



International Institute for  
Applied Systems Analysis  
Schlossplatz 1  
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342  
Fax: +43 2236 71313  
E-mail: [publications@iiasa.ac.at](mailto:publications@iiasa.ac.at)  
Web: [www.iiasa.ac.at](http://www.iiasa.ac.at)

---

## **Interim Report**

**IR-07-051**

### **On State-Space Reduction in Multi-Strain Pathogen Models, With an Application to Antigenic Drift in Influenza A**

Sergey Kryazhimskiy ([skryazhi@princeton.edu](mailto:skryazhi@princeton.edu))  
Ulf Dieckmann ([dieckmann@iiasa.ac.at](mailto:dieckmann@iiasa.ac.at))  
Simon Levin ([slevin@princeton.edu](mailto:slevin@princeton.edu))  
Jonathan Dushoff ([dushoff@eno.princeton.edu](mailto:dushoff@eno.princeton.edu))

---

#### **Approved by**

Leen Hordijk  
Director, IIASA

December 2007

---

*Interim Reports* on work of the International Institute for Applied Systems Analysis receive only limited review. Views or opinions expressed herein do not necessarily represent those of the Institute, its National Member Organizations, or other organizations supporting the work.

## Contents

Abstract.....	1
Author summary.....	1
Introduction .....	2
Model.....	5
Types of equation-based multi-strain SIR models .....	5
Model description.....	7
Expansion in immunity variables .....	9
Implementation details .....	13
Results .....	13
Comparison of closures for a pathogen with four strains.....	13
Application to antigenic drift in influenza A.....	16
Discussion.....	19
Supporting Information .....	21
Acknowledgments .....	21
References .....	22
Table caption .....	24
Figure captions .....	25
Table .....	26
Figures .....	27
Supporting information .....	33

# On state-space reduction in multi-strain pathogen models, with an application to antigenic drift in influenza A

Sergey Kryazhimskiy<sup>1,\*</sup>, Ulf Dieckmann<sup>2</sup>,  
Simon A. Levin<sup>3</sup>, Jonathan Dushoff<sup>3,4</sup>

**1** Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, USA, **2** Evolution and Ecology Program, International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria, **3** Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA, **4** Fogarty International Center, National Institute of Health, Bethesda, MD 20892, USA

## Abstract

Many pathogens exist in phenotypically distinct strains that interact with each other through competition for hosts. General models that describe such multi-strain systems are extremely difficult to analyze because their state spaces are enormously large. Reduced models have been proposed, but so far all of them necessarily allow for coinfections and require that immunity be mediated solely by reduced infectivity, a potentially problematic assumption. Here we suggest a new state-space reduction approach that allows immunity to be mediated by either reduced infectivity or reduced susceptibility and that can naturally be used for models with or without coinfections. Our approach utilizes the general framework of status-based models. The cornerstone of our method is the introduction of immunity variables, which describe multi-strain systems more naturally than the traditional tracking of susceptible and infected hosts. Models expressed in this way can be approximated in a natural way by a truncation method that is akin to moment closure, allowing us to sharply reduce the size of the state space, and thus to consider models with many strains in a tractable manner. Applying our method to the phenomenon of antigenic drift in influenza A, we propose a potentially general mechanism that could constrain viral evolution to a one-dimensional manifold in a two-dimensional trait space. Our framework broadens the class of multi-strain systems that can be adequately described by reduced models. It permits computational, and even analytical, investigation and thus serves as a useful tool for understanding the evolution and ecology of multi-strain pathogens.

## Author summary

Many important human pathogens, including HIV and influenza viruses, consist of many different “strains”, which elicit distinct immune responses in their hosts. Infection by one

---

\*To whom correspondence should be addressed. E-mail: skryazhi@princeton.edu

variant usually triggers partial cross-immunity against several other variants. This process leads to a complicated and dynamic immunity structure in the host population. Most existing models of a population with a multi-strain pathogen are either very complex, or rely on specific simplifying assumptions. Here, the authors suggest a new way of simplifying such models that allows for greater flexibility in underlying assumptions. This approach could lead to deeper understanding of the ecology and evolution of multi-strain pathogens. The authors apply their approach to a simple model of evolution of influenza A that illustrates one hypothesis about how influenza A evolution may be structured.

## Introduction

Microbial pathogens are tremendously diverse. Pathogens that cause one and the same disease may differ remarkably in both their genotype and their phenotype, like in HIV/AIDS [1], influenza [2], malaria [3], and meningitis [4]. Phenotypically different variants of the same pathogen are called strains. If several strains exist in a host population, they interact with each other in two ways.

The first type of interaction may be referred to as ecological interference [5,6]. For many infectious diseases, a host infected with one strain is removed, for the duration of the disease, from the population of hosts susceptible to the pathogen. This is because (a) the immune system of the host becomes activated upon infection by the first strain, so that it is hard for a second strain to enter and/or replicate in this host, and (b) the infected host may be physically removed from the susceptible population, by dying or staying at home. Ecological interference takes place even between unrelated pathogens [6].

The second type of interaction, referred to as cross-immunity interference, is specific to different strains of the same pathogen: these can confer full or partial immunity to each other. This means that a host infected with one strain becomes substantially less susceptible to certain other strains of the pathogen for a prolonged period of time after the initial infection is cleared. Cross-immunity is highest between phenotypically similar strains. Since phenotypic similarity usually implies recent common ancestry, a pathogen's ecology is thus intrinsically entangled with its evolution.

Understanding the dynamics of multi-strain pathogens at a general theoretical level turns out to be extremely difficult. Numerous models have been proposed during the past twenty years (e.g., [3,7–9]). Although these models share many similarities, they substantially differ in particulars, often resulting in conflicting model predictions. In consequence, there is little agreement as to how best to gain insights into the ecology and evolution of multi-strain pathogens. Models of multi-strain pathogens can be either equation- or agent-based. Agent- or individual-based models have recently become increasingly elaborate and interesting [10–13], largely due to an increase in computational capabilities. Since these models, however, are not designed for analytical tractability, we do not dwell on this type of model here.

Virtually all equation-based models of disease dynamics can be traced back to the compartment model introduced by Kermack and McKendrick in 1927 [14]. These models

are also known as SIR models, reflecting a host population’s partitioning into susceptible, infected, and recovered individuals. A problem that arises immediately when attempting to extend this classical SIR framework to multiple strains is that the number of state variables, and typically also of parameters, increases exponentially with the number of strains [8, 9]. This presents not only computational challenges but also draws attention to a fundamental conceptual difficulty: even for a moderately large number of strains, the resultant number of state variables quickly surpasses any realistic host population size. Most compartments in such a model therefore consist of few individuals, if they are occupied at all: effects of demographic stochasticity must then not be neglected. To avoid this complication, existing approaches to modeling multi-strain pathogens have attempted to reduce the number of model compartments. Usually, such reductions are valid only under certain sets of assumptions that may or may not be adequate depending on the modeled phenomenon. Thus, it is important to expand the set of assumptions under which reduced models are applicable. Our work presented here contributes to this goal.

Traditionally, full models have been developed based on the assumption of reduced susceptibility, which implies that immune hosts are able to block off an infection completely, with a certain probability [8, 9]. On the other hand, all existing reduced models rely on the assumption of reduced infectivity that implies that all hosts, immune or not, get infected with the same probability, but those that possess immunity become less infectious than those who do not [3, 15]. The reality, most likely, lies somewhere between these two abstractions. Nevertheless, as we discuss in the Model section, the reduced susceptibility assumption seems more plausible. In this study, we develop a state-space reduction approach that can be applied under either of these assumptions, in models with or without coinfections. Our approach differs from the existing ones in that it produces a collection of models that approximate the full models with the desired degree of accuracy. The number of variables needed for the resulting approximations grows algebraically with the number  $n$  of strains, rather than exponentially: when  $n$  is large, the difference between, e.g.,  $n^2$  and  $2^n$  is enormous, with the former growing much more slowly than the latter. If coinfections and reduced infectivity are assumed, our approach produces a model equivalent to that of Gog and Grenfell [15].

To illustrate the utility of our approach, and that of reduced models in general, we demonstrate its application to the phenomenon of drift in influenza A. Using reduced models we are able to simulate up to 400 strains. Influenza A is a multi-strain pathogen whose epidemiology and evolution display an intricate interaction pattern. Because the human immune system can produce protective antibodies against influenza’s surface glycoprotein hemagglutinin, individuals gain lifelong immunity against each strain of the virus with which they have been infected [2, 16]. This results in a complex partitioning of the human host population according to the immunity of individuals to different influenza strains. The ensuing frequency-dependent selection is thought to drive the evolution of influenza A, giving rise to a process known as antigenic drift [17]. Lapedes and Farber [18] have shown that the antigenic space of influenza is approximately five-dimensional. Subsequently, Smith et al. [19] argued that the first two principal dimensions are most important. Moreover, as follows from results by Smith et al., the temporal evo-

lution of influenza’s H3N2 subtype proceeds along a single line in the antigenic space, i.e., antigenic clusters corresponding to different years are well separated along the first principal dimension. This agrees with the observation that the phylogenetic tree of subtype H3N2 possesses a single trunk [20] (but see an alternative hypothesis by Recker et al. [21]). In other words, even though the H3N2 subtype experiences substantial genetic diversity during each epidemic season, only one progeny strain survives in the longer run. Accordingly, the number of coexisting H3N2 strains does not grow from year to year.

A few recent studies have attempted to model, and thereby explain, the phenomenon of antigenic drift in influenza A. Apart from individual-based models, most of these studies consider a one-dimensional strain space in which some sort of traveling-wave behavior is observed [15, 22–25]. In order to constrain the evolution of a virus to one dimension in a two-dimensional strain space, it has been necessary to require that the strain space was essentially unviable except for a relatively thin region along one axis [15].

In a recent study Koelle et al. [26] took a different approach and succeeded in constraining the diversity of a virus living in a high-dimensional sequence space. The authors explicitly mapped viral genotypes to phenotypes and showed that the single-trunk phylogeny of influenza A may be a consequence of the neutral network structure of the influenza genotype space. However, it is an open question which properties of the phenotype space are sufficient to constrain viral diversity in the course of its evolution. Recker et al. [21] suggest one explanation. They argue that the succession of antigenically distinct variants may be an intrinsic feature of the dynamics of a limited set of antigenic types that are always present in the host population and, thus, is decoupled from the genetic evolution of the virus. Here we suggest an alternative conceptual scenario that follows the more traditional view that antigenic drift and genetic evolution are tightly connected. However, we deliberately avoid the problem of mapping genotypes to phenotypes and, instead, assume a relatively simple structure of the phenotype space – a rectangular lattice. Our model offers a straightforward explanation of what could be happening in such a phenotype space in order for the diversity of a virus to be constrained in the long run. Our work is, thus, complementary to that of Koelle et al. [26]. In our two-dimensional phenotype space each coordinate captures changes in the conformations of an epitope – local region on the surface of the hemagglutinin molecule that interacts with the immune system [27–29]. We then investigate a scenario in which the immune response of hosts depends on two epitopes and full immune protection is gained against all strains sharing an epitopic conformation with a previous infection. In this respect our model is closely related to models studied by Gupta and colleagues [3, 21, 30]. We show that the evolutionary trajectory of the influenza A virus in our model follows a line, even though the model’s strain space is two-dimensional. This finding agrees with the observed single-trunk phylogeny of influenza’s H3N2 subtype [20].

# Model

## Types of equation-based multi-strain SIR models

As mentioned above, SIR models are based on partitioning a host population into susceptible, infected, and recovered classes. Abundances in the resultant compartments are then tracked through time by means of ordinary differential equations. There are three major dichotomies according to which multi-strain SIR models can be classified. The first dichotomy refers to the treatment of individuals with respect to cross-immunity: there are history-based and status-based approaches.

In the history-based approach introduced by Castillo-Chavez et al. [31] and generalized by Andreasen et al. [8] the hosts are grouped into classes by their disease histories, which are defined in terms of all strains with which an individual has ever been infected. Disease histories determine the rates at which these compartments are populated and depopulated when hosts acquire infections with new strains.

In the status-based approach introduced by Gog and Swinton [9] the individuals are grouped together by their immune status, which is defined in terms of all strains against which an individual is immune. This set is at least as large as the set of all strains with which an individual has been infected, provided one assumes that infection results in complete and persistent protection. By definition, an individual is fully susceptible to all strains not included in its immune status. The immune status determines the rate at which the compartment is populated and depopulated when individuals acquire immunity against new strains. In particular, after an infection with a new strain, individuals move to new immune status classes with rates that are determined by the probabilities of acquiring cross-immunity against other strains. Thus, after an infection with a particular strain, different individuals with the same immune status change their immune status in different ways. This approach to capturing the probabilistic aspect of cross-immunity is called polarized immunity.

The second dichotomy is the permission or prohibition of coinfections. Coinfection is an event through which an individual, while already being infected with one strain, gets simultaneously infected with a second strain. Unrestricted permission or complete prohibition of coinfections are, of course, mathematical abstractions. One or the other may be more plausible for any particular pathogen.

The third dichotomy refers to the way protective immunity works. As mentioned above, either the chance for an immune host to get infected or the infectivity of an immune host during a secondary infection are reduced. Within the history-based approach, models constructed under both assumptions were shown to behave qualitatively similar, at least in simple systems [32,33]. A mathematical formulation of a status-based multi-strain model usually forces the modeler to make a choice between the reduced-susceptibility assumption and the reduced-infectivity assumption. The latter is more convenient from the mathematical standpoint, as such models easily yield themselves to state-space reduction. This is because the rate at which hosts acquire infection and, hence, immunity is assumed to be independent of the immune status (or disease history, in history-based models).

However, within the status-based framework this assumption is somewhat problematic from the biological standpoint.

In general, infection of a host with a virus results in two effects: (a) the host becomes sick and transmits the virus, and (b) the immune system of the host develops protective antibodies against the infecting variant. If the host is already partially immune to the infecting strain, that is, protective antibodies had been developed prior to the infection, then (a) the severity and the duration of infection are reduced, and (b) the production of new types of antibodies is slowed down. In the limit case, when the host has full immune protection against the infecting strain, it is capable of completely fending off the infection with the existing arsenal of antibodies, resulting in no infectiousness and no production of new types of antibodies. In the simplest version of status-based models, individuals are either fully susceptible to a strain or fully immune against it, so this limit case should apply. The assumption of reduced infectivity within this framework correctly captures the first effect but neglects the second effect. In other words, hosts that are fully immune against a particular variant do not transmit it but, paradoxically, still increase their repertoire of antibodies in exactly the same way as do susceptible individuals, upon an infection with this variant (see supplementary information text for mathematical implications of this assumption). Evidently, this results in an over-immunization of the host population. In this light, the assumption of reduced susceptibility seems more plausible, at least within the status-based framework. In the remainder of this study we will therefore work under this assumption, even though our approach is applicable to models with reduced infectivity as well.

Full history-based or status-based multi-strain SIR models are cumbersome in terms of their analytical treatment [33,34] and numerical analysis [35], mainly due to the enormous size of their state spaces. Therefore, reduced models are necessary and have been proposed in the past. However, as mentioned above, only models that allow for coinfections and work under the reduced-infectivity assumption yielded themselves so far to reduction [15, 30, 32], whereas models that prohibit coinfections and those that work under the assumption of reduced susceptibility remained intractable. With a view towards filling this gap within the status-based framework, we propose a new state-space reduction approach that is based on two key observations:

1. When attempting to infect a potential host, an infecting strain does not in any way “perceive” the host’s entire immune status or disease history. What it does perceive is simply whether or not this potential host possesses any immunity against the focal strain. This consideration suggests a natural set of state variables: it is helpful to keep track of the proportions of a population that are immune to each strain, or combinations thereof. It turns out to be possible to reformulate any full status-based model in terms of such new immunity variables. Below we refer to this transformation as an “expansion in immunity variables”.
2. The utility of this transformation becomes clear when we recognize that, at any moment in time and for most diseases, many hosts will be immune to only a few of the strains currently circulating, while only a few hosts will be immune to many of

these strains. Consequently, the immunity variables that describe the latter small proportions of the host population need not be tracked exactly and independently but can instead be approximated, without much disturbing the overall disease dynamics. Higher-order immunity variables can thus be approximated by functions of lower-order immunity variables. Below we refer to this approximation as truncating or “closing” the disease dynamics at a desired order.

The approach presented in this article introduces a general representation of status-based multi-strain models in terms of immunity variables. This representation is useful because it produces a hierarchical structure of equations describing the dynamics of a multi-strain system. The equations at any given order  $\ell$  of this hierarchy are decoupled from equations at all orders above  $\ell + 1$ , under the assumption of reduced susceptibility, or even above  $\ell$ , under the assumption of reduced infectivity. Thus, this hierarchy can easily be truncated at any order, either by approximating higher-order immunity variables with functions of lower-order variables under the former assumption, or by simply ignoring higher-order variables under the latter assumption. The resultant truncated models provide either approximate or exact reduced descriptions of the original system.

## Model description

We now proceed to the mathematical formulation of our framework. For the sake of clarity, we develop our reasoning for the model with coinfections; the model in which coinfections are prohibited is outlined in the supporting information text. The following notations are used throughout (see Table 1 for a full summary):  $K$  is the set of all  $n$  strains; individual strains from this set are referred to by their index  $i$ ;  $S_A$  is the proportion of individuals in the host population that possess immune status  $A \subset K$  and that, therefore, are currently fully susceptible to all strains in the subset  $K \setminus A$  (we refer to these individuals as being *in state A*);  $S_\emptyset$  thus denotes the class of naive individuals, i.e., individuals that have no immunity whatsoever; and  $I_i$  is the proportion of individuals in the host population that are currently infectious with strain  $i$ . Since all host individuals naturally fall in exactly one of the  $S_A$  classes, we have  $\sum_{A \subset K} S_A = 1$ . Here, the summation is taken over all subsets of the set of strains, including the empty set.

Based on this notational framework, we can specify the disease dynamics of a multi-strain pathogen in three steps, (a) to (c) below, by considering the three processes that cause host individuals to enter and exit classes defined by their immune status.

(a) *Births and deaths.* We assume that hosts are born at per capita rate  $\mu$ , being uninfected and susceptible to all strains. Accordingly, all newly born hosts enter the population through class  $S_\emptyset$ . The birth rate into class  $S_A$  is thus given by  $\mu\delta_{A,\emptyset}$ , where

$$\delta_{A,\emptyset} = \begin{cases} 1 & \text{if } A = \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

We further assume that infections do not alter the death rate of hosts and that the host population is at its demographic equilibrium. This implies a constant per capita death rate  $\mu$  for all classes.

(b) *Acquisition of immunity.* The crucial ingredient in this type of model is to specify how an infection changes the immune status of the infected host. We assume that hosts recover from an infection at per capita rate  $\nu$ . The proportion of hosts that recover to a state  $B \subset K$ , having been in state  $A \subset K$  when they were infected by strain  $i \in K$ , is denoted by  $C(A, i, B)$ . Following Gog and Swinton [9], we assume that  $C(A, i, B)$  is allowed to take non-zero values only if:

1.  $i \in B$ , which means that each strain confers total immunity to itself;
2.  $A \subseteq B$ , which means that immunity is only gained;
3.  $i \notin A$ , which means that, to get infected, hosts must be susceptible to the infecting strain.

Any chosen set of functions  $C$  also has to fulfill the consistency condition

$$\sum_{B \subset K} C(A, i, B) = 1 \quad \text{for } i \notin A,$$

since there are no deaths from infection. Note that point 3 above is the mathematical formulation of the reduced-susceptibility assumption (compare the three assumptions to those of Gog and Grenfell [15] provided in the supporting information).

(c) *Acquisition of infections.* The force of infection for strain  $i$ , describing the per capita rate at which host individuals susceptible to strain  $i$  get infected by strain  $i$ , is

$$\Lambda_i = \beta_i I_i,$$

where  $\beta_i$  is the transmission coefficient for strain  $i$ . Note that mutations between strains can be easily incorporated into this dynamics, by adjusting the force of infection,

$$\Lambda_i = \beta_i \left( (1 - m) I_i + \sum_{j \in M_i} \frac{m}{|M_j|} I_j \right),$$

where  $M_i \subset K$  is the mutational neighborhood of strain  $i$ , i.e., the set of all strains that can mutate into  $i$ . Depending on the chosen description of the pathogen,  $M_i$  can be, for example, the set of all point-mutation neighbors or a set of neighbors in a phenotype space.  $|M_i|$  denotes the number of strains in this neighborhood. We assume that all neighborhoods are mutual, i.e.,  $i \in M_j$  if and only if  $j \in M_i$  for any  $i, j \in K$ . We assume that mutations can occur with probability  $m$  during the infection period.

Based on (a) to (c), we obtain the following system of equations [9],

$$\dot{S}_A = \mu(\delta_{A,\emptyset} - S_A) + \sum_{k \in K} \sum_{B \subset K \setminus \{k\}} S_B \Lambda_k C(B, k, A) - \sum_{k \in K \setminus A} \Lambda_k S_A \quad \text{for all } A \subset K, \quad (1)$$

$$\dot{I}_i = \Lambda_i \sum_{B: i \in K \setminus B} S_B - (\nu + \mu) I_i \quad \text{for all } i \in K. \quad (2)$$

where dots denote derivatives with respect to time. Altogether, this system contains  $2^n + n - 1$  equations, where  $n$  is the number of strains.

To complete the definition of our multi-strain model, we further specify the process of immunity acquisition. We introduce the probability  $C_{i_1 i_2 \dots i_\ell}^*(A, k)$ ,  $\ell = 1, 2, \dots, n$  of acquiring immunity against strains  $i_1, i_2, \dots, i_\ell$  (all different) after an infection with strain  $k$  for a host that had immune status  $A$  prior to the infection,

$$C_{i_1 i_2 \dots i_\ell}^*(A, k) = \sum_{\substack{B \subset K: \\ i_1, i_2, \dots, i_\ell \in B}} C(A, k, B).$$

We assume that (a) the chance of obtaining immunity against strain  $i$  after the infection with strain  $k$  does not depend on the previous immune status of the individual, and (b) if strain  $k$  can potentially confer cross-immunity to strains  $i$  and  $j$ , then obtaining immunity against  $i$  and obtaining immunity against  $j$  are independent. According to (b), the chance that a host becomes immune to strains  $i$  and  $j$  after being infected with strain  $k$  is simply  $\sigma_{ki}\sigma_{kj}$ . Mathematically, these assumptions determine the shape of the  $C^*$ -functions defined above. All functions  $C_{i_1 i_2 \dots i_\ell}^*(A, k)$  vanish if  $k \in A$  and, if  $k \notin A$  we have

$$C_{i_1 i_2 \dots i_\ell}^*(A, k) = \prod_{\substack{j=1, 2, \dots, \ell: \\ i_j \notin A}} \sigma_{ki_j},$$

where, by definition,  $\prod_{j \in \emptyset} \sigma_{kj} = 1$ . Based on these assumptions, cross-immunity is entirely characterized by the matrix of pairwise cross-immunity coefficients,  $\sigma = (\sigma_{ij})$ ,  $i, j \in K$  [9]. The elements of  $\sigma$  are all probabilities, so that  $\sigma_{ij} \in [0, 1]$  and  $\sigma_{ii} = 1$ .

## Expansion in immunity variables

To derive the approximations of model (1)–(2), we rewrite that system in terms of the immunity variables,

$$\xi_i = \sum_{\substack{A \subset K: \\ i \in A}} S_A, \quad \xi_{ij} = \sum_{\substack{A \subset K: \\ i, j \in A}} S_A, \quad \dots$$

Each immunity variable has a clear intuitive interpretation:  $\xi_{i_1 i_2 \dots i_\ell}$  describes the proportion of hosts that currently have immunity against strains  $i_1, i_2, \dots, i_\ell \in K$ . We will refer to immunity variables  $\xi_{i_1 i_2 \dots i_\ell}$  as being of order  $\ell$ . Evidently, all immunity variables are symmetric with respect to the permutation of their indices. Immunity variables with duplicate indices  $i$  remain unchanged when the duplicate index is removed,  $\xi_{\dots i \dots i \dots} = \xi_{\dots i \dots}$ . By definition, the immunity variables satisfy monotonicity conditions,

$$1 \geq \xi_{i_1} \geq \xi_{i_1 i_2} \geq \dots$$

for all pairwise different  $i_1, i_2, \dots \in K$ .

Recalling that  $\sum_{A \subset K} S_A = 1$ , we easily obtain the new equations for  $\dot{I}_i$ ,

$$\dot{I}_i = \Lambda_i(1 - \xi_i) - (\nu + \mu)I_i \quad \text{for all } i \in K. \quad (3)$$

Derivation of the equations for  $\dot{\xi}_i$ ,  $\dot{\xi}_{ij}$ , etc. is more technical. Here we restrict ourselves to an explicit derivation of the equation for  $\dot{\xi}_i$ . For this purpose, we apply the time

derivative to the definition of  $\xi_i$  and use (1), to obtain

$$\dot{\xi}_i = \mu \sum_{\substack{A \subset K: \\ i \in A}} (\delta_{A, \emptyset} - S_A) + \sum_{\substack{A \subset K: \\ i \in A}} \sum_{k \in K} \sum_{B \subset K \setminus \{k\}} S_B \Lambda_k C(B, k, A) - \sum_{\substack{A \subset K: \\ i \in A}} \sum_{k \in K \setminus A} \Lambda_k S_A. \quad (4)$$

The first term in (4) accounts for the depletion, due to deaths, of healthy host individuals that are immune to strain  $i$ ,

$$\mu \sum_{\substack{A \subset K: \\ i \in A}} (\delta_{A, \emptyset} - S_A) = -\mu \xi_i.$$

The second term in (4) can be simplified as follows,

$$\begin{aligned} & \sum_{\substack{A \subset K: \\ i \in A}} \sum_{k \in K} \sum_{B \subset K \setminus \{k\}} S_B \Lambda_k C(B, k, A) = \sum_{k \in K} \sum_{B \subset K \setminus \{k\}} S_B \Lambda_k C_i^*(B, k) \\ &= \sum_{k \in K} \sum_{\substack{B \subset K \setminus \{k\}: \\ i \in B}} S_B \Lambda_k C_i^*(B, k) + \sum_{k \in K} \sum_{B \subset K \setminus \{k, i\}} S_B \Lambda_k C_i^*(B, k) \\ &= \sum_{k \in K} \sum_{\substack{B \subset K \setminus \{k\}: \\ i \in B}} S_B \Lambda_k + \sum_{k \in K} \sum_{B \subset K \setminus \{i, k\}} S_B \Lambda_k \sigma_{ki} \\ &= \sum_{k \in K} \Lambda_k (\xi_i - \xi_{ik}) + \sum_{k \in K} \Lambda_k \sigma_{ki} (1 - \xi_i - \xi_k + \xi_{ik}). \end{aligned}$$

The last equality is satisfied due to the inclusion-exclusion principle [36]. Finally, the third term in (4) can be rewritten as

$$\sum_{\substack{A \subset K: \\ i \in A}} \sum_{k \in K \setminus A} \Lambda_k S_A = \sum_{k \in K \setminus \{i\}} \sum_{\substack{A \subset K \setminus \{k\}: \\ i \in A}} \Lambda_k S_A = \sum_{k \in K} \Lambda_k (\xi_i - \xi_{ik}).$$

Collecting the three results above, we obtain the equation for  $\dot{\xi}_i$  expressed in terms of the new variables,

$$\dot{\xi}_i = \sum_{k \in K} \Lambda_k \sigma_{ki} (1 - \xi_i - \xi_k + \xi_{ik}) - \mu \xi_i \quad \text{for all } i \in K. \quad (5)$$

The equations for  $\dot{\xi}_{ij}$  are obtained analogously,

$$\begin{aligned} \dot{\xi}_{ij} &= \sum_{k \in K} \left\{ \Lambda_k \sigma_{kj} (\xi_i - \xi_{ik} - \xi_{ij} + \xi_{ijk}) + \Lambda_k \sigma_{ki} (\xi_j - \xi_{jk} - \xi_{ij} + \xi_{ijk}) + \right. \\ &\quad \left. + \Lambda_k \sigma_{ki} \sigma_{kj} (1 - \xi_i - \xi_j - \xi_k + \xi_{ij} + \xi_{ik} + \xi_{jk} - \xi_{ijk}) \right\} - \mu \xi_{ij} \end{aligned} \quad (6)$$

for all  $i, j \in K$ ,  $i \neq j$ .

Fortunately, it is not necessary to rewrite the full system of equations (1) and (2) in terms of immunity variables, since our goal is to reduce it. The equations for the immunity variables of order  $\ell$  depend on the immunity variables of order  $\ell$  and  $\ell + 1$ , but not on any immunity variables of higher orders. To truncate this hierarchy of equations at order  $\ell$ , we

thus need to approximate the  $\xi$ -variables of order  $\ell + 1$  by a function of immunity variables of lower orders. This procedure is similar to moment-closure techniques widely used in, e.g., spatial ecology and epidemiology [37, 38]. To distinguish the true  $\xi$ -variables from their approximations, we add a hat to the latter ones, as, e.g., in  $\hat{\xi}_{ij}$ . Each approximation must satisfy three conditions that directly follow from the definition of the  $\xi$ -variables:

1. Symmetry condition:

$$\begin{aligned}\hat{\xi}_{ij} &= \hat{\xi}_{ji} && \text{for all } i, j \in K, \\ \hat{\xi}_{ijk} &= \hat{\xi}_{ikj} = \hat{\xi}_{jik} = \hat{\xi}_{jki} = \hat{\xi}_{kij} = \hat{\xi}_{kji} && \text{for all } i, j, k \in K, \\ &\dots\end{aligned}$$

2. Monotonicity condition:

$$\begin{aligned}\hat{\xi}_{ij} &\leq \min(\xi_i, \xi_j) && \text{for all } i, j \in K, \\ \hat{\xi}_{ijk} &\leq \min(\xi_{ij}, \xi_{ik}, \xi_{kj}) && \text{for all } i, j, k \in K, \\ &\dots\end{aligned}$$

3. Redundancy condition: the approximate immunity variables with duplicate indices must be equal to the corresponding immunity variables of the previous order when the duplicate index is removed,  $\hat{\xi}_{\dots i \dots i \dots} = \xi_{\dots i \dots}$ .

Below we introduce and discuss two simple closures of order 1 and one simple closure of order 2.

- (a) *Order-1 independence closure:*

$$\hat{\xi}_{ij} = \begin{cases} \xi_i \xi_j & \text{if } i \neq j, \\ \xi_i & \text{if } i = j, \end{cases} \quad \text{for all } i, j \in K. \quad (7)$$

The symmetry and redundancy conditions are evidently met, and the monotonicity condition is satisfied because, by definition,  $\xi_k \leq 1$  for all  $k \in K$ . The motivation underlying this closure is the following. If the cross-immunity interference between strains is small, and if the time that hosts spend in the infectious state is small compared with the time they spend in the healthy state, then hosts immune to each strain will be almost independently distributed among all hosts. We expect this closure to underestimate the level of immunity in the host population especially for high values of cross-immunity when correlations in the population immunity structure are high. However, even when cross-immunity between strains is absent altogether, the independence closure is not exact because of host aging (this can easily be checked for a simple system with two strains). It would be interesting to know whether there exists an exact closure in this special case.

- (b) *Order-1 interpolation closure:*

$$\hat{\xi}_{ij} = \xi_i \xi_j \left( 1 - \frac{\sigma_{ij} + \sigma_{ji}}{2} \right) + \frac{\sigma_{ij} + \sigma_{ji}}{2} \min(\xi_i, \xi_j) \quad \text{for all } i, j \in K. \quad (8)$$

Again, the symmetry and redundancy conditions are evidently met, and the monotonicity condition is satisfied because

$$\hat{\xi}_{ij} \leq \max \{ \xi_i \xi_j, \min(\xi_i, \xi_j) \},$$

and  $0 \leq (\sigma_{ij} + \sigma_{ji})/2 \leq 1$  for all  $i, j \in K$ . If we rewrite this approximation for the case of a symmetric cross-immunity matrix  $\sigma$ ,

$$\hat{\xi}_{ij} = \xi_i \xi_j (1 - \sigma_{ij}) + \sigma_{ij} \min(\xi_i, \xi_j),$$

it is easy to see the motivation underlying this closure. It results from the linear interpolation between two extreme cases: absent cross-immunity and full cross-immunity. If, in the one extreme,  $\sigma_{ij} = 0$ , the order-1 interpolation closure will coincide with the order-1 independence closure. If, in the other extreme,  $\sigma_{ij} = 1$ , all hosts who have been infected with strain  $i$  (strain  $j$ ) will be immune also to strain  $j$  (strain  $i$ ). We approximate the fraction of hosts that have been infected with strain  $i$  (strain  $j$ ) by  $\xi_i$  ( $\xi_j$ ), which motivates the minimum term. This approximation is crude because we neglect cases in which immunity to strain  $i$  is mediated through a third strain  $k$  that may not provide cross-immunity to strain  $j$ .

(c) *Order-2 independence closure:*

$$\hat{\xi}_{ijk} = \begin{cases} \frac{1}{3} (\xi_{ij}\xi_k + \xi_{ik}\xi_j + \xi_{jk}\xi_i) & \text{if } i \neq j \neq k \neq i \\ \xi_{ij} & \text{if } i = k \text{ or } j = k \\ \xi_{ik} & \text{if } i = j \end{cases} \quad (9)$$

The motivation underlying this closure is analogous to that of the order-1 independence closure. All conditions are fulfilled.

Models truncated at first order have  $2n$  remaining variables, while models truncated at second order have  $n(n+3)/2$  variables. Thus, our approximation has enabled a switch from an exponential to an algebraic scaling of the number of state variables with the number  $n$  of strains.

It is interesting to compare the obtained first-order equations under, say, the independence closure, with the reduced model proposed by Gog and Grenfell [15]. The latter is an exact reduced representation of the full status-based model with coinfections and reduced infectivity (we demonstrate this in the supporting information text), and, as such, can be used as an approximation for the full model with coinfections and reduced susceptibility. Gog and Grenfell's equations are [15]

$$\dot{I}_i = \Lambda_i S_i - (\nu + \mu) I_i, \quad (10)$$

$$\dot{S}_i = \mu - S_i \Lambda_i - S_i \sum_{k \in K \setminus \{i\}} \sigma_{ki} \Lambda_k - \mu S_i. \quad (11)$$

Now, we define the fraction of susceptibles to strain  $i$  as  $S_i = 1 - \xi_i$  and rewrite equations (3), (5),

$$\begin{aligned} \dot{I}_i &= \Lambda_i S_i - (\nu + \mu) I_i, \\ \dot{S}_i &= \mu - S_i \Lambda_i - S_i \sum_{k \in K \setminus \{i\}} \sigma_{ki} \Lambda_k S_k - \mu S_i. \end{aligned}$$

These equations are very similar to (10)–(11) with the only difference occurring in the summation term in the equation for  $\dot{S}_i$ . As expected from the initial assumptions, immunization of the host population happens at a faster rate in Gog and Grenfell’s model.

We have thus demonstrated how our framework of approximation – based on transformation to, and expansion in, immunity variables – provides simplified disease dynamics of multi-strain pathogens, in particular under the assumption of reduced susceptibility. In the next section we will compare, for a pathogen with four strains, the dynamics of the full system with that of the proposed approximations and Gog and Grenfell’s model.

## Implementation details

All numerical analyses were carried out using the MATLAB computing environment (The Mathworks, Natick, Massachusetts, USA) with numerical accuracy  $10^{-7}$ . Code is available upon request.

## Results

### Comparison of closures for a pathogen with four strains

We consider a simple system of four strains along a line. The strain space is given by  $K = \{1, 2, 3, 4\}$ , and strains adjacent on the line constitute the mutational neighborhood,

$$M_1 = \{2\}, \quad M_2 = \{1, 3\}, \quad M_3 = \{2, 4\}, \quad M_4 = \{3\}.$$

Adjacent strains also confer cross-immunity to each other, resulting in a tridiagonal cross-immunity matrix,

$$\sigma = \begin{pmatrix} 1 & s & 0 & 0 \\ s & 1 & s & 0 \\ 0 & s & 1 & s \\ 0 & 0 & s & 1 \end{pmatrix}.$$

The functions  $C(A, i, B)$ , describing the probability that a host’s immune status changes from  $A$  to  $B$  owing to infection with strain  $i$ , are determined according to Gog and Swinton [9],

$$C(A, i, B) = \begin{cases} \prod_{j \in B \setminus A} \sigma_{ij} \prod_{j \notin B} (1 - \sigma_{ij}) & \text{if } i \notin A \text{ and } A \subset B, \\ 0 & \text{otherwise.} \end{cases}$$

With such a small number of strains, the behavior of the full SIR model is tractable and can be used as a baseline reference. We can also examine how Gog and Grenfell’s reduced infectivity model performs if we consider it as an approximation to the full model with reduced susceptibility. We numerically solve the following equations for the time interval  $[0, T]$  and for a range of parameters:

1. The full SIR system: equations (1)–(2).

2. The Gog and Grenfell model: equations (10)–(11).
3. The approximation based on the order-1 independence closure: equations (3) and (5), where the  $\xi_{ij}$  are substituted according to (7).
4. The approximation based on the order-1 interpolation closure: equations (3) and (5), where the  $\xi_{ij}$  are substituted according to (8).
5. The approximation based on the order-2 independence closure: equations (3), (5), and (6), where the  $\xi_{ijk}$  are substituted according to (9).

The following parameters are kept fixed:  $\nu = 1$ ,  $\mu = 0$  and  $T = 40$ . We choose  $\beta_i = R_0$  for all  $i$ . We vary the transmission coefficient  $R_0$ , the cross-immunity coefficient  $s$ , and the mutation probability  $m$  in the following ranges:  $R_0 \in \{2, 3, 4, 5\}$ ,  $s \in \{0, 0.1, 0.2, \dots, 1\}$ , and  $m \in \{10^{-8}, 10^{-6}, 10^{-4}\}$ . Initially, 99% of the host population is fully susceptible to all strains, while 1% is infected with strain 1, is immune against it and fully susceptible to all other strains.

In order to assess how well the reduced models approximate the full model, we introduce one qualitative and one quantitative accuracy measure. We consider an epidemic detection threshold  $\varepsilon = 10^{-4}$ , describing the smallest proportion of infected hosts at which the disease can still be detected in the population. The results presented below are not particularly sensitive to the exact value of this parameter within three orders of magnitude (results not shown). We denote by  $f_i(t)$  the fraction of individuals infected with strain  $i$  as predicted by the full model, and by  $g_i(t)$  the same fraction as predicted by one of the reduced models. We consider four situations:

1. Neither  $f_i(t)$  nor  $g_i(t)$  exceed the epidemic detection threshold within the considered time interval. In this case we say that the approximation correctly captures the dynamics of the real system, both qualitatively and quantitatively.
2. Only  $f_i(t)$  exceeds the epidemic detection threshold within the considered time interval, while  $g_i(t)$  always stays below it. In this case, there is a qualitative difference in predictions, because the full model predicts an epidemic, while the reduced model does not. Then, if  $Y_i \subset [0, T]$  is the set of moments in time at which  $f_i(t) > \varepsilon$ , we define the qualitative accuracy measure as  $\rho_i = L(Y_i)/T$ , where  $L(Y_i)$  is the sum of the lengths of time intervals during which such qualitative difference is observed for the  $i$ th strain; in other words,  $L(Y_i)$  is the Lebesgue measure of the set  $Y_i$ .
3. Only  $g_i(t)$  exceeds the epidemic detection threshold within the considered time interval, while  $f_i(t)$  always stays below it. In this case, again, there is a qualitative difference in predictions, because the reduced model predicts an epidemic, while the full model does not. Then, if  $Y_i \subset [0, T]$  is the set of moments in time at which  $g_i(t) > \varepsilon$ , we define the qualitative measure analogously to the second case,  $\rho_i = L(Y_i)/T$ .

4. Both  $f_i(t)$  and  $g_i(t)$  exceed the epidemic detection threshold within the considered time interval. In this case we say that there is no qualitative difference in the model predictions, but there still could be a quantitative difference between them. We quantify the latter by the absolute error of  $g_i(t)$  with respect to  $f_i(t)$ ,

$$\Delta_i = \int_0^T |f_i(t) - g_i(t)| dt.$$

To obtain the overall qualitative error,  $\rho$ , we average  $\rho_i$  over all strains,  $\rho = 1/4 \sum_{i=1}^4 \rho_i$ . This accuracy measure is suitable when we care most about determining which strains will cause epidemics and which will not, while ignoring quantitative differences in epidemic size and timing.

To obtain the overall quantitative accuracy measure,  $\Delta$ , in a way that we could compare across models with different parameter values, we normalize  $\Delta_i$  by the total size of all epidemics,  $Z = \sum_{i=1}^4 \int_0^T f_i(t) dt$ , and sum over all strains,  $\Delta = \sum_{i=1}^4 \Delta_i / Z$ . This accuracy measure is suitable when we care most about correctly capturing the shape and timing of major epidemics, while ignoring possible mistakes in minor epidemics.

The results of this comparison are shown in Figures 1 through 3. As could have been expected, the order-1 interpolation closure performs better than the order-1 independence closure, both qualitatively and quantitatively. The order-1 independence closure is quite accurate for small cross-immunity coefficients, but is problematic for medium and large cross-immunity coefficients. The order-1 interpolation closure, in turn, is most problematic for intermediate cross-immunity coefficients, which is again consistent with expectations. Indeed, owing to the nature of the interpolation that we used to determine  $\hat{\xi}_{ij}$ , we expect this closure to be more accurate for extreme values of cross-immunity and less accurate for intermediate values. As expected, the order-2 approximation performs substantially better than both order-1 approximations and Gog and Grenfell's model, making no qualitative errors at all and incurring quantitative errors on the order of a few percent. Gog and Grenfell's model is problematic for small and intermediate values of the cross-immunity coefficient but, quite surprisingly, its error converges to zero for high values of cross-immunity where it becomes superior to the order-1 independence closure. This may have to do with the fact that the independence closure underestimates the level of cross-immunity of the population, especially for large values of cross-immunity.

We have conducted the same type of analysis for a circular four-strain system and for a two six-strain systems (see the supporting information text) – the results turn out to be qualitatively similar to those presented here. Based on these numerical investigations, we conclude that the approximate model resulting from the order-1 interpolation closure offers a good compromise between accuracy and computational effort. The proposed order-2 approximation can be used if higher accuracy is required. Although Gog and Grenfell's model can also be used to approximate the dynamics of systems developed under the reduced-susceptibility assumption, if the degree of cross-protection between strains is high, we find that models developed under different assumptions lead, in general, to qualitatively different dynamics (see also the supporting information text). Despite the

robustness of our conclusions on the quality of proposed closures with respect to key parameters within a sensible parameter region, further research is necessary to establish rigorous error bounds throughout the parameter space. In particular, we have not investigated the dependence of closure quality on the birth/mortality rate  $\mu$  – performance may worsen for large  $\mu$ . A more difficult, but equally important, question is how the topology of the strain space and of the cross-immunity structure alter the dynamics and the resultant performance of approximations in comparison with the full model.

## Application to antigenic drift in influenza A

The application presented in this section illustrates how one could capture some aspects of the phenomenon of antigenic drift in influenza A’s subtype H3N2 using models capable of tracing hundreds of variables. This is an attempt to extend the work by Gog and Grenfell [15] to the case of a two-dimensional viral phenotype space: those authors had shown that the dominant phenotype of the virus would move along a one-dimensional line in a two-dimensional strain space if it was assumed that viable phenotypes are concentrated in a narrow region along one of the axes. The structure of the influenza A phenotype space is not known, and it would be very surprising if all viable phenotypes were located only along a one-dimensional manifold. It is more likely that immuno-selection reduces the relative fitness of certain phenotypes at certain points in time [17]. Here, we construct an immunity-based mechanism that restricts the virus from exploring the whole phenotype space even if this space is mostly viable. We did not attempt to fit any parameters to real data, so that a quantitative correspondence, or even a close qualitative match, are not intended here.

We assume that the strain space of the considered influenza-like virus is a two-dimensional rectangular lattice, i.e., each variant of the virus is characterized by a pair of integers  $(i, j) \in Z^2$  and the phenotypic mutational neighborhood is given by the next-neighbor relation on the lattice,  $M_{(i,j)} = \{(i-1, j), (i+1, j), (i, j-1), (i, j+1)\}$ . While it is clear that the true strain space spanned by all viable conformations of the hemagglutinin protein does not resemble anything as simple as a two-dimensional rectangular lattice, by considering this example we still hope to shed some light on the mechanisms that govern the complicated evolutionary process of antigenic drift. This hope is not entirely unfounded. Plotkin et al. [39] have observed that changes in the influenza A hemagglutinin occur “along a different epitopic axis every 2-5 years”. Earlier, Wilson and Cox [29] have noticed that amino acid substitutions in two epitopes were on average necessary to produce a new antigenic type. In addition to that, Smith et al. [19] have argued that the antigenic space of influenza A’s subtype H3N2 is principally two-dimensional. Bringing these ideas together, we investigate influenza A evolution in an abstract two-dimensional phenotype space. We associate the movement along each dimension of this space with changes in states of a viral “effective epitope” (referred to as A and B below). The effective epitopes may not correspond to actual epitopes on the surface of the influenza A hemagglutinin protein (of which five are currently known) because not all epitopes may be immunogenic at all times. Below we drop the word “effective”,

but it is always implied in the context of our model. In our interpretation, each pair of indices  $(i, j) \in Z^2$  characterizing a particular strain implies that epitope A is in its  $i$ th conformation while epitope B is in its  $j$ th conformation.

By the same token, it is natural to incorporate another idea, introduced by Gupta and colleagues [3, 30] – the assumption that each individual that has been infected with strain  $(i, j)$  develops immunity against both epitopes, and thus gains full protection against all strains that possess the same conformation of either epitope. In other words,  $\sigma_{(i,j)(k,\ell)} = 1$  if either  $i = k$  or  $j = \ell$ , resulting in a cross-shaped cross-immunity structure (Figure 4). The specificity of antibodies determines how fast cross-immunity decays when the epitopes of the challenging strain differ from the epitopes of the immunogenic strain. We follow [15] in choosing a Gaussian function for the decay of cross-immunity, with mean zero and standard deviation  $a$ , implying that  $a$  is the cumulative number of conformation changes in one epitope that reduces cross-immunity to  $e^{-1/2} = 60.7\%$ ,

$$\sigma_{(i,j)(k,\ell)} = \exp \left\{ -\frac{1}{2} \left( \frac{\min(|i - k|, |j - \ell|)}{a} \right)^2 \right\}. \quad (12)$$

The cross-immunity matrix  $\sigma$  thus specified is sufficient for generating a one-dimensional trajectory of strain evolution: variants along the diagonal of strain space cause epidemics, whereas other variants do not (Figure S16). This conforms with intuition. Indeed, once a strain causes an epidemic, the host population acquires immunity not only against it but also against all the strains that share at least one epitope conformation with it, i.e., against all strains that have either the same  $x$ -coordinate or the same  $y$ -coordinate. Thus, the phenotypically closest mutant that can cause the next epidemic must differ from the current epidemic strain in both epitopes – so that this mutant can only be its diagonal neighbor.

In this simple model, strictly one strain dominates the host population at each epidemic season. In a slightly more general setting, several strains can coexist within an epidemic season [19]. To account for this possibility, we introduce some heterogeneity in the transmission coefficients of strains. This is indeed as expected in reality, since some strains are likely to be slightly more virulent than others, because the conformation of the hemagglutinin protein influences how effectively a virus can penetrate target cells [28, 29]. The amount of heterogeneity determines how many strains can coexist. In this version of our model, one to three strains usually coexist during an epidemic, while the principal component of the evolutionary trajectory remains one-dimensional (Figure 5). However, we conjecture that, if we considered an infinite rectangular lattice and started from the same initial condition, we would have observed either two lineages evolving in opposite directions along one of the two diagonals, or even four lineages evolving along both diagonals. We did not investigate this scenario in detail because it presumes that there is no immunity in the population whatsoever at the initial time point. This may be true for real influenza right after a reassortment event, but the evolutionary forces active during and right after reassortment are probably quite different from those governing the subsequent antigenic drift, and go far beyond the subject of our concern here. Thus, considering only

the first quadrant of the phenotype space when starting viral evolution at the origin is equivalent to assuming that antigenic drift has already been going on for some time. The question thus addressed is how antigenic drift with limited diversity can be sustained.

To show that this result is a consequence of the immunity structure suggested here, rather than just a peculiarity of the considered equations, we have simulated exactly the same system for Gog and Grenfell’s equations and for the model with no coinfections. We thus can show that the results reported above are qualitatively robust to model choice (see the supporting information text). Moreover, the assumptions of our model can be relaxed along two directions: (a) we may allow for mutations to more and more distant neighbors, with decreasing probabilities; and (b) cross-immunity may be allowed to decay slowly along the epitopic axes. As long as the immunity neighborhood extends further along the epitopic axes than the mutational neighborhood, our qualitative conclusions hold (results not shown). On the contrary, if the local cross-immunity structure suggested in [15] is used instead of the epitope-based one, viral evolution is no longer contained to a one-dimensional manifold (see the supporting information text).

The real influenza virus, which is likely to live in an approximately two-dimensional strain space, appears to experience a selection regime that, while allowing for the temporary coexistence of a small number of variants, constrains the long-term evolution of the virus to a single branch. Our model suggests a possible mechanism for explaining this surprising reduction. In particular, we conjecture that two ingredients are responsible for the associated evolutionary dynamics:

1. The first ingredient is the non-local nature of the immune response after an infection. This results from the fact that cross-immunity protects hosts not only against strains that are very similar to the infecting strain, but also against strains that are quite distant from it in strain space, as long as at least one of their epitope conformations resembles that of the infecting strain. Non-locality of the immune response prevents the virus from conquering the entire trait space.
2. The second ingredient, which enables temporary coexistence of several strains, is the heterogeneity of transmission coefficients in trait space. Paradoxically, this heterogeneity occasionally leads to temporary parity among the effective reproduction ratios of different strains. The effective reproduction ratio of a particular strain is the quantity that determines whether this strain takes off and causes an epidemic or dies out without ever reaching the epidemic threshold [40]. In our model, the effective reproduction ratio  $R_{(i,j)}$  of strain  $(i, j)$  equals  $\beta_{(i,j)}S_{(i,j)}/\nu$ . When all strains possess the same transmission coefficient, effective reproduction ratios thus depend only on the fractions of susceptible individuals. Clearly, the pool of susceptibles to the diagonal strains is larger than the pool of susceptibles to nearby off-diagonal strains, because past infections have induced low cross-immunity against the former and high cross-immunity against the latter. Hence, the diagonal strains successively cause epidemics, while the off-diagonal strains do not – accordingly, no polymorphism can emerge, not even in the short term. By contrast, in a trait space that is heterogeneous with respect to the transmission coefficient, relatively small values

of  $S_{(i,j)}$  for off-diagonal strains may occasionally be compensated by high values of  $\beta_{(i,j)}$ . In this manner, the effective reproduction ratios for off-diagonal strains may become comparable to those for diagonal strains. If, in addition, strains with comparable effective reproduction ratios start from comparable initial conditions, they reach epidemic values around the same time. This leads to short-term polymorphisms.

## Discussion

The approach introduced here enables systematic reductions in the complexity of status-based models of multi-strain pathogens. It is applicable to models with or without coinfections, with reduced susceptibility or reduced infectivity. If coinfections are allowed and reduced infectivity is assumed, our approach coincides with that of Gog and Grenfell [15]. The key is to depart from the traditional compartment models by adopting the viewpoint of the pathogen. This allows restating full status-based models in terms of immunity variables and truncating the hierarchy of resultant equations at the desired order. Overcoming the exponential explosion of state variables in traditional models of multi-strain pathogens, we have shown that the complexity of our approximation grows only algebraically, i.e., proportionally to powers of the number of strains in the system. Numerically solving the resulting approximate equations, we show that they relatively well mimic the dynamics of full systems, even when the first-order closures are used.

It would be interesting to perform a rigorous mathematical analysis of the behavior of our approximations and compare it to the behavior of the full system [9]. In theory, the behavior of the full model should be mimicked more and more accurately when the chain of equations for immunity variables is truncated at higher and higher orders. Indeed, when sufficiently many of them are satisfied exactly, one obtains the original full model expressed in terms of immunity variables.

In this work we presented a state-space reduction approach that applies only to the class of status-based models. We focused on this class of models for two reasons. Apart from easier mathematical treatment, there is an important conceptual difference that favors status-based models, at least under the reduced susceptibility assumption. Consider a situation when a host is repeatedly challenged with a strain. In the history-based approach, the probability for a host to acquire an infection remains the same across successive challenges. Thus, a host that has successfully used cross-reacting antibodies to repel one or more challenges from a particular pathogen is just as likely to be infected at the next challenge as is a host with the same infection history that has never seen this particular pathogen. To us, the status-based assumption – that, if antibodies fend off the first challenge, subsequent challenges will fail too – seems more realistic. Nevertheless, it would be interesting to know whether a state-space reduction approach similar to the one presented here could also be applied to history-based models.

Using our framework, we have investigated a potentially general mechanism for constraining viral evolution to one-dimensional manifolds when the underlying strain space

is two-dimensional. Based on general knowledge about the antigenic space of the real influenza virus, we considered a hypothetical influenza-like virus whose phenotype space was given by a regular lattice with the same dimensionality as the number of principal components of the virus' antigenic space. We associate the movement along the axes of the resultant phenotype space with changes in the conformation of the viral "effective epitopes". We based our analysis on two plausible qualitative assumptions: (a) during an infection, immunity is independently generated against all effective epitopes, and (b) immunity against one effective epitope suffice for full protection against viruses with similar epitopes. The resultant cross-shaped cross-immunity neighborhood drives the evolution of the virus along the diagonal of the phenotype space. This observation offers a conceptually simple approach to understanding single-trunk phylogenies of infectious pathogens.

Qualitatively different hypotheses for explaining the single-trunk phylogeny of influenza were introduced earlier on by Ferguson et al. [10] and, recently, by Koelle et al. [26]. Ferguson et al.'s hypothesis was later given a theoretical justification by Andreasen and Sasaki [24]. Ferguson et al. suggested that a presumed strain-independent short-lived form of cross-immunity could be the decisive factor for restricting the diversity of influenza strains. In their model, strains were described by genetic sequences and an ad hoc rule was considered to translate the genetic distance between strains into an antigenic distance. In reality, however, this translation must be expected to be (a) highly degenerate – the antigenic space has only two principal dimensions – and (b) highly nonlinear – small genetic changes will sometimes cause large antigenic changes and vice versa [19]. Failure to take into account these two factors results in an inflation in the dimensionality of the strain space of influenza, which, in turn, may lead to a necessity for introducing additional assumptions like the existence of strain-independent short-lived cross-immunity. Koelle et al.'s recent work [26] is an excellent demonstration of this point. The authors explicitly modeled the highly nonlinear and highly degenerate mapping from genotypes to phenotypes and thus were able to obtain single-trunk phylogenies without the assumption of strain-independent cross-immunity. Their results constitute an important step toward understanding the evolution of influenza A. However, it still remains to be understood how exactly the influenza virus moves through the phenotype space in such a way that its diversity remains restricted. In particular, what are the crucial properties of the phenotype space's topology that allow for such a peculiar evolution? Our work outlines one possible set of such properties. Namely, if each independent direction of the antigenic space is associated with an effective epitope, and if the effective epitopes are immunogenically independent, then a one-dimensional trajectory of antigenic drift naturally emerges from the epitope-based cross-immunity structure. Our model is, of course, simplistic, and much more needs to be done to improve our understanding of the phenotype evolution of influenza. In particular, large-scale immunological experiments, as well as further studies of the protein folding process, will be necessary to uncover the actual topology of the influenza virus strain space, its cross-immunity structure, and the types of immunity that need to be considered.

# Supporting Information

**Text S1.** Supporting information text and figures.

## Acknowledgments

The authors express thanks to Sarah Cobey, Arkadii Kryazhimskiy, and Pleuni Pennings for fruitful discussions during the initial stages of this work, to Burton Singer for helpful comments on an early version of this manuscript, and to Julia Gog and two anonymous reviewers for valuable comments. Miriam Dushoff also made comments on the manuscript. SK gratefully acknowledges financial support by the Burroughs Wellcome Fund Training Program in Biological Dynamics (#1001782), by the DARPA grant #HR0011-05-1-0057, and by IIASA's US National Member Organization during his stay at the International Institute for Applied Systems Analysis. UD gratefully acknowledges financial support by the Vienna Science and Technology Fund, WWTF. SL and JD gratefully acknowledge financial support by the NIH grant #P50 GM071508 and by the DARPA grant #HR0011-05-1-0057.

## References

- [1] Nowak MA, May RM, Phillips RE, Rowland-Jones S, Lalloo DG, et al. (1995) Antigenic oscillations and shifting immunodominance in HIV-1 infections. *Nature* 375:606–611.
- [2] Earn DJD, Dushoff J, Levin SA (2002) Ecology and evolution of the flu. *Trends Ecol Evol* 17:297–345.
- [3] Gupta S, Maiden MCJ, Feavers IM, Nee S, May R, et al. (1996) Maintenance of strain structure in populations of recombining infectious agents. *Nature Med* 2:437–442.
- [4] Richardson AR, Yu Z, Popovic T, Stojiljkovic I (2002) Mutator clones of *neisseria meningitidis* in epidemic serogroup A disease. *Proc Natl Acad Sci USA* 99:6103–6107.
- [5] Levin SA, Dushoff J, Plotkin JB (2004) Evolution and persistence of influenza A and other diseases. *Math Biosci* 188:17–28.
- [6] Rohani P, Green CJ, Matilla-Beniers NB, Grenfell BT (2003) Ecological interference between fatal diseases. *Nature* 422:885–888.
- [7] Pease CM (1987) An evolutionary epidemiological mechanism with applications to type A influenza. *Theor Pop Biol* 31:422–452.
- [8] Andreasen V, Lin J, Levin SA (1997) The dynamics of cocirculating strains conferring partial cross-immunity. *J Math Biol* 35:825–842.
- [9] Gog JR, Swinton J (2002) A status-based approach to multiple strain dynamics. *J Math Biol* 44:169–184.
- [10] Ferguson NM, Galvani AP, Bush RM (2003) Ecological and immunological determinants of influenza evolution. *Nature* 422:428–433.
- [11] Cisternas J, Gear CW, Levin SA, Kevrekidis IG (2004) Equation-free modelling of evolving diseases: coarse-grained computations with individual-based models. *Proc R Soc London A* 460:2761–2779.
- [12] Tria F, Lässig M, Peliti L, Franz S (2005) A minimal stochastic model for influenza evolution. *J Stat Mech* P07008.
- [13] Buckee CO, Koelle K, Mustard MJ, Gupta S (2004) The effects of host contact network structure on pathogen diversity and strain structure. *Proc Natl Acad Sci USA* 101:10839–10844.
- [14] Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. *Proc R Soc London A* 115:700–721.

- [15] Gog JR, Grenfell BT (2002) Dynamics and selection of many-strain pathogens. *Proc Natl Acad Sci USA* 99:17209–17214.
- [16] Couch RB, Kasel JA (1983) Immunity to influenza in man. *Annu Rev Microbiol* 37:529–549.
- [17] Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, et al. (2004) Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* 303:327–332.
- [18] Lapedes A, Farber R (2001) The geometry of shape space: application to influenza. *J Theor Biol* 212:57–69.
- [19] Smith DJ, Lapedes AS, de Jong JC, Bestebroer TM, Rimmelzwaan GF, et al. (2004) Mapping the antigenic and genetic evolution of influenza virus. *Science* 305:371–376.
- [20] Fitch WM, Bush RM, Bender CA, Cox NJ (1997) Long term trends in the evolution of H(3)NA1 human influenza type A. *Proc Natl Acad Sci USA* 94:7712–7718.
- [21] Recker M, Pybus OG, Nee S, Gupta S (2007) The generation of influenza outbreaks by a network of host immune responses against a limited set of antigenic types. *Proc Natl Acad Sci USA* 104:7711–7716.
- [22] Lin J, Andreasen V, Casagrandi R, Levin SA (2003) Traveling wave in a model of influenza A drift. *J Theor Biol* 222:437–445.
- [23] Andreasen V (2003) Dynamics of annual influenza A epidemics with immunoselection. *J Math Biol* 46:504–536.
- [24] Andreasen V, Sasaki A (2006) Shaping the phylogenetic tree of influenza by cross-immunity. *Theor Pop Biol* 70:164–173.
- [25] Gökaydin D, Oliveira-Martins JB, Gordo I, Gomes MGM (2007) The reinfection threshold regulates pathogen diversity: the case of influenza. *J R Soc Interface* 4:137–142.
- [26] Koelle K, Cobey S, Grenfell BT, Pascual M (2006) Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. *Science* 314:1898–1903.
- [27] Wiley DC, Wilson IA, Skehel JJ (1981) Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature* 289:373–378.
- [28] Wiley DC, Skehel JJ (1987) The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. *Annu Rev Biochem* 56:365–394.
- [29] Wilson IA, Cox NJ (1990) Structural basis of immune recognition of influenza virus hemagglutinin. *Annu Rev Immunol* 8:737–771.

- [30] Gupta S, Ferguson N, Anderson R (1998) Chaos, persistence, and evolution of strain structure in antigenically diverse infectious agents. *Science* 280:912–915.
- [31] Castillo-Chavez C, Hethcote HW, Andreasen V, Levin SA, Liu WM (1989) Epidemiological models with age structure, proportionate mixing, and cross-immunity. *J Math Biol* 27:233–258.
- [32] Ferguson NM, Andreasen V (2002) The influence of different forms of cross-protective immunity on the population dynamics of antigenically diverse pathogens. In: Blower S, Castillo-Chavez C, editors, *Mathematical approaches for emerging and re-emerging infectious diseases: models, methods and theory*, New York: Springer, volume 125. pp. 157–169.
- [33] Dawes JHP, Gog JR (2002) The onset of oscillatory dynamics in models of multiple disease strains. *J Math Biol* 45:471–510.
- [34] Lin J, Andreasen V, Levin SA (1999) Dynamics of influenza A drift: the linear three-strain model. *Math Biosci* 162:33–51.
- [35] Gomes MGM, Medley GF, Nokes DJ (2002) On the determinants of population structure in antigenically diverse pathogens. *Proc R Soc London A* 269:227–233.
- [36] Erickson MJ (1996) *Introduction to combinatorics*. Wiley-Interscience. John Wiley & Sons, Inc.
- [37] Bolker BM, Pacala SW, Levin SA (2000) Moment methods for ecological processes in continuous space. In: Dieckmann U, Law R, Metz JAJ, editors, *The geometry of ecological interactions: simplifying spatial complexity*, Cambridge University Press. pp. 388–411.
- [38] Keeling MJ (2000) Multiplicative moments and measures of persistence in ecology. *J Theor Biol* 205:269–281.
- [39] Plotkin J, Dushoff J, Levin SA (2002) Hemagglutinin sequence clusters and the antigenic evolution of influenza A virus. *Proc Natl Acad Sci USA* 99:6263–6268.
- [40] Anderson RM, May RM (1991) *Infectious diseases of humans: dynamics and control*. Oxford University Press.

## Tables

**Table 1.** Notations used throughout this study.

## Figures

**Figure 1.** Qualitative accuracy measure  $\rho$  (vertical axis) in dependence on the cross-immunity coefficient  $s$  (horizontal axis), the transmission coefficient  $R_0$  (rows) and the mutation probability  $m$  (columns). Each panel shows the performance of four models: Gog and Grenfell’s model (circles), order-1 independence closure (pluses), order-1 interpolation closure (squares), order-2 independence closure (triangles). The highlighted region indicates the parameter combination  $m = 10^{-6}$ ,  $R_0 = 5$ , and  $s = 0.6$ , for which the system’s dynamics are shown in Figure 3.

**Figure 2.** Quantitative accuracy measure  $\Delta$  for the performance of four models. Details as in Figure 1.

**Figure 3.** Dynamics of the full four-strain SIR model (black curves) compared with its approximations (thick gray curves). Solution of the full model, Gog and Grenfell’s model, order-1 independence closure model, order-1 interpolation closure model, and order-2 independence closure model are denoted by “F”, “GG”, “O1A”, “O1B”, and “O2”, respectively. Parameter values:  $m = 10^{-6}$ ,  $R_0 = 5$ , and  $s = 0.6$ .

**Figure 4.** Cross-immunity structure. Strains to which strain  $(0, 0)$  confers full, 37%, 2% or no cross-immunity are shown in white, light gray, dark gray, and black, respectively. **A.** Our model. **B.** Gog and Grenfell’s model [15].

**Figure 5.** Approximate dynamics of antigenic drift in influenza A, based on the order-1 interpolation closure. Parameter values:  $\mu = 0$ ,  $\nu = 1$ ,  $m = 10^{-4}$ , and  $a = 1/\sqrt{2}$ ; the heterogeneous transmission coefficients  $\beta_{(i,j)}$  were drawn from a normal distribution with mean 3 and standard deviation 0.5. The numerical solution for the time interval  $t \in [0, 100]$  was obtained for a strain space given by a 20-by-20 rectangular lattice. The initial condition was given by all state variables being zero except for  $I_{(1,1)}(0) = 0.01$  and  $\xi_{(1,1)}(0) = 0.01$ , corresponding to a healthy and fully susceptible host population with 1% of hosts infected with strain  $(1, 1)$ . **A.** Strains whose maximum epidemic size exceeded 0.01 are shown. The gray shade indicates the maximum epidemic size; the number above each shaded square indicates the time when the maximum of the epidemic for that particular strain was reached. Circles indicate strains whose transmission coefficients are less than 3; crosses indicate strains with transmission coefficients greater than 3. **B.** The sum of all proportions of infectious hosts as a function of time.

Parameters	
$K$	Set of all pathogen strains
$\beta_i$	Transmission coefficient of pathogen strain $i \in K$
$m$	Probability of pathogen mutation during infection
$M_i$	Mutational neighborhood of pathogen strain $i \in K$ , i.e., the set of all strains that can mutate into $i$ (and to which $i$ can mutate)
$\nu$	Per capita recovery rate of host
$\mu$	Per capita birth and death rate of host
$C(A, i, B)$	Fraction of hosts that, when being immune against strains in $A \subset K$ and becoming infected with strain $i \in K$ become immune against strains in $B \subset K$ after the infection
$\sigma_{ij}$	Probability that hosts obtain immunity against strain $j \in K$ after an infection with strain $i \in K$
Variables	
$S_A$	Fraction of hosts that are immune against strains in the set $A \subset K$
$I_i$	Fraction of hosts infectious with strain $i$
$\Lambda_i$	Force of infection for strain $i$
$\xi_{i_1 i_2 \dots i_\ell}$	Fraction of hosts that are immune against strains $i_1, i_2, \dots, i_\ell \in K$

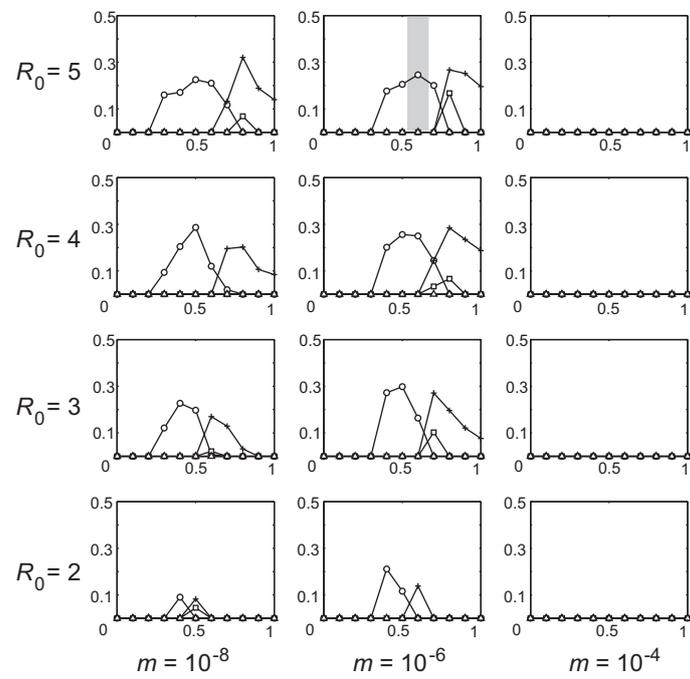


Figure 1

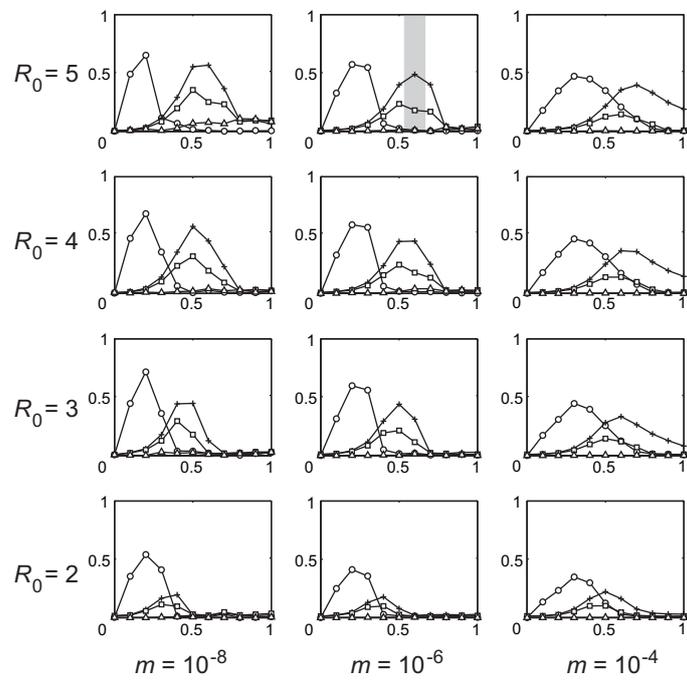


Figure 2

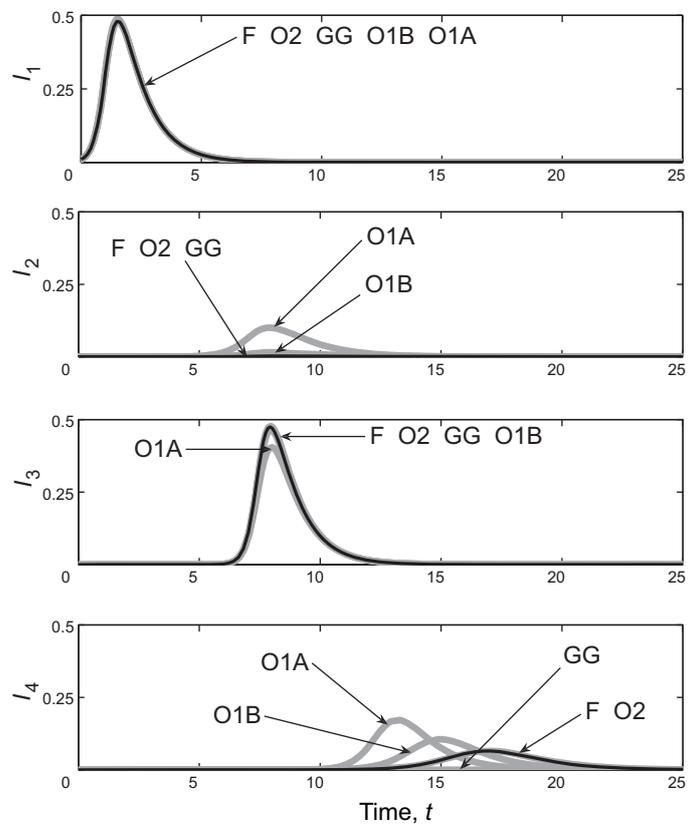


Figure 3

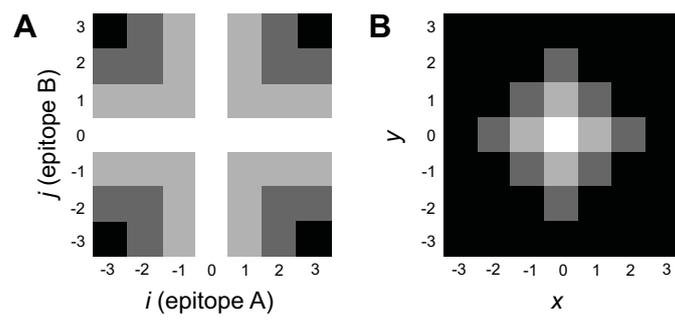


Figure 4

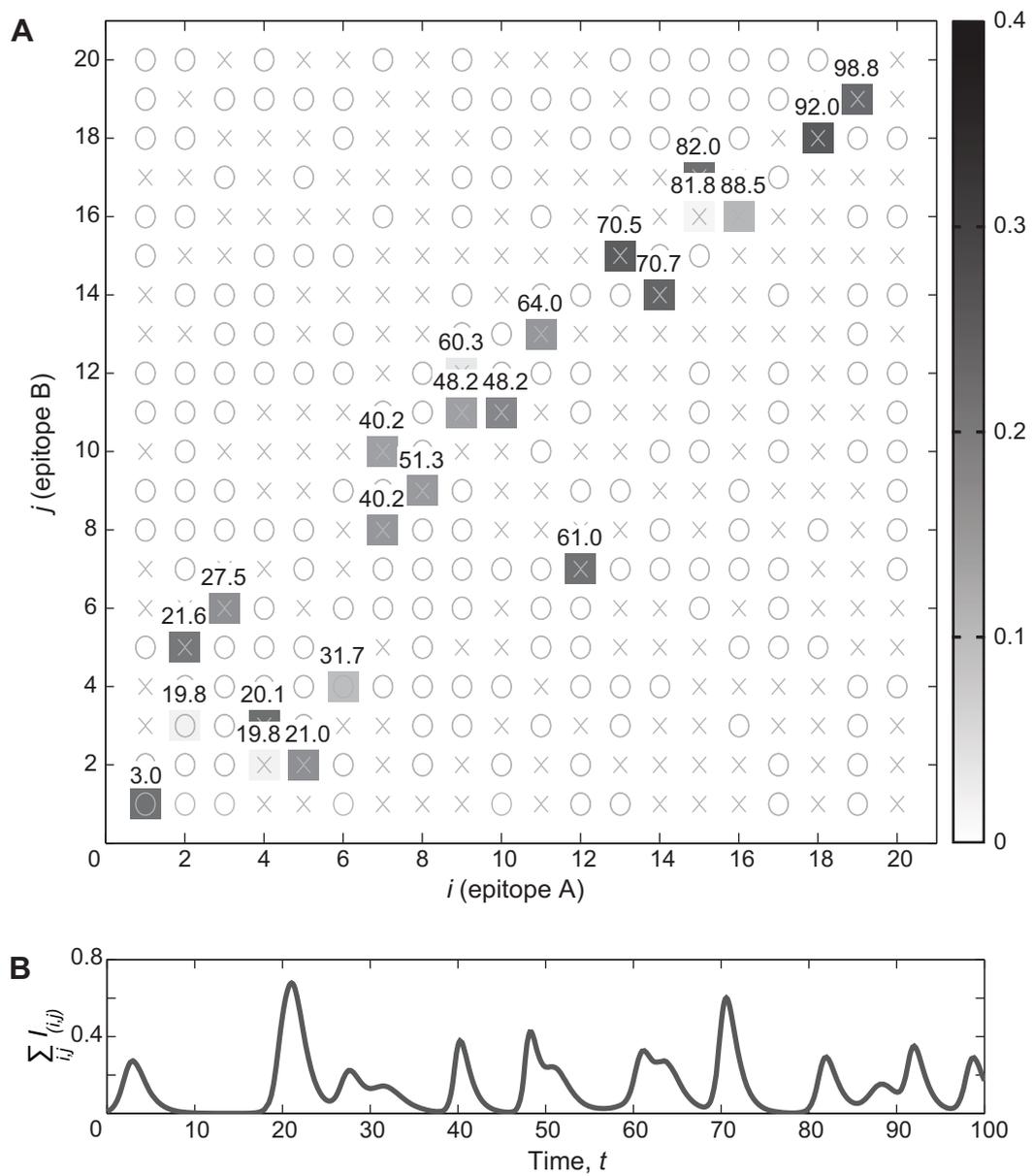


Figure 5

# On state-space reduction in multi-strain pathogen models, with an application to antigenic drift in influenza A

Sergey Kryazhimskiy\*, Ulf Dieckmann,  
Simon A. Levin, Jonathan Dushoff

## Supporting information

**This PDF file includes:**  
Supporting information text  
Figures S1 – S19

---

\*To whom correspondence should be addressed. E-mail: [skryazhi@princeton.edu](mailto:skryazhi@princeton.edu)

This supporting information document contains four sections. The first section provides a derivation of the expansion in immunity variables of the model with coinfections and reduced infectivity and demonstrates its connection to the reduced model presented in [S1]. In the second section we elucidate the expansion in immunity variables for a model without coinfections. In the third section we present an extended analysis of accuracy for Gog and Grenfell’s model [S1], as well as for order-1 and order-2 closures in models with and without coinfections. The last section contains an additional analysis of the influenza A drift model.

## State space reduction in a model with coinfections and reduced infectivity

In this section we derive the expansion in immunity variables for the full status-based model with coinfections and reduced infectivity and show that the truncation of this expansion at the first order leads to the model studied by Gog and Grenfell [S1]. Our approach provides a generalization of that model since immunity variables allow us to truncate the chain of equations at higher orders if a more detailed description of the immunity structure of the population is desired.

The general status-based model with coinfection was formulated by Gog and Swinton [S2]. However, it was constructed under the assumption of reduced susceptibility. The assumption of reduced infectivity enters in the definition of the function  $C(A, i, B)$  that determines the proportion of the host population that, after an infection with strain  $i$ , changes its immune status from  $A$  to  $B$ , as well as in the properties of the function  $C_i^*(A, k)$  (see below). The reduced infectivity assumption implies that  $C(A, i, B)$  is allowed to take non-zero values only if:

1.  $i \in B$ ;
2.  $A \subseteq B$ .

The difference between this and the assumptions on  $C$  made in [S2] is that  $C(A, i, B)$  can take non-zero values if  $i \in A$ . Biologically, this implies that individuals are allowed to change their immune status due to a mere exposure to a strain even if this strain is already included in their immune status (see also the discussion in the main text). Mathematically, this assumption implies that the equations for the class  $S_A$  are different from equations (1) in the main text of our paper in that the second sum is taken over all subsets of the set of strains and the last sum is taken over all strains. Equations for  $I_i$  remain the same.

$$\dot{S}_A = \mu(\delta_{A,\emptyset} - S_A) + \sum_{k \in K} \sum_{B \subset K} S_B \Lambda_k C(B, k, A) - \sum_{k \in K} \Lambda_k S_A, \quad \text{for all } A \subset K, \quad (\text{S1})$$

$$\dot{I}_i = \Lambda_i \sum_{B: i \in K \setminus B} S_B - (\nu + \mu) I_i, \quad \text{for all } i \in K. \quad (\text{S2})$$

Analogously to the derivation in the main text, we define the probability  $C_{i_1 i_2 \dots i_\ell}^*(A, k)$  of obtaining immunity against strains  $i_1, i_2, \dots, i_\ell$  (all different) after an infection with strain  $k$  for a host that had immune status  $A$  prior to the infection. But now, according to the assumption of reduced infectivity, we assume that

$$C_{i_1 i_2 \dots i_\ell}^*(A, k) = \prod_{\substack{j=1,2,\dots,\ell: \\ i_j \notin A \cup \{k\}}} \sigma_{k i_j}.$$

This implies that (a) the chance of obtaining immunity against strain  $i$  after the infection with the strain  $k$  does not depend on the presence or absence of immunity against other strains, and (b) if the host is already immune to strain  $k$  (i.e.,  $k \in A$ ), the chance of getting immunity to strain  $i$  through cross-protection between strains  $k$  and  $i$ , is proportional to  $\sigma_{k i}$ . This is a biologically subtle point of the reduced-infectivity assumption (see the discussion in the main text).

Analogously to the derivation in the main text, to obtain the reduced version of the model (S1)–(S2), we introduce the immunity variables

$$\xi_i = \sum_{A:i \in A} S_A, \quad \xi_{ij} = \sum_{\substack{A \subset K: \\ i, j \in A}} S_A, \quad \dots,$$

in terms of which we rewrite the system (S1)–(S2). The equation for  $I_i$  is transformed straightforwardly once one recalls that  $\sum_{A \subset K} S_A = 1$ .

$$\dot{I}_i = \Lambda_i \left( \sum_{B \subset K} S_B - \sum_{B:i \in B} S_B \right) - (\nu + \mu) I_i = \Lambda_i (1 - \xi_i) - (\nu + \mu) I_i. \quad (\text{S3})$$

Differentiating the definition of  $\xi_i$  with respect to time and using (S1), we obtain

$$\dot{\xi}_i = \mu \sum_{A:i \in A} (\delta_{A, \emptyset} - S_A) + \sum_{A:i \in A} \sum_{k, B} S_B \Lambda_k C(B, k, A) - \sum_{A:i \in A} \sum_{k \in K} \Lambda_k S_A, \quad (\text{S4})$$

We shall consider the three sums one after another. The first sum yields

$$\sum_{A:i \in A} (\delta_{A, \emptyset} - S_A) = -\xi_i$$

since the sum of  $\delta_{A, \emptyset}$ -terms gives zero. The second term can be transformed in the following way,

$$\begin{aligned} & \sum_{A:i \in A} \sum_{k, B} S_B \Lambda_k C(B, k, A) = \sum_k \sum_B S_B \Lambda_k C_i^*(B, k) \\ & = \sum_{k \in K} \left( \sum_{B:i \in B} S_B \Lambda_k C_i^*(B, k) + \sum_{B:i \in K \setminus B} S_B \Lambda_k C_i^*(B, k) \right) \end{aligned}$$

$$\begin{aligned}
&= \sum_{k \in K} \left( \sum_{B:i \in B} S_B \Lambda_k + \sum_{B:i \in K \setminus B} S_B \Lambda_k \sigma_{ki} \right) \\
&= \sum_{k \in K} (\xi_i \Lambda_k + (1 - \xi_i) \Lambda_k \sigma_{ki})
\end{aligned}$$

The transformation of the last sum in (S4) is also straightforward:

$$\sum_{A:i \in A} \sum_{k \in K} \Lambda_k S_A = \sum_{k \in K} \Lambda_k \xi_i.$$

Gathering all the terms above, we obtain:

$$\dot{\xi}_i = \sum_{k \in K} \Lambda_k \sigma_{ki} (1 - \xi_i) - \mu \xi_i. \tag{S5}$$

The equation for  $\dot{\xi}_{ij}$  is obtained analogously,

$$\begin{aligned}
\dot{\xi}_{ij} &= \sum_{k \in K} \Lambda_k [\sigma_{ki} (\xi_j - \xi_{ij}) + \sigma_{kj} (\xi_i - \xi_{ij}) + \sigma_{ki} \sigma_{kj} (1 - \xi_i - \xi_j + \xi_{ij})] - \\
&\quad - \mu \xi_{ij} \quad \text{for all } i, j \in K, i \neq j. \tag{S6}
\end{aligned}$$

Continuing this chain, we would obtain the full system (S1)–(S2) expressed in terms of immunity variables. Note, however, that equations for immunity variables of any particular order are uncoupled from equations for immunity variables of higher orders. Therefore, the chain of equations truncated at a particular order would exactly represent the dynamics of the immunity variables up to that order. In particular, if we truncate the chain of equations at order 1, we can express (S3), (S5) in terms of  $I_i$  and  $S_i = 1 - \xi_i$ . This leads to the model studied by Gog and Grenfell [S1].

$$\begin{aligned}
\dot{S}_i &= \mu(1 - S_i) - \sum_{k \in K} \Lambda_k \sigma_{ki} S_i, \\
\dot{I}_i &= \Lambda_i S_i - (\nu + \mu) I_i,
\end{aligned}$$

for all  $i \in K$ .

## State space reduction in models with no coinfections

In this section we derive the expansion in immunity variables for a model with no coinfections under the reduced-susceptibility assumption. At the end of the section we also provide, without a derivation, the expansion in immunity variables for the model with no coinfections and reduced infectivity. We start out from a model with no coinfection that is analogous to the model with coinfections considered by Gog and Swinton [S2]. Here, we have to slightly change the meaning of our notations.  $S_A$  represents the proportion of

currently *non-infected* hosts that possess immune status  $A \subset K$  and that, therefore, are currently fully susceptible to all strains in the subset  $K \setminus A$  (we refer to these individuals as being *in state A*); and  $I_A^i$  is the proportion of individuals in the host population that are currently infected with strain  $i$  and had immune status  $A \subset K$  before the current infection (which implies  $i \notin A$ ). Since all host individuals naturally fall in exactly one of these classes, we have

$$\sum_{A \subset K} S_A + \sum_{i \in K} \sum_{A \subset K \setminus \{i\}} I_A^i = 1. \quad (\text{S7})$$

The proportion of hosts that recover to state  $B \subset K$ , having been in state  $A \subset K$  when they were infected by strain  $i \in K$ , is given by  $C(A, i, B)$  that has the properties described in the main text. The following is the full system of equations for the model with no coinfections.

$$\begin{aligned} \dot{S}_A &= \mu(\delta_{A, \emptyset} - S_A) + \nu \sum_{k \in K} \sum_{B \subset K \setminus \{k\}} I_B^k C(B, k, A) \\ &\quad - \sum_{k \in K \setminus A} \Lambda_k S_A, \quad \text{for all } A \subset K, \end{aligned} \quad (\text{S8})$$

$$\dot{I}_A^i = \Lambda_i S_A - (\nu + \mu) I_A^i, \quad \text{for all } i \in K \text{ and } A \subset K \setminus \{i\}, \quad (\text{S9})$$

where

$$\Lambda_i = \beta_i \left( (1 - m) \sum_{A \in K \setminus \{i\}} I_A^i + \sum_{j \in M_i} \sum_{A \in K \setminus \{j\}} \frac{m}{|M_j|} I_A^j \right). \quad (\text{S10})$$

Altogether, this system consists of  $2^n + n2^{n-1} - 1$  equations, where  $n$  is the number of strains.

We now rewrite this system in terms of the proportions of hosts infected with strain  $i$ ,

$$I_i = \sum_{A \subset K \setminus \{i\}} I_A^i,$$

and the immunity variables,

$$\begin{aligned} \xi_i &= \sum_{\substack{A \subset K: \\ i \in A}} S_A, & \xi_{ij} &= \sum_{\substack{A \subset K: \\ i, j \in A}} S_A, & \dots, \\ \eta_i^k &= \sum_{\substack{A \subset K \setminus \{k\}: \\ i \in A}} I_A^k, & \eta_{ij}^k &= \sum_{\substack{A \subset K \setminus \{k\}: \\ i, j \in A}} I_A^k, & \dots \end{aligned}$$

where  $\xi_{i_1 i_2 \dots i_\ell}$  describes the proportion of hosts that are currently not infected and have immunity against strains  $i_1, i_2, \dots, i_\ell \in K$ ; and  $\eta_{i_1 i_2 \dots i_\ell}^k$  describes the proportion of hosts that are currently infected with strain  $k$  and have immunity against strains  $i_1, i_2, \dots, i_\ell \in K$ . We will refer to immunity variables  $\xi_{i_1 i_2 \dots i_\ell}$  and  $\eta_{i_1 i_2 \dots i_\ell}^k$  as being of order  $\ell$ . As in the model with coinfections, all immunity variables are symmetric with respect to the permutation of

their subindices. Immunity variables with duplicate subindices  $i$  remain unchanged when the duplicate subindex is removed,  $\xi_{\dots i \dots i \dots} = \xi_{\dots i \dots}$  and  $\eta_{\dots i \dots i \dots}^k = \eta_{\dots i \dots}^k$ . When a strain index  $k$  appears in both an  $\eta$ -variable's superscript and subscript, the  $\eta$ -variable must be zero,  $\eta_{\dots k \dots}^k = 0$ . By definition, the immunity variables satisfy monotonicity conditions,

$$1 \geq H \geq \xi_{i_1} \geq \xi_{i_1 i_2} \geq \dots$$

and

$$1 \geq I_k \geq \eta_{i_1}^k \geq \eta_{i_1 i_2}^k \geq \dots$$

for all pairwise different  $k, i_1, i_2, \dots \in K$ . Here  $H = 1 - \sum_{i \in K} I_i$  is the proportion of healthy (i.e., not infected) host individuals.

Recalling that  $\sum_{A \subset K} S_A + \sum_{i \in K} I_i = 1$ , we easily obtain the equations for  $\dot{I}_i, \dot{\eta}_i^k, \dot{\eta}_{ij}^k$  etc.,

$$\dot{I}_i = \Lambda_i S_i - (\nu + \mu) I_i \quad \text{for all } i \in K, \quad (\text{S11})$$

$$\dot{\eta}_i^k = \Lambda_k (\xi_i - \xi_{ik}) - (\nu + \mu) \eta_i^k \quad \text{for all } i, k \in K; i \neq k, \quad (\text{S12})$$

$$\dot{\eta}_{ij}^k = \Lambda_k (\xi_{ij} - \xi_{ijk}) - (\nu + \mu) \eta_{ij}^k \quad \text{for all } i, j, k \in K; i, j, k \text{ pairwise different,} \quad (\text{S13})$$

...

where  $S_i = H - \xi_i$  is the fraction of individuals that are currently healthy and susceptible to strain  $i$ . The force of infection for strain  $i$ , expressed in terms of new variables, is

$$\Lambda_i = \beta_i \left( (1 - m) I_i + \sum_{j \in M_i} \frac{m}{|M_j|} I_j \right). \quad (\text{S14})$$

The derivation of the equations for  $\dot{\xi}_i, \dot{\xi}_{ij}$  etc. is more technical. Here, we restrict ourselves to an explicit derivation of the equation for  $\dot{\xi}_i$ . For this purpose, we apply the time derivative to the definition of  $\xi_i$  and use (S8), to obtain

$$\dot{\xi}_i = \mu \sum_{\substack{A \subset K: \\ i \in A}} (\delta_{A, \emptyset} - S_A) + \nu \sum_{\substack{A \subset K: \\ i \in A}} \sum_{k \in K} \sum_{B \subset K \setminus \{k\}} I_B^k C(B, k, A) - \sum_{\substack{A \subset K: \\ i \in A}} \sum_{k \in K \setminus A} \Lambda_k S_A. \quad (\text{S15})$$

The first term in (S15) accounts for the depletion, due to deaths, of healthy host individuals that are immune to strain  $i$ ,

$$\mu \sum_{\substack{A \subset K: \\ i \in A}} (\delta_{A, \emptyset} - S_A) = -\mu \xi_i.$$

The second term in (S15) can be simplified as follows,

$$\begin{aligned} & \sum_{\substack{A \subset K: \\ i \in A}} \sum_{k \in K} \sum_{B \subset K \setminus \{k\}} I_B^k C(B, k, A) = \sum_{k \in K} \sum_{B \subset K \setminus \{k\}} I_B^k C_i^*(B, k) \\ &= \sum_{k \in K} \sum_{\substack{B \subset K \setminus \{k\}: \\ i \in B}} I_B^k C_i^*(B, k) + \sum_{k \in K} \sum_{B \subset K \setminus \{k, i\}} I_B^k C_i^*(B, k) \\ &= \sum_{k \in K} \sum_{\substack{B \subset K \setminus \{k\}: \\ i \in B}} I_B^k + \sum_{k \in K} \sum_{B \subset K \setminus \{i, k\}} I_B^k \sigma_{ki} = \sum_{k \in K} \eta_i^k + \sum_{k \in K} \sigma_{ki} (I_k - \eta_i^k). \end{aligned}$$

The last equality is satisfied because

$$\sum_{k \in K} \sum_{B \subset K \setminus \{i, k\}} I_B^k \sigma_{ki} = \sum_{k \in K} \sigma_{ki} \left( \sum_{B \subset K \setminus \{k\}} I_B^k - \sum_{\substack{B \subset K \setminus \{k\}: \\ i \in B}} I_B^k \right).$$

Note that the second term in (S15) can be rewritten as

$$\sum_{k \in K} \eta_i^k + \sum_{k \in K \setminus \{i\}} \sigma_{ki} (I_k - \eta_i^k) + I_i,$$

showing that it accounts for the replenishment of healthy host individuals that are immune to strain  $i$  due to (a) recovery of hosts that have previously been immune to strain  $i$ , but have been infected with some other strain  $k$ ; (b) recovery of hosts that, after having been infected with some strain  $k$ , have gained cross-immunity to strain  $i$ ; and (c) recovery of hosts that have been infected with strain  $i$ . Finally, the third term in (S15) accounts for the infection of individuals, immune to strain  $i$ , with some other strain  $k$ , and can be rewritten as

$$\sum_{\substack{A \subset K: \\ i \in A}} \sum_{k \in K \setminus A} \Lambda_k S_A = \sum_{k \in K \setminus \{i\}} \sum_{\substack{A \subset K \setminus \{k\}: \\ i \in A}} \Lambda_k S_A = \sum_{k \in K} \Lambda_k (\xi_i - \xi_{ik}).$$

Collecting the three results above, we obtain the equation for  $\dot{\xi}_i$  expressed in terms of new variables,

$$\dot{\xi}_i = \sum_{k \in K} [\nu (\eta_i^k + \sigma_{ki} (I_k - \eta_i^k)) - \Lambda_k (\xi_i - \xi_{ik})] - \mu \xi_i \quad \text{for all } i \in K. \quad (\text{S16})$$

The equation for  $\dot{\xi}_{ij}$  is obtained analogously,

$$\begin{aligned} \dot{\xi}_{ij} = & \sum_{k \in K} \left[ \nu (\eta_{ij}^k + \sigma_{ki} (\eta_j^k - \eta_{ij}^k) + \sigma_{kj} (\eta_i^k - \eta_{ij}^k) + \sigma_{ki} \sigma_{kj} (I_k - \eta_j^k - \eta_i^k + \eta_{ij}^k)) - \right. \\ & \left. - \Lambda_k (\xi_{ij} - \xi_{ijk}) \right] - \mu \xi_{ij} \quad \text{for all } i, j \in K, i \neq j. \end{aligned} \quad (\text{S17})$$

Observe that the equations for the immunity variables of order  $\ell$  depend on the  $\eta$ -variables of order  $\ell$  and on the  $\xi$ -variables of order  $\ell + 1$ , but not on any immunity variables of higher orders. To truncate this hierarchy of equations at order  $\ell$ , in full analogy to the model with coinfections, we need to approximate the  $\xi$ -variables of order  $\ell + 1$  by a function of immunity variables of lower orders. Each approximation must, again, satisfy the symmetry, monotonicity and redundancy conditions outlined in the main text.

We suggest the following simple closures:

(a) *Order-1 independence closure:*

$$\hat{\xi}_{ij} = \begin{cases} \frac{\xi_i \xi_j}{H} & \text{if } i \neq j \\ \xi_i & \text{if } i = j \end{cases} \quad \text{for all } i, j \in K. \quad (\text{S18})$$

(b) *Order-1 interpolation closure:*

$$\hat{\xi}_{ij} = \frac{\xi_i \xi_j}{H} \left( 1 - \frac{\sigma_{ij} + \sigma_{ji}}{2} \right) + \frac{\sigma_{ij} + \sigma_{ji}}{2} \min(\xi_i, \xi_j) \quad \text{for all } i, j \in K. \quad (\text{S19})$$

(c) *Order-2 independence closure:*

$$\hat{\xi}_{ijk} = \begin{cases} \frac{1}{3H} (\xi_{ij}\xi_k + \xi_{ik}\xi_j + \xi_{jk}\xi_i) & \text{if } i \neq j \neq k \neq i \\ \xi_{ij} & \text{if } i = k \text{ or } j = k \\ \xi_{ik} & \text{if } i = j \end{cases} \quad \text{for all } i, j, k \in K. \quad (\text{S20})$$

Models truncated at first order have  $n(n+1)$  remaining variables, while models truncated at second order have  $n(n^2+3)/2$  variables.

Finally, we provide, without a derivation, the reduced version of the model with coinfections under the assumption of reduced infectivity. First, the full model with this assumption differs from the model with reduced susceptibility in that the class  $I_A^i$ , where  $i \in A$ , is no longer empty, since we assume that hosts can be infected with a variant even if it is in their immune status. Such individuals will not, however, contribute to infectivity with strain  $i$  and, therefore, the force of infection is still given by expression (S10). Equation (S7) is transformed to  $\sum_{A \subset K} S_A + \sum_{i \in K} \sum_{A \subset K} I_A^i = 1$ , and instead of (S8)–(S9) we have

$$\dot{S}_A = \mu(\delta_{A,\emptyset} - S_A) + \nu \sum_{k \in K} \sum_{B \subset K} I_B^k C(B, k, A) - \sum_{k \in K} \Lambda_k S_A, \quad \text{for all } A \subset K, \quad (\text{S21})$$

$$\dot{I}_A^i = \Lambda_i S_A - (\nu + \mu) I_A^i, \quad \text{for all } i \in K \text{ and } A \subset K. \quad (\text{S22})$$

To obtain the reduced version of this model, we rewrite (S21)–(S22) in terms of

$$\begin{aligned} I_i &= \sum_{A \subset K} I_A^i, \\ \xi_i &= \sum_{\substack{A \subset K: \\ i \in A}} S_A, \quad \xi_{ij} = \sum_{\substack{A \subset K: \\ i, j \in A}} S_A, \quad \dots, \\ \eta_i^k &= \sum_{\substack{A \subset K: \\ i \in A}} I_A^k, \quad \eta_{ij}^k = \sum_{\substack{A \subset K: \\ i, j \in A}} I_A^k, \quad \dots \end{aligned}$$

It can be shown that system (S21)–(S22) is equivalent to the following system.

$$\begin{aligned} \dot{I}_i &= \Lambda_i H - (\nu + \mu) I_i \\ \dot{\eta}_i^k &= \Lambda_k \xi_i - (\nu + \mu) \eta_i^k, \\ \dot{\eta}_{ij}^k &= \Lambda_k \xi_{ij} - (\nu + \mu) \eta_{ij}^k, \\ \dot{\xi}_i &= \sum_{k \in K} [\nu (\eta_i^k + \sigma_{ki} (I_k - \eta_i^k)) - \Lambda_k \xi_i] - \mu \xi_i \end{aligned}$$

$$\begin{aligned}
\dot{\xi}_{ij} &= \sum_{k \in K} \left[ \nu (\eta_{ij}^k + \sigma_{ki}(\eta_j^k - \eta_{ij}^k) + \sigma_{kj}(\eta_i^k - \eta_{ij}^k) + \sigma_{ki}\sigma_{kj}(I_k - \eta_i^k - \eta_j^k + \eta_{ij}^k)) - \right. \\
&\quad \left. - \Lambda_k \xi_{ij} \right] - \mu \xi_{ij} \\
&\vdots
\end{aligned}$$

for all  $i, j, k \in K$ ,  $i \neq j$ , where  $H = 1 - \sum_{k \in K} I_k$  and  $\Lambda_i$  is defined by (S14). As it is common for models with reduced infectivity, the dynamics of the immunity variables of a particular order does not depend on the immunity variables of higher orders. Therefore, the truncation at a given order reproduces the true behavior of the immunity variables up to that order.

## Extended analysis of accuracy for reduced models

In this section we present the results of the same type of analysis as in the main text, for additional strain space topologies and for the model with no coinfections. We use the following topologies:

*Topology 1.* Linear four-strain system as described in the main text.

*Topology 2.* Circular four-strain system. The mutational neighborhood and the cross-immunity matrix are given by

$$M_1 = \{2, 4\}, \quad M_2 = \{1, 3\}, \quad M_3 = \{2, 4\}, \quad M_4 = \{3, 1\},$$

$$\sigma = \begin{pmatrix} 1 & s & 0 & s \\ s & 1 & s & 0 \\ 0 & s & 1 & s \\ s & 0 & s & 1 \end{pmatrix}.$$

*Topology 3.* Linear six-strain system. The mutational neighborhood and the cross-immunity matrix are given by

$$M_1 = \{2\}, \quad M_2 = \{1, 3\}, \quad M_3 = \{2, 4\}, \quad M_4 = \{3, 5\}, \quad M_5 = \{4, 6\}, \quad M_6 = \{5\},$$

$$\sigma = \begin{pmatrix} 1 & s & 0 & 0 & 0 & 0 \\ s & 1 & s & 0 & 0 & 0 \\ 0 & s & 1 & s & 0 & 0 \\ 0 & 0 & s & 1 & s & 0 \\ 0 & 0 & 0 & s & 1 & s \\ 0 & 0 & 0 & 0 & s & 1 \end{pmatrix}.$$

*Topology 4.* Torus-like six-strain system (Figure S1). The mutational neighborhood and the cross-immunity matrix are given by

$$\begin{aligned}
M_1 &= \{2, 3, 4\}, & M_2 &= \{1, 3\}, & M_3 &= \{1, 2, 6\}, \\
M_4 &= \{1, 5, 6\}, & M_5 &= \{4, 6\}, & M_6 &= \{3, 4, 5\},
\end{aligned}$$

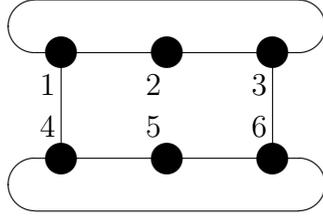


Figure S1: Topology 4.

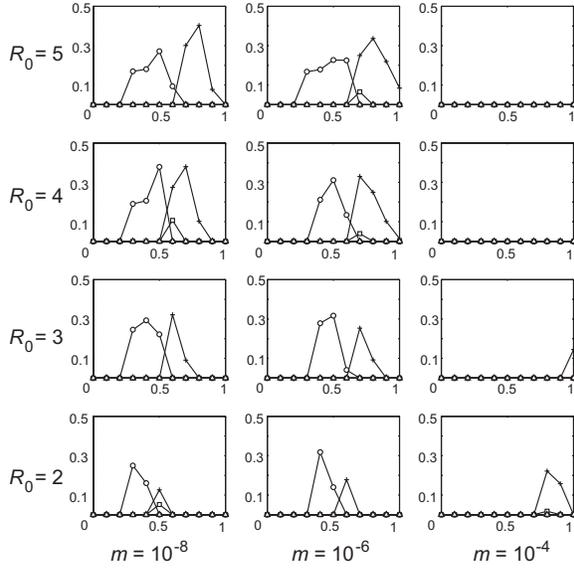


Figure S2: Qualitative accuracy measure  $\rho$  for the performance of four models with coinfections on the Topology 2 system: Gog and Grenfell's model (circles), order-1 independence closure (pluses), order-1 interpolation closure (squares), order-2 independence closure (triangles). Plots in each row have the value of  $R_0$  that is indicated on the left; plots in each column have the value of  $m$  that is indicated at the bottom.

$$\sigma = \begin{pmatrix} 1 & s & s & s & 0 & 0 \\ s & 1 & s & 0 & 0 & 0 \\ s & s & 1 & 0 & 0 & s \\ s & 0 & 0 & 1 & s & s \\ 0 & 0 & 0 & s & 1 & s \\ 0 & 0 & s & s & s & 1 \end{pmatrix}.$$

## Model with coinfections

According to the procedure outlined in the main text, we numerically find solutions for topologies 2, 3 and 4. The results are shown in Figures S2 – S7.

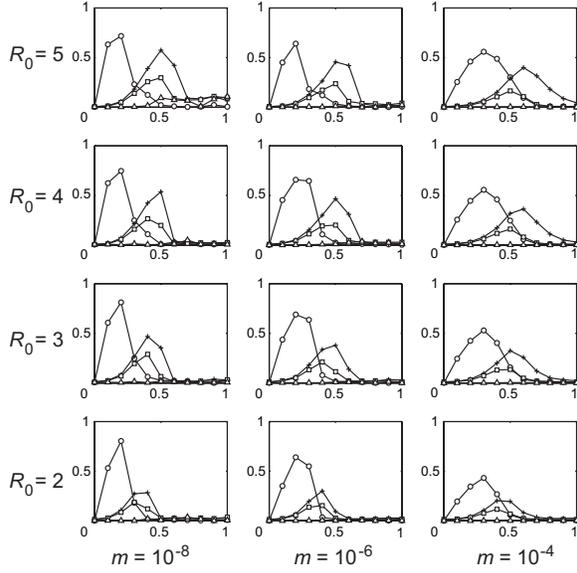


Figure S3: Quantitative accuracy measure  $\Delta$  for the performance of four models with coinfections on the Topology 2 system. Details as in Figure S2.

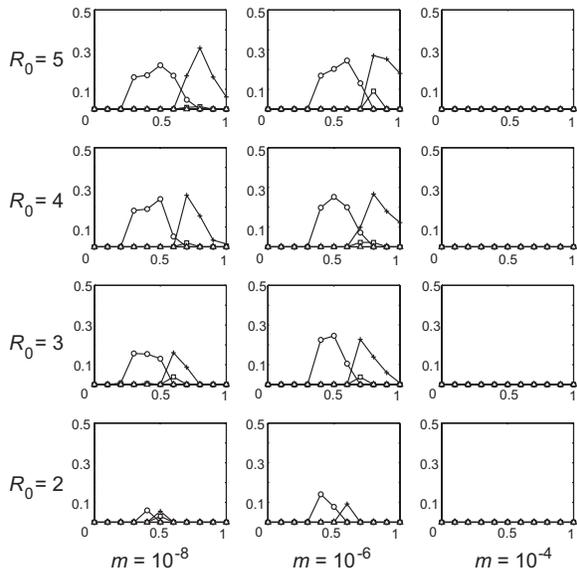


Figure S4: Qualitative accuracy measure  $\rho$  for the performance of four models with coinfections on the Topology 3 system. Details as in Figure S2.

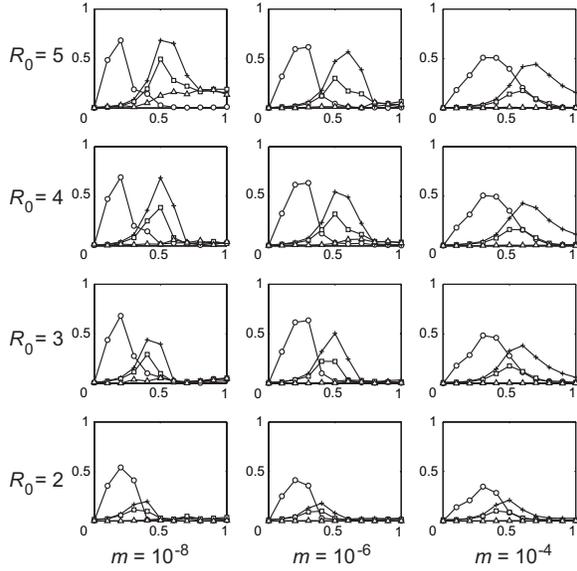


Figure S5: Quantitative accuracy measure  $\Delta$  for the performance of four models with coinfections on the Topology 3 system. Details as in Figure S2.

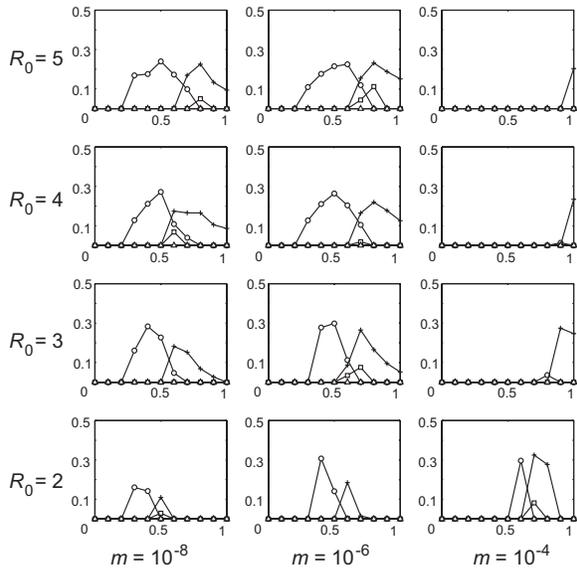


Figure S6: Qualitative accuracy measure  $\rho$  for the performance of four models with coinfections on the Topology 4 system. Details as in Figure S2.

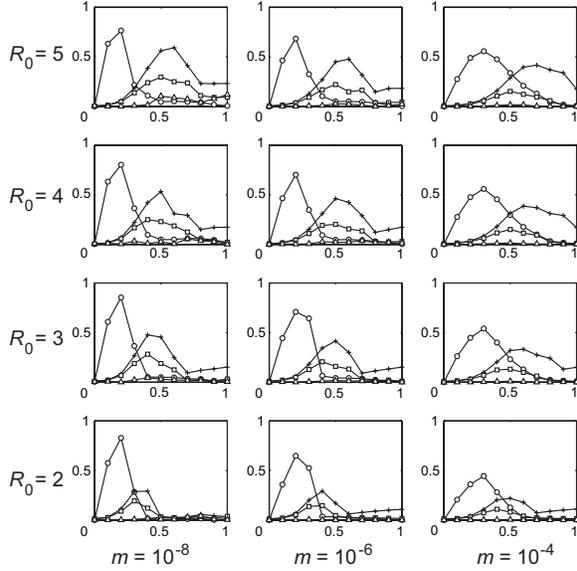


Figure S7: Quantitative accuracy measure  $\Delta$  for the performance of four models with coinfections on the Topology 4 system. Details as in Figure S2.

## Model with no coinfections

For the model with no coinfections, we numerically solve the following equations for the time interval  $[0, T]$  and for a range of parameters:

1. The full SIR system: equations (S8) and (S9).
2. The approximation based on the order-1 independence closure: equations (S11), (S12), and (S16), where the  $\xi_{ij}$  are substituted according to (S18).
3. The approximation based on the order-1 interpolation closure: equations (S11), (S12), and (S16), where the  $\xi_{ij}$  are substituted according to (S19).
4. The approximation based on the order-2 independence closure: equations (S11), (S12), (S13), (S16), and (S17), where the  $\xi_{ijk}$  are substituted according to (S20).

Initially, 99% of the host population are healthy and fully susceptible to all strains, while 1% is infected with strain 1. All other parameter values are the same as in the model with coinfections. The results for topologies 1 – 4 are shown in Figures S8 – S15.

## Additional analysis of the influenza A drift model

In this section we present results of additional simulations for our influenza A drift model.

First, we show that, a model with homogeneous transmission coefficients exhibits a succession of diagonal variants but no coexistence (Figure S16). Therefore, heterogeneity in transmission coefficients is necessary in our model to ensure coexistence of similar

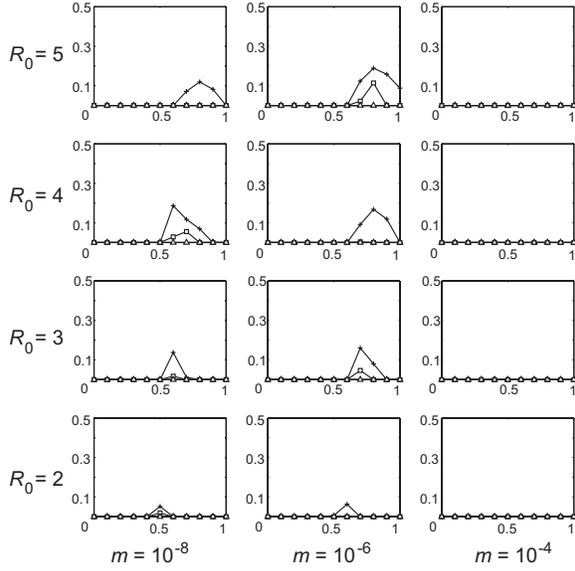


Figure S8: Qualitative accuracy measure  $\rho$  for the performance of three models with no coinfections on the Topology 1 system: order-1 independence closure (pluses), order-1 interpolation closure (squares), order-2 independence closure (triangles). Plots in each row have the value of  $R_0$  that is indicated on the left; plots in each column have the value of  $m$  that is indicated at the bottom.

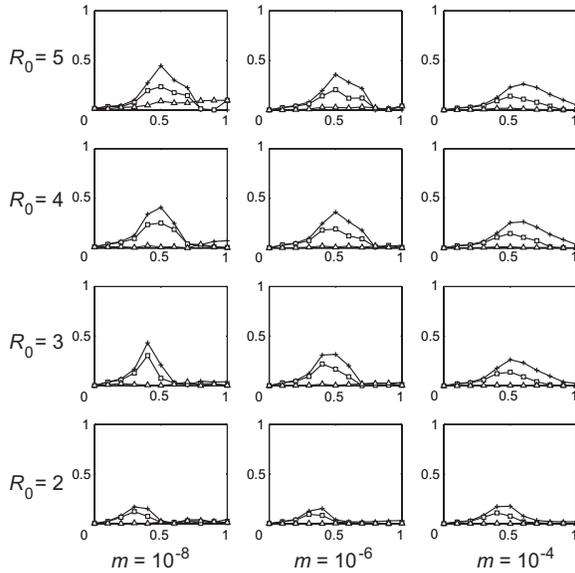


Figure S9: Quantitative accuracy measure  $\Delta$  for the performance of three models with no coinfections on the Topology 1 system. Details as in Figure S8.

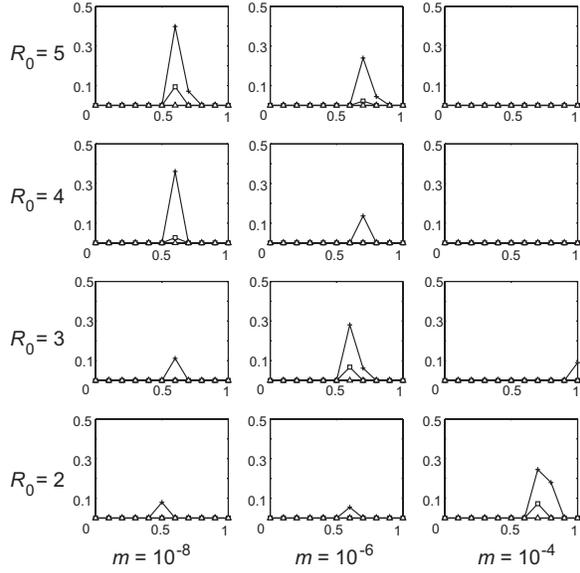


Figure S10: Qualitative accuracy measure  $\rho$  for the performance of three models with no coinfections on the Topology 2 system. Details as in Figure S8.

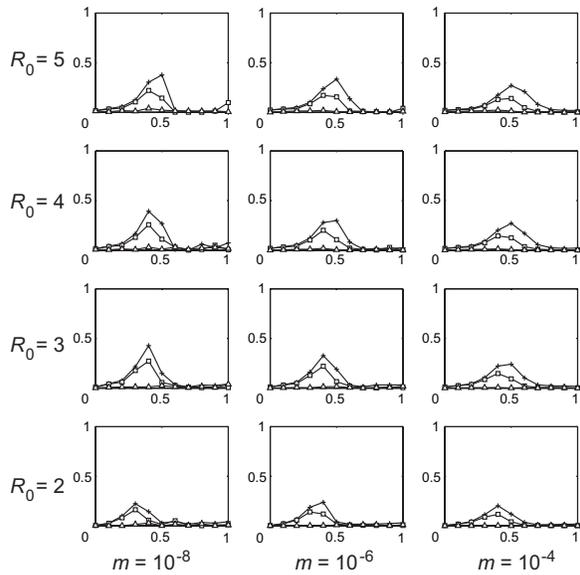


Figure S11: Quantitative accuracy measure  $\Delta$  for the performance of three models with no coinfections on the Topology 2 system. Details as in Figure S8.

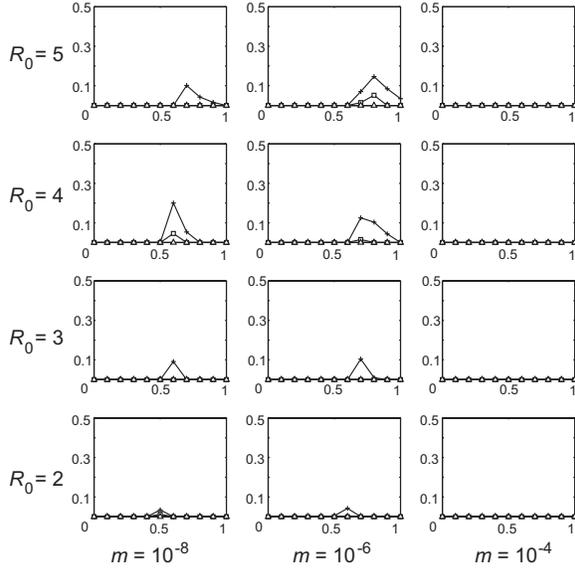


Figure S12: Qualitative accuracy measure  $\rho$  for the the performance of three models with no coinfections on the Topology 3 system. Details as in Figure S8.

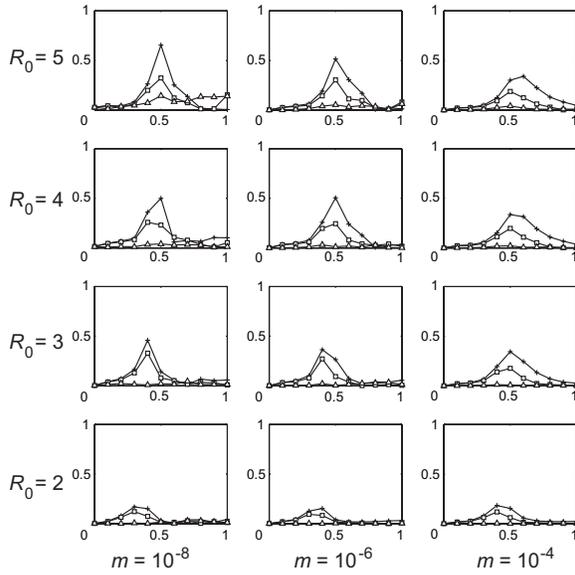


Figure S13: Quantitative accuracy measure  $\Delta$  for the performance of three models with no coinfections on the Topology 3 system. Details as in Figure S8.

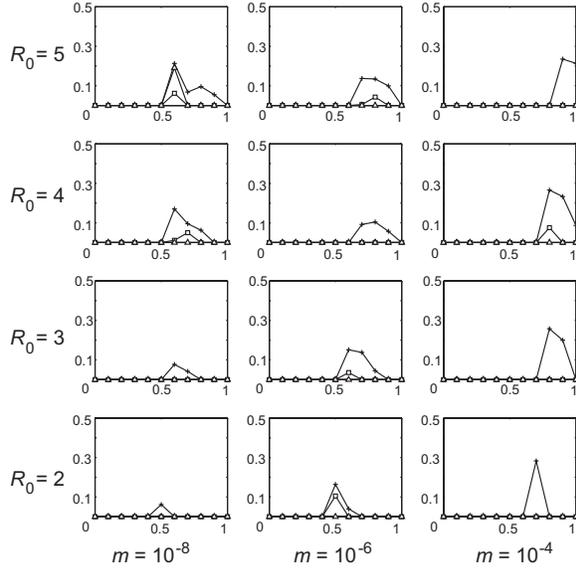


Figure S14: Qualitative accuracy measure  $\rho$  for the the performance of three models with no coinfections on the Topology 4 system. Details as in Figure S8.

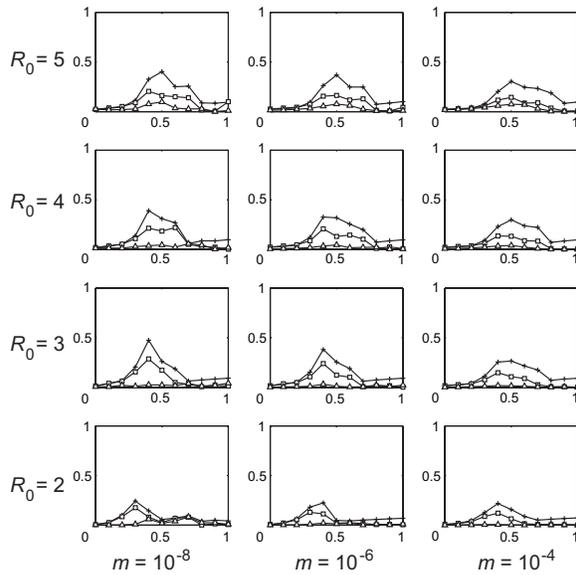


Figure S15: Quantitative accuracy measure  $\Delta$  for the performance of three models with no coinfections on the Topology 4 system. Details as in Figure S8.

variants. It is interesting to notice that in the model by Tria et al. [S3] heterogeneity in transmission coefficients had an opposite effect – it was necessary to constrain viral diversity. Although that model is qualitatively different from ours in many respects and cannot be compared to our model directly, we think that the main reason for this discrepancy is the fact that Tria et al. assume local cross-immunity structure (in their model the strength of cross-immunity monotonically depends on the Hamming distance between variant sequences) while we assume highly non-local cross-immunity structure. We conjecture that the effect of heterogeneity of transmission coefficients in Gog and Grenfell’s model [S1] with strain-independent cross-immunity would coincide with that by Tria et al.

The second simulation differs from the model described in main text only by the shape of the cross-immunity structure between strains. We demonstrate that, if cross-immunity structure is local, as shown in Figure 4B in the main text, constrained evolution is impossible – the virus explores the whole strain space, unless, of course, some part of it is unviable, i.e., transmission coefficients are equal to or less than unity, as in Gog and Grenfell’s work [S1]. Here we use essentially the same local cross-immunity structure as Gog and Grenfell [S1]:

$$\sigma_{(i,j)(k,\ell)}^{(\text{loc})} = \exp \left\{ -\frac{1}{2} \left( \frac{|i-k| + |j-\ell|}{a} \right)^2 \right\}.$$

The results of this simulation are shown in Figure S17. Note that the sum of fractions of infected hosts surpasses 1 and reaches its maximum near 5 before it starts dropping. This implies that many hosts survive multiple coinfections. Obviously, this does not happen in a model where coinfections are prohibited, despite an explosion in diversity that is still captured (results not shown). Also note that the decline in the number of coinfections in the second half of the simulation is probably due to the fact that the whole 20 by 20 strain space was explored within the simulation time.

The final set of simulations aims at demonstrating that the main qualitative result of our paper – the one-dimensionality of influenza A drift – does not depend on the details of the model. We simulate the influenza evolution in a setting with the immunity structure and the distribution of transmission coefficients as described in the main text using now Gog and Grenfell’s model (Figure S18) and the approximate model with no coinfections with the order-1 interpolation closure (Figure S19). As expected, the evolution of the virus is principally one-dimensional irrespectively of the model. However, it is instructive to notice some similarities and differences between the simulation results in different models:

1. The sets of strains that cause epidemics coincide for both our models but slightly differ from the set predicted by Gog and Grenfell’s model (for example, in the latter model, strains (12,7) or (16,16) do not cause epidemics). In general, strains cause less severe epidemics in Gog and Grenfell’s model than in either of our models. This is consistent with the fact that Gog and Grenfell’s model overestimates the level of immunity in the population.

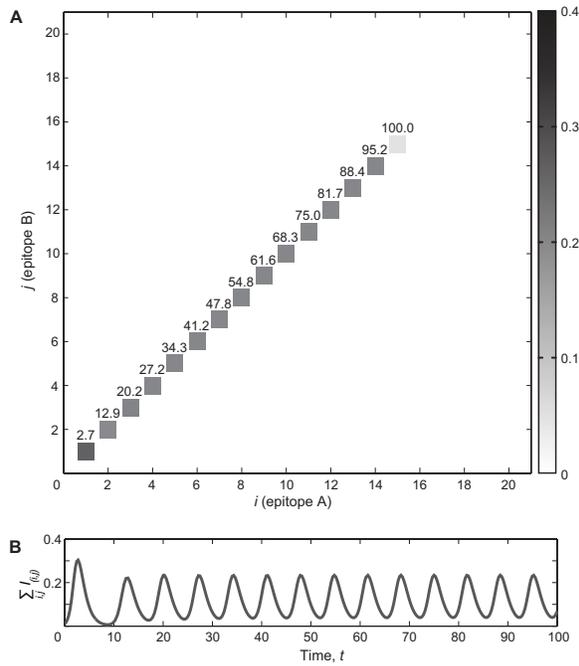


Figure S16: Approximate dynamics of antigenic drift in influenza A, based on the order-1 interpolation closure in a model with coinfections. All transmission coefficients  $\beta_{(i,j)}$  are equal to 3. Other parameter values are the same as in Figure 5 in the main text.

2. The epidemic peak times in our model with coinfections generally coincide with those in Gog and Grenfell’s model, while the corresponding epidemics occur later in the model where coinfections are excluded. This suggests that the evolution of the virus proceeds slower in a system with no coinfections.

These observations provide additional evidence for the fact that, even though the coarse qualitative behavior of status-based models based on different assumptions may be quite similar even in complex settings, substantial quantitative differences in predictions of such models do exist.

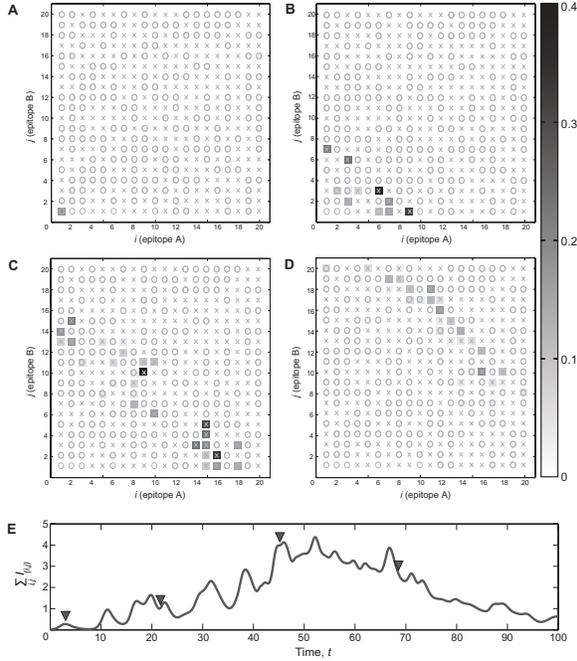


Figure S17: Approximate dynamics of antigenic drift in influenza A, based on the order-1 interpolation closure in a model with coinfections.  $\sigma = \sigma^{(\text{loc})}$ . All other parameters, including parameter  $a$  and the landscape of transmission coefficients are the same as in Figure 5 in the main text. Four snapshots are taken at times  $t = 3$  (A),  $t = 23$  (B),  $t = 46$  (C), and  $t = 67$  (D). E. The sum of all proportions of infectious hosts as a function of time.

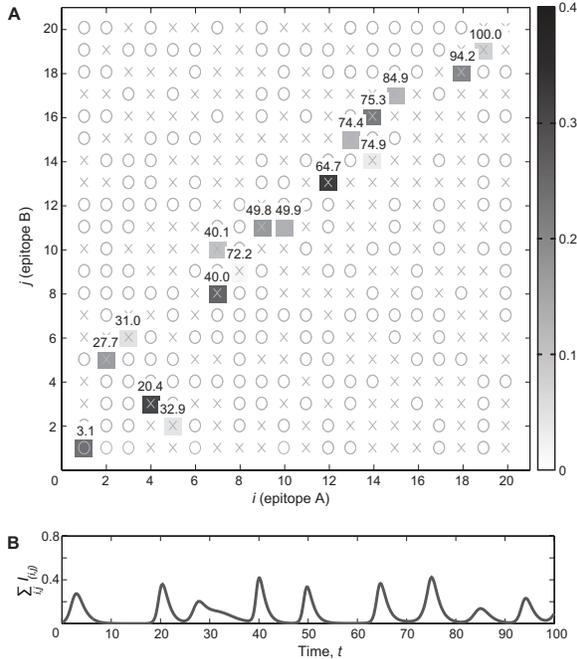


Figure S18: Approximate dynamics of antigenic drift in influenza A, based on Gog and Grenfell's model. All parameters, including the landscape of transmission coefficients are the same as in Figure 5 in the main text.

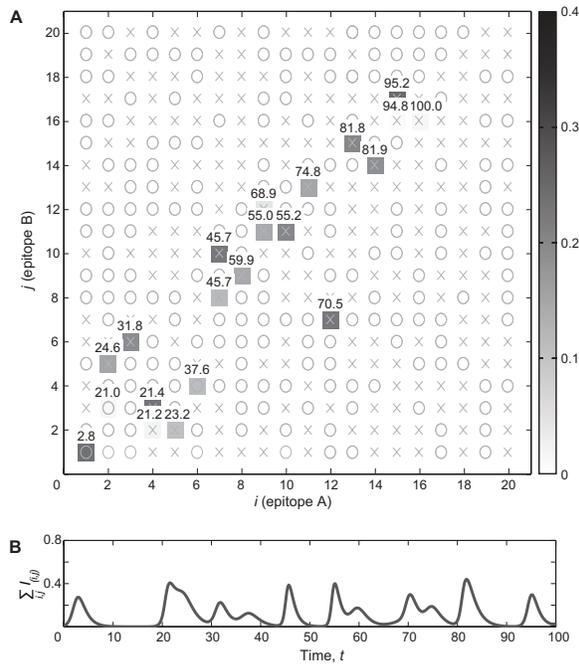


Figure S19: Approximate dynamics of antigenic drift in influenza A, based on the order-1 interpolation closure in a model with no coinfections. The initial condition was given by all state variables being zero except for  $I_{(1,1)}(0) = 0.01$ , corresponding to a healthy and fully susceptible host population with 1% of hosts infected with strain (1, 1). All other parameters including the landscape of transmission coefficients are the same as in Figure 5 in the main text. **B.** The total fraction of infected hosts as a function of time.

## References

- [S1] Gog JR, Grenfell BT (2002) Dynamics and selection of many-strain pathogens. *Proc Natl Acad Sci USA* 99:17209–17214.
- [S2] Gog JR, Swinton J (2002) A status-based approach to multiple strain dynamics. *J Math Biol* 44:169–184.
- [S3] Tria F, Lässig M, Peliti L, Franz S (2005) A minimal stochastic model for influenza evolution. *J Stat Mech* P07008.