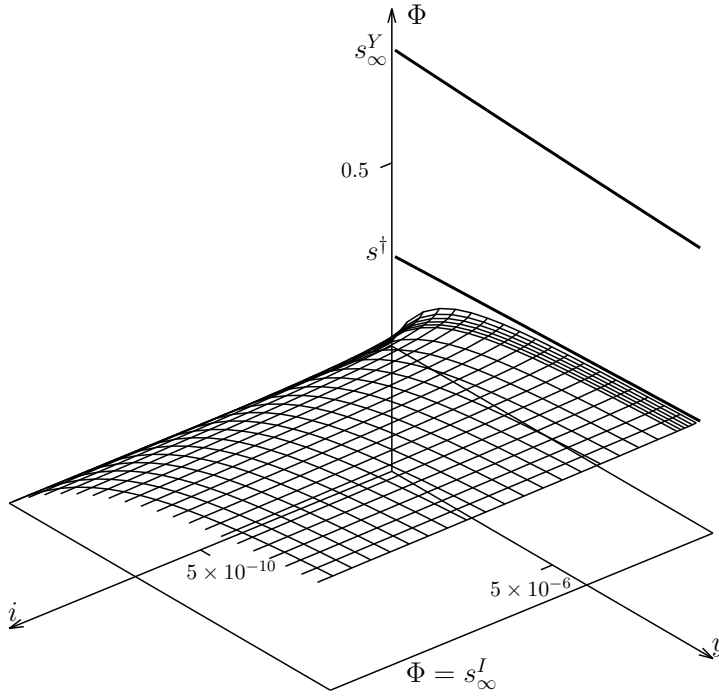
Finally we notice that

$$\|X^* - X_0^*\| \leq \|X^* - X(T)\| + \|X(T) - X_0(T)\| + \|X_0(T) - X_0^*\|,$$

and that we have shown that each term is smaller than  $\varepsilon$ .

This concludes our discussion of the proof of Theorem 2.

To finish our characterization of the possible outcomes of the mixed epidemic for the case where  $i, y \ll 1$ , we combine Theorem 2 with our observations in Section



**Fig. 3** The map  $\Phi$  for values of  $(i, y)$  close to the origin. Along the  $i$ -axis  $\Phi(i, 0) \rightarrow s_\infty^I$  - the size of an epidemic consisting only of type Y pathogens. On the  $y$ -axis  $\Phi(0, y) \rightarrow s_\infty^Y$  - the size of an epidemic consisting only of type I. If sequential epidemics are possible - that is if  $\mathfrak{R}_0^I s_\infty^Y > 1$ , the value of double limit is  $\lim_{y \rightarrow 0^+} \lim_{i \rightarrow 0^+} \Phi(i, y) = s^\dagger$  - the final size of the susceptible population in an  $i$ -epidemic that occurs in a host population that has already been exposed to a  $y$ -epidemic (case shown). If sequential epidemics can not occur (case not shown) we have  $\lim_{y \rightarrow 0^+} \lim_{i \rightarrow 0^+} \Phi(i, y) = s_\infty^Y$ . In both case the map  $\Phi$  is discontinuous at  $(0, 0)$ . Parameter values used  $\mathfrak{R}_0^I = 2$ ,  $\mathfrak{R}_0^Y = 1.2$

2. Figure 3 shows the nature of  $\Phi$  in a neighborhood of origin and in particular we notice that

$$\begin{aligned} \lim_{i \rightarrow 0^+} \Phi(i, 0) &= s_\infty^I \\ \lim_{y \rightarrow 0^+} \Phi(0, y) &= s_\infty^Y \\ \lim_{y \rightarrow 0^+} \lim_{i \rightarrow 0^+} \Phi(i, y) &= s^\dagger \end{aligned}$$

where  $s^\dagger$  is the final size of the susceptible population after sequential epidemics if  $\mathfrak{R}_0^I s_\infty^Y > 1$ , and  $s^\dagger = s_\infty^Y$  if sequential epidemics are not possible. Thus the map  $\Phi$  is discontinuous at the origin - and if sequential epidemics are possible, it is in addition also discontinuous along the  $y$ -axis. Therefore every neighborhood of the origin will contain 1) a point that is mapped into a value close to  $s_\infty^I$  and 2) a point that is mapped into a value close to  $s^\dagger$  and 3) a point that is mapped into a value close to  $s_\infty^Y$  - where the two latter points may be identical in the case of no sequential epidemics. Since

$\Phi$  is continuous in the interior of  $\Gamma$  and  $\Gamma$  is connected, we can find points that are mapped into every value in the interval  $[s_\infty^I, s^\dagger)$  while  $s_\infty^Y$  may be an isolated point in the image when sequential epidemics can occur. This result explains the range of outcomes that is indicated in Figure 2b.

Although it seems obvious, it is hard to prove that  $\Phi$  does not have extreme values outside the interval  $[s_\infty^I, s^\dagger)$ . Kendall and Saunders (1983) provide a quite complicated proof of the claim - we will return to the issue in the discussion.

We conclude that it is not possible to uniquely determine a final size of the susceptible population in a mixed epidemic unless one makes further assumptions about the size and composition of the initial infection. We did, however, show that the final size of the susceptible population falls in the set  $[s_\infty^I, s^\dagger) \cup \{s_\infty^Y\}$ . Once  $s_\infty$  is determined the corresponding attack rates for the two pathogens can be found using Equations (6–7).

#### 4 Partial cross-immunity

A somewhat more complicated situation arises if infection with one strain confers only partial protection against infection with the other strain. Here we assume that cross-immunity acts by reducing infectivity by a factor  $\tau$ ,  $0 < \tau < 1$  during the second infection. Furthermore we will assume that the two infections are independent in the sense that a host can become infected by the second strain while still being infected with the first strain, and that recovery from the two infections is independent.

Let  $I_0$  denote the fraction of hosts that are currently infected with strain  $I$  and who have no immunity to strain  $Y$  while  $I_1$  denotes the fraction of hosts who are infected with strain  $I$  and who are immune to strain  $Y$  at the time of infection. With our assumptions the force of infection of strain  $I$  is

$$\Lambda_I = \beta_I(I_0 + \tau_Y I_1), \quad (11)$$

with similar expressions for strain  $Y$ .

Let  $R_K$ ,  $K = I, Y$  denote the fraction of hosts that are or have been infected with strain  $K$  and who have not yet been infected with strain  $L = Y, I$ . Since we assume that hosts are susceptible to a second infection immediately after acquiring their first infection, hosts infected with strain  $I$  flow directly from the susceptible class  $S$  into the  $R_I$ -class and we have

$$\dot{S} = -(\Lambda_I + \Lambda_Y)S \quad (12)$$

$$\dot{R}_K = \Lambda_K S - \Lambda_L R_K, \quad \text{for } K = I, Y \quad L = Y, I. \quad (13)$$

In addition we can keep track of infected hosts and their immune status at the time of infection

$$\dot{K}_0 = \Lambda_K K_0 - \nu_K K_0 \quad \text{for } K = I, Y$$

$$\dot{K}_1 = \Lambda_K K_1 - \nu_K K_1 \quad \text{for } K = I, Y$$

which – using equation (11) – simplifies to

$$\dot{\Lambda}_K = \beta_K \Lambda_K (S + \tau_L) - \nu_K \Lambda_K. \quad \text{for } K = I, Y \quad L = Y, I. \quad (14)$$

Equations (12,13,14) give a total of 5 equations yielding a closed system, describing the transmission dynamics of the two strains.

In addition we need to specify initial conditions  $S(0) = s, R_I(0) = r_I, R_Y(0) = r_Y$ . Here  $0 \leq s + r_I + r_Y \leq 1$  and  $0 < \Lambda_K(0) = l_K$  as in the previous section.

In order to obtain simple first integrals we change to variables that account for the fraction of totally susceptible individuals  $S$  and the fraction of hosts susceptible to each of the two strains  $P_K = S + R_L$ . In this representation the model is

$$\begin{aligned} \dot{S} &= -(\Lambda_I + \Lambda_Y)S \\ \dot{P}_K &= -\Lambda_K P_K \quad \text{for } K = I, Y \quad L = Y, I \\ \dot{\Lambda}_K &= \beta_K((1 - \tau_L)S + \tau_L P_K) \Lambda_K - \nu_K \Lambda_K \end{aligned} \quad (15)$$

with initial conditions  $(s, p_K, l_K)$  satisfying  $0 \leq p_I + p_Y - s \leq 1$ ,  $0 < s \leq p_K \leq 1$  and  $0 \leq l_K$ . An extension of Theorem 1 shows that the trajectories all have  $\omega$ -limits of the form  $S(\infty) = s_\infty > 0$ ,  $P_K(\infty) = p_K^\infty \geq 0$  and  $\Lambda_K(\infty) = 0$ .

One may now look for first integrals as in the previous sections. There are three trivial integrals:  $\dot{S}/S$ ,  $\dot{P}_I/P_I$ ,  $\dot{P}_Y/P_Y$ . We next look for a suitable linear combination of the five dynamical equations that yields a generalization of the balance equation (7). To find such a combination, we seek a linear combination of the five dynamical equations such that the four quadratic terms  $\Lambda_K S$  and  $\Lambda_K P_K$  vanish. Thus the coefficients of the quadratic terms must satisfy the (under-determined) homogeneous linear equations

$$\begin{array}{ccccc} & \dot{S} & \dot{P}_I & \dot{P}_Y & \dot{\Lambda}_I & \dot{\Lambda}_Y \\ \Lambda_I S & \begin{pmatrix} -1 & 0 & 0 & \beta_I(1 - \tau_Y) & 0 \\ -1 & 0 & 0 & 0 & \beta_Y(1 - \tau_I) \\ 0 & -1 & 0 & \beta_I \tau_Y & 0 \\ 0 & 0 & -1 & 0 & \beta_Y \tau_I \end{pmatrix} & \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_5 \end{pmatrix} & = & \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ \Lambda_Y S & & & & \\ \Lambda_I P_I & & & & \\ \Lambda_Y P_Y & & & & \end{array}$$

where a row corresponds to a quadratic term and a column to a dynamical equation. Since the coefficient matrix has rank 4, the solution

$$\left( (1 - \tau_I)(1 - \tau_Y), \tau_Y(1 - \tau_I), \tau_I(1 - \tau_Y), \frac{1 - \tau_I}{\beta_I}, \frac{1 - \tau_Y}{\beta_Y} \right)^T$$

is unique up to a scaling factor. We conclude that

$$\begin{aligned} (1 - \tau_I)(1 - \tau_Y)\dot{S} + \tau_Y(1 - \tau_I)\dot{P}_I + \tau_I(1 - \tau_Y)\dot{P}_Y + \frac{1 - \tau_I}{\beta_I}\dot{\Lambda}_I + \frac{1 - \tau_Y}{\beta_Y}\dot{\Lambda}_Y = \\ - \frac{\nu_I(1 - \tau_I)}{\beta_I}\Lambda_I - \frac{\nu_Y(1 - \tau_Y)}{\beta_Y}\Lambda_Y, \end{aligned} \quad (16)$$

which gives us our fourth and last first integral.

We can now formulate a necessary condition that must be satisfied by the outcome of the mixed epidemic in terms of  $\sigma = s_\infty/s$  (the fraction of those hosts that were fully susceptible at onset of the epidemic who are still fully susceptible at the end of the epidemic)  $\pi_K = P_K(\infty)/p_K$ , (with similar interpretation) and  $\alpha_K = \int_0^\infty \Lambda_K dt$ , (the cumulated force of infection). We find that the outcome of the epidemic must satisfy

$$\begin{aligned} \log \sigma &= -(\alpha_I + \alpha_Y) \\ \log \pi_K &= -\alpha_K \quad K = I, Y \\ 0 &= s(1 - \tau_I)(1 - \tau_Y)(1 - \sigma) + p_I \tau_Y(1 - \tau_I)(1 - \pi_I) \\ &\quad + p_Y \tau_I(1 - \tau_Y)(1 - \pi_Y) - \frac{v_I(1 - \tau_I)}{\beta_I} \alpha_I - \frac{v_Y(1 - \tau_Y)}{\beta_Y} \alpha_Y \\ &\quad - \frac{1 - \tau_I}{\beta_I} (\Lambda_I(\infty) - l_I) - \frac{1 - \tau_Y}{\beta_Y} (\Lambda_Y(\infty) - l_Y) \end{aligned} \quad (17)$$

Equations (17) give us four equations in the five unknowns  $(\sigma, \pi_I, \pi_Y, \alpha_I, \alpha_Y)$  and we conclude that the outcome of the competing epidemics is determined up to one parameter, the value of which will depend on the details of the initial conditions — as in Section 2.

To further simplify the discussion, we assume that the initial infection is small  $0 \leq l_k \ll 1$  such that the last line of 17 vanishes. Substituting  $\sigma = \pi_I \pi_Y$  and  $\alpha_K = -\log \pi_K$  into the last equation in (17) we obtain after a suitable rearrangement, the condition

$$\begin{aligned} F(\pi_I, \pi_Y) &= -(1 - \tau_I)(1 - \tau_Y)(1 - \pi_I)(1 - \pi_Y) \\ &\quad + \frac{1 - \tau_I}{s\mathfrak{R}_0^I} (\mathfrak{R}_e^I(1 - \pi_I) + \log \pi_I) + \frac{1 - \tau_Y}{s\mathfrak{R}_0^Y} (\mathfrak{R}_e^Y(1 - \pi_Y) + \log \pi_Y) = 0, \end{aligned} \quad (18)$$

where  $\mathfrak{R}_e^K = \mathfrak{R}_0^K((1 - \tau_L)s + \tau_L p_K)$  is the effective reproduction number of the strain considering the preexisting immunity in the population. Equation (18) gives one equation in the two unknowns  $(\pi_I, \pi_Y)$  and the corresponding curve in  $(\pi_I, \pi_Y)$ -space gives us the range of possible outcomes of the competing epidemics in the sense that once  $(\pi_I, \pi_Y)$  is determined,  $\sigma$  and  $(\alpha_I, \alpha_Y)$  may be found from (17).

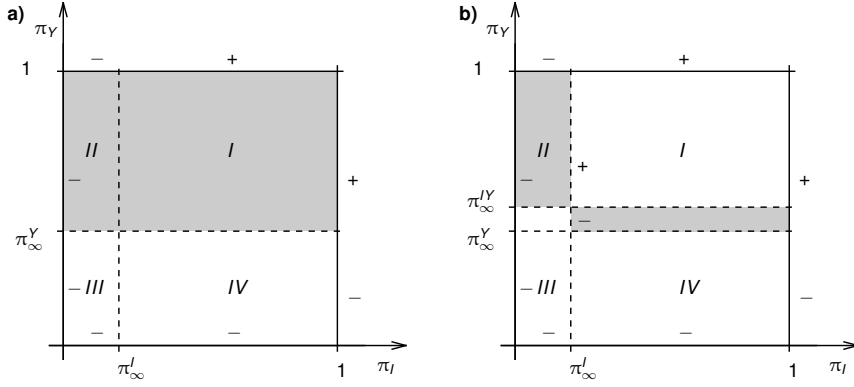
Thus to characterize the possible outcomes of the competing epidemics it suffices to describe the solutions to Equation (18). We will assume that both strains can grow at the onset of the epidemic, so that  $\mathfrak{R}_e^K > 1$  for both strains. In addition we must have that  $0 \leq \pi_K \leq 1$  to ensure that the solution is biologically meaningful.

The structure of Equation (18) allows for an elementary analysis of its solutions as shown in Figure 4. Let  $\pi_\infty^K$  denote the non-trivial solution to the one-strain final size of the susceptible population i.e.  $0 < \pi_\infty^K < 1$  is the (unique) solution to

$$\mathfrak{R}_e^K(1 - \pi) + \log \pi = 0.$$

Clearly the equation  $F(\pi_I, \pi_Y) = 0$  has three simple solutions:

$$(\pi_\infty^I, 1) \quad (1, \pi_\infty^Y) \quad (1, 1).$$



**Fig. 4** Geometric analysis of the number of solutions to  $F(\pi_I, \pi_Y) = 0$ , Equation (18). Since  $F$  is concave for fixed  $\pi_K$ , the equation has at most two solutions along a horizontal line (or along a vertical line). **a.** (Sequential epidemics do not occur.) The sign of  $F$  at the boundary of the unit-square gives the number of solutions in the different regions. In the grey region there is exactly one solution along each horizontal line; in the white area there is between 0 and 2 solutions. The points  $\pi_\infty^K$ ,  $K = I, Y$  denote the probability of remaining uninfected to strain  $K$  at the end of an epidemic consisting only of strain  $K$ . **b.** (A sequential epidemic with a  $Y$  epidemic followed by an  $I$  epidemic may occur.) The sign variation along the lines  $\pi_I = \pi_\infty^I$  depends on the existence and magnitude of a  $IY$ -sequential epidemic  $\pi_\infty^{IY}$ . In the grey areas of region  $I + II$  there is one solution along a horizontal line – in the white areas of  $I + II$  there is none. The same analysis may be applied to vertical lines.

The lines  $\pi_K = \pi_\infty^K$ ,  $K = I, Y$  divide the unit square into four regions as indicated on Figure 4. Along the line  $\pi_I = 1$ ,  $F$  changes sign as indicated on the Figure 4, while  $F(x, y) \rightarrow -\infty$  for  $x \rightarrow 0^+$  ensuring that  $F$  is negative close to the axis. Since  $\partial^2 F / \partial \pi_I^2 = -\Re_0^I / \pi_I^2 < 0$ , the function  $\pi_I \mapsto F(\pi_I, \pi_Y)$  is concave and along a horizontal line we have that

**Lemma 2** For fixed  $\pi_Y^* \in (\pi_\infty^Y, 1)$ , Equation (18) has exactly one solution  $\pi_I \in (0, 1)$ .

For fixed  $\pi_Y^* \in (0, \pi_\infty^Y)$ , Equation (18) has 0, 1 or 2 solutions. The case of one solution occurs iff  $\partial F / \partial \pi_I$  vanishes at the solution point.

To determine more precisely the location of the solutions in the interval  $(\pi_\infty^Y, 1)$ , consider the function

$$\begin{aligned} \pi_Y &\mapsto F(\pi_\infty^I, \pi_Y) \\ &= \frac{1 - \tau_Y}{s\Re_0^Y} ((-(1 - \tau_I)(1 - \pi_\infty^I)s\Re_0^Y + \Re_e^Y)(1 - \pi_Y) + \log \pi_Y) \\ &= \frac{1 - \tau_Y}{s\Re_0^Y} (\Re_e^{IY}(1 - \pi_Y) + \log \pi_Y), \end{aligned} \quad (19)$$

where

$$\Re_e^{IY} = \Re_0^Y ((1 - \tau_I)\pi_\infty^I s + \tau_I p_Y) \quad (20)$$

is the effective sequential reproduction number for strain  $Y$  if the population has already experienced the full course of an epidemic with strain  $I$ . If  $\Re_e^{IY} > 1$ , an  $IY$  sequential epidemic is possible and  $F = 0$  has a solution of the form  $(\pi_\infty^I, \pi_\infty^{IY})$  where  $\pi_\infty^{IY}$  solves the final size equation  $F(\pi_\infty^I, y) = 0$ .

If  $\mathfrak{R}_e^{IY} < 1$  then  $F$  is negative on the line segment  $(\pi_\infty^I, y)$ . Looking along a horizontal line segment,  $y \in (0, 1)$  there is one solution in region  $I$  and none in region  $II$ .

If  $\mathfrak{R}_e^{IY} > 1$ , then  $F$  is positive on the line segment  $(\pi_\infty^I, y)$ ,  $y \in (1, \pi_\infty^{IY})$  and for each fixed  $\pi_Y$  in this interval,  $F = 0$  has exactly one solution in region  $II$  and none in region  $I$ . For  $y < \pi_\infty^{IY}$ ,  $F$  is negative on the line  $\pi_I = \pi_\infty^I$  and the location over the solution is reversed.

We can now use the implicit function theorem to see that the set of solutions consists of one or more smooth curves in the unit-square and exclude the possibility of singular points:

**Lemma 3** *If  $(\pi_I, \pi_Y) \in (0, 1) \times (0, 1)$  then*

1.  $\partial F / \partial \pi_I = 0 \Rightarrow \pi_I > 1 / \mathfrak{R}_e^I$
2.  $F = 0 \wedge \partial F / \partial \pi_I = 0 \Rightarrow \pi_Y < \pi_\infty^Y$

*Proof* We have that

$$\partial F / \partial \pi_I = (1 - \tau_I)(1 - \tau_Y)(1 - \pi_Y) + \frac{1 - \tau_I}{s\mathfrak{R}_0^I} \left( -\mathfrak{R}_e^I + \frac{1}{\pi_I} \right) > \frac{1 - \tau_I}{s\mathfrak{R}_0^I} \left( -\mathfrak{R}_e^I + \frac{1}{\pi_I} \right).$$

The last term can be negative only if  $\pi_I > 1 / \mathfrak{R}_e^I$ .

To see the last claim we observe that by the elementary inequality  $-\log u + u - 1 > 0$  we have that

$$\begin{aligned} F + (1 - \pi_I) \frac{\partial F}{\partial \pi_I} &= \frac{1 - \tau_I}{s\mathfrak{R}_0^I} \left( \log \pi_I + \frac{1}{\pi_I} - 1 \right) + \frac{1 - \tau_Y}{s\mathfrak{R}_0^Y} (\mathfrak{R}_e^Y (1 - \pi_Y) + \log \pi_Y) \\ &> \frac{1 - \tau_Y}{s\mathfrak{R}_0^Y} (\mathfrak{R}_e^Y (1 - \pi_Y) + \log \pi_Y). \end{aligned}$$

The last term can be negative only if  $\pi_Y < \pi_\infty^Y$ .

Since  $1 / \mathfrak{R}_e^K > \pi_\infty^K$ , we observe that the curve of solutions can have a horizontal tangent only in region  $IV$  and a vertical tangent only in region  $II$  and that there are no singular points.

Considering the curve in terms of a function of  $\pi_Y$ , we have that at the point  $(\pi_\infty^I, 1)$

$$0 = F'_I(\pi_\infty^I, 1)\pi'_I + F'_Y(\pi_\infty, 1) = \frac{1 - \tau_I}{s\mathfrak{R}_0^I} \left( -\mathfrak{R}_e^I + \frac{1}{\pi_\infty^I} \right) \pi'_I + \frac{1 - \tau_Y}{s\mathfrak{R}_0^Y} (-\mathfrak{R}_e^{IY} + 1),$$

showing that  $\pi_I$  is negative when  $\mathfrak{R}_e^{IY} > 1$  and positive when the condition is not satisfied.

We can now summarize this observation and Lemma 2 and 3 in the following

**Theorem 3** *Assume that  $\mathfrak{R}_e^K > 1$ ,  $K = I, Y$ , let  $\pi_\infty^K$  denote the final size of the susceptible population for each strain and let  $\mathfrak{R}_e^{IY}$  given by Equation (20) denote the effective reproduction number of strain  $Y$  for a  $IY$ -sequential epidemic. When  $\mathfrak{R}_e^{IY} > 1$ , the size of the sequential epidemic is denoted  $\pi_\infty^{IY}$ .*

The lines  $\pi_K = \pi_\infty^K$  divide the unit square into four regions. The region  $(\pi_\infty^I, 1) \times (\pi_\infty^Y, 1)$  is named region I, and the other three regions are named II, III, and IV counter-clockwise as indicated in Figure 4.

The solutions to Equation (18) in the unit-square consist of the isolated point  $(1, 1)$  and a curve  $\gamma(u) = (\pi_I(u), \pi_Y(u))$  with the following properties:

1. The curve  $\gamma$  enters the top of the unit-square at  $(\pi_\infty^I, 1)$ .
  - a) If  $R_e^{IY} > 1$ , the tangent of  $\gamma$  at the point of entry lies in region II. In region II,  $\gamma$  is concave and it leaves region II at the point  $(\pi_\infty^I, \pi_\infty^{IY})$  (cf. Figure 5 c,d).
  - b) If  $R_e^{IY} < 1$ , the tangent of  $\gamma$  at the entry point lies in region I and  $\gamma$  does not intersect region II (cf. Figure 5 a,b).
2. In region I,  $\gamma$  is a monotonically decreasing single-valued function of  $\pi_I$ .
3. The properties of  $\gamma$  are symmetric in I and Y.
4. No point in region III satisfies  $F = 0$ .

The curve  $\gamma$  describes a necessary condition that must be satisfied by solutions to Equations (12,13,14) or equivalently to Equation (15). More precisely, any  $\omega$ -limit point of a solution to the differential equation with initial conditions  $l_I, l_Y \ll 1$  satisfies

$$(S(\infty)/S(0), P_I(\infty)/P_I(0), P_Y(\infty)/P_Y(0), \Lambda_I(\infty), \Lambda_Y(\infty)) = (\pi_i \pi_y, \pi_i, \pi_y, 0, 0),$$

where  $(\pi_I, \pi_Y)$  lies on  $\gamma$ . One can use methods from the proof of Theorem 2 to prove that not all points on  $\gamma$  are limit points for solutions to (15). In fact we find that the size of the sequential epidemic bounds the possible outcomes for epidemics where both strains are present, such that only those points in  $\gamma$  that lie in Region I of Figure 4 plus the final outcome of single-strain epidemics are possible final sizes.

This gives our final characterization of the outcome of "epidemics in competition".

**Theorem 4** *Let*

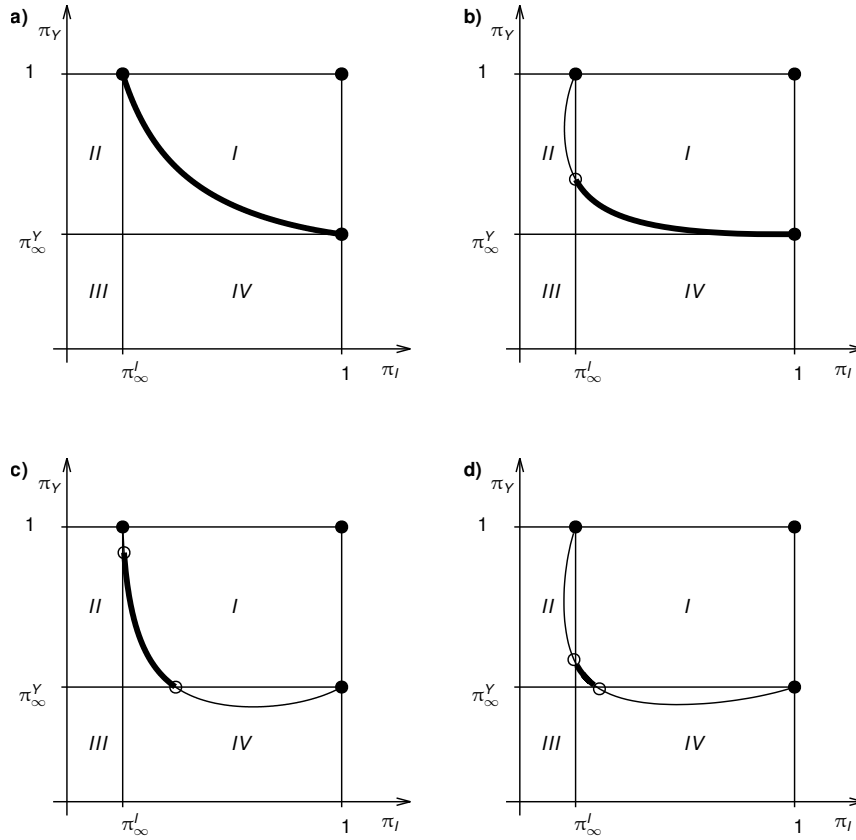
$$\Gamma_\delta = \left\{ (s, p_I, p_Y, l_I, l_Y) \mid 0 \leq p_I + p_Y - s \leq 1, \right. \\ \left. 0 < s \leq p_K \leq 1, 0 \leq l_k < \delta \right\}$$

*denote the set of biologically feasible initial conditions to Equation (15) with initial infection less than  $\delta > 0$  and let*

$$\Phi : \Gamma_\delta \rightarrow [0, 1] \times [0, 1]$$

*denote the map from initial conditions to  $(\pi_I, \pi_Y)$ , the limit of  $(P_I(t)/P_I(0), P_Y(t)/P_Y(0))$  for  $t \rightarrow \infty$ . Then for  $\delta \rightarrow 0$  the set  $\Phi(\Gamma_\delta)$  approaches the set consisting of the segment of the curve  $\gamma$  that lies in Region I on Figure 4 combined with the three points  $(\pi_\infty^I, 1)$ ,  $(1, \pi_\infty^Y)$  and  $(1, 1)$ .*





**Fig. 5** The possible outcomes of a two-strain epidemic. The solutions to the necessary condition (18) are shown in thin line while the heavy line indicate those points that can arise as solutions to the epidemic model (15). Open circles indicate sequential epidemics while filled circles indicate epidemics where at least one strain is absent. The panels a) and b) correspond to the two panels in Figure 4. Parameter values common to all graphs are  $\mathfrak{R}_0^I = 2.0$ ,  $\mathfrak{R}_0^Y = 1.5$  and  $s = p_I = p_Y = 1$ . In a)  $\tau_I = 0.0$ ,  $\tau_Y = 0.0$ ; in b)  $\tau_I = 0.8$ ,  $\tau_Y = 0.2$ ; in c)  $\tau_I = 0.5$ ,  $\tau_Y = 0.6$ ; in d)  $\tau_I = 0.9$ ,  $\tau_Y = 0.8$ .

## 5 The effect of cross-immunity on the final epidemic

To describe the biological implications of Theorem 4, let us consider how the strength of cross-immunity affects the possible outcomes in the situation where all hosts are initially susceptible to both strains such that  $s = p_I = p_Y = 1$ . The two extreme situations are well known: If there is no cross-immunity  $\tau_K = 1$  the two diseases act independently and if both strains are initially present in the population we expect the final size of the susceptible population to the strains to be  $(\pi_\infty^I, \pi_\infty^Y)$ . The case of full cross-immunity  $\tau_K = 0$  was discussed in Sections 2–3.

For intermediate values of  $\tau_K$  we first observe that with initial conditions  $s = p_I = p_Y = 1$  both strains have effective reproduction numbers  $\mathfrak{R}_e^K > 1$  that are independent of the strength of the cross-immunity  $\tau_K$ . Consequently the final size of the susceptible population for a one-strain epidemic  $\pi_\infty^K$  is also independent of  $\tau_K$  and the effec-

tive reproduction number for the sequential epidemic  $\mathfrak{R}_e^{IY} = \mathfrak{R}_0^Y s((1 - \tau_I)\pi_\infty^I) + \tau_I$  is linear in  $\tau_I$ .

When there is no cross-immunity,  $\tau_k = 1$ , we have  $\mathfrak{R}_e^{IY} = \mathfrak{R}_0^Y$  – corresponding to the fact that an initial epidemic with strain  $I$  does not affect strain  $Y$ . When cross-immunity is perfect  $\tau_K = 0$ , we obtain (as expected) the same condition for a sequential epidemic as we found in Section 2. Thus for weaker cross-immunity ( $\tau_K$  larger) the effective reproduction number of a  $KL$ -sequential epidemic increases and for some value of  $1 > \tau_K \geq 0$ , a  $KL$ -sequential epidemic is possible.

This scenario is depicted in Figure 5. For strong cross-immunity (small  $\tau_K$ ) the possible outcomes lie as shown in Figure 5 a) – or if one type of sequential epidemic is possible either in b) or c). As cross-immunity weakens the outcomes look like b) or c) and for even weaker cross-immunity as d). As both  $\tau_k$  approaches zero the two sequential final sizes moves down towards the point  $(\pi_\infty^I, \pi_\infty^Y)$  describing the final sizes of two independent epidemics.

## 6 Discussion

In this article we have described the outcome of a mixed epidemic where two related strains of a pathogen compete for susceptible hosts during a single epidemic episode. The final result of such "epidemics in competition" depends critically on details of the initial conditions and possibly on stochastic variation during the early period of low disease prevalence. Rather than making specific assumptions about conditions during the onset, we have focused on those aspects of the epidemic that may be assessed by deterministic methods.

In general the final outcome may be characterized by the number of hosts that have experienced each of the possible histories of infection during the epidemic. In the cases we have studied, these numbers fall on a single curve in the space of possible outcomes, with the caveat that epidemics where only one of the strains is present may lead to an outcome that is an isolated point which is not part of the curve for mixed epidemics. Such isolated points arise when conditions support a sequential epidemic, i. e. when a strain can produce an epidemic in a population that has already experienced the full course of an epidemic with the other strain.

These observations hold also for the case where strains exert only partial cross-immunity in the sense that infection with one strain reduces the transmission of the alternate strain during a subsequent infection. We finally showed that – as cross-immunity weakens – this model spans the full range of outcomes from the case of full cross-immunity to the case of completely unrelated strains.

This type of competition between two closely related strains, represents the first step in the natural selection among new variants of a disease during outbreaks or pandemics. In focusing on the interaction during the course of the epidemic, the analysis offers a somewhat unusual way to study natural selection among micro-parasites as most studies of natural selection assumes that the disease is in a quasi-steady state determined by the balance between disease transmission and demographic turn-over in the host population. The present approach may be especially useful for rapidly

evolving pathogens such as many RNA-viruses where the demographic balance is of little importance.

From a mathematical view point the paper raises several interesting questions. As all models of a single epidemic, the present models are structurally unstable in the technical sense that small changes of the right hand side of the equations may lead to qualitatively different final states. This fact is somewhat obvious because even small levels of demographic turn-over alter the long term behavior of the solutions. Still the single epidemic is worth studying as it is of major interest in applications to zoonotic outbreaks and pandemic influenza as well as forming the basic structure for much current theoretical work in particular in network theory. It would be interesting to know which types of perturbations – if any – the models are robust to.

The structural instability of the model also limits the available mathematical tools as most results from dynamical systems require that the model is generic. Thus our basic results concerning the continuity of the map from initial conditions to outcome rely on quite basic theory, namely the fact that the value of the solution to an ODE after *finite time* depends continuously on initial conditions combined with explicit evaluation of the approach to the semi-stable equilibrium. How and when may this method be generalized?

Kendall and Saunders (1983) showed that the map from initial conditions to final size is monotonic in the sense that if the inoculum of one strain is reduced by a certain amount while the inoculum of the second strain is increased by the same amount, then the final number of hosts infected with the second type increases while the number infected with the first type decreases. To prove monotonicity, Kendall and Saunders used a complicated coupling argument, see Ball (1995). It may be possible to generalize this proof to the case of partial cross-immunity.

The most important mathematical challenge lies in the determination of the set of possible outcomes of two-strain epidemics. Locating the outcome involves two distinct limits: the  $\omega$ -limit for the solutions of the model's differential equations and subsequently the limit where initial conditions for the infected classes go to zero mimicking the epidemic situation of a newly introduced pathogen. To handle the last limit we relied on the fact that ad hoc methods yield a suitable number of first integrals of the model so that the relation between initial conditions and the  $\omega$ -limit was analytically available. The model of infectivity-reduction was carefully chosen to allow for this process. It is straightforward to see that the methods we have developed in Section 4 may also be applied to the case of polarized immunity. Polarized immunity refers to the situation where after infection with the first strain, a fraction of the hosts recognizes perfectly a challenge from the a second strain while the remainder of the hosts experience the second strain as a novel infection (Gog and Swinton, 2002). However it is not clear if our results – and in particular the one-dimensional nature of the outcome space – will generalize to the case where cross-immunity leads to a reduction in susceptibility.

Throughout the paper we have assumed that the host population mixes homogeneously. It is well known that for a single pathogen in a heterogeneously mixing population, there exists a unique final size but the size is in general not analytically available (Magal et al, 2016; Miller, 2012; Andreasen, 2011; Rass and Radcliffe, 2003). The generalization to multi-strain epidemics is not clear. In fact for heteroge-

neously mixing populations it is easy to find examples where the space of outcomes for a two-strain epidemic is multi-dimensional.

## References

- Ackerman E, I M Longini J, Seaholm S, Hedin AS (1990) Simulation model for viral interference in influenza. *Int J Epidemiol* 19:444–454
- Alexander HK, Day T (2010) Risk factors for the evolutionary emergence of pathogens. *J Roy Soc Interface* 7:1455–1474
- Andreasen V (2011) The final size of an epidemic and its relation to the basic reproduction number. *Bull Math Biol* 73:2305–2321
- Andreasen V, Sasaki A (2006) Shaping the phylogenetic tree of influenza by cross-immunity. *Theor Pop Biol* 70:164–173
- Antia R, Regoes RR, Koella JC, Bergstrom CT (2003) The role of evolution in the emergence of infectious diseases. *Nature (London)* 426:658–661
- Ball F (1995) Coupling methods in epidemic theory. In: Mollison D (ed) *Epidemic Models: Their Structure and Relation to Data*, Cambridge University Press, Cambridge, pp 34–52
- Boni MF, Gog JR, Andreasen V, Feldman MW (2006) Epidemic dynamics and antigenic drift in a single season of influenza A. *Proc Roy Soc London B* 273:1307–1316
- Brauer F (2008) Epidemic models with heterogeneous mixing and treatment. *Bull Math Biol* 70:1869–1885
- Bremermann H, Thieme H (1989) A competitive-exclusion principle for pathogen virulence. *J Math Biol* 27:179–190
- Coddington E, Levinson N (1955) *Theory of Ordinary Differential Equations*. McGraw-Hill
- Ferguson N, Andreasen V (2001) The influence of different forms of cross-protective immunity on pathogen strain structure. In: Blower S, Castillo-Chavez C, Kirschner D, van den Driessche P, Yakubu A (eds) *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods and Theory*, IMA Volumes Mathematics and its Applications **125**, Springer-Verlag, New York, pp 157–169
- Fleming-Davies AE, Dukic V, Andreasen V, Dwyer G (2015) Effects of host heterogeneity on pathogen diversity and evolution. *Ecol Let* 18(11):1252–1261
- Funk S, Jansen VAA (2010) Interacting epidemics on overlay networks. *Phys Rev E* 81
- Funk S, Gilad E, Watkins C, Jansen VAA (2009) The spread of awareness and its impact on epidemic outbreaks. *Proc Natl Acad Sci USA* 106:6872–6877
- Gjini E, Valente C, Sá-Leão R, Gomes MGM (2016) How direct competition shapes coexistence and vaccine effects in multi-strain pathogen systems. *J theor Biol* 388:50–60
- Gog JR, Swinton J (2002) A status-based approach to multiple strain dynamics. *J Math Biol* 44:169–184
- Gomes MGM, Medley GF (2001) Dynamics of multiple strains of infectious agents coupled by cross-immunity: A comparison of models. In: Blower S, Castillo-

- Chavez C, Cooke KL, Krischner D, van der Driessche P (eds) *Mathematical Approaches for emerging and reemerging diseases: Models, Methods, and Theory*, IMA Volumes Mathematics and its Applications **125**, Springer-Verlag, New York, pp 171–191
- Gupta S, Ferguson N, Anderson R (1998) Chaos, persistence, and evolution of strain structure in antigenically diverse infectious agents. *Science* 280:912–915
- Hebert-Dufresne L, Patterson-Lomba O, Goerg GM, Althouse BM (2013) Pathogen Mutation Modeled by Competition Between Site and Bond Percolation. *Phys Rev Lett* 110(10)
- Ho BS, Chao KM (2017) Data-driven interdisciplinary mathematical modelling quantitatively unveils competition dynamics of co-circulating influenza strains. *J Translational Med* 15(1):163
- Holmes EC (2009) The Evolutionary Genetics of Emerging Viruses. *Ann Rev Ecol Syst* 40:353–372
- Karrer B, Newman MEJ (2011) Competing epidemics on complex networks. Preprint arXiv:1105.3424v1
- Kendall W, Saunders I (1983) Epidemics in competition-II - the general epidemic. *J R Statist Soc B* 45:238–244
- Kiss IZ, Cassell J, Recker M, Simon PL (2010) The impact of information transmission on epidemic outbreaks. *Math Biosci* 225:1–10
- Koelle K, Cobey S, Grenfell B, Pascual M (2006) Epochal evolution shapes the phylodynamics of interpanemic influenza A (H3N2) in humans. *Science* 314:1898
- Kubiak RJ, Arinaminpathy N, McLean AR (2010) Insights into the Evolution and Emergence of a Novel Infectious Disease. *PLoS Comp Biol* 6
- Lipsitch M, Cohen T, Murray M, Levin BR (2007) Antiviral resistance and the control of pandemic influenza. *PLOS Med* 4(1):1–11
- Lipsitch M, Colijn C, Cohen T, Hanage WP, Fraser C (2009) No coexistence for free: Neutral null models for multistrain pathogens. *Epidemics* 1(1):2–13
- Magal P, Seydi O, Webb G (2016) Final size of an epidemic for a two-group SIR model. *SIAM J Appl Math* 76(5):2042–2059
- Marceau V, Noel P, Hebert-Dufresne L, Allard A, Dube LJ (2011) Modeling the dynamical interaction between epidemics on overlay networks. *Phys Rev E* 84:026105
- Miller J (2013) Cocirculation of infectious diseases on networks. *Phys Rev E* 87:060801(R)
- Miller JC (2012) A Note on the Derivation of Epidemic Final Sizes. *Bull Math Biol* 74(9):2125–2141
- Newman MEJ (2005) Threshold effects for two pathogens spreading on a network. *Phys Rev Lett* 95(10):108701
- Ohtsuki A, Sasaki A (2006) Epidemiology and disease-control under gene-for-gene plant-pathogen interaction. *J theor Biol* 238:780–794
- Omori R, Adams B, Sasaki A (2010) Coexistence conditions for strains of influenza with immune cross-reaction. *J theor Biol* 262:48–57
- Poletti P, Caprile B, Ajelli M, Pugliese A, Merler S (2009) Spontaneous behavioural changes in response to epidemics. *J theor Biol* 260:31–40

- Rass L, Radcliffe J (2003) Spatial deterministic epidemics. Mathematical Surveys and Monographs, vol. 102, Amer. Math. Soc., Providence, RI
- Saunders I (1981) Epidemics in competition. J Math Biol 11:311–318
- Svensson Å, Scalia Tomba G (2001) Competing epidemics on closed populations. Research Report in Mathematical Statistics 2001:8, Department of Mathematics, Stockholm University. <http://www2.math.su.se/matstat/reports/seriea/>