



# Virus-immune dynamics determined by prey-predator interaction network and epistasis in viral fitness landscape

Cameron J. Browne<sup>1</sup> · Fadoua Yahia<sup>1</sup>

Received: 16 June 2021 / Revised: 10 July 2022 / Accepted: 22 November 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

## Abstract

Population dynamics and evolutionary genetics underly the structure of ecosystems, changing on the same timescale for interacting species with rapid turnover, such as virus (e.g. HIV) and immune response. Thus, an important problem in mathematical modeling is to connect ecology, evolution and genetics, which often have been treated separately. Here, extending analysis of multiple virus and immune response populations in a resource—prey (consumer)—predator model from Browne and Smith (2018), we show that long term dynamics of viral mutants evolving resistance at distinct epitopes (viral proteins targeted by immune responses) are governed by epistasis in the virus fitness landscape. In particular, the stability of persistent equilibrium virus-immune (prey-predator) network structures, such as nested and one-to-one, and bifurcations are determined by a collection of *circuits* defined by combinations of viral fitnesses that are minimally additive within a hypercube of binary sequences representing all possible viral epitope sequences ordered according to *immunodominance* hierarchy. Numerical solutions of our ordinary differential equation system, along with an extended stochastic version including random mutation, demonstrate how pairwise or multiplicative epistatic interactions shape viral evolution against concurrent immune responses and convergence to the multi-variant steady state predicted by theoretical results. Furthermore, simulations illustrate how periodic infusions of subdominant immune responses can induce a bifurcation in the persistent viral strains, offering superior host outcome over an alternative strategy of immunotherapy with strongest immune response.

**Keywords** Virus-immune response model · Predator–prey network · Fitness landscape · HIV quasispecies · Epistasis · Eco-evolutionary dynamics

**Mathematics Subject Classification** 92D10 · 92D15 · 92D25 · 92C42 · 34D20

---

✉ Cameron J. Browne  
cambrowne@louisiana.edu

<sup>1</sup> Mathematics Department, University of Louisiana at Lafayette, Lafayette, LA, USA

## 1 Introduction

The evolution of ecological networks depends on the underlying population dynamics, genetics, and structure of the composite species. Interactions between populations, for example prey-predator or competitive forces, constrain and shape the network, in concert with evolution also diversifying and adapting species variants. The complexity of these eco-evolutionary dynamics have challenged researchers to classify patterns in rapidly evolving communities. In the single species context, theoretical models of fitness landscapes have simplified the study of adaptation by reducing individuals to either genotypes or phenotypes, whose reproductive success is determined by a single trait, namely fitness. Although a multitude of evolutionary pathways exist, evolution predictability can be driven by genetic variant constraints. A more analytically challenging scenario is the evolution or coevolution of prey-predator systems whereby the predator range and selection of prey resistance balanced by constraints on reproduction together form a dynamic fitness landscape. Examples include phage-microbe and virus-immune response networks, with the latter, specifically HIV, being a primary motivation for this work.

During HIV infection, a diverse collection of viral strains, often called a quasispecies, compete for a target cell population (mainly  $CD4^+$  T-cells) while the host immune response population (e.g.  $CD8^+$  T-cells) predate and proliferates upon pathogen recognition. HIV can also rapidly evolve resistance to immune response attack at different epitopes (proteins in virus genome displayed on infected cells), inducing a dynamic network of interacting viral and immune variants. Deciphering patterns in the trajectories of virus and immune response populations, along with their interactions, can advance biological theory and have applications for vaccine or immunotherapy development Walker and Xu (2013), Chakraborty and Barton (2017). Analogous questions in other biological systems, such as phage-microbe communities, have mostly led to models of species compositions in the face of ecological interactions *independent* of explicit genetic mutations. The properties of these ecosystem models have classically been studied using dynamical systems, where concepts such as stability, equilibria and population persistence are used to characterize feasible species assemblages. Recently, generalized Lotka-Volterra (L-V), chemostat and ecosystem models have been utilized to understand how different motifs, such as nested or one-to-one networks, are built through invading species and convergence to stable equilibria (Jover et al. 2013; Korytowski and Smith 2015; Browne 2017). Additionally, several works have developed polymorphic evolution sequences, where an individual based stochastic model converges to solutions of L-V equations in the limit of small mutation rates and large populations (Champagnat and Méléard 2011; Costa et al. 2016). However, how population dynamics, genetics and evolution together determine network structures for rapidly evolving ecosystems is not generally established.

From an evolutionary genetics perspective, a high mutation rate allows HIV populations to change and explore sequence space on short timescales, lending themselves to being studied as model biological systems, along with the significant clinical interest. Disease progression, escape pathways, and treatment fate depend on viral fitness. To estimate *in vivo* fitness landscapes, several evolutionary models have linked fitness to viral genotype frequencies, for example the quasispecies model (Seifert et al. 2015)

and multi-strain versions of a standard within-host virus model (Ganusov et al. 2011). These models can be mathematically tractable, allowing for analysis of equilibria and stability in terms of mutation rates and fitness quantities. In particular, the common setting of finite binary sequences, the assumed form of viral genotypes in this current paper, enables geometric or algebraic properties of the binary hypercube space to be exploited for characterizing equilibrium distributions (Bratus et al. 2019). Inclusion of viral mutation from multiple dynamic immune response populations complicates matters, as neither the virus strain fitness or immune response strength simply determine epitope escape (Ganusov and De Boer 2006; Levityang and Ganusov 2015). However, correlation analysis (Liu et al. 2013) and a statistical physics model of viral sequences with *epistasis* (discussed further below) (Barton et al. 2016) applied to HIV patient datasets have found determinants of viral evolution based on viral fitness landscapes and immunodominance hierarchies (relative expansion levels of the responding immune populations).

Epistasis refers to nonlinearity in the fitness landscape or dependence of fitness change from a mutation on the genetic background. Epistatic interactions play a critical role in fitness landscape features, and ultimately evolutionary trajectories, thus measuring epistasis has received much attention when studying evolution. However, the large amount of interactions within a genome challenge both theoretical and experimental quantification of epistasis. Several methods for computing epistasis have been proposed (Mani et al. 2008; Ferretti et al. 2016). Here we focus on the concept of circuits introduced by Beerenwinkel (2007) as fundamental measures of epistatic interactions and underlying geometry of the fitness landscape. Circuits have been utilized to characterize single species fitness landscapes in both theoretical and data-driven studies (Hallgrímsdóttir and Yuster 2008; Crona 2017; Gould et al. 2018).

In this paper, we investigate how epistasis impacts evolution of prey-predator interacting species, specifically how virus fitness landscapes affect the overall virus (prey) and host immune response (predator) ecosystem dynamics. We show that connecting population genetics and dynamics offers a way to extract biological meaningful relationships from the equilibria stability conditions of a complex network differential equation for interacting species' variants. We build off of our previous analysis of a multi-variant virus-immune model (Browne and Smith 2018), which established different regimes of attractors, each with a distinct set of viral strains persisting by extending Lyapunov function methods first applied to generalized L-V equations (Goh 1978; Hofbauer and Sigmund 1998). In particular, the stability of certain equilibria structures and associated bifurcations are sharply determined by relevant circuits, which recast strain invasion rates as algebraic combinations of binary sequences shaping viral fitness landscape epistasis. Furthermore, we simulate eco-evolutionary dynamics showing that our theoretical calculations can carry over to an extended stochastic model with mutations, and also illustrate how distinct immunotherapies can be incorporated in our system to shed light on potential strategies. We conclude with a discussion on how our study supports the utility of evolutionary genetics concepts, in particular the construction of circuits for measuring epistasis of fitness landscapes, applied to characterizing bifurcations in virus-immune response population dynamics, which represents a specific example of a prey-predator ecosystem model.

## 2 General model and binary sequence case

We begin by considering the following rescaled model introduced to describe a network of viral and immune response variants during host infection (Browne and Smith 2018):

$$\begin{aligned}\dot{x} &= 1 - x - x \sum_{i=1}^m \mathcal{R}_i y_i, \\ \dot{y}_i &= \gamma_i y_i \left( \mathcal{R}_i x - 1 - \sum_{j=1}^n a_{ij} Z_j \right), \quad i = 1, \dots, m \\ \dot{Z}_j &= \frac{\sigma_j}{\rho_j} Z_j \left( \sum_{i=1}^m a_{ij} y_i - \rho_j \right), \quad j = 1, \dots, n.\end{aligned}\quad (1)$$

Here  $x$  denotes the population of target cells, along with  $m$  competing virus strains ( $y_i$  denotes strain  $i$  infected cells), and  $n$  variants of immune response ( $Z_j$ ). The parameter  $\mathcal{R}_i$  represents the basic reproduction number of virus strain  $i$ . The  $m \times n$  nonnegative matrix  $A = (a_{ij})$  describes the virus-immune *interaction network*, which determines each immune response population's avidity to the distinct viral strains. Then  $\rho_j$  represents the reciprocal of the immune response fitness excluding the (rescaled) avidity to each strain  $j$ . Additionally,  $\gamma_i$  and  $\sigma_j$  represent scaling factors for corresponding viral and immune variant growth rates.

Each virus strain  $i$  (cells infected with strain  $i$ ),  $y_i$ , has a set of immune responses,  $Z_j$ , that recognize and attack  $y_i$ . We call this set the *epitope set of  $y_i$* , denoted by  $\Lambda_i$  where  $\Lambda_i := \{j \in [1, n] : a_{ij} > 0\}$ . Here  $j \in \Lambda_i$  if  $y_i$  is *not* completely resistant to immune response  $Z_j$ . We remark that the system generally models a tri-trophic ecosystem with a single resource consumed by  $m$  prey (or consumer) populations subject to potential attack by  $n$  distinct predators (prey  $i$  subject to attack by any predator  $j$  in  $\Lambda_i$ ). For example, this model can describe bacteria-phage communities in a chemostat (or single resource environment), where the set  $\Lambda_i$  classifies the infection network (whom infects who).

In this article, we specialize system (1) to the case where each virus strain is represented by a binary sequence of length  $n$ , exactly coding the loci (epitopes) for which  $n$  specific immune responses can recognize and attack. Note that consideration of binary sequences is perhaps the most common way to represent distinct variants which can differ at some loci of their genome (e.g. quasispecies, haploid models). A major goal of this work is to connect concepts in evolution and genetics with population dynamics, so this special case is an appropriate setting. Here the  $n$  viral epitopes have two possible alleles: the wild type (0) and the mutated type (1) which has escaped recognition from the cognate immune response. For each virus strain  $y_i$ , we associate a binary sequence of length  $n$ ,  $y_i \sim \mathbf{i} = (i_1, i_2, \dots, i_n) \in \{0, 1\}^n$ , coding the allele type at each epitope. We assume that each immune response ( $Z_j$ ) targets its specific epitope at the specific rate  $a_j$  for virus strains containing the wild-type (allele 0) epitope  $j$ , whereas  $Z_j$  completely loses ability to recognize strains with the mutant (allele 1) epitope  $j$ , i.e.

$$\forall i \in [1, m], y_i \sim \mathbf{i} \Rightarrow a_{ij} = (1 - i_j)a_j \quad (\text{or } a_{ij} = a_j \text{ if } j \in \Lambda_i, a_{ij} = 0 \text{ otherwise}), \quad (2)$$

with  $\Lambda_i$  the *epitope set* defined earlier for model (1) (see Fig. 1). We also use binary sequence vector notation for viral strains,  $y_i$ , and associated sets. With assumption (2), we can define an *immune reproduction number* corresponding to each  $Z_j$ :

$$\mathcal{I}_j := \frac{a_j}{\rho_j}. \quad (3)$$

Then there are  $m = 2^n$  possible viral mutant strains, each distinguished by binary sequence  $\mathbf{i} = (i_1, i_2, \dots, i_n)$  and denoted  $y_i$ , governed by the following system:

$$\begin{aligned} \dot{x} &= 1 - x - x \sum_{\mathbf{i} \in \{0,1\}^n} \mathcal{R}_i y_i, \\ \dot{y}_i &= \gamma_i y_i \left( \mathcal{R}_i x - 1 - \sum_{j=1}^n (1 - i_j) z_j \right), \quad \mathbf{i} \in \{0, 1\}^n, \\ \dot{z}_j &= \frac{\sigma_j}{s_j} z_j \left( \sum_{\mathbf{i} \in \{0,1\}^n} (1 - i_j) y_i - s_j \right), \quad j = 1, \dots, n \end{aligned} \quad (4)$$

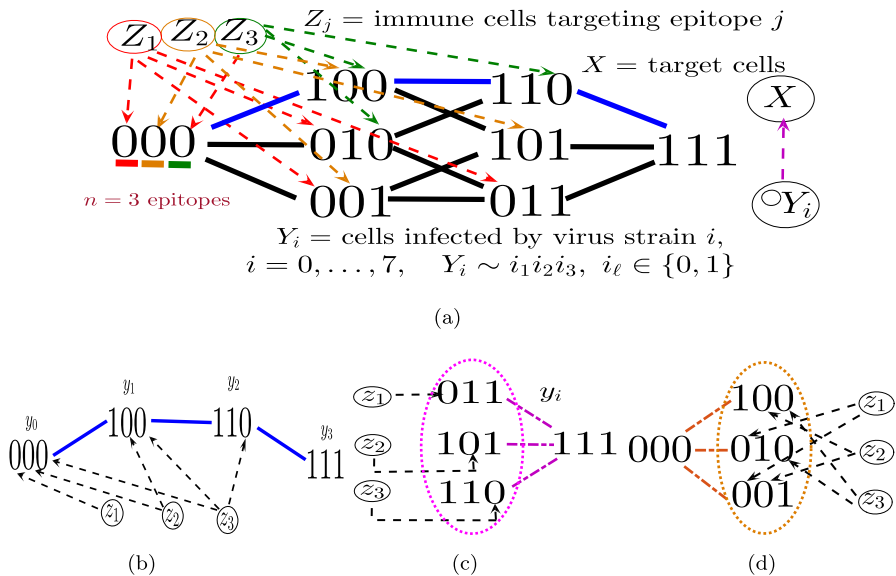
where  $z_j = a_j Z_j$  and  $s_j = 1/\mathcal{I}_j$ .

The  $2^n$  potential virus strains can be viewed in a fitness landscape; each strain  $\mathbf{i}$  is a vertex in an  $n$ -dimensional *hypercube graph* with fitness  $\mathcal{R}_i$ , as shown in Fig. 1a and d in the case of  $n = 3$  and  $n = 2$  epitopes. Viral strains  $y_i$  and  $y_k$  are connected by an edge, if the sequences  $\mathbf{i}$  and  $\mathbf{k}$  differ in exactly one bit, i.e. their *Hamming distance* – denoted by  $d(y_i, y_k)$  – is one. We define the *wild-type (founder) virus strain* as the unique strain which is susceptible to attack by all immune responses, denoted  $y_0$ , characterized by the sequence of all zeroes,  $\mathbf{0} = (0, 0, \dots, 0)$ , or equivalently the epitope set  $\Lambda_0 = \{1, \dots, n\}$ . Each mutation from 0 to 1 of an epitope incurs a fitness cost associated with gaining resistance to the corresponding immune response, so we constrain the collection of viral reproduction numbers (fitness landscape) to satisfy the following condition:

$$\text{If } d(y_i, y_k) = 1 \text{ and } d(y_i, y_0) < d(y_k, y_0), \text{ then } \mathcal{R}_i > \mathcal{R}_k. \quad (5)$$

The occurrence of fitness costs (in reproduction rate) for gaining resistance to predation is a general concept in eco-evolutionary systems, for example in bacteria-phage networks. Finally we say that an immune response  $z_j$  is *immunodominant* over another immune response  $z_k$  if  $\mathcal{I}_j > \mathcal{I}_k$  and assume, without loss of generality, the ordered *immunodominance hierarchy*;

$$\mathcal{I}_1 \geq \mathcal{I}_2 \geq \dots \geq \mathcal{I}_n, \quad \text{i.e. } s_1 \leq s_2 \leq \dots \leq s_n. \quad (6)$$



**Fig. 1** **a** The virus-immune network on  $n=3$  epitopes for model (4) overlying the hypercube. Here each viral strain  $y_i$ ,  $i=0, \dots, 7$ , is associated with a unique binary string  $\mathbf{i} \in \{0, 1\}^3$  coding their allele type, susceptible (0) or resistant (1), at each epitope. Immune response  $z_j$  attacks  $y_i \sim \mathbf{i}$  if  $i_j = 0$ , or equivalently if  $j$  is in epitope set of  $y_i$  ( $j \in A_i$ ). The wild-type virus,  $y_0 \sim 000$ , can evolve resistance to each epitope-specific immune response  $z_j$  by successive single epitope mutations forming a path in the hypercube graph to the completely resistant viral strain (111). The number of epitope mutations which viral strain  $y_i$  has accumulated is  $d(y_i, y_0)$  (Hamming distance between  $\mathbf{i}$  and 000). Note that system (4) does not explicitly include mutation between viral strains. **b** The perfectly nested network, as a subgraph of the hypercube. In this case, sequential mutations of epitopes appear in immunodominance order with specialist to generalist virus (prey) resistance and immune (predator) attack. **c** The one-to-one network, with strain-specific immune responses, is representative of a completely modular ecosystem. **d** The network with at most one mutation signifies constrained evolution. All three subgraphs appear as feasible equilibrium structures of the system and are analyzed in Sect. 4

System (4) generalizes many previous model structures in the sense that they can be seen as subgraphs of our “hypercube network”. For instance, the “strain-specific” (virus-immune response) network (Nowak and Bangham 1996) (also called “one-to-one network” in phage-bacteria models (Jover et al. 2013; Korytowski and Smith 2015)) is equivalent to restricting (4) to the  $m=n$  viral strains which have mutated  $n-1$  epitopes (Fig. 1c). The “perfectly nested network” restricts (4) to the  $m=n+1$  viral strains which have sequential epitope mutations in the order of the immunodominance hierarchy (Fig. 1b). Nested networks were considered in HIV models (Browne 2017), along with phage-bacteria models (Korytowski and Smith 2015), and may be a common persistent structure in ecological communities (Gurney et al. 2017). The network with at most one mutation, or “at most one mutation” network, denoted by  $\tilde{\mathcal{S}}_1$  with  $n+1$  strains  $y_i$  such that  $d(\mathbf{i}, \mathbf{0}) \leq 1$ , is an example of constrained immune escape steady state (Fig. 1d) considered in Browne and Smith (2018) and the current study. The “full hypercube network” has been considered for modeling immune escape patterns in HIV infected individuals (Althaus and Boer 2008; Deutekom et al. 2013).

### 3 Necessary population dynamics/genetics definitions and results

#### 3.1 Stability and persistence

First we review some relevant definitions and results on the equilibria and asymptotic dynamics valid in the general model (1) that are further detailed in prior work (Browne and Smith 2018). For a non-negative equilibrium point,  $\mathcal{E}^* = (x^*, y^*, Z^*) \in \mathbb{R}_+^{1+m+n}$ , define the “persistent variant sets” associated with  $\mathcal{E}^*$  as:

$$\Omega_y = \Omega_y(\mathcal{E}^*) = \{i \in [1, m] : y_i^* > 0\} \quad \text{and} \quad \Omega_z = \Omega_z(\mathcal{E}^*) = \{j \in [1, n] : Z_j^* > 0\}. \quad (7)$$

In addition, define the following subsets of  $\mathbb{R}_+^{1+m+n}$ :

$$\begin{aligned} \Omega &= \Omega(\mathcal{E}^*) = \{(x, y, Z) \mid y_i, z_j > 0, i \in \Omega_y, j \in \Omega_z\}, \\ \Gamma_\Omega &= \Omega \cap \{y_i, z_j = 0, i \notin \Omega_y, j \notin \Omega_z\}. \end{aligned}$$

Here  $\Gamma_\Omega$ , consisting of only those state vectors having the same set of positive and zero components as equilibrium  $\mathcal{E}^*$ , is called the *positivity class* of  $\mathcal{E}^*$ . Notice that the dimension of the subset  $\Gamma_\Omega$  is  $1 + |\Omega_y| + |\Omega_z|$ , where the notation  $|\Omega_y|$  ( $|\Omega_z|$ ) denotes the *cardinality* of the set  $\Omega_y$  ( $\Omega_z$ ). We denote  $A'$  as the  $m' \times n'$  interaction matrix consisting of the  $m'$  virus strains and  $n'$  immune responses which have positive values at equilibrium  $\mathcal{E}^*$ , where  $|\Omega_y| = m'$  and  $|\Omega_z| = n'$ . The equilibrium  $\mathcal{E}^*$  must satisfy the following equations:

$$\sum_{i \in \Omega_y} a_{ij} y_i^* = \rho_j, \quad j \in \Omega_z \quad (8)$$

$$\sum_{j \in \Omega_z} a_{ij} Z_j^* = \mathcal{R}_i x^* - 1, \quad i \in \Omega_y \quad (9)$$

$$1 + \sum_{i \in \Omega_y} \mathcal{R}_i y_i^* = \frac{1}{x^*} \quad (10)$$

Note that if  $i \in \Omega_y$  ( $y_i^* > 0$ ), then  $\mathcal{R}_i > 1$  must hold, even in the absence of immune response, although  $\mathcal{R}_i$  is not sufficient for positivity of the component.

The following proposition provides the condition for uniqueness of an equilibrium within a positivity class, and shows that in such equilibria the number of virus strains either is equal to or exactly one more than the number of immune responses.

**Proposition 1** (Browne and Smith 2018) *Suppose the equilibrium  $\mathcal{E}^* = (x^*, y^*, Z^*)$  exists in positivity class  $\Gamma_\Omega$ , where  $(y^*, Z^*)$  satisfy the linear system of Eq. (8)-(9). Then  $\mathcal{E}^*$  is the unique equilibrium in  $\Gamma_\Omega$ , i.e.  $\mathbf{v} = (y^*, Z^*)^T$  is the unique solution to (8)-(9), if and only if  $\text{Ker}(A')^T \cap \mathcal{R}'^\perp = \{0\}$  and  $\text{Ker}(A') = \{0\}$ . Moreover, if  $\mathcal{E}^*$  is not unique in its positivity class  $\Gamma_\Omega$ , then  $\Gamma_\Omega$  contains an infinite number (a*

continuum) of equilibria. Conversely, if  $\mathcal{E}^*$  is unique in a positivity class  $\Gamma_\Omega$  (with  $m'$  and  $n'$  persistent virus and immune responses), then one of the following holds:

- (i)  $m' = n'$ , and  $x^* = 1 / (1 + (\rho)^T (A')^{-1} \mathcal{R}')$ .
- (ii)  $m' = n' + 1$ , and  $x^* = \mathbf{1}^T C_{(n'+1)}^{-1}$ , where  $C_{(n'+1)}^{-1}$  is the last column in the  $(n' + 1) \times (n' + 1)$  matrix inverse of  $C = (A' \ \mathcal{R}')^T$ .

This proposition, along with prior results on competitive exclusion, demonstrate that virus (prey) or ecosystem diversity in our model is entirely mediated by the immune response (predator) populations. Thus the model provides a good system for exploring how prey-predator ecosystems can diversify and patterns in their underlying structure.

Next we are concerned with the stability of equilibria, and which populations persist in the long run. First, based on the idea of being “weakly stable” against missing species (Hofbauer and Sigmund 1998), we call an equilibrium  $\mathcal{E}^* = (x^*, y^*, Z^*)$  of (1) *saturated* if the following holds:

$$\begin{aligned} \frac{\dot{y}_i}{y_i y_i} \Big|_{\mathcal{E}^*} &= \mathcal{R}_i x^* - 1 - \sum_{j \in \Omega_z} a_{ij} Z_j^* \leq 0, \quad \forall i \notin \Omega_y, \quad \frac{s_j \dot{z}_j}{\sigma_j z_j} \Big|_{\mathcal{E}^*} \\ &= \sum_{i \in \Omega_y} a_{ij} y_i^* - \rho_j \leq 0, \quad \forall j \notin \Omega_z \end{aligned} \quad (11)$$

Here each term in (11) gives the sign of the “invasion rate” of a missing species. For a notion of persistent populations, define  $\Omega_{yz}$  *persistence* as

$$\begin{aligned} &\exists \epsilon > 0 \text{ and } T(\mathbf{w}_0) \text{ such that } y_i(t), Z_j(t) > \epsilon, \quad i \in \Omega_y, j \in \Omega_z, \quad \forall t > T(\mathbf{w}_0), \text{ and} \\ &\lim_{t \rightarrow \infty} y_i(t), Z_j(t) = 0, \quad i \notin \Omega_y, j \notin \Omega_z, \quad \text{for every solution with initial condition } \mathbf{w}_0 \in \Omega. \end{aligned}$$

We describe the individual populations  $i \in \Omega_y, j \in \Omega_z$  in the above definition of  $\Omega_{yz}$  *persistence* as being *uniformly persistent*. Now we state a main theorem of Browne and Smith (2018) concerning the stability of equilibria and persistence of viral and immune variants of model (1).

**Theorem 1** (Browne and Smith 2018) *Suppose that  $\mathcal{E}^* = (x^*, y^*, Z^*)$  is a non-negative equilibrium of system (1) with positivity class  $\Gamma_\Omega$ . Suppose further that  $\mathcal{E}^*$  is saturated, i.e. the inequalities (11) hold. Then  $\mathcal{E}^*$  is locally stable and  $x(t) \rightarrow x^*$  as  $t \rightarrow \infty$ .*

Furthermore, if  $\mathcal{E}^*$  is the unique equilibrium in its positivity class  $\Gamma_\Omega$  and the inequalities (11) are strict, then  $y_i, Z_j \rightarrow 0$  for all  $i \notin \Omega_y, j \notin \Omega_z$ . If  $i \in \Omega_y$  and  $a_{ij} = 0 \ \forall j \in \Omega_z$ , i.e.  $\Lambda_i \cap \Omega_z = \emptyset$ , then  $y_i \rightarrow y_i^*$  and  $x^* = 1/\mathcal{R}_i$ . In addition, assuming positive initial conditions, for each  $i \in \Omega_y, j \in \Omega_z$ ,  $y_i$  and  $Z_j$  persist (the system is  $\Omega_{yz}$  persistent) with asymptotic averages converging to equilibria values, i.e.

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t y_i(s) ds = y_i^*, \quad \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t Z_j(s) ds = Z_j^*,$$

*In the case that there are less than or equal to two persistent viral strains with non-empty epitope sets (restricted to  $\Omega_z$ ), i.e.  $|\{i \in \Omega_y : \Lambda_i \cap \Omega_z \neq \emptyset\}| \leq 2$ , then  $\mathcal{E}^*$  is globally asymptotically stable.*

Note that the global convergence of the persistent variants to equilibria values is still an open question when there are more than two persistent immune responses. In Browne and Smith (2018), we conjecture the global stability of an equilibrium  $\mathcal{E}^*$  of (1) which is strictly saturated (11) and unique in its positivity class  $\Gamma_\Omega$ , regardless of  $\Gamma_\Omega$  dimension, based on the following observations. First, we proved Theorem 1 by showing solution dynamics asymptotically satisfy Lotka-Volterra (L-V) predator-prey differential equations, which in isolation need not have global convergence to  $\mathcal{E}^*$ , but there is an additional algebraic constraint of solutions to a hyperplane which may force global convergence to  $\mathcal{E}^*$ . Furthermore, all of our numerical simulations of (1) support global stability of  $\mathcal{E}^*$ , although convergence can be rather oscillatory and slow.

### 3.2 Fitness and epistasis

In the rest of this article we consider the “binary sequence” case of model (1), which leads to the simplified system (4) through assumption (2). For our virus-immune ecosystem, we are considering the situation where  $n$  immune response populations  $z_j$  each target the corresponding epitope  $j$  in the virus strains at a rate solely dependent on the allele type of epitope  $j$ ; the rate being positive for (0) wild-type or zero for (1) mutated form conferring full resistance to  $z_j$ . The avidity of immune response  $z_j$  and (wild-type) epitope  $j$  is described by the immune reproduction number  $\mathcal{I}_j$  given by (3), and are ordered according to the immunodominance hierarchy (6). As opposed to this simple immune fitness ordering, the collection of virus reproduction numbers (fitnesses) in our model can have much more complex relationships among each other depending on the fitness landscape, formally defined below.

Consider the space of binary sequences of length  $n$ ,  $\{0, 1\}^n$ , which contain all possible  $2^n$  virus strains. For a given strain  $i$  with sequence  $\mathbf{i} \in \{0, 1\}^n$ , we also denote its reproduction number in terms of binary sequence;  $\mathcal{R}_i$ . The reproduction numbers can be described in terms of the fitness cost (relative to wild-type fitness  $\mathcal{R}_0$ ) associated with the corresponding combinations of epitope mutations. The *fitness landscape* is defined as the precise map between the virus sequences and their reproduction numbers:

$$w : \{0, 1\}^n \rightarrow \mathbb{R}, \quad w(\mathbf{i}) = \mathcal{R}_i.$$

The set of all reproduction numbers is the image of the fitness landscape,

$$\mathcal{F} := w(\{0, 1\}^n) = \{\mathcal{R}_i\}_{i \in \{0, 1\}^n} = \{\mathcal{R}_i\}_{i=0}^{2^n-1},$$

where we can utilize either the sequence or integer indices for viral strains. An important special case of a fitness landscape is when  $w$  is *additive*. In an additive fitness landscape,

$$\mathcal{R}_i = \mathcal{R}_0 - \mathbf{c} \cdot \mathbf{i}, \quad (12)$$

where  $\mathbf{c} = \langle c_1, c_2, \dots, c_n \rangle$  is the vector of individual fitness costs for mutating each epitope, with the assumption that  $\mathbf{c} \cdot \mathbf{1} < \mathcal{R}_0$  and  $c_i > 0$ , so that all viral reproduction numbers remain positive and decrease with mutation satisfying (5).

Whereas an additive fitness landscape is solely determined linearly by the wild-type and single-mutant fitness values, the concept of epistasis allows for combinations of mutations to have more general nonlinear fitness landscapes. Informally, a system has epistasis if the effect of a mutation depends on genetic background. Here we generally define epistasis as a deviation from additivity. A common way to incorporate epistasis is via pairwise interactions between loci, as in the quadratic Ising or Pott's model (Stadler 2002) which has been used in applications to HIV-immune data (Barton et al. 2016). Let  $B$  be a strictly upper triangular matrix encoding (possibly random) pairwise interactions and define

$$\mathcal{R}_i = \mathcal{R}_0 - \mathbf{c} \cdot \mathbf{i} + \sum_{j=1}^n i_j \sum_{k>j} i_k B_{jk}, \quad \mathbf{i} \in \{0, 1\}^n, \quad (13)$$

where  $B, \mathbf{c}$  are suitable to fit our requirements for the viral fitness (cost) landscape (5), i.e. all viral reproduction numbers remain positive and decrease with mutation.

To study epistasis in general, first consider a subset of the sequence space  $\mathcal{S} \subset \{0, 1\}^n$  and associated fitness landscapes,  $w(\mathcal{S}) = \{\mathcal{R}_i\}_{i \in \mathcal{S}}$ . Motivated by the concept of interaction coordinates (Beerenwinkel 2007; Crona 2017), we measure genetic interactions among the sequences in  $\mathcal{S}$  by specifying a type of linear functional from the space of all possible fitness landscapes to  $\mathbb{R}$ :

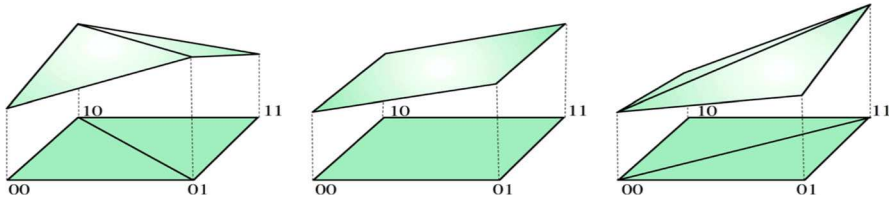
**Definition 1** Let a *vanishing linear form* on  $\mathcal{S}$  be a linear form  $g = \sum_{\mathbf{k} \in \mathcal{S}} \alpha_{\mathbf{k}} \mathcal{R}_{\mathbf{k}}$  with integer coefficients  $\alpha_{\mathbf{k}}$ , which is zero for any fitness landscape  $w$  that is additive, and satisfies  $\sum_{\mathbf{k} \in \mathcal{S}} \alpha_{\mathbf{k}} = 0$  with some  $\alpha_{\mathbf{k}} \neq 0$ .

Note that an equivalent definition, can be formed from the following observation upon consideration of additive fitness (12):

$$\begin{aligned} g &= \sum_{\mathbf{k} \in \mathcal{S}} \alpha_{\mathbf{k}} \mathcal{R}_{\mathbf{k}} = \sum_{\mathbf{k} \in \mathcal{S}} \alpha_{\mathbf{k}} (\mathcal{R}_0 - \mathbf{c} \cdot \mathbf{k}) = - \sum_{\mathbf{k} \in \mathcal{S}} \alpha_{\mathbf{k}} \mathbf{c} \cdot \mathbf{k} = -\mathbf{c} \cdot \sum_{\mathbf{k} \in \mathcal{S}} \alpha_{\mathbf{k}} \mathbf{k} \\ \Rightarrow g &= 0 \quad \forall \mathbf{c} \in \mathbb{R}^n \Leftrightarrow \sum_{\mathbf{k} \in \mathcal{S}} \alpha_{\mathbf{k}} \mathbf{k} = \mathbf{0}. \end{aligned}$$

So a vanishing linear form on  $\mathcal{S} \subset \{0, 1\}^n$  equivalently satisfies  $\sum_{\mathbf{k} \in \mathcal{S}} \alpha_{\mathbf{k}} \mathbf{k} = \mathbf{0}$  and  $\sum_{\mathbf{k} \in \mathcal{S}} \alpha_{\mathbf{k}} = 0$  with not all  $\alpha_{\mathbf{k}} = 0$ . The two conditions can be combined by adding to every binary sequence in  $\{0, 1\}^n$  a 1 at the end of the sequence. Considering each extended binary sequence  $\mathbf{11} = (i_1 \dots i_n 1)$  as a vector in  $\mathbb{R}^{n+1}$ , existence of a vanishing linear form on  $\mathcal{S} \subset \mathbb{R}^{n+1}$  simply signifies  $\mathcal{S}$  to be a linearly dependent set of vectors.

Vanishing linear forms are similar to the notion of additive dependence relations, along with interaction coordinates and Walsh coefficients; more commonly used epistasis measures, which are Fourier transforms on the group  $(\mathbb{Z}_2)^n$  measuring total



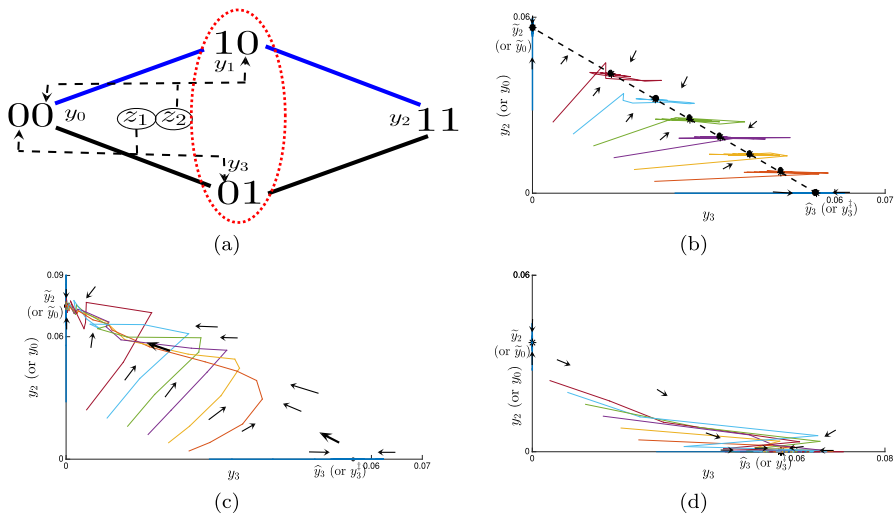
**Fig. 2** In the case  $n = 2$  epitopes, the circuit of all binary sequences with corresponding vanishing linear form  $\mathcal{A}$  delineates cases (diagrams from left to right) of (i) negative epistasis ( $\mathcal{A} < 0$ ), (ii) additive ( $\mathcal{A} = 0$ ), (iii) positive epistasis ( $\mathcal{A} > 0$ ), which determine fitness landscape shape (figure considers positive fitness contributions due to mutations from 0 allele to 1 allele in loci, as opposed to negative fitness costs (5) assumed in our model)

$k$ -way epistasis for  $2 \leq k \leq n$  (Weinreich et al. 2013; Crona 2017). In this work, Definition 1 and the provided equivalences readily allow us to introduce circuits (Beerenwinkel 2007), which give a fundamental description of epistatic interactions in fitness landscapes and relate to bifurcations in our model.

**Definition 2** A circuit  $\mathcal{C} \subset \{0, 1\}^n$  is a minimal set which has a vanishing linear form. In other words, there exists a vanishing linear form on a circuit  $\mathcal{C}$  and no proper subset of  $\mathcal{C}$  has a vanishing linear form.

Considering the extended binary sequences in  $\mathbb{R}^{n+1}$ , a circuit is a minimally linearly dependent collection of vectors, i.e. a linearly dependent collection of vectors in which any proper subset is linearly independent (Crona 2020). Circuits allow for detection of sign epistasis relative to a vanishing linear form. In particular, suppose  $\mathcal{C}$  is a circuit with vanishing linear form  $g = \sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} \mathbf{R}_{\mathbf{k}}$ , then the circuit  $\mathcal{C}$  has *sign epistasis* for fitness landscape  $w$  if  $\sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} \mathbf{R}_{\mathbf{k}} \neq 0$ . In a strictly additive fitness landscape the vanishing linear forms on each circuit would all be zeros, i.e. vanish. How to classify *positive epistasis* (or negative) on the circuit  $\mathcal{C}$  depends on the non-unique assignment of signs of the coefficients for the linear form.

The simplest class of circuits measure the *conditional epistasis* of a pair of loci against a fixed background of the other loci. For example, the classical example of two loci ( $n = 2$ ), has a single circuit distinguishing between positive, additive and negative epistatic fitness landscapes, depicted in Fig. 2. Here, the vanishing linear form  $\mathcal{A} := \mathcal{R}_{00} - \mathcal{R}_{10} - \mathcal{R}_{01} + \mathcal{R}_{11}$  is positive for a fitness landscape whenever the pairwise interaction between epitopes 1 and 2 are synergistic, so that the double mutant has larger reproduction number than it would have under additivity. As noted in Crona (2017), circuits are possibly a more fundamental epistasis measure than say interaction coordinates. For example in the case  $n = 3$ , the “ $u_{110}$ ” interaction coordinate,  $u_{110} = \mathcal{A}_{12} + \mathcal{A}'_{12} = \mathcal{R}_{000} - \mathcal{R}_{100} - \mathcal{R}_{010} + \mathcal{R}_{110} + \mathcal{R}_{001} - \mathcal{R}_{101} - \mathcal{R}_{011} + \mathcal{R}_{111}$ , is the sum of circuit linear forms,  $\mathcal{A}_{12}$  and  $\mathcal{A}'_{12}$ , giving the conditional epistasis of the first two loci against the third locus fixed at 0 and fixed at 1, respectively, and  $u_{110}$  is a vanishing linear form, but does not form a circuit because it does not satisfy minimal dependence property. We can identify how the “pairwise interaction fitness landscape” (13) directly relates to these conditional epistasis circuits in any dimension  $n$ . Consider loci  $1 \leq j < k \leq n$ . A circuit measuring the conditional epistasis of  $j, k$  against any



**Fig. 3** For  $n = 2$  epitopes, a single circuit of all viral strain binary sequences with corresponding linear combination of reproduction numbers determines epistasis and steady state in the case of persistent immune responses  $z_1$  and  $z_2$ . **a** The 2-d hypercube is displayed with distinct feasible persistent virus strain sequences depicted, where the vanishing linear form  $\mathcal{A} := \mathcal{R}_{00} - \mathcal{R}_{10} - \mathcal{R}_{01} + \mathcal{R}_{11}$  (positive  $\mathcal{A} > 0$  versus negative  $\mathcal{A} < 0$ ) decides stable nested equilibrium ( $\tilde{\mathcal{E}}_2$  or  $\mathcal{E}_2$  containing  $\{00, 10\}$  persistent strain set) versus stable one-to-one or network with at most one mutation equilibrium ( $\mathcal{E}_2^\pm$ ,  $\mathcal{E}_2^\pm$  or  $\tilde{\mathcal{E}}_2$  containing  $\{10, 01\}$  persistent strain set). **b** Bifurcation at  $\mathcal{A} = 0$  (additive fitness landscape) presents line of equilibria connecting nested  $\tilde{\mathcal{E}}_2$  and one-to-one  $\mathcal{E}_2^\pm$  (or at most on mutation  $\tilde{\mathcal{E}}_2$ ), projected on  $y_3, y_2$  ( $y_0$ ) axis. Here, the (colored) curves represent solution paths of (4) with different initial conditions converging to distinct points on line of equilibria, predicted by Proposition 2) for  $\mathcal{A} = 0$ . **c**  $\mathcal{A} > 0 \Rightarrow$  solutions converge to  $\tilde{\mathcal{E}}_2$ . **d**  $\mathcal{A} < 0 \Rightarrow$  solutions converge to  $\mathcal{E}_2^\pm$  (or  $\tilde{\mathcal{E}}_2$ )

background will resolve as follows;  $\mathcal{A}_{jk} := -\mathcal{R}_{\cdot 0 \cdot 0} + \mathcal{R}_{\cdot 1 \cdot 0} + \mathcal{R}_{\cdot 0 \cdot 1} - \mathcal{R}_{\cdot 1 \cdot 1} = B_{jk}$ , where the changing alleles occur in the  $j, k$  positions.

Another class of circuits relates marginal epistases of two pairs of loci (defined by interaction coordinates Beerenwinkel 2007). An example is the circuit which can be given by linear form  $\mathcal{A}(w) = -\mathcal{R}_{000} + \mathcal{R}_{001} + \mathcal{R}_{110} - \mathcal{R}_{111}$ , and also can be thought of as calculating conditional epistasis of the first 2 loci (as a block) and the third locus. Both conditional and “marginal” epistasis classes of circuits are comprised, against a background where a subset of loci are fixed, of two distinct pairs of (ones’) complement sequences, defined to be sequences  $\tilde{\mathbf{k}}$  and  $\bar{\mathbf{k}}$  where  $\tilde{\mathbf{k}} + \bar{\mathbf{k}} = \tilde{\mathbf{1}}$  for a subset of loci  $\tilde{J} \subset [1, n]$ . Note that because the sum of coefficients and weighted sum of sequences must vanish, along with a circuit being minimal, the number of binary sequences in a “(ones’) complement” circuit must be four. Other types of circuits measure higher order epistasis, and will be discussed later. In general, the number of circuits rapidly grows with  $n$  (there are 20 circuits for  $n = 3$ , 1348 circuits for  $n = 4$  Eble et al. 2020) and can be interpreted geometrically in terms of shapes formed by vertices of the  $n$ -cube (Beerenwinkel 2007).

## 4 Main results

In this section, we present our main theorems and their ramifications. Proofs to new results appear in the Appendix. Our major goal is to rigorously connect the concept of circuits with bifurcations and stable equilibria in model (4). First, in order to demonstrate a general link between circuits and the dynamical system, we establish that persistent viral strains comprise a circuit only in a critical case. In particular, we show that a circuit has positive components in a feasible equilibrium only when this circuit is additive with respect to fitness landscape, in which case a degenerate infinite dimensional subspace of equilibria appears. Indeed, the following proposition generalizes a previous result in Browne and Smith (2018) on degeneracy of equilibria forming a cycle in virus sequence hypercube.

**Proposition 2** *Consider the binary sequence model (4) with  $2^n$  viral strains identified in  $\{0, 1\}^n$ . Suppose that  $\mathcal{C} \subset \{0, 1\}^n$  is a circuit, has vanishing linear form  $g = \sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} w(\mathbf{k})$  for any additive fitness landscape  $w$ , and consider the fixed fitness landscape with image (reproduction numbers) denoted by  $\mathcal{R}_{\mathbf{k}}$  for  $\mathbf{k} \in \{0, 1\}^n$ . If  $\sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} \mathcal{R}_{\mathbf{k}} \neq 0$ , then there does not exist an equilibrium  $\mathcal{E}^* = (x^*, \mathbf{y}^*, \mathbf{z}^*)$  with  $y_{\mathbf{k}}^* > 0$  for all  $\mathbf{k} \in \mathcal{C}$ . On the other hand if  $\sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} \mathcal{R}_{\mathbf{k}} = 0$  and there exists an equilibrium with  $y_{\mathbf{k}}^* > 0$  for all  $\mathbf{k} \in \mathcal{C}$ , then there are infinitely many equilibria,  $\bar{\mathbf{y}}$ , in the positivity class of  $\mathbf{y}^*$ , with components parametrized by  $y_{\mathbf{k}} = y_{\mathbf{k}}^* + \beta \alpha_{\mathbf{k}}$  for some  $\beta \in \mathbb{R}$ .*

The proposition implies that any equilibrium with persistent strains forming a circuit must be unstable, in particular as part of a continuum of equilibria. The dimension of the infinite dimensional subspace of equilibria is the number of linearly independent vanishing forms corresponding to the circuit, where the dimension can be greater than one if the circuit contains distinct (sub-) circuits as subsets. Although unstable, the lines of equilibria will be seen in the ensuing sections as bifurcations where certain types of stable equilibria are invaded with strain replacement and stability being sharply determined by signed epistasis of the corresponding circuits.

### 4.1 Nested network determined by epistasis

Next, we focus on (perfectly) nested equilibria, which describe sequential mutations of epitopes in the order of the immunodominance hierarchy and persistence of all strains along this pathway. The successive rise of more broadly resistant prey (coming with a fitness cost) and weaker but more generalist predators, in a nested fashion, has been proposed in bacteria-phage communities (Jover et al. 2013; Korytowski and Smith 2015; Weitz et al. 2013), and there is some evidence that nestedness is a feature of HIV and immune response dynamics (Kessinger et al. 2015; Liu et al. 2013; Deutekom et al. 2013). Furthermore, this specialist-generalist structure is a well studied pattern in a variety of ecosystems, in particular nested networks are of interest in explaining the biodiversity and structure of mutualistic (e.g. plant-pollinator) communities (Bascompte et al. 2003).

First, we describe equilibria of model (4), where the persistent network is constrained to be nested, which were described in Jover et al. (2013), Korytowski and

Smith (2015), Browne (2017). We introduce a “nested priority” indexing for the viral strains, which allows convenient definition of threshold quantities for nested networks. The  $n + 1$  binary sequences contained in nested equilibria are of the form  $1^k 0^{n-k}$  (in power notation for the length  $n$  binary string), where  $0 \leq k \leq n$ . Let  $y_k$  denote the viral strain with binary sequence  $1^k 0^{n-k}$ ,  $0 \leq k \leq n$ . Then, for each  $k \in [1, n]$ , define the following *nested* equilibria:

$$\tilde{\mathcal{E}}_k = (\tilde{x}, \tilde{y}, \tilde{z}), \quad \tilde{x} = \frac{1}{\mathcal{R}_k}, \quad \tilde{y}_j = s_{j+1} - s_j \text{ for } 0 \leq j < k, \quad \tilde{y}_k = 1 - \frac{\mathcal{Q}_k}{\mathcal{R}_k}, \quad (14)$$

$$\tilde{z}_j = \frac{\mathcal{R}_{j-1} - \mathcal{R}_j}{\mathcal{R}_k} \text{ for } 1 \leq j < k, \quad \tilde{z}_k = 0, \quad \tilde{y}_j = \tilde{z}_j = 0 \text{ for } k < j \leq n$$

$$\bar{\mathcal{E}}_k = (\bar{x}, \bar{y}, \bar{z}), \quad \bar{x} = \frac{1}{\mathcal{Q}_k}, \quad \bar{y}_j = s_{j+1} - s_j \text{ for } 0 \leq j < k, \quad (15)$$

$$\bar{z}_j = \frac{\mathcal{R}_{j-1} - \mathcal{R}_j}{\mathcal{Q}_k} \text{ for } 1 \leq j < k, \quad \bar{z}_k = \frac{\mathcal{R}_{k-1}}{\mathcal{Q}_k} - 1, \quad \bar{y}_j = \bar{z}_j = 0 \text{ for } k < j \leq n,$$

$$\text{where } \mathcal{Q}_k = \mathcal{Q}_{k-1} + (s_k - s_{k-1})\mathcal{R}_{k-1}, \quad \mathcal{Q}_0 = 1, s_0 = 0, s_k = 1/\mathcal{I}_k. \quad (16)$$

Equilibrium  $\tilde{\mathcal{E}}_k$  represents the appearance of escape mutant  $y_k$  from the equilibrium  $\bar{\mathcal{E}}_k$  containing  $k$  viral strains  $y_0, \dots, y_{k-1}$  and immune responses  $z_1, \dots, z_k$ . Notice that  $\mathcal{Q}_k = \mathcal{Q}_{k-1} + \mathcal{R}_{k-1}y_{k-1}^*$ , and thus can be interpreted as the additional reproductive fitness contributed by the  $k^{\text{th}}$  nested strain scaled by its frequency at equilibrium. The stability of these equilibria *restricted within the nested network* (non-nested strains  $y_{n+1}, \dots, y_{2^n-1}$  are set to zero) was proved to be determined by the largest  $k$  such that  $\mathcal{R}_{k-1} > \mathcal{Q}_k$  and if equilibrium (14) is positive (if  $\mathcal{R}_k > \mathcal{Q}_k$ ) (Browne 2017), which we will expand upon below in our main result of this section, Theorem 2.

Before we present our new result, observe that, along with the specialist to generalist ordering in nested equilibria, nested networks are evolutionary pathways in the full fitness landscape hypercube. As opposed to some other feasible equilibria, such as the one-to-one network, the persistent strains in the nested equilibria form a path from the wild-type to the most resistant strain as single mutations accumulate in stepwise fashion. In a single (quasi-)species system, the underlying viral fitness landscape, which is generally shaped by epistatic interactions, determines evolutionary trajectories. When another trophic level is added, as immune response (predators) here, the overall viral fitnesses are expected to be dynamic since they depend upon the immune response populations. However, here we show that the nested trajectory in our system is solely dependent on the relevant epistasis in the viral fitness landscape.

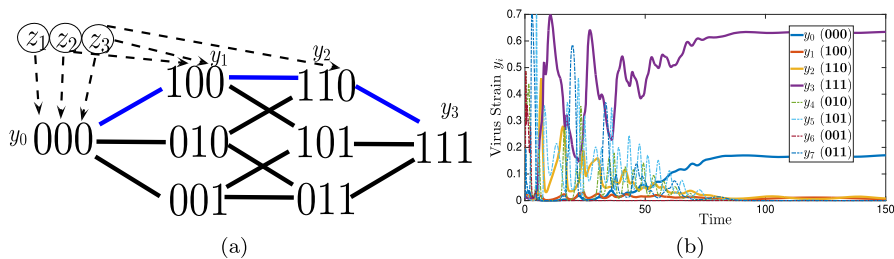
In order to define the appropriate viral fitness landscape epistasis corresponding to equilibrium stability, we introduce “invasion circuits” below.

**Definition 3** For an equilibrium  $\mathcal{E}^*$  and missing strain  $y_i$  (where  $y_i^* = 0$ , i.e.  $i \notin \Lambda_i$ ), let an *invasion circuit* be a circuit (def. 2),  $\mathcal{C}_i$ , containing exactly the  $y_i$  binary sequence,  $\mathbf{i}$ , and some positive  $\mathcal{E}^*$  component strain sequences (subset  $\mathcal{S}_i \subset \{\mathbf{k} \in \{0, 1\}^n : \text{corresponding component } y_k^* > 0\}$ ). The vanishing linear form,  $\mathcal{A}_i$ , is taken as negative of the scaled relative invasion rate in (11), namely  $\mathcal{A}_i \propto -\frac{\dot{y}_i}{y_i}|_{\mathcal{E}^*}$ , and we say  $\mathcal{C}_i$  has *positive epistasis* if  $\mathcal{A}_i > 0$ .

The interpretation of this definition is that each of these circuits represent potential alternate pathways, which correspond to strain invasion in the model. Thus, we say that a certain equilibrium assemblage of strains on the hypercube of binary sequences has *positive epistasis* if every invasion circuit corresponding to a missing strain has positive epistasis. Theorem 2 (below) proves that the nested network is stable and persistent if and only if it has positive epistasis as a pathway in the viral fitness landscape. In particular, we decode the general saturated equilibria inequalities (11) conferring stability and persistence by Theorem 1 into biological meaningful conditions on sign epistasis of associated invasion circuits. The utility of introducing invasion circuits is two-fold. First, because circuits measure epistasis, our ensuing equilibrium stability/persistence results connect evolutionary genetics with ecological dynamics principles. Second, we will see that by formulating as sign epistasis of circuits, a missing strain's scaled invasion rate at equilibrium with  $n$  or  $n + 1$  persistent strains simplifies to a linear combination of the much smaller subset of persistent strain fitnesses and invader fitness which are connected through the circuit. Moreover, which equilibrium strain can be replaced resolves further based on corresponding coefficients of the circuit linear form. Although our model does not explicitly include mutation, the persistent variants of stable equilibria can still represent evolutionary outcomes, as later simulations show. Thus the following theorem suggests a necessary and sufficient condition based on epistasis in the viral fitness landscape for a nested trajectory in a generalized eco-evolutionary version of model (4).

**Theorem 2** Consider the binary sequence model (4) with  $2^n$  viral strains and  $n$  immune responses ordered by immunodominance (6). Assume that  $\mathcal{R}_0 > \mathcal{Q}_1$  (at least one virus strain and immune response persists). Let  $k$  be the largest integer in  $[1, n]$  such that  $\mathcal{R}_{k-1} > \mathcal{Q}_k$ . The conclusions of Theorem 1 at nested equilibria (14) hold if each of  $2^k - k - 1$  invasion circuits has positive epistasis. In particular,  $\tilde{\mathcal{E}}_k$  (or  $\tilde{\mathcal{E}}_k$  if  $\mathcal{R}_k > \mathcal{Q}_k$ ) is stable with uniformly persistent strains  $y_0, y_1, \dots, y_k$  (and  $y_{k+1}$  if  $\mathcal{R}_k > \mathcal{Q}_k$ ) if each invasion circuit consisting of a non-nested strain ( $\{y_i\}$ ,  $i \in [n+1, \dots, 2^n-1]$ ) contained in the  $k$ -dimensional (sub-)hypercube (binary sequence has  $i_{k+1} = \dots i_n = 0$ ) union a subset of nested strains ( $\mathcal{S}_i \subset \{y_0, y_1, \dots, y_k\}$ ), has positive vanishing linear form,  $\mathcal{A}_i$ , where  $\mathcal{A}_i$  is proportional to  $-\frac{y_i}{y_i|_{\mathcal{E}^*}}$ .

We provide two proofs of the above theorem, given in the appendix. First, we prove the stability condition pattern by adopting a linear algebra approach where each binary sequence is extended by an additional fixed bit. This leads to a solvable system of equations for the linear forms and circuits determining nested equilibria stability. Second, we apply a combinatorial technique to find the strains in the nested network forming the circuit and linear form for each possible invading strain not in the nested network. In particular, we distinguish a “non-nested sequence”  $\mathbf{i}$  by existence of a (01) string, and utilize an induction argument on the number of such strings. Each method yields equivalent, yet distinct, characterizations of the critical circuits  $\mathcal{C}_i$  and linear forms  $\mathcal{A}_i$ . Furthermore, when  $\mathcal{R}_n > \mathcal{Q}_n$ , in the case of the full nested network ( $n + 1$  strains persistent), the equilibrium  $\tilde{\mathcal{E}}_n$  becomes unstable with the non-positivity of an invasion circuit, yielding an “if and only if” condition for stability. These extended results are summarized below in a corollary to Theorem 2.



**Fig. 4** Convergence to nested network assuming multiplicative viral fitness landscape. **a** Persistent viral strains is reduced to nested network (in blue) as  $t \rightarrow \infty$  in Prop. 5. **b** Viral strain components  $y_0(t), \dots, y_2(t)$  persist as system (4) converges to equilibrium  $\tilde{\mathcal{E}}_3$

**Corollary 1** Suppose  $\mathcal{R}_n > \mathcal{Q}_n$ , so that, the equilibrium  $\tilde{\mathcal{E}}_n$  is feasible. A necessary and sufficient condition for stability of  $\tilde{\mathcal{E}}_n$  is the positivity of  $2^n - n - 1$  linear forms  $\mathcal{A}_i$  corresponding to circuits  $\mathcal{C}_i$ , each containing a single missing strain  $y_i$ ,  $i \in [n+1, 2^n - 1]$ , along with strains in the nested network dependent on the  $y_i$  sequence  $\mathbf{i} = (i_1 \dots i_n)$  in the following equivalent ways:

- i* Define the sequence  $(\alpha_j)$ ,  $j = 0, 1, \dots, n$ , where  $\alpha_0 = 1 - i_1$ ,  $\alpha_j = i_j - i_{j+1}$  for  $j = 2, \dots, n-1$ ,  $\alpha_n = i_n$ . Let  $\mathcal{J}_i$  be the nonzero terms in sequence  $(\alpha_j)$ , i.e.  $\mathcal{J}_i := \{j \in [0, n] : \alpha_j \neq 0\}$ , where  $\alpha_j = \pm 1$  for  $\alpha_j \in \mathcal{J}_i$ . Then

$$\mathcal{C}_i = y_i \cup \{y_j\}_{j \in \mathcal{J}_i}, \quad \mathcal{A}_i = -\mathcal{R}_i + \sum_{j \in \mathcal{J}_i} \alpha_j \mathcal{R}_j. \quad (17)$$

- ii* Let  $0 \leq m_1 < p_1 < m_2 < \dots < p_s < m_{s+1} \leq n$  denote the positions  $p_1, \dots, p_s$  beginning the  $s$  (01) strings and positions  $m_1, \dots, m_{s+1}$  of the last “1” before and after the (01) strings. In other words, the sequence  $\mathbf{i}$  in “power notation” is given by

$$\mathbf{i} = 1^{m_1} 0^{p_1 - m_1} 1^{m_2 - p_1} \dots 0^{p_s - m_s} 1^{n - m_{s+1}}. \text{ Then}$$

$$\begin{aligned} \mathcal{C}_i &= y_i \cup \{y_{m_j}, y_{p_j}\}_{j=1}^s \cup y_{m_{s+1}} = \mathbf{i} \cup \{1^{m_j} 0^{n-m_j}, 1^{p_j} 0^{n-p_j}\}_{j=1}^s \cup 1^{m_{s+1}} 0^{n-m_{s+1}}, \\ \mathcal{A}_i &= -\mathcal{R}_i + \sum_{j=1}^{s+1} \mathcal{R}_{m_j} - \sum_{j=1}^s \mathcal{R}_{p_j}. \end{aligned} \quad (18)$$

Thus, the  $n+1$  strain nested network equilibrium is stable if and only if it has positive epistasis.

In order to illustrate Theorem 2 and accompanying Corollary 1, we first discuss the model dynamics in the case  $n = 2$ , which is depicted in Fig. 3a and was found to have precisely 10 distinct feasible persistent variant sets (global asymptotic stability in 8 of these regimes) in Browne and Smith (2018). In this case, there is just one “non-nested” strain,  $y_{01}$ , with the single mutation escaping the second (subdominant) immune response  $z_2$ . The single circuit consists of this strain together with the nested strains, totaling the whole sequence space, i.e.  $\mathcal{C} = \{01, 00, 10, 11\}$ , along with the

corresponding linear form  $\mathcal{A} = \mathcal{R}_{00} - \mathcal{R}_{10} - \mathcal{R}_{01} + \mathcal{R}_{11}$ . Thus the sign of the single quantity  $\mathcal{A}$  determines the stability and persistence of the nested network. Here  $\mathcal{A} > 0$  implies that the persistent strains and positive components of the stable equilibria lie within the nested network  $\mathcal{N} = \{00, 10, 11\}$ . The precise persistence structure when  $\mathcal{A} > 0$  depends upon which of equilibria (14) and (15) are positive. In particular, the diversity increases stepwise from just the wild-type virus  $y_0$  to both immune responses  $z_1, z_2$  and three nested strains  $y_0, \dots, y_k$  based upon the largest  $k$  such that  $\mathcal{R}_{k-1} > \mathcal{Q}_k$  and whether  $\mathcal{R}_k > \mathcal{Q}_k, k = 1, 2$ , where  $z_1, z_2$  persist when  $\mathcal{R}_1 > \mathcal{Q}_2$ . On the other hand when  $\mathcal{A} < 0$  (which implies  $\mathcal{R}_1 > \mathcal{Q}_2$  and  $z_1, z_2$  persist), the nested equilibrium is invaded by  $y_3$  (01). Yet  $y_1$  (10) always persists when any immune escape occurs, independent of the sign of  $\mathcal{A}$  and even when  $y_3$  would have a larger escape rate in the single epitope case. Thus, we suggested in Browne and Smith (2018) that immunodominance may be the most important factor in multi-epitope escape, which was also inferred from data analysis in a previous study of HIV (Liu et al. 2013).

The feasible strain invasions obtained for  $n = 2$  in previous work (Browne and Smith 2018) can be seen as the simplest example of a more general pattern for bifurcations from nested equilibria obtained from Theorem 2 and Proposition 2. When the (sign) epistasis in one of the circuits defining the nested pathway becomes negative, the nested network becomes unstable and a transcritical bifurcation occurs. In particular, a missing strain invades the nested network when the corresponding circuit goes from positive to negative epistasis. In the critical case of zero epistasis, or circuit additivity, there is a line of equilibria, given by Proposition 2, which connects the nested equilibrium with the invasion equilibrium. Indeed consider the nested equilibrium  $\tilde{\mathcal{E}}_n$ . We arrange the (persistent) nested virus components, together with the invading strain, in the vector  $\tilde{\mathbf{v}} = (\tilde{\mathbf{y}}, 0)^T$ , where the  $\tilde{\mathbf{y}}$  is from (14) and the last component is the invading strain,  $y_i$ , which is zero when at equilibrium  $\tilde{\mathcal{E}}_n$ . In the critical case, where the linear form  $\mathcal{A}_i$  corresponding to circuit  $\mathcal{C}_i$  is zero, there is a line of equilibria given by  $\mathbf{v}^* = \tilde{\mathbf{v}} - \beta \boldsymbol{\alpha}$  where  $\boldsymbol{\alpha}$  is the (circuit) coefficients of  $\mathcal{A}_i$  and  $0 \leq \beta \leq C$  with  $C = \min \{\alpha_k \tilde{y}_k : \alpha_k > 0, k = 0, 1, \dots, n\}$ . Thus, in the bifurcation where  $y_i$  invades  $\tilde{\mathcal{E}}_n$ , the invading strain replaces one of the nested strains in the circuit with positive coefficient ( $\alpha_k > 0$ ), in particular the above “ $C$ -minimizing” nested strain,  $\arg \min (\{\alpha_k \tilde{y}_k : \alpha_k > 0, k = 0, 1, \dots, n\})$ . By the proof of Theorem 2, the positive coefficients correspond to a subset of nested strains given in order as:  $y_{m_i}, i = 1, \dots, s$ , where  $0 \leq m_1 < p_1 < m_2 < \dots < p_s < m_{s+1} \leq n+1$  count the maximal position of a 1 before each of  $s$  01 strings (each at position  $p_1, \dots, p_s$ ) in the sequence of the missing strain. Which of these feasible strains are replaced depends on the model parameters. Notice that if  $\alpha_k = 1$  for all  $k$  such that  $\alpha_k > 0$ , then in a feasible equilibrium after invasion by  $y_i$ , the replaced strain would be the “circuit positive coefficient” nested strain with smallest value at the nested equilibrium. Thus, the replaced strain must have the property of being the inferior competitor in the nested hierarchy with a positive coefficient in circuit linear form. In the following subsection, we will see a similar principle in invasion of another equilibria structure besides the nested structure, namely the one-to-one network.

A major advantage of investigating the critical case of virus strain  $y_i$  invading a known equilibria structure (here the nested network) is that new equilibria can be obtained by application of Proposition 2, the circuit coefficients, and known equi-

libria values. The line of equilibria (virus and immune components denoted by  $\mathbf{y}$  and  $\mathbf{z}$ ) remain positive in some neighborhood around the bifurcation parameter set where the circuit linear form,  $\mathcal{A}_i$ , is zero. Indeed, the values of  $\mathbf{z}$  remain constant throughout the line of equilibria for  $\mathcal{A}_i = 0$ , so the positive components in the boundary nested equilibrium carry over to the boundary equilibrium of the new invasion equilibrium. For the simple case of  $n = 2$  described above, the loss of stability of  $\tilde{\mathcal{E}}_2$  when  $\mathcal{A} = 0$  results in strain  $y_3$  (01) replacing either  $y_0$  (00) if  $\tilde{y}_0 < \tilde{y}_2$ , or  $y_2$  (11) if  $\tilde{y}_0 > \tilde{y}_2$  (displayed in Fig. 3b). By (14), the strain which is replaced depends upon the sign of  $s_1 - 1 + \frac{Q_2}{R_2}$ . In the case of  $n = 3$ , there are  $2^3 - 3 - 1 = 4$  circuits corresponding to a non-nested invading strain. Explicitly the circuits, characterized by the corresponding linear form (with the non-nested strain term appearing first), are as follows: (i)  $\mathcal{A}_4 = -\mathcal{R}_{010} + \mathcal{R}_{000} - \mathcal{R}_{100} + \mathcal{R}_{110}$ , (ii)  $\mathcal{A}_5 = -\mathcal{R}_{001} + \mathcal{R}_{000} - \mathcal{R}_{110} + \mathcal{R}_{111}$ , (iii)  $\mathcal{A}_6 = -\mathcal{R}_{101} + \mathcal{R}_{100} - \mathcal{R}_{110} + \mathcal{R}_{111}$ , (iv)  $\mathcal{A}_7 = -\mathcal{R}_{011} + \mathcal{R}_{000} - \mathcal{R}_{100} + \mathcal{R}_{111}$ . Thus, Theorem 2 implies the nested equilibrium is stable if and only if all of the quantities (i)-(iv) are positive. Furthermore, in each case that a single inequality fails, the following bifurcation occurs where the missing strain replaces a nested strain  $y_j$  where  $j$  is determined by (i)  $\arg \min_{j=0,2}(\tilde{y}_j)$ , (ii)  $\arg \min_{j=0,3}(\tilde{y}_j)$ , (iii)  $\arg \min_{j=1,3}(\tilde{y}_j)$ , (iv)  $\arg \min_{j=0,3}(\tilde{y}_j)$ , where  $\tilde{y}_j$  are defined in terms of viral and immune response fitness quantities in (14). For example, if a bifurcation from nested equilibrium  $\tilde{\mathcal{E}}_3$  occurs through inequality (iii) switching sign, then  $y_6$  (101) replaces either  $y_1$  or  $y_3$ , depending on whether  $\tilde{y}_1 < \tilde{y}_3$ , i.e.  $s_1 < 1 - \frac{Q_3}{R_3}$ . In the case this inequality holds and  $y_1$  is replaced, the new stable equilibrium will consist of persistent strain (sequence) set  $\{101, 000, 110, 111\}$ . For  $n = 4$ , there are 11 circuits determining stability of nested network, 10 of which consist of 4 strains (ones' complement circuits) and one that has 6 strains in the circuit,  $\mathcal{A}_{0101} := -\mathcal{R}_{0101} + \mathcal{R}_{0000} - \mathcal{R}_{1000} + \mathcal{R}_{1100} - \mathcal{R}_{1110} + \mathcal{R}_{1111}$ . Thus in the case of invasion of the nested equilibrium  $\tilde{\mathcal{E}}_4$  by strain 0101, there are 3 possible strain replacements and (in terms of integer indexing)  $\arg \min_{j=0,2,4}(\tilde{y}_j)$  determines which nested strain is replaced.

We can expand upon our observation of the importance of immunodominance in determining viral evolution. We notice that in any of the invasion scenarios, a viral strain containing minimal sequential mutations to the most immunodominant responses will remain in the equilibrium, no matter the fitness costs. For  $n = 2$ , we had observed that  $y_1$  (strain 10) always persists. For  $n = 3$ , the only invasion scenario where  $y_1$  does not persist can be the case of 101 invasion with invasion equilibrium consisting of strain sequences  $\{000, 110, 101, 111\}$ . For the nested equilibrium with  $n+1$  strains,  $\tilde{\mathcal{E}}_n$ , replacement of the immunodominant resistant strain  $y_1$  only can occur with invasion by a non-nested strain with resistance at the first epitope (sequence of form  $10 \dots$  with at least two "1" alleles), so that all strains will have at least 2 mutations.

## 4.2 One-to-one network determined by epistasis

Now we turn to another possible persistent equilibrium assemblage of virus and immune response variants; the one-to-one (or strain-specific) network. Consider the viral strains that have gained resistance to  $n$  or  $n - 1$  immune response, forming a

subsystem of (4) with the  $m = n + 1$  strains containing more than  $n - 1$  mutations ( $n - 1$  ones in binary sequence). For convenience, we index the strains according to the position of the susceptible epitope (zero in binary sequence), so that in more general Eq. (1),  $y_i$ ,  $i = 1, \dots, n + 1$  has epitope set  $\Lambda_i = \{i\}$  or  $\Lambda_{n+1} = \emptyset$  and  $A$  is a  $n + 1 \times n$  matrix comprised of the diagonal matrix  $\text{diag}(a_1, \dots, a_n)$  and a row of zeros. This subsystem of a “one-to-one” interaction network, where each immune response population attacks a unique specific viral strain, has been considered in Wolkowicz (1989), Korytowski and Smith (2015), Bobko and Zubelli (2015). Stability and persistence results, analogous to Browne (2017) for the nested subsystem, were proved in Wolkowicz (1989) for the one-to-one network (strains  $y_i$ ,  $i = 1, \dots, n + 1$ , without hypercube) under the assumption of strictly decreasing strain reproduction numbers. Although relaxing their assumption of decreasing reproduction numbers allows for multiple saturated degenerate equilibria (with “persistent strains” as subsets of  $[1, n + 1]$ ) in the  $n + 1$  strain one-to-one network, we previously proved in Browne and Smith (2018) that the fitness assumptions (5) of our  $2^n$  strain model (4) permit one of two feasible “one-to-one network” equilibria to be saturated (or stable) in the larger hypercube. Indeed, define the following equilibria with persistent strains  $y_1, \dots, y_n$  ( $\mathcal{E}_n^\dagger$ ), and with persistent strains  $y_1, \dots, y_{n+1}$  ( $\mathcal{E}_{n+1}^\ddagger$ ), where:

$$\begin{aligned} x_i^\ddagger &= \frac{1}{\mathcal{R}_{n+1}}, \quad y_i^\ddagger = s_i, \quad z_i^\ddagger = \frac{\mathcal{R}_i}{\mathcal{R}_{n+1}} - 1, \quad i = 1, \dots, n, \quad y_{n+1}^\ddagger = 1 - \frac{\mathcal{P}_n}{\mathcal{R}_{n+1}}, \\ x_i^\dagger &= \frac{1}{\mathcal{P}_n}, \quad y_i^\dagger = s_i, \quad z_i^\dagger = \frac{\mathcal{R}_i}{\mathcal{P}_n} - 1, \quad i = 1, \dots, n, \quad \text{with } \mathcal{P}_n = 1 + \sum_{k=1}^n s_k \mathcal{R}_k, \quad s_k = 1/\mathcal{I}_k. \end{aligned} \quad (19)$$

Expanding upon our prior results, we prove the stability of the one-to-one network is determined by  $2^n - n - 1$  circuits corresponding to potential invading strains.

**Theorem 3** Consider system (4) on the full network with  $n$  epitopes ( $m = 2^n$  virus strains) and fitness costs (5). Suppose the viral strains,  $y_i$   $i = 0, \dots, 2^n - 1$ , are ordered so that  $\Lambda_j = \{j\}$  for  $j = 1, \dots, n$  and  $\Lambda_{n+1} = \emptyset$  (where  $\Lambda_j$  denotes strain  $j$  epitope set (2)). If  $\mathcal{E}_n^\dagger$  or  $\mathcal{E}_{n+1}^\ddagger$  with corresponding positive epistasis invasion circuits,  $A_i > 0$ , then  $\mathcal{E}_n^\dagger$  (if  $\mathcal{R}_{n+1} \leq \mathcal{P}_n$ ) or  $\mathcal{E}_{n+1}^\ddagger$  (if  $\mathcal{R}_{n+1} > \mathcal{P}_n$ ) is stable with persistent strains  $y_1, \dots, y_n$  ( $y_{n+1}$  also if  $\mathcal{R}_{n+1} > \mathcal{P}_n$ ). The vanishing linear forms  $A_i$  and invasion circuits  $C_i$  corresponding to missing strains  $y_i$ ,  $i = 0, n + 2, \dots, 2^n - 1$ , are characterized below:

$$C_i = y_i \cup \{y_j\}_{i_j=0}, \quad A_i = -\mathcal{R}_i - (|\Lambda_i| - 1) \mathcal{R}_{n+1} + \sum_{j \in \Lambda_i} \mathcal{R}_j \quad (j \in [1, n]). \quad (20)$$

Furthermore, these are the only scenarios of stable “one-to-one network” equilibria and the above positive epistasis of  $n + 1$  strain positive equilibrium,  $\mathcal{E}_{n+1}^\ddagger$ , is a necessary and sufficient condition for its stability.

Note the proof of this theorem is in Appendix, and here we make a few remarks to interpret the result. First, observe that the reproduction number  $\mathcal{R}_i$  of a potential invading strain  $y_i$ , depends on its epitope set  $\Lambda_i$ . Because each mutation comes with a fitness cost (5),  $\mathcal{R}_i$  roughly correlates with number of susceptible (non-mutated) epitopes,  $|\Lambda_i|$ , and thus both negative terms and the positive summation in (20) increase with  $|\Lambda_i|$ . Therefore, there is no general rule for determining the sign of invading strain circuits corresponding to the one-to-one network, each depending on the relevant combinations of fitness costs, i.e. epistasis. We can discuss possible strain replacements for invasion of  $\mathcal{E}_{n+1}^\pm$  as before. In this case, we find that the replaced strain is  $y_{\min(j \in \Lambda_i)}$ , i.e. the strain susceptible to strongest immune response among the susceptible epitopes of strain  $\mathbf{i}$ , since this strain has lowest value in equilibrium corresponding to positive coefficient in circuit. Compared to the  $n + 1$  strain nested network ( $\tilde{\mathcal{E}}_n$ ), the “invasion circuit” and strain replacement of the  $n + 1$  strain one-to-one network ( $\mathcal{E}_{n+1}^\pm$ ) is simpler to determine. Note that invasion of the  $n$  strain  $\mathcal{E}_n^\pm$  can result in addition of the new strain rather than replacement, and the critical case does not correspond to a line of equilibria as with the  $n + 1$  strain equilibria. As an example of the invasion circuits (20), consider the case  $n = 3$ , where the missing strains with a single ( $0 \rightarrow 1$ ) mutation have vanishing linear forms:  $\mathcal{A}_{100} = -\mathcal{R}_{100} - \mathcal{R}_{111} + \mathcal{R}_{101} + \mathcal{R}_{110}$ ,  $\mathcal{A}_{010} = -\mathcal{R}_{010} - \mathcal{R}_{111} + \mathcal{R}_{011} + \mathcal{R}_{110}$ ,  $\mathcal{A}_{001} = -\mathcal{R}_{001} - \mathcal{R}_{111} + \mathcal{R}_{011} + \mathcal{R}_{101}$ . Each corresponds to an embedded 2-cube measuring marginal epistasis with their single 1-allele fixed. Note that  $\mathcal{A}_{100} = -\mathcal{A}_6$ , where  $\mathcal{A}_6$  also is the circuit corresponding to invasion of nested equilibrium by (101). Now consider potential invasion by the wild-type strain (000) given by  $\mathcal{A}_{000} = -\mathcal{R}_{000} - 2\mathcal{R}_{111} + \mathcal{R}_{011} + \mathcal{R}_{101} + \mathcal{R}_{110}$ , which biologically tells us how the three ( $0 \rightarrow 1$ ) two-mutation sequences combine with respect to fitness in comparison to the three-mutation combination (111). Of note, the sign of this circuit does not have a two-locus interpretation, making them truly of higher-order (Gould et al. 2018).

### 4.3 Other equilibrium network structures and open questions

The full utility of the circuit analysis comes with bifurcations of equilibria with  $n + 1$  strains, as our above examples illustrate, because the critical state corresponds to persistent strains forming a circuit in Proposition 2. How far can we go with this analysis? Can we generalize to all equilibrium structures? Observe from the proofs of Theorem 2 and Theorem 3 that the two equilibrium networks considered, nested and one-to-one, with  $n + 1$  strains ( $\tilde{\mathcal{E}}_n$  and  $\mathcal{E}_{n+1}^\pm$ ) form a basis of  $\mathbb{R}^{n+1}$  when the strains are considered as binary sequences with a one added at the the end of sequences, and moreover every binary sequence has integer coordinates with respect to this basis. This directly leads to the “invasion circuits”, and this is generalized to any assemblage of  $n + 1$  strain sequences in the following proposition (proof in Appendix):

**Proposition 3** *Suppose  $\mathcal{S} \subset \{0, 1\}^n$  is the set of binary sequences of an equilibrium,  $\mathcal{E}^*$ , with  $n + 1$  strains ( $|\Lambda_y| = n + 1$ ). Assume that  $\mathcal{S} \times \{1\}$  is a basis of  $\mathbb{R}^{n+1}$  and any added binary sequence  $\mathbf{i}1 \in \{0, 1\}^n \times \{1\}$  has integer coordinates with respect to this basis. Then for all  $\mathbf{i} \in \{0, 1\}^n \setminus \mathcal{S}$ ,  $\mathcal{C} = \{\mathbf{i}\} \cup \mathcal{S}$  forms a circuit with vanishing*

linear form is  $\mathcal{A}_i = -\mathcal{R}_i - \sum_{\mathbf{k} \in \mathcal{S}} \alpha_{\mathbf{k}} \mathcal{R}_{\mathbf{k}}$ , where  $(\alpha_{\mathbf{k}})_{\mathbf{k} \in \mathcal{S}}$  are coordinates of  $\mathbf{i}$  with respect to  $\mathcal{S} \times \{1\}$ . Furthermore, for any missing strain,  $y_i$ ,  $i \notin \Lambda_y$ , the invasion rate is given by  $\frac{\dot{y}_i}{y_i y_i} = -x^* \mathcal{A}_i$ , and thus the stability of  $\mathcal{E}^*$  is determined by the sign of  $\mathcal{A}_i$ .

Now consider the scenario that strain  $\mathbf{k} \in \mathcal{S}$  is replaced by  $\mathbf{i}$ , then the new equilibrium sequences  $\mathcal{S}' = \{\mathbf{i}\} \cup \mathcal{S} \setminus \{\mathbf{k}\}$  forms a basis of  $\mathbb{R}^{n+1}$  since any proper subset of a circuit is linearly independent. Thus the strain replacement with invader  $\mathbf{i}$  will result in this new equilibrium structure  $\mathcal{S}'$  also forming a circuit if any sequence has integer coordinates with respect to  $\mathcal{S}' \times \{1\}$ . In this fashion, we might observe a sequence of strain invasions determined by circuits. Notice that strain invasions of the two  $n + 1$  strain equilibria structures explored here, nested and one-to-one networks, would result in a strain replacement whose new equilibrium has stability determined by linear form on circuit. Indeed, because the coordinate of any potential invader  $\mathbf{i}$  was shown to be  $+1$  corresponding to the strain it can replace, it is not hard to show that the new basis will also yield integer coordinates for any other sequence. Once we move past this initial invasion though, it would not be clear if the circuit stability pattern continues though.

Another consideration is whether a strain can be added to an  $n$  strain equilibrium (where  $n$  is number of persistent immune responses) in order to have a positive  $n + 1$  strain equilibrium which satisfies Proposition 3, i.e. forms a set  $\mathcal{S}$  corresponding to a basis with integer coordinates in the extended  $n + 1$  dimensional binary sequence space. In our examples, we add the completely resistant strain (with sequence  $\mathbf{1}$ ) to the  $n$  strain nested or one-to-one networks (with  $n$  persistent immune responses) to get an  $n + 1$  strain equilibrium satisfying the hypotheses of Proposition 3. In general, this might not always be the case. First, we recall that determining the feasibility of a  $n + 1$  strain positive equilibrium is dependent on calculation of  $C = (A' \mathcal{R}')^T$  by Proposition 1, with  $A'$  as the virus-immune interaction network of the  $n + 1$  strains where the rows of  $A'$  correspond to the complements  $(\mathbf{1} - \mathbf{i})$  of the viral sequences in  $\mathcal{S}$ . If there is a feasible  $n$  strain equilibrium with network  $A$  and reproduction numbers  $\mathbf{R}$ , then the complete resistance strain  $\mathbf{1}$  can be added if  $\mathcal{R}_1 > 1 + \mathbf{a} A^{-1} \mathbf{R}$ . However, the calculation for adding other strain sequences is more complicated, thus the problem of both determining feasibility and whether an equilibrium satisfies Proposition 3 may be difficult.

As an example, consider another possible equilibrium type, the  $n$  strain 1-mutation network:  $\mathcal{S}_1 = \{y_i^1 \mid i = 1, \dots, n\}$  in which  $y_i^1$  has only escaped  $z_i$  so that its binary sequence is  $\mathbf{i}^1 = (\delta_{\ell i})_{\ell=1}^n$  where  $\delta_{\ell i}$  is Kronecker delta function. If we add  $y_w(\mathbf{0})$  to  $\mathcal{S}_1$ , then circuits determine stability, however adding  $\mathbf{1}$  does not yield circuits determining stability (in particular stability condition for invasion by  $\mathbf{0}$ ) is not a circuit. Indeed, we can derive some conditions for positivity of an equilibrium consisting of viral strains  $\tilde{\mathcal{S}}_1 = \{\mathbf{0}\} \cup \mathcal{S}_1$  (see Appendix 1). Consider the case  $n = 3$ , where the circuit for invasion of  $\tilde{\mathcal{S}}_1$  by strain  $\mathbf{i} = \mathbf{1}$  can be calculated according to coordinate basis description in extended sequence space:

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \alpha \Rightarrow \mathcal{A} = -\mathcal{R}_i - \sum_{\mathbf{k} \in S} \alpha_{\mathbf{k}} \mathcal{R}_{\mathbf{k}} = -\mathcal{R}_{111} - 2\mathcal{R}_{000} + \mathcal{R}_{100} + \mathcal{R}_{010} + \mathcal{R}_{001}$$

Similar, to the example circuit given in the one-to-one network, this measures higher-order epistasis, in particular whether the one-mutation associations predict the three-mutation combination. Here, the strain replacement would be 111 replacing 001 because this sequence would have the smallest equilibrium value of positive coefficient strains in  $\tilde{S}_1$ . It can be shown the other invasion circuits correspond to conditional epistasis (embedded 2-cubes), where the single non-mutated epitope of the invader remains fixed. Indeed, using the coordinate basis method above, we have the following proposition for invasion of the “ $\leq 1$  mutation” network:

**Proposition 4** Consider the network with at most one mutation,  $\tilde{S}_1$ , consisting of wild-type and 1-mutation viral strains  $y_0, y_1, \dots, y_n$  where the sequence of  $y_j$  is  $\mathbf{j} = (\delta_{\ell j})_{\ell=1}^n$  for  $j = 1, \dots, n$ . Suppose that there is a positive equilibrium,  $\hat{E}_n$ , with  $\tilde{S}_1$  as persistent viral strain set  $\Omega_y$ .  $\hat{E}_n$  is stable if and only if  $\mathcal{A}_i > 0$ , where  $i = n + 1, \dots, 2^n - 1$ , and linear forms  $\mathcal{A}_i$  correspond to invasion circuits  $C_i$ , as characterized below:

$$C_i = y_i \cup y_0 \cup \{y_j\}_{j=1}^n, \quad \mathcal{A}_i = -\mathcal{R}_i - (n - |A_i| - 1) \mathcal{R}_0 + \sum_{j \notin \Lambda_i} \mathcal{R}_j \quad (j \in [1, n]). \quad (21)$$

Observe that for the case of  $n = 3$ , we have now highlighted all the circuits determining stability of three equilibria structures: the nested, one-to-one, and one-mutation network. While there are 4 corresponding linear forms for each network dictating invasion by each missing strain, together this results in 10 distinct circuits since  $C = \{000, 100, 110, 010\}$  and  $\bar{C} = \{100, 110, 111, 101\}$  are invasion circuits that the nested network shares with the one-mutation and one-to-one network, respectively. There are 20 total circuits for  $n = 3$  (Beerenwinkel 2007), and we leave it to future work as to whether the any of the other 10 circuits correspond to stability of feasible “transitional equilibria” between the highlighted networks. However, the immunodominance hierarchy will impose an effective fitness ordering on the virus genotypes so that for example the “reverse nested” network  $\{000, 001, 011, 111\}$  would never be feasible. Therefore, some circuits should not correspond to any meaningful bifurcation under the assumptions of our model.

#### 4.4 Special cases of fitness landscapes

While fitness landscapes on the  $n$ -dimensional hypercube generally yield a multitude of circuits determining bifurcations and stability of equilibria, there are some simple landscapes that can be analyzed. First, consider the pairwise interaction case as described by Eq. (13), where  $\mathcal{R}_i = \mathcal{R}_0 - \mathbf{c} \cdot \mathbf{i} + \sum_{j=1}^n i_j \sum_{k>j} i_k B_{jk}$  for a (strictly) upper triangular matrix  $B$ . If the matrix  $B$  is positive, then the fitness of any sequence

with at least 2 mutations will always be larger than the additive case, whereas if  $B$  is negative, the resulting fitness from a pair of mutations is less than expected under additivity. Thus, in the former case of  $B$  positive, synergistic interactions should favor double mutants, while in the latter antagonistic interactions might discourage consecutive mutations. The exact translation of these informal notions to expected results in our model with sign-definite pairwise interactions is not obvious due to there being a dynamic overall fitness landscape when taking into account immune response (predator) populations and other variables/parameters which might influence the viral escape pathway. Nevertheless, we prove here that the nested network is generally stable when pairwise loci interaction matrix  $B$  is positive, whereas a non-nested network, such as one-to-one or “at most one mutation” network, is stable when  $B$  is negative.

**Theorem 4** *Consider binary sequence model (4) having pairwise interaction fitness landscape (13) with upper triangular matrix  $B$  that is sign-definite. Assume that  $\mathcal{R}_0 > \mathcal{Q}_1$  (so that at least one virus strain and immune response persists). If  $B_{jk} > 0$  for all  $k > j$ , then the nested network is stable with uniformly persistent strains (as in Theorem 2). On the other hand, if  $B_{jk} < 0$  for all  $k > j$ , then one-to-one network (or network with at most one mutation) is stable against invasion and persistent if components of associated equilibrium are positive.*

Another basic example of a fitness landscape is *multiplicative*, where each mutation at a fixed locus reduces the reproduction number of a strain by a fraction regardless of the of sequence background at other loci. Thus the loci act independently, but not additively. This *multiplicative* fitness landscape has been assumed in several studies of HIV-immune evolution at multiple epitopes, e.g. Althaus and Boer (2008), Deutekom et al. (2013). We prove the following proposition, generalizing a theorem in Browne and Smith (2018) showing multiplicative equal fitness costs evolve a nested network.

**Theorem 5** *Assume that fitness costs of mutating locus  $j$  come with a multiplicative reproductive loss  $f_j$ , i.e.  $\mathcal{R}_i = \mathcal{R}_0 \prod_{j=1}^n f_j$  where  $0 < f_j < 1$ ,  $j = 1, \dots, n$ . Then the nested network is stable and persistent.*

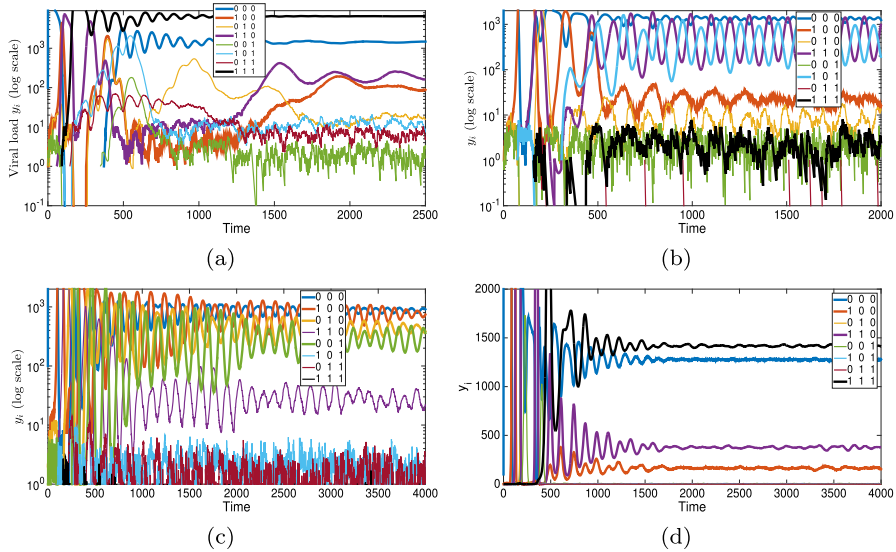
## 5 Simulations of and predictions for virus-immune evolution

In this section, we conduct simulations of model (4), along with a hybrid stochastic/deterministic version, in order to illustrate our results. The model was coded in MATLAB, where the built-in ODE solver ODE45 was utilized for simulations. For the deterministic model, we find numerical solutions to (4) under the multiplicative viral fitness landscape for  $n = 3$  epitopes, initiating the simulation with positive concentrations of all virus and immune variants,  $y_i$ ,  $i = 0, \dots, 7$  and  $z_j$ ,  $j = 1, 2, 3$ , where we adopt the nested priority indexing from Sect. 4.1. The immunodominance hierarchy utilized in the simulation is  $\mathcal{I}_1 = 6, \mathcal{I}_2 = 5.7, \mathcal{I}_3 = 5.4$ . We assume each epitope mutation imparts equal independent multiplicative fitness costs, i.e. if  $(i_1 \dots i_n)$  represents the epitope sequence of strain  $i$  and  $\mathcal{R}_i = \mathcal{R}_0(1 - \kappa)^{i_1 + \dots + i_n}$  where  $\mathcal{R}_0 = 11.8$  and  $\kappa = 0.1$  is fitness cost in our simulation. The scaling factors

for viral and immune variant growth rates in (4) are set to:  $\gamma_i = 3.5, i = 0, \dots, 3$  and  $\gamma_i = 18.5, i = 4, \dots, 7$ . The corresponding calculations lead to positive epistasis in the invasion circuits of the nested equilibrium  $\tilde{\mathcal{E}}_3$  (Theorem 2 and Proposition 5) and, as shown in prior work (Browne and Smith 2018), result in a sequential nested immune escape trajectory (Fig. 4).

Do the predicted patterns from our theoretical results on differential equation system (4) hold when random mutation is included (as in HIV infection)? To answer this question, we conduct numerical simulations of a stochastic extension of the model using parameters representative of HIV. However, since this is a preliminary simulation effort, we choose a rather large viral wild-type (basic) reproduction number  $\mathcal{R}_0$  and low death rate of immune response to better mimic virus-immune evolution for the stochastic model, as in Rife Magalis et al. (2021). Similar to the methods in Deutekom et al. (2013), we simulate mutations of the  $n$  loci by drawing from a binomial distribution in a hybrid ODE-stochastic algorithm. With a mutation rate of  $\epsilon = 1.67 \times 10^{-3}$  per site per day, we compute the number of mutations during replication as follows. We update mutations at fixed time steps, taken as  $\Delta t = 1$  day, where we approximate the daily number of cells that become de novo infected per viral variant as  $M_i = \beta_i X Y_i$  cells. To improve computation speed, we assume that only one of the  $n$  loci mutates per replication, i.e. the small probability of simultaneous mutations are neglected. Then for each viral variant  $i = 1, \dots, m$  and locus  $\ell = 1, \dots, n$ , the number of mutations is given by  $\text{Bin}(M_i, \epsilon)$ . The viral populations are updated accordingly, and the ODE solver is run for  $\Delta t$  time units and then the process repeats. In the following simulations, we assume that initially there is just the wild-type virus,  $y_0(0) > 0$ , all other strains are absent  $y_i(0) = 0, i = 1, \dots, 2^n - 1$ , and each immune response is present,  $z_j(0) > 0, j = 1, \dots, n$ . Thus the extended model allows for random mutation and deterministic selection evolving from initial infection by the founder (wild-type) strain.

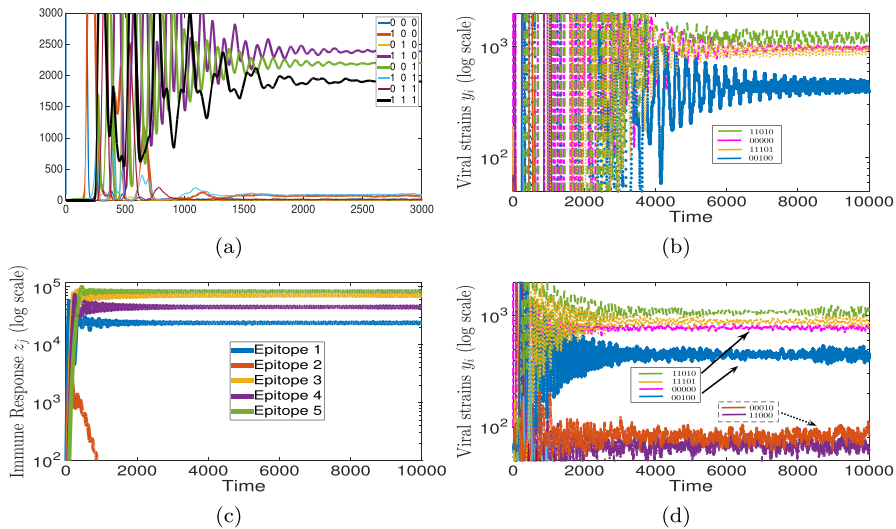
First for the stochastic extension of (4), consider  $n = 3$  epitopes, which for simplicity is much less than an actual HIV genome and taken to be a representative cluster or sample of loci. We utilize variables and parameters from the unscaled version of (4), system (1) in Browne and Smith (2018)  $X = \frac{b}{c}x, Y_i = \frac{b}{\delta}y_i, Z_j = \frac{\rho_j}{\mathcal{I}_j}z_j, \rho_j = \frac{\mu_j}{bq_j}, \sigma_j = \frac{\mu_j}{c}$  in order to represent concentrations ( $ml^{-1}$ ) of target cells, virus and immune response, along with immune decay and scaling factor. Let  $b = 5 \times 10^3 (ml \cdot d)^{-1}, c = 0.01 d^{-1}, \delta = 0.5 d^{-1}, \mu_j = 0.01 d^{-1}, q_j = 1.5$ . For the immunodominance hierarchy, we assume that  $\mathcal{I}_j, j = 1, \dots, n$  are ordered uniform random variables in the range  $[3.75, 7.875]$ . First, assume that the viral fitnesses are calculated as  $\mathcal{R}_i = \left[ \sum_{j=1}^n (1 - \kappa_j) + \sum_{i,j=1, i \neq j}^n B_{jk} \right] \mathcal{R}_0$ , where additive fitness costs  $\kappa$  were uniformly distributed in the range  $[0, 0.5]$ . and pairwise interaction  $B_{jk}$  is uniformly distributed (random positive epistasis) in the range  $[0, 1]$ . Then, all pairwise interactions,  $B_{jk}$ , are positive, along with the invasion circuits which we index  $i = 1, \dots, 4$  in ascending order with respect to the invading binary sequence conversion to decimal representation. The system is expected to converge to the nested network by Proposition 4, with asymptotic stability of equilibrium  $\tilde{\mathcal{E}}_3$ , persistence of nested strains  $y_0, \dots, y_3$  and extinction of remaining viral strains  $y_4, \dots, y_7$  subject to small perturbations caused by random mutations, as displayed in Fig. 5a. Next, we



**Fig. 5** Simulations of extended model with stochastic mutation and pairwise epistatic interactions illustrate eco-evolutionary dynamics. **a** Trajectories of virus strains in the case of  $n = 3$  epitopes with uniformly distributed viral and fitness quantities, and (random) positive pairwise interactions,  $B_{jk}$ , which implies positive epistasis with respect to “nested circuits” and convergence to nested steady state containing  $\{000, 100, 110, 111\}$ . **b** Increasing viral strain  $y_6$  (101) reproduction number ( $\mathcal{R}_6$ ) changes the sign of its invasion circuit so that epistasis is no longer positive, resulting in it replacing  $y_3$  (111) and non-nested persistent strains. **c** Assuming negative pairwise interactions also leads to non-nested convergence, here to network with at most one mutation containing  $\{000, 100, 010, 001\}$ . Note that (110) strain persists at low levels due to invasion circuit being close to zero, along with random mutation. Gaussian distributed pairwise interactions ( $B_{jk}$  random sign) result in convergence to nested network in **d** because positive  $B_{jk}$  randomly drawn

increase the reproduction number of  $y_6$  (101), so that the corresponding invasion circuit linear form  $\mathcal{A}_3$  switches from positive to negative. From our feasible bifurcations based on the circuit coefficients, we predict that (101) can replace (100) or (111). Observe in Fig. 5b, that equilibrium  $\mathcal{E}_3$  is altered by (101) invading (111), although the mutations allow to (010) to be only at slightly lower levels than (100) in the new strain hierarchy.

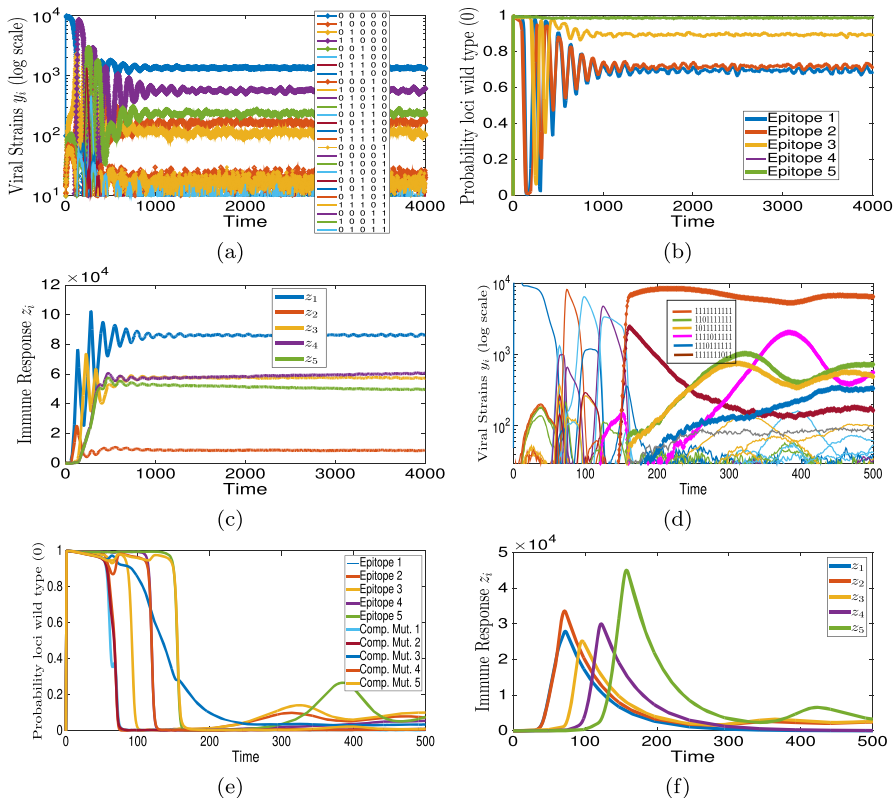
When epistatic interactions become negative by subtracting the pairwise matrix terms,  $B_{jk}$ , from additive fitnesses, we project a non-nested pattern according to Proposition 4. Indeed, in Fig. 5c, simulations converge to the network with at most one mutation, and hence the antagonism of negative interactions between epitopes thwarts the escape of virus at multiple epitopes. Finally, we consider Gaussian distributed pairwise interactions, where  $B_{jk}$  are random normal variables with mean zero (random signs) and variance of 0.1 affecting magnitude of epistasis. Observe that the system may (Fig. 5d) or may not converge (Fig. 6a) to the nested network depending on the sign of the invasion circuits determining the overall epistasis encoded in the nested pathway. Furthermore, in the latter case, simulations converge to an equilibrium structure that is not “close” to being nested, one-to-one, or “at most one mutation” network,



**Fig. 6** Simulations of model with stochastic mutation consistent with analysis of viral fitness epistasis in deterministic system. Generally, Gaussian distributed pairwise interactions can result in trajectories converging to steady states other than nested, one-to-one or “at most one mutation”, as illustrated in simulations (a), and in (b,c,d), with  $n = 3$  and  $n = 5$  epitopes, respectively. Observe that the dynamics in original (deterministic) ODE solution displayed in (b) are consistent with stochastic mutation simulations (g,h), except for low level persistence of two strains with small negative invasion rates

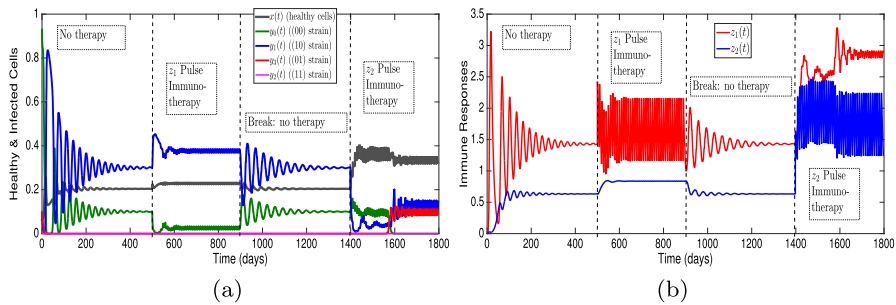
indicating the presence of additional stable equilibrium structures and corresponding circuits not analyzed in this study for the  $n = 3$  epitope setting. We also consider  $n = 5$  under epitopes with the same fitness landscape structure, although a variance of 0.05 in the normally distributed pairwise epistasis is set to counteract accumulated fitness cost from strains with more mutated epitopes. Simulations displayed for this case show that numerical solutions of the (deterministic) model (4) (Fig. 6b) are consistent with the stochastic extension (Fig. 6c and d), supporting our argument that theoretical results in the differential equations carry over to the eco-evolutionary dynamics with random mutation. Here, the fitness costs and non-positive epistasis circuits (with respect to nested network) prevent the dominance of strains with several mutations, and lead to the extinction of the weakest immune response  $z_5$ , along with persistence of only 4 strains, despite the 5 epitopes.

In Fig. 7, we simulate eco-evolutionary dynamics again for 5 epitopes under Gaussian distributed pairwise interactions, where  $B_{jk}$  are zero-mean normal random variables with variance of 0.05, and all other parameter assumptions remaining the same. The balance between immune response pressure selecting for resistance and the fitness costs occurring with each epitope mutation results in the virus mutant strains evolving to escape some immune responses, but the ancestral strains, including wild-type  $y_0$  can still persist (Fig. 7a). In addition, “backward” mutations allow mutated epitopes to revert back to wild-type (0) in a large proportion of viral population (Fig. 7b), even after invasion by mutant allele (1), as the sign of the invading circuit linear form and rise of more immune response populations (Fig. 7c) deter-



**Fig. 7** Fixation of resistant alleles and nestedness increases when incorporating compensatory mutations. **a, b** and **c** Simulations of model (4) with random viral mutations for  $n = 5$  epitopes under Gaussian distributed pairwise epitope interaction fitness cost landscape show that **a** viral strains  $y_i(t)$ , **(b)** allele frequency at each epitope and **(c)** immune responses  $z_j(t)$  converge to steady state with large prevalence of wild-type (0) allele in viral population at each epitope. **d, e** and **f** Adding complementary loci for each epitope to model which can compensate for 95% of fitness cost of resistance mutations. The compensatory mutations drive **d** viral strains  $y_i(t)$  to rapidly converge to “nearly nested” structure as **e** sequential epitope and corresponding compensatory mutations sequentially become ascendant in population, and **f** immune responses  $z_j(t)$  are escaped in order of immunodominance hierarchy

mine strain additions or replacements which result in the persistent strain structure of the equilibrium. In HIV infection, resistance mutations often to become more dominant in viral population with several escapes persisting in the population without reversion because of compensatory mutations in linked loci which allow the virus to regain most of the fitness cost associated with an epitope mutation (Althaus and Boer 2008). We simulate compensatory mutations by adding a complementary loci for each epitope  $j = 1, \dots, 5$ , which is either neutral (0), not impacting fitness or if mutated (1) can result in the virus restoring 95% of its original fitness value if the strain has mutated epitope  $j$  from wild-type (0) to resistant (1). Indeed, consider loci  $5 + j$ ,  $j = 1, \dots, 5$ , and viral sequence  $i'$  with  $i_{5+j} = 0$  which has undergone mutation and fitness cost in epitope  $j$  from neighboring strain



**Fig. 8** Simulating pulse immunotherapies in two-epitope model shows priming subdominant response  $z_2$  is more effective than therapy with dominant response  $z_1$ . Viral strain  $y_i(t)$  and healthy cell  $x(t)$  (a), along with immune response  $z_j(t)$  (b), trajectories in model (4) under no treatment initially, then periodic  $z_1$  immune infusions, followed by treatment interruption, and finally periodic  $z_2$  immune infusions. Even though subdominant resistant strain  $y_3$  (01) has higher reproduction number, the  $z_2$  therapy restores larger healthy cell count as system bifurcates from nested (strains {00, 10}) to “at most one mutation” (strains {00, 10, 01}) persistent network

$i$  ( $i_j = 0 \rightarrow i'_j = 1 \Rightarrow \mathcal{R}_i > \mathcal{R}_{i'}$ ). Then assuming all other epitopes remain fixed, we suppose the strain  $i''$  gaining compensatory mutation has the following update in fitness:  $i'_{5+j} = 0 \rightarrow i''_{5+j} = 1 \Rightarrow \mathcal{R}_{i''} = \mathcal{R}_{i'} + .95 \left(1 - \frac{\mathcal{R}_{i'}}{\mathcal{R}_i}\right)$ . In contrast to the case of reproduction numbers solely dependent on epitope sequence, the addition of these complementary loci allows for sequential epitope escapes with concurrent compensatory mutations dominating the viral population (Fig. 7d,e) and suppressing the immune response (Fig. 7f).

Finally, we numerically illustrate implications of our results for designing potential immunotherapy strategies against an immune escaping virus such as HIV. We consider the deterministic ODE (4) with  $n = 2$  epitopes (diagram shown in Fig. 3a), and add periodic infusions of the immune response populations,  $z_1(t)$  and  $z_2(t)$ . In particular, we incorporate periodic infusion times,  $t_k = t^{(j)} + k\tau$ ,  $k = 1, \dots, N$ , of the immune population  $z_j$  by applying an impulsive increase of  $D$  units to the model, i.e. Dirac delta distributions ( $D\delta(t - t_k)$ ) are added to the  $\dot{z}_j$  component in (4), and numerically solve in the cases of no treatment and distinct immunotherapies (see Fig. 8). The viral fitness parameters utilized are  $\mathcal{R}_0 = 15$ ,  $\mathcal{R}_1 = 8$ ,  $\mathcal{R}_2 = 3$ ,  $\mathcal{R}_3 = 11.5$ ,  $\mathcal{I}_1 = 10$ , and  $\mathcal{I}_2 = 2.5$  so that without therapy the system converges to nested equilibrium  $\tilde{E}_2$  with  $y_0, y_1, z_1, z_2$  persisting. Upon convergence to this rest point after perturbing the wild-type (immune-free) virus equilibrium  $\tilde{E}_0$  by introducing mutant strains and immune responses, at  $t^{(1)} = 500$  days we begin to pulse the dominant immune response  $z_1$  by adding  $D = 1$  units of cells every  $T = 10$  days (Fig. 8a and b). The persistent variants remain in the same nested structure and the system settles into a periodically forced solution with an increase in the “ $z_1$ -resistant” viral strain ( $y_1$  or 10) prevalence, decrease in  $y_0$ , and modest 12.4% jump in healthy cell count. After removing the  $z_1$ -therapy and solutions returning to original state  $\tilde{E}_0$ , at  $t^{(2)} = 1400$  days we test the periodic  $z_2$ -therapy with the same impulse magnitude of  $D = 1$  and frequency  $T = 10$  days. Contrary to the first therapy, the periodic infusion of  $z_2$  immune cells causes a bifurcation from the nested to the “at most one mutation” network with

addition of the subdominant  $z_2$ -resistant strain  $y_3$  (01) into the viral quasi-species. Furthermore, both  $z_1$  and  $z_2$  populations are enhanced and the healthy cells increase by around 67%. In each case, the stability condition given by inequality (11) is altered, so that even though the viral fitnesses  $\mathcal{R}_i$  are constant, the pulsed  $z_j$  levels can be thought to induce effective reproduction numbers which may change the sign of epistasis in the circuits (17) (or (18)) corresponding to the nested equilibria. Here, the strategy of priming the subdominant response  $z_2$  tilts this effective fitness landscape toward negative epistasis, convergence to “at most one mutation” network, and, although invasion by the higher mutant fitness strain  $y_3$  occurs, an improved outcome for host is obtained.

## 6 Discussion

In this paper, we rigorously connect population dynamics thresholds with concepts from evolutionary genetics, which allows us to characterize distinct regimes of multi-strain persistence, stability, and resistance pathways in a virus-immune ecosystem in a biologically meaningful manner. The complexity of the viral (binary sequence) genetic structure, along with dynamic virus fitness landscape and immune response populations, lead to a multitude of equilibria and general stability conditions which challenge interpretation, classification or simplification in terms of fundamental parameters such as reproduction number. By finding equivalent sharp thresholds based on an appropriate definition of epistasis in the fitness landscape governing persistent equilibrium network structures, we are able to gain insight on eco-evolutionary dynamics. In particular, the prediction of the virus escape pathway against immune attack on multiple epitopes is determined by epistasis in the “invasion circuits” controlling the bifurcations in our dynamical system.

Our theoretical results lend support to *circuits*, the minimal additive combinations of binary sequences (Beerenwinkel 2007), as the fundamental measure of epistasis in a fitness landscape. Other ways to quantify epistasis may be simpler or offer other advantages, but circuits underly fitness landscape shape, and here we show that they also dictate prey-predator dynamics on top of building the phenotypic/genetic structure of the prey (virus) population. This connection between population dynamics and genetics naturally comes from applying linear algebra to formulate the invasion rates of missing virus strains at an equilibrium as minimal combinations of virus reproduction numbers. Moreover, the invasion circuit and corresponding linear form encode the resident strains which can be replaced by a mutant strain, and together with their equilibrium strain densities, determine the bifurcations resulting in new feasible steady states.

The persistent network structures of virus and immune response populations analyzed in this work represent distinct patterns formed by the forces of viral resistance and fitness costs, and immunodominance. The nested network equilibria admits a diverse ecosystem with generalist to specialist ordering in prey-predator interactions, as opposed to the modularity of the one-to-one (strain-specific) and “at most one mutation” network. In terms of viral escape from the immune response, the nested pathway offers the most efficient evolution as mutant strains sequentially gain resis-

tance to immune populations strongest to weakest. That the special case of positive (or synergistic) pairwise interactions between epitopes presents a nested trajectory (Proposition 4) highlights how convergence to this network coincides with the classical definition of positive epistasis favoring double mutants. While this proposition may be expected, both the dominant epitope escape being favored even when exacting a larger fitness cost than other epitopes and the viral (prey) fitness epistasis determining fate of the virus-immune (prey-predator) ecosystem, are less intuitive features of the result, along with our more general Theorem 2 on nested network equilibrium stability. In contrast, the one-to-one and “at most one mutation” network are instances of resulting dynamics for negative (antagonistic) pairwise interactions, and particularly the “at most one mutation” structure is ideal from the host perspective of containing multi-epitope resistance.

Numerical solution of the ordinary differential Eq. (4), along with an extended version that includes randomly drawn mutations, demonstrate how eco-evolutionary trajectories are determined by epistasis in the viral fitness landscape, as predicted by our analytical results. Indeed efficient viral escape in a nested fashion occurs when our necessary and sufficient conditions regarding positive epistasis are satisfied, and becomes more complex as negative epistatic interactions allow different combinations of resistance mutations to persist in the virus population. Under random epistatic pairwise interactions, any number of equilibria structures can be realized which may hinder multi-epitope resistance, but compensatory mutations may allow for sequential viral escape of immune responses, as shown in Fig. 7. Furthermore, our model and results may inform upon immunotherapy for HIV. In most clinical trials of therapeutic vaccines, potentially favorable T cell responses were of limited success due to viral escape from epitopes used in vaccine (Pantaleo and Lévy 2013), but one possible strategy is to immunize with a set of the most conserved (associated with high fitness cost of resistance), subdominant epitopes (McMichael 2006; Ahmed et al. 2019). Thus, it may be desirable to guide the virus-immune trajectory toward a non-nested network structure by priming subdominant immune responses. Here, we illustrate that this strategy can work even when resistance to subdominant response comes with less fitness cost, as a bifurcation is induced to a state with viral mutant competition and optimal healthy cells compared to an immunotherapy with the dominant response (see Fig. 8).

Future work can build upon our results in several directions. While the dynamics for  $n = 2$  epitopes is resolved for model (4), the case  $n \geq 3$  has not been completely classified, and our work shows that feasible stable equilibria may be discovered through analysis of relevant circuits, although even  $n = 3$  is challenging due to large number of strain combinations. One way to explore how a particular ecosystem structure evolves is to follow the convergence of stepwise mutations and selection from wild-type strain in the hybrid stochastic/deterministic approach of polymorphic evolution sequences (Champagnat and Méléard 2011). However, simulations conducted (not shown here) revealed that the attracting (saturated) equilibrium was not obtained by a sequence of viral strain and immune response invasions starting from initial infection by the **0** strain, thus multi-loci mutations and invasions are necessary, perhaps in the spirit of the “adaptive walks” jumping between equilibria of Lotka-Volterra systems developed in Kraut and Bovier (2019). This approach of obtaining Lotka-Volterra dynamics from

limits of stochastic models relies on strong conditions guaranteeing global stability for the ODE, and so it is an open problem for our system. Finally, by incorporating data on the vial fitness landscape at multiple epitopes in the face of epistatic interactions and concurrent immune response attack, model parameterization with calculation of “invasion circuits” may verify theoretical results, predict eco-evolutionary trajectory, and inform upon potential immunotherapies.

**Acknowledgements** We would like to thank the editor and anonymous reviewers for comments which have led to an improved manuscript. We also thank Hal Smith for insightful discussions.

**Funding** CJB and FY acknowledge support by a U.S. National Science Foundation grant (DMS-1815095).

## Declarations

**Conflict of interest** The authors have no conflicts of interest to declare that are relevant to the content of this article.

## Appendix

### Proofs of Theorems

**Proof of Proposition 2** Let  $\mathcal{C} \subset \{0, 1\}^n$  be a circuit and suppose by way of contradiction that  $\mathcal{E}^* = (x^*, y^*, z^*)$  is an equilibrium of (4) with  $y_k^* > 0$  for corresponding sequences  $\mathbf{k} \in \mathcal{C}$ . Adding up the relative growth rates of these nonzero components at equilibrium in differential Eq. (4), we find the following:

$$\begin{aligned}
 0 &= \sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} \frac{\dot{y}_{\mathbf{k}}}{y_{\mathbf{k}} y_{\mathbf{k}}^*} \\
 &= x^* \sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} \mathcal{R}_{\mathbf{k}} - \sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} - \sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} \sum_{j=1}^n (1 - \mathbf{k}_j) z_j^* \\
 &= x^* \sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} \mathcal{R}_{\mathbf{k}} - \sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} - \sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} (\mathbf{1} - \mathbf{k}) \cdot \mathbf{z}^* \\
 &= x^* \sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} \mathcal{R}_{\mathbf{k}} - (1 + \mathbf{1} \cdot \mathbf{z}^*) \sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} + \left( \sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} \mathbf{k} \right) \cdot \mathbf{z}^* \\
 &= x^* \sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} \mathcal{R}_{\mathbf{k}},
 \end{aligned}$$

because  $\sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} \mathbf{k} = 0$  and  $\sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} = 0$  for a circuit. This contradicts assumption  $\sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} \mathcal{R}_{\mathbf{k}} \neq 0$  and thus proves the first statement. The next statement follows from Proposition 1 upon assuming  $\sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} \mathcal{R}_{\mathbf{k}} = 0$ . Indeed, uniqueness of equilibrium in a certain positivity class is equivalent to  $\text{Ker}(A')^T \cap \mathcal{R}'^{\perp} = \{\mathbf{0}\}$ , which is equivalent to the condition that the augmented matrix  $C = (A' \ \mathcal{R}')^T$  has trivial kernel (Browne and Smith 2018). Here  $A'$  is the  $m' \times n'$  interaction matrix consisting of the  $m'$  strains

comprising the circuit and  $n'$  (positive component) immune responses. Consider the vector  $\alpha$  consisting of the circuit weights. Then from the previous points, we find that  $C\alpha = \mathbf{0}$ . Thus, there cannot be a unique equilibrium with  $y_{\mathbf{k}}^* > 0$  for all  $\mathbf{k} \in \mathcal{C}$ . If such an equilibrium exists, then the virus component vector denoted by  $\bar{\mathbf{y}}$  which satisfies  $\bar{\mathbf{y}} - \mathbf{y}^* \in \text{Ker}(A')^T \cap \mathcal{R}'^\perp$ , also satisfies equilibrium equations, and thus there are infinitely many equilibria with component  $\bar{\mathbf{y}} = \mathbf{y}^* + \beta\alpha$  for  $\beta \in \mathbb{R}$ .  $\square$

**Proof of Theorem 2** In order to prove the theorem, we show that the general saturated equilibria inequalities (11) reduce to vanishing linear forms of invasion circuits (Definition 3) when evaluated at the nested network equilibria (14). Then we will apply Theorem 1 to show stability of nested equilibrium and uniform persistence of associated positive component strains.

Let  $\mathcal{R}_0 > \mathcal{Q}_0 := 1$  and  $k$  be the largest integer in  $[1, n]$  such that  $\mathcal{R}_{k-1} > \mathcal{Q}_k$ . First, we look at the case  $k = n$  and  $\mathcal{R}_n > \mathcal{Q}_n$  when  $\mathcal{E}_n$  is non-negative equilibrium with positive components at the  $n + 1$  nested strains  $y_0, y_1, \dots, y_n$ . Consider a given missing viral strain  $y_i$  ( $i \in [n + 1, 2^n - 1]$ ) with sequence  $\mathbf{i}$ . We define a linear form,  $\mathcal{A}_i$ , based on its invasion rate as follows:

$$\frac{\dot{y}_i}{y_i y_i} = -\frac{\mathcal{A}_i}{\mathcal{R}_n}, \quad \text{where} \quad -\mathcal{A}_i := \mathcal{R}_i - \mathcal{R}_n - \sum_{j=1}^n (1 - i_j) (\mathcal{R}_{j-1} - \mathcal{R}_j). \quad (22)$$

The telescoping sum above is determined by the following sequence:  $(\alpha_j)$ ,  $j = 0, 1, \dots, n$ , where  $\alpha_0 = 1 - i_1$ ,  $\alpha_j = i_j - i_{j+1}$  for  $j = 2, \dots, n - 1$ ,  $\alpha_n = i_n$ . In this way,  $-\mathcal{A}_i := \mathcal{R}_i - \sum_{j=0}^n \alpha_j \mathcal{R}_j$ . In order to prove that this is a vanishing linear form of a circuit, we show that it is the linear form of a minimally linearly dependent collection of extended binary sequences. Denote the binary sequences of nested network as  $\mathbf{k}_0, \dots, \mathbf{k}_n$  corresponding to ordered strains  $y_0, \dots, y_n$ . Let  $\mathcal{N} \subset \{0, 1\}^{n+1}$  denote the subset of nested extended binary sequences, where  $\hat{\mathbf{i}} = \mathbf{i}1 \in \{0, 1\}^{n+1} \setminus \mathcal{N}$  and  $\hat{\mathbf{k}} = \mathbf{k}1 \in \mathcal{N}$  represent binary sequences extended by digit 1. Notice that  $\mathcal{N}$  forms a basis of  $\mathbb{R}^{n+1}$  (since the  $n + 1 \times n + 1$  matrix  $(\mathbf{k}_n, \mathbf{k}_{n-1}, \dots, \mathbf{k}_0)$  has a triangular row reduced echelon form with values  $\pm 1$  on diagonal). Thus for  $\hat{\mathbf{i}} \in \{0, 1\}^{n+1} \setminus \mathcal{N}$ , there is a unique set of coefficients  $\alpha_j$ ,  $j = 0, 1, 2, \dots, n$ , yielding  $\hat{\mathbf{i}}$  as a linear combination of the nested network vectors:

$$\hat{\mathbf{i}} = \alpha_0 \hat{\mathbf{k}}_0 + \alpha_1 \hat{\mathbf{k}}_1 + \dots + \alpha_n \hat{\mathbf{k}}_n.$$

The above linear system resolves as follows:

$$\begin{aligned} \alpha_1 + \alpha_2 + \dots + \alpha_n &= i_1 \\ \alpha_2 + \alpha_3 + \dots + \alpha_n &= i_2 \\ &\vdots \\ \alpha_k + \alpha_{k+1} + \dots + \alpha_n &= i_k \\ &\vdots \end{aligned}$$

$$\begin{aligned}\alpha_n &= i_n \\ \alpha_0 + \alpha_1 + \cdots + \alpha_n &= 1\end{aligned}$$

which leads us to the set of coefficients  $\alpha_k$  where  $k = 0, 1, \dots, n$  defined by the following:

$$\begin{aligned}\alpha_0 &= 1 - i_1 \\ \alpha_k &= i_k - i_{k+1} \quad \text{for } k = 1, \dots, n-1 \\ \alpha_n &= i_n\end{aligned}$$

Therefore the set  $\hat{\mathbf{i}} \cup \{\hat{\mathbf{k}}\}_{\hat{\mathbf{k}} \in \mathcal{N}}$  is linearly dependent. Let  $\alpha_i$  be the nonzero terms in sequence  $(\alpha_j)$ , i.e.  $\Theta_i := \{j \in [0, n] : \alpha_j \neq 0\}$ , where  $\alpha_j = \pm 1$  for  $\alpha_j \in \Theta_i$ . Since  $(\alpha_j)$  is unique linear combination with respect to basis  $\mathcal{N}$ , the set  $\hat{\mathbf{i}} \cup \{\hat{\mathbf{k}}\}_{\hat{\mathbf{k}} \in \Theta_i}$  is a minimal linearly dependent set. Thus we obtain the following circuit and corresponding vanishing linear form:

$$\mathcal{C}_i = y_i \cup \{y_j\}_{j \in \Theta_i}, \quad \mathcal{A}_i = -\mathcal{R}_i - \sum_{j \in \Theta_i} \alpha_j \mathcal{R}_j.$$

Furthermore, if  $\mathcal{A}_i > 0$ , then  $\frac{\dot{y}_i}{y_i y_i} < 0$  for  $i \in [n+1, 2^n - 1]$ . For this case  $k = n$ , the nested equilibrium (14) include all immune response  $z_1, \dots, z_n$  as positive components. If  $\mathcal{R}_n > \mathcal{Q}_n$ , then all missing species of non-negative equilibrium,  $\tilde{\mathcal{E}}_n$  are the  $y_i$  with index  $i \in [n+1, 2^n - 1]$ . Thus, inequalities (11) hold strictly and the conclusions of Theorem 1 follow, in particular,  $\tilde{\mathcal{E}}_n$  is stable,  $y_0, \dots, y_n$  are uniformly persistent, and missing strains  $y_i, i \in [n+1, 2^n - 1]$ , go extinct. Notice in this case of  $\mathcal{R}_n > \mathcal{Q}_n$ , that  $\tilde{\mathcal{E}}_n$  becomes unstable if any of the invasion circuits' linear forms become non-positive. In particular, if  $\mathcal{A}_i \leq 0$ , then  $\frac{\dot{y}_i}{y_i y_i} \geq 0$ , yielding the necessary condition of stability of  $\tilde{\mathcal{E}}_n$  based on sign of  $\mathcal{A}_i$  which is stated in Corollary 1. Furthermore, by Proposition 2, in the critical case of  $\mathcal{A}_i = 0$ , there exists a continuum of equilibria with positive components including the circuit elements consisting of invading strain,  $y_i$  and subset of nested strains  $\mathcal{S}_i, \mathcal{C}_i = \{y_i\} \cup \mathcal{S}_i$ , which connects to nested equilibrium  $\tilde{\mathcal{E}}_n$ .

Next, consider the case  $k = n$  but  $\mathcal{R}_n \leq \mathcal{Q}_n$ , then we consider non-negative equilibrium  $\bar{\mathcal{E}}_n$ . Notice that

$$\begin{aligned}\frac{\dot{y}_i}{y_i y_i} &= -\frac{1}{\mathcal{Q}_n} \left[ \mathcal{R}_i - \mathcal{Q}_n - \sum_{j=1}^{n-1} (1 - i_j) (\mathcal{R}_{j-1} - \mathcal{R}_j) - (1 - i_n) (\mathcal{R}_{n-1} - \mathcal{Q}_n) \right] \\ &= -\frac{1}{\mathcal{Q}_n} \left[ \mathcal{R}_i - \mathcal{R}_{n-1} - \sum_{j=1}^{n-1} (1 - i_j) (\mathcal{R}_{j-1} - \mathcal{R}_j) + i_n (\mathcal{R}_{n-1} - \mathcal{Q}_n) \right]\end{aligned}$$

$$\begin{aligned} &\leq -\frac{1}{Q_n} \left[ \mathcal{R}_i - \mathcal{R}_{n-1} - \sum_{j=1}^{n-1} (1 - i_j) (\mathcal{R}_{j-1} - \mathcal{R}_j) + i_n (\mathcal{R}_{n-1} - \mathcal{R}_n) \right] \\ &= -\frac{\mathcal{A}_i}{Q_n}, \end{aligned}$$

where  $\mathcal{A}_i$  is the same vanishing linear form (22) with corresponding circuit  $\mathcal{C}_i$  as prior case. Therefore, if  $\mathcal{A}_i > 0$  (invasion circuit has positive epistasis), then  $\frac{\dot{y}_i}{y_i y_i} \leq -\frac{\mathcal{A}_i}{Q_n} < 0$ . The only other missing species additional to  $y_i, i \in [n+1, 2^n-1]$ , is  $y_n$ . It is not hard to see that  $\frac{\dot{y}_n}{y_n y_n} = \frac{\mathcal{R}_n - Q_n}{Q_n} \leq 0$ . Thus, all saturated inequalities (11) are satisfied strictly (except when  $\mathcal{R}_n = Q_n$ ), and stability of  $\tilde{\mathcal{E}}_n$  (and other conclusions) from Theorem 1 follow. In the case  $\mathcal{R}_n = Q_n$ , arguments from Browne (2017) give same result. Finally, suppose  $k < n$ , so there are  $k$  positive component immune responses,  $z_1, \dots, z_k$ , in feasible nested equilibrium  $\tilde{\mathcal{E}}_k$  or  $\tilde{\mathcal{E}}_k$ . Consider the missing strains  $y_i, i \in [n+1, 2^n-1]$ , which are non-nested strains. Let  $y_i$  be contained on the  $k$  dimensional hypercube. It necessarily has binary sequence where  $i_{k+1} = \dots = i_n = 0$ . Observe from above that the invasion circuit involving  $y_i$  will contain nested strains on the  $k$  dimensional (sub-)hypercube;  $y_0, \dots, y_k$ . It follows that if  $\mathcal{A}_i > 0$ ,  $\frac{\dot{y}_i}{y_i y_i} \leq -\frac{\mathcal{A}_i}{C_k} < 0$ , where  $C_k = \mathcal{R}_k$  when  $\mathcal{R}_k > Q_k$  and  $C_k = Q_k$   $\mathcal{R}_k \leq Q_k$ . For a strain  $y_i$  not on  $k$  dimensional hypercube, then we can find another strain  $y_\ell$  with same sequence in first  $k$  bits, and less mutations overall, so that  $\mathcal{R}_\ell > \mathcal{R}_k$ . Then it is not hard to see that

$$\frac{\dot{y}_i}{y_i y_i} = \mathcal{R}_i x^* - 1 - \sum_{j=1}^k z_k^* < \mathcal{R}_j x^* - 1 - \sum_{j=1}^k z_k^* = \frac{\dot{y}_j}{y_j y_j} \leq 0.$$

Arguments from Browne (2017) show that  $\frac{s_j \dot{z}_j}{\sigma_j z_j |_{\mathcal{E}^*}} < 0$  for  $j > k$ . Applying Theorem 1 gives the desired result.  $\square$

**Proof II of Theorem 2** We prove the case where  $\mathcal{R}_n > Q_n$ , so that  $\tilde{\mathcal{E}}_n$  is feasible equilibrium and, by same arguments in Proof I above, the other cases follow. Consider a given missing viral strain  $y_i$  ( $i \in [n+1, 2^n-1]$ ) with sequence  $\mathbf{i}$ . Define the following linear form based on its invasion rate:

$$\frac{\dot{y}_i}{y_i y_i} = -\frac{\mathcal{A}_i}{\mathcal{R}_n}, \quad \text{where} \quad -\mathcal{A}_i := \mathcal{R}_i + \mathcal{R}_n + \sum_{j=1}^n (1 - i_j) (\mathcal{R}_{j-1} - \mathcal{R}_j),$$

and  $C_n = \mathcal{R}_n$  when  $\mathcal{R}_n > Q_n$  and  $C_n = Q_n$   $\mathcal{R}_n \leq Q_n$ . We claim that  $\mathcal{A}_i = 0$  in additive case, and furthermore  $\mathcal{A}_i \neq 0$  if any (non-zero) viral fitness is removed from  $\mathcal{A}_i$  in the resulting sum. In other words we claim that  $\mathcal{A}_i$  defines a circuit  $\mathcal{C}$  containing strain  $i$  and other strains on nested network. To test additivity, it suffices to consider the linear form on the binary sequences:

$$-f_i := \mathbf{i} - 1^n - \sum_{j=1}^n (1 - i_j) (1^{j-1} 0^{n-j+1} - 1^j 0^{n-j})$$

Since  $\mathbf{i}$  is not in nested network ( $i \in [n+1, 2^n - 1]$ ), there exists  $p \in [1, n-1]$  such that  $i_p = 0, i_{p+1} = 1$ . In other words, there exists a 01 string in the binary sequence  $\mathbf{i}$ . We prove that  $\mathcal{A}_i$  defines a circuit by induction on the number of 01 strings,  $s$ . First suppose that  $s = 1$ . Let  $0 \leq m_1 < p$  be maximal such that  $i_p = 1$  and  $p+1 \leq m_2 \leq n$  be maximal such that  $i_{m_2} = 0$ . With these conditions,  $\mathbf{i} = 1^{m_1}0^{p-m_1}1^{m_2-p}0^{n-m_2}$ . Then

$$\begin{aligned} -f_i &:= \mathbf{i} - 1^n - \sum_{j=1}^n (1 - i_j) \left( 1^{j-1}0^{n-j+1} - 1^j0^{n-j} \right) \\ &= \mathbf{i} - 1^n - (1^{m_1}0^{n-m_1} - 1^p0^{n-p}) - (1^{m_2}0^{n-m_2} - 1^n) \\ \Rightarrow f_i &= \mathbf{i} - 1^{m_1}0^{n-m_1} + 1^p0^{n-p} - 1^{m_2}0^{n-m_2} \\ &= 0^p1^{m-p-1}0^{n-m_2} - 0^p1^{m-p-1}0^{n-m_2} \\ &= 0. \end{aligned} \quad (23)$$

Furthermore  $f_i = \mathbf{i} - 1^{m_1}0^{n-m_1} + 1^p0^{n-p} - 1^{m_2}0^{n-m_2}$  contains the viral sequences corresponding the non-zero fitness quantities in  $\mathcal{A}_i$ . Thus  $\mathcal{A}_i$  defines a circuit since the minimal circuit size is 4. Now for the induction step, consider  $s > 1$ . Assume that  $\mathcal{A}_\ell$  defines a circuit for any sequence  $\ell$  with  $s-1$  or less (01) strings, and suppose the sequence  $\mathbf{i}$  has  $s$  (01) strings. Let  $p_1 < p_2 < \dots < p_s$  be locations of the 01 strings (with  $i_{p_j} = 0, i_{p_j+1} = 1$ ). Let  $0 \leq m_1 < p_1$  be maximal such that  $i_{m_1} = 1$  and  $p_1+1 \leq m_2 \leq p_2$  be maximal such that  $i_{m_2} = 1$ . So  $\mathbf{i} = 1^{m_1}0^{p_1-m_1}1^{m_2-p_1}i_{p_2} \dots i_n$ . Then

$$\begin{aligned} f_i &:= \mathbf{i} - 1^n - (1^{m_1}0^{n-m_1} - 1^{p_1}0^{n-p_1}) - \sum_{j=p_1+2}^n (1 - i_j) \left( 1^{j-1}0^{n-j+1} - 1^j0^{n-j} \right) \\ &= 1^{m_1-1}0^{p_2-m_1}i_{p_2} \dots i_n - 1^n - \sum_{j=p_1+2}^n (1 - i_j) \left( 1^{j-1}0^{n-j+1} - 1^j0^{n-j} \right) \\ &= \tilde{\mathbf{i}} - 1^n - \sum_{j=1}^n (1 - \tilde{i}_j) \left( 1^{j-1}0^{n-j+1} - 1^j0^{n-j} \right) \\ &= \tilde{f}_i, \end{aligned} \quad (24)$$

where  $\tilde{\mathbf{i}} = 1^{m_1}0^{p_2-m_1}1i_{p_2+2} \dots i_n$  has  $s-1$  (01) strings. Thus by induction hypothesis, we obtain  $f_i = \tilde{f}_i = 0$ . Let  $\mathcal{C}_i$  denote the collection of viral sequences corresponding the non-zero fitness quantities in  $\mathcal{A}_i$ . Notice that it is not hard to ascertain from the above calculations that

$$\begin{aligned} \mathcal{C}_i &= \mathbf{i} \cup \{1^{m_j}0^{n-m_j}, 1^{p_j}0^{n-p_j}\}_{j=1}^s \cup 1^{m_{s+1}}0^{n-m_{s+1}}, \\ \mathcal{A}_i &= \mathcal{R}_i - \sum_{j=1}^{s+1} \mathcal{R}_{m_j} + \sum_{j=1}^s \mathcal{R}_{p_j}, \quad (0 \leq m_1 < p_1 < m_2 < \dots < p_s < m_{s+1} \leq n+1). \end{aligned}$$

Consider an arbitrary proper subset  $\mathcal{B}$  of  $\mathcal{C}_i$ . First, we claim that there can not be a circuit consisting solely of sequences in the nested network. Suppose by way of contradiction that there exists a linear form with  $g := \sum_{j=1}^{n+1} b_j \mathbf{j} = 0$ . Let  $k = \max \{1 \leq j \leq n+1 \mid b_j \neq 0\}$ . Then for the  $k^{\text{th}}$  digit in the binary sequence of  $g_N$ , we find  $(g)_k \neq 0$ . So there are no vanishing linear forms on the nested network. Thus it suffices to consider the case where  $\mathbf{i} \in \mathcal{B}$ . Motivated from calculations above, define

$$\tilde{\mathbf{i}} = \mathbf{i} + \sum_{\mathcal{C}_i \setminus \mathcal{B}} (-1^{m_j} 0^{n-m_j} + 1^{p_j} 0^{n-p_j}),$$

where  $\tilde{\mathbf{i}}$  is not in nested network since  $\mathcal{B} \neq \emptyset$ . Furthermore because  $\mathcal{C}_i \setminus \mathcal{B} \neq \emptyset$ , we obtain that  $\tilde{\mathbf{i}}$  has less than  $s$  (01) strings. By induction hypothesis,  $\mathcal{A}_{\tilde{i}}$  defines a circuit  $\mathcal{C}_{\tilde{i}}$  for the sequence  $\tilde{\mathbf{i}}$ , where  $\mathcal{C}_{\tilde{i}} = \{\tilde{\mathbf{i}}\} \cup \mathcal{B} \setminus \{\mathbf{i}\}$ . Denote the vanishing linear form as  $f_{\tilde{i}} = \sum_{\ell \in \mathcal{C}_{\tilde{i}}} a_{\ell} \ell$ . Now for arbitrary coefficients  $b_j$ ,

$$\begin{aligned} \sum_{j \in \mathcal{B}} b_j \mathbf{j} &= b_i \mathbf{i} + \sum_{\mathcal{B} \setminus \{\mathbf{i}\}} b_j \mathbf{j} \\ &= b_i \left( \tilde{\mathbf{i}} - \sum_{\mathcal{C}_i \setminus \mathcal{B}} (-1^{m_j} 0^{n-m_j} + 1^{p_j} 0^{n-p_j}) \right) + \sum_{\mathcal{B} \setminus \{\mathbf{i}\}} b_j \mathbf{j} \\ &= \sum_{\mathcal{B} \setminus \{\mathbf{i}\}} (b_j - b_i a_j) \mathbf{j} - b_i \sum_{j \in \mathcal{C}_{\tilde{i}}} a_j \mathbf{j}, \end{aligned}$$

The above sum consists solely of sequences in the nested network and thus there are no vanishing linear forms. This implies that the above sum is zero only if  $b_i = 0$ , which further leads to conclusion that  $b_j = 0$  for  $j \in \mathcal{B} \setminus \{\mathbf{i}\}$ . Thus the proper subset  $\mathcal{B}$  can not be a circuit for any linear form.  $\square$

**Proof of Theorem 3** By Proposition 6 in Browne and Smith (2018), equivalent conditions for stability of (i)  $\mathcal{E}_n^{\dagger}$  or (ii)  $\mathcal{E}_{n+1}^{\dagger}$  are the following:

- i.  $\mathcal{R}_{n+1} \leq \mathcal{P}_n$  and  $(|\mathcal{A}_i| - 1) \mathcal{P}_n + \mathcal{R}_i \leq \sum_{j \in \mathcal{A}_i} \mathcal{R}_j \quad \forall i \in [n+2, 2^n]$ , in which case  $\Omega_y = \Omega_z = [1, n]$ .
- ii.  $\mathcal{R}_{n+1} > \mathcal{P}_n$  and  $(|\mathcal{A}_i| - 1) \mathcal{R}_{n+1} + \mathcal{R}_i \leq \sum_{j \in \mathcal{A}_i} \mathcal{R}_j \quad \forall i \in [n+2, 2^n]$ , in which case  $\Omega_y = [1, n+1]$  and  $\Omega_z = [1, n]$ .

Furthermore, (i) and (ii) are the only possible stable equilibria  $\mathcal{E}^*$  with a strain-specific subgraph, i.e.  $\Omega_y \subseteq [1, n+1]$ , and strains  $y_1, \dots, y_n$  ( $y_{n+1}$  also for (ii)) are uniformly persistent. We remark that these prior results follow from Theorem 1. Now to prove Theorem 3, we will show that positive epistasis of corresponding invasion circuits are sufficient (also necessary in case (ii)) for inequalities in (i) and (ii) to hold. Fix an invading strain  $y_i$ ,  $i \in [n+2, 2^n]$ , with binary sequence. First, note that a sufficient condition for inequalities in cases (i) and (ii) to be satisfied is  $\frac{\mathcal{A}_i}{K_n} \geq 0$  where  $\mathcal{A}_i = -\mathcal{R}_i - (|\mathcal{A}_i| - 1) \mathcal{R}_{n+1} + \sum_{j \in \mathcal{A}_i} \mathcal{R}_j$ , and  $K_n = \mathcal{P}_n$  if  $\mathcal{R}_{n+1} \leq \mathcal{P}_n$  and  $K_n = \mathcal{R}_{n+1}$  if

$\mathcal{R}_{n+1} > \mathcal{P}_n$ . In particular, inequality in case (ii) is equivalent to  $\frac{\mathcal{A}_i}{K_n} \geq 0$ . To show that  $\mathcal{C}_i = y_i \cup \{y_j\}_{j \in \mathcal{A}_i}$  is a circuit with linear form  $\mathcal{A}_i$ , we proceed with a similar approach to our first proof of Theorem 2. Denote the binary sequences of one-to-one network as  $\mathbf{k}_1, \dots, \mathbf{k}_{n+1}$  corresponding to ordered strains  $y_1, \dots, y_{n+1}$ . Let  $\mathcal{S} \subset \{0, 1\}^{n+1}$  denote the subset of strain-specific extended binary sequences, where  $\hat{\mathbf{i}} = \mathbf{i}\mathbf{1} \in \{0, 1\}^{n+1} \setminus \mathcal{S}$  and  $\hat{\mathbf{k}}_j = \mathbf{k}_j\mathbf{1} \in \mathcal{S}$  represent binary sequences extended by digit 1. Notice that  $\mathcal{S}$  forms a basis of  $\mathbb{R}^{n+1}$ . Indeed, it is not hard to show the row reduced echelon form of  $n+1 \times n+1$  matrix is triangular. Thus for  $\hat{\mathbf{i}} \in \{0, 1\}^{n+1} \setminus \mathcal{S}$ , there is a unique set of coefficients  $\alpha_j, j = 1, 2, \dots, n+1$ , yielding  $\hat{\mathbf{i}}$  as a linear combination of the nested network vectors:

$$\hat{\mathbf{i}} = \alpha_1 \hat{\mathbf{k}}_1 + \dots + \alpha_{n+1} \hat{\mathbf{k}}_{n+1}.$$

The above linear system resolves as follows:

$$\begin{aligned} \alpha_2 + \dots + \alpha_{n+1} &= i_1 \\ &\vdots \\ \sum_{j \neq k} \alpha_j &= i_k \\ &\vdots \\ \alpha_1 + \dots + \alpha_{n+1} &= 1 \end{aligned}$$

which leads us to the set of coefficients  $\alpha_k$  where  $k = 1, \dots, n+1$  defined by the following:

$$\begin{aligned} \alpha_k &= 1 - i_k \quad \text{for } k = 1, \dots, n \\ \alpha_{n+1} &= 1 - (n - \sum_{k=1}^n i_k) = -(|\mathcal{A}_i| - 1) \end{aligned}$$

Thus, with analogous argument as before, we obtain the indicated circuit  $\mathcal{C}_i$  and corresponding linear form  $\mathcal{A}_i$ .  $\square$

**Proof of Proposition 3** Let  $\mathbf{i} \in \{0, 1\}^n \setminus \mathcal{S}$  with integer coordinates  $(\alpha_{\mathbf{k}})_{\mathbf{k} \in \mathcal{S}}$  with respect to  $\mathcal{S} \times \{1\}$  as a basis of  $\mathbb{R}^{n+1}$  and addended binary sequence  $\mathbf{i}\mathbf{1} \in \{0, 1\}^n \times \{1\}$ . Clearly  $\mathcal{C} \times \{1\} = \mathcal{S} \times \{1\} \cup \{\mathbf{i}\mathbf{1}\}$  is a linearly dependent set  $\mathbb{R}^{n+1}$  with linear form on fitnesses given by  $\mathcal{A}_i = -\mathcal{R}_i - \sum_{\mathbf{k} \in \mathcal{S}} \alpha_{\mathbf{k}} \mathcal{R}_{\mathbf{k}}$ . Furthermore any proper subset is linearly independent since  $\mathcal{S} \times \{1\}$  is a basis of  $\mathbb{R}^{n+1}$ . Thus  $\mathcal{C}$  is a circuit with linear form  $\mathcal{A}_i$ . By proof of Prop 2,

$$\frac{\dot{y}_i}{\gamma_i y_i} = \frac{\dot{y}_i}{\gamma_i y_i} + \sum_{\ell=1}^{n+1} \alpha_{\ell} \frac{\dot{y}_{\ell}}{\gamma_{\ell} y_{\ell}}$$

$$\begin{aligned}
&= \sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} \frac{\dot{y}_{\mathbf{k}}}{\gamma_{\mathbf{k}} y_{\mathbf{k}}} \\
&= x^* \sum_{\mathbf{k} \in \mathcal{C}} \alpha_{\mathbf{k}} \mathcal{R}_{\mathbf{k}} \\
&= x^* \left[ \mathcal{R}_i + \sum_{\ell=1}^{n+1} \alpha_{\ell} \mathcal{R}_{\ell} \right] \\
&= -x^* \mathcal{A}_i.
\end{aligned}$$

Thus, by Theorem 1, the stability of  $\mathcal{E}^*$  is determined by the sign of  $\mathcal{A}_i$ .  $\square$

**Proof of Proposition 4** We apply Proposition 3. Consider the network with at most one mutation,  $\tilde{\mathcal{S}}_1$ , consisting of wild-type and 1-mutation viral strains  $y_0, y_1, \dots, y_n$  where the sequence of  $y_j$  is  $\mathbf{j} = (\delta_{\ell j})_{\ell=1}^n$  for  $j = 1, \dots, n$ . First, notice that binary sequences of  $y_0, y_1, \dots, y_n$  form basis of  $\mathbb{R}^{n+1}$ . For missing strain,  $y_i, i = n+1, \dots, 2^n - 1$ ,  $\mathbf{i} \in \{0, 1\}^n \setminus \tilde{\mathcal{S}}_1$ , the coordinates of addended binary sequence  $\mathbf{i}1$  with respect to this basis can be seen to be  $\alpha_j = i_j, j = 1, \dots, n, \alpha_0 = 1 - \sum_{j=1}^n i_j$ . Thus, by Proposition 3,  $\frac{\dot{y}_i}{y_i y_i} = -x^* \mathcal{A}_i$ , where  $\mathcal{A}_i = -\mathcal{R}_i - \sum_j = 1^n i_j \mathcal{R}_j - \left(1 - \sum_{j=1}^n i_j\right) \mathcal{R}_0 = -\mathcal{R}_i - (n - |\mathcal{A}_i| - 1) \mathcal{R}_0 + \sum_{j \notin \mathcal{A}_i} \mathcal{R}_j$ .  $\square$

**Proof of Theorem 4** First assume that pairwise interaction matrix  $B$  is positive and consider the stability of the nested equilibrium,  $\tilde{\mathcal{E}}_n$  (or  $\mathcal{E}_n$ ), as characterized by circuits in Corollary 1 (ii). We proceed by induction on the number of (01) strings denoted by  $s$  for the invading strain. Suppose  $s = 1$  and the invading strain is written as in prior proof as  $\mathbf{i} = 1^{m_1} 0^{p-m_1} 1^{m_2-p} 0^{n-m_2}$  and the collection of strains in the circuit is given by  $\mathcal{C}_i = \mathbf{i} \cup \{1^{m_1} 0^{n-m_1}, 1^{p_1} 0^{n-p_1}\} \cup 1^{m_2} 0^{n-m_2}$ . Then since the additive elements will sum to zero in the linear form  $\mathcal{A}_i$ , the only remain terms come from pairwise interactions in  $B$  and can be calculated as:

$$\begin{aligned}
-\mathcal{A}_i &= \sum_{j=1}^{m_1} \sum_{k>j}^{m_1} B_{jk} + \sum_{j=1}^{m_1} \sum_{k=p_1+1}^{m_2} B_{jk} + \sum_{j=p_1+1}^{m_2} \sum_{k>j}^{m_2} B_{jk} \\
&\quad - \sum_{j=1}^{m_1} \sum_{k>j}^{m_1} B_{jk} + \sum_{j=1}^{p_1} \sum_{k>j}^{p_1} B_{jk} - \sum_{j=1}^{m_2} \sum_{k>j}^{m_2} B_{jk} \\
&= \sum_{j=1}^{p_1} \sum_{k>j}^{p_1} B_{jk} - \sum_{j=1}^{m_1} \sum_{k>j}^{p_1} B_{jk} - \sum_{j=m_1+1}^{p_1} \sum_{k>j}^{m_2} B_{jk} \\
&= - \sum_{j=m_1+1}^{p_1} \sum_{k=p_1+1}^{m_2} B_{jk} < 0
\end{aligned}$$

Now for the induction step, suppose that  $\mathbf{i}$  has  $s$  (01) strings. It is not hard to see that  $\mathcal{A}_i = \mathcal{A}_{\tilde{\mathbf{i}}}$ , for invading strain  $\tilde{\mathbf{i}}$ , where  $\tilde{\mathbf{i}} = 1^{m_1} 0^{p_2-m_1} 1_{p_2+2} \dots i_n$  has  $s-1$  (01) strings. Thus by induction hypothesis  $-\mathcal{A}_i < 0$ , or  $\mathcal{A}_i > 0$  giving positive epistasis and stability of nested network.

Next suppose that matrix  $B$  is negative and consider the stability of  $\mathcal{E}_n^\dagger$  and  $\mathcal{E}_{n+1}^\ddagger$  consisting of strains  $y_1, \dots, y_n, y_{n+1}$  with binary sequences  $\mathbf{k}_1, \dots, \mathbf{k}_{n+1}$ , where  $\Lambda_j = \{j\}$  for  $j = 1, \dots, n$  and  $\Lambda_{n+1} = \emptyset$ . We inspect the invasion circuit of a strain with sequence  $\mathbf{i}$  outside the one-to-one network. Let  $s$  be the number of 1s in sequence  $\mathbf{i}$ , located at loci  $\ell_1, \dots, \ell_s$ , where  $0 \leq s \leq n-2$ . Again the additive terms in  $\mathcal{A}_i$  are zero and thus we have:

$$\begin{aligned} \mathcal{A}_i &= -\mathcal{R}_i - (|\Lambda_i| - 1) \mathcal{R}_{n+1} + \sum_{j \in \Lambda_i} \mathcal{R}_j \\ &= -\sum_{j=1}^s \sum_{k > \ell_j} B_{\ell_j k} - (n-1-s) \sum_{j=1}^n \sum_{k > j} B_{jk} + (n-s) \sum_{j=1}^n \sum_{k > j} B_{jk} \\ &\quad - \sum_{m=1}^n \left[ \sum_{k > m} B_{mk} + \sum_{k=1}^{m-1} B_{km} \right] + \sum_{j=1}^s \left[ \sum_{k > \ell_j} B_{\ell_j k} + \sum_{k=1}^{\ell_j-1} B_{k \ell_j} \right] \\ &= -\sum_{j=1}^n \sum_{k=1}^{j-1} B_{kj} + \sum_{j=1}^s \sum_{k=1}^{\ell_j-1} B_{k \ell_j} \\ &> 0 \quad \text{since } B < 0, s \leq n-2. \end{aligned}$$

Finally, for the network with at most one mutation, only the invading strain  $\mathbf{i}$  will have  $\geq 2$  mutations, so

$$\begin{aligned} \mathcal{A}_i &= -\mathcal{R}_i - (n - |\Lambda_i| - 1) \mathcal{R}_0 + \sum_{j \notin \Lambda_i} \mathcal{R}_j \\ &= -\sum_{j=1}^n \sum_{k > j} B_{jk} > 0 \quad \text{since } B < 0. \end{aligned}$$

□

**Proof of Theorem 5** Let  $0 < f_j < 1$ ,  $j = 1, \dots, n$  represent the multiplicative fitness costs for each epitope. Similar to proof of Proposition 4, we use stability formulation by circuits in Corollary 1 (ii) and prove by induction on the number of (01) strings denoted by  $s$  for the invading strain. Suppose  $s = 1$  and the invading strain is written as in prior proof as  $\mathbf{i} = 1^{m_1} 0^{p-m_1} 1^{m_2-p} 0^{n-m_2}$  and the collection of strains in the circuit is given by  $\mathcal{C}_i = \mathbf{i} \cup \{1^{m_1} 0^{n-m_1}, 1^{p_1} 0^{n-p_1}\} \cup 1^{m_2} 0^{n-m_2}$ . Then the linear form  $\mathcal{A}_i$  can be calculated as:

$$\begin{aligned} -\mathcal{A}_i &= \mathcal{R}_0 (f_1 \cdots f_{m_1} f_{p_1+1} \cdots f_{m_2} - f_1 \cdots f_{m_1} + f_1 \cdots f_{p_1} - f_1 \cdots f_{m_2}) \\ &= \mathcal{R}_0 f_1 \cdots f_{m_1} (1 - f_{p_1+1} \cdots f_{m_2}) (f_{m_1+1} \cdots f_{p_1} - 1) \\ &< 0 \end{aligned}$$

Now for the induction step, suppose that  $\mathbf{i}$  has  $s$  (01) strings. It is not hard to see that  $\mathcal{A}_i = \mathcal{A}_{\tilde{i}}$ , for invading strain  $\tilde{i}$ , where  $\tilde{i} = 1^{m_1} 0^{p_2 - m_1} 1 i_{p_2+2} \dots i_n$  has  $s - 1$  (01) strings. Thus by induction hypothesis  $-\mathcal{A}_i < 0$ , proving nested network is stable.  $\square$

### The “at most one mutation” network equilibria

Consider the network with at most one mutation,  $\tilde{S}_1$ , consisting of wild-type and 1-mutation viral strains  $y_0, y_1, \dots, y_n$  where the sequence of  $y_j$  is  $\mathbf{j} = (\delta_{\ell j})_{\ell=1}^n$  for  $j = 1, \dots, n$ . First it is simpler to look at the  $n$  strain equilibrium  $\mathcal{E}^{1*}$  containing positive components for  $y_1^*, \dots, y_n^*$ , where  $y_0^* = 0$ , i.e. leaving out the wild-type strain. By Proposition 1 and (10), such a positive equilibrium  $\mathcal{E}^{1*} = (x^*, y^{1*}, z^*)$  of system (4) satisfies

$$x^* = \frac{1}{\sum_{j=1}^n \mathcal{R}_j - (n-1)\mathcal{R}_0}, \quad A y^{1*} = \mathbf{s}, \quad A z^* = \mathcal{R}^1 x^* - \mathbf{1}, \quad \text{where } A = \mathbf{1}(\mathbf{1})^T - I_n, \\ A^{-1} = \frac{1}{n-1} \mathbf{1}(\mathbf{1})^T - I_n, \quad y^1 = (y_1, y_2, \dots, y_n)^T, \quad \mathcal{R}^1 = (\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_n)^T$$

with  $I_n$  is the  $n \times n$  identity matrix. Here we find that:

$$y_i^* = \frac{1}{n-1} \left( -(n-2)s_i + \sum_{j \neq i} s_j \right), \quad x^* = \frac{n-1}{n-1 + \sum_i \mathcal{R}_i s_i}, \quad z_i^* = \frac{1}{n-1} (\mathcal{R}_i x^* - 1)$$

With the immunodominance hierarchy  $s_i \leq s_{i+1}$ , then  $y_i^* > 0$  if  $s_1 > \sum_{i>1} (s_n - s_i)$  and  $z_i^* > 0$  if  $\mathcal{R}_i (n-1 - \sum_i s_i) > n-1$ . If these conditions are satisfied, then the equilibrium  $\mathcal{E}^{1*}$  is saturated in the subsystem restricted to  $S_1$ . In Browne and Smith (2018) we showed that in the larger network of viral strains, the equilibrium  $\mathcal{E}^{1*}$  is always unstable in the case with equal reproduction numbers  $\mathcal{R}_1 = \mathcal{R}_2 = \dots = \mathcal{R}_n$ .

Now consider invasion by the wild strain  $y_0$ , which can result in an  $n+1$  strain equilibrium  $\tilde{\mathcal{E}}^{1*}$  consisting of the viral strain network  $\tilde{S}_1$ . By Proposition 1, the positive components  $x^*, \tilde{y}^{1*}, z^*$  of  $\tilde{\mathcal{E}}^{1*}$  satisfies:

$$x^* = \mathbf{1}^T C_{(n+1)}^{-1}, \quad \text{here } C = \begin{pmatrix} A & \mathcal{R} \\ \mathbf{1}^T & \mathcal{R}_0 \end{pmatrix}, \\ \Rightarrow x^* = \frac{1}{\sum_{j=1}^n \mathcal{R}_j - (n-1)\mathcal{R}_0}, \\ \tilde{y}^{1*} = (\mathbf{s}^T \frac{1}{x^*} - 1) C^{-1} \\ A z^* = \mathcal{R} x^* - \mathbf{1}, \quad \sum_{i=1}^n z_i^* = \mathcal{R}_0 x^* - 1.$$

The above equations are difficult to analyze in general, but when  $x^* > 0, \tilde{y}^{1*} > 0, z^* > 0$ , the  $n+1$  strain “at most one mutation” equilibrium will be positive. Fur-

thermore, if the linear forms of invasion circuits (21) are positive, then by Proposition 4,  $\tilde{\mathcal{E}}^{1*}$  will be stable.

## References

- Ahmed SF, Quadeer AA, Morales-Jimenez D, McKay MR (2019) Sub-dominant principal components inform new vaccine targets for hiv gag. *Bioinformatics* 35(20):3884–3889
- Althaus CL, Boer RD (2008) Dynamics of immune escape during HIV/SIV infection. *PLoS Comput Biol* 4:e1000103
- Barton JP, Goonetilleke N, Butler TC, Walker BD, McMichael AJ, Chakraborty AK (2016) Relative rate and location of intra-host HIV evolution to evade cellular immunity are predictable. *Nat Commun* 7:1–10
- Bascompte J, Jordano P, Melián CJ, Olesen JM (2003) The nested assembly of plant-animal mutualistic networks. *Proc Natl Acad Sci* 100(16):9383–9387
- Beerenwinkel N, Pachter L, Sturmfels B (2007) Epistasis and shapes of fitness landscapes. *Stat Sinica* pp 1317–1342
- Bobko N, Zubelli JP (2015) A singularly perturbed HIV model with treatment and antigenic variation. *Math Biosci Eng MBE* 12(1):1–21
- Bratus AS, Novozhilov AS, Semenov YS (2019) Rigorous mathematical analysis of the quasispecies model: from Manfred Eigen to the recent developments. *Adv Math Methods Biosci Appl* pp 27–51
- Browne, C.: Global properties of nested network model with application to multi-epitope HIV/CTL dynamics. *J Math Biol* pp 1–22 (2017)
- Browne CJ, Smith HL (2018) Dynamics of virus and immune response in multi-epitope network. *J Math Biol* 77(6–7):1833–1870
- Chakraborty AK, Barton JP (2017) Rational design of vaccine targets and strategies for HIV: a crossroad of statistical physics, biology, and medicine. *Rep Progress Phys* 80(3):032601
- Champagnat N, Méléard S (2011) Polymorphic evolution sequence and evolutionary branching. *Probab Theory Relat Fields* 151(1–2):45–94
- Costa M, Hauzy C, Loeuille N, Méléard S (2016) Stochastic eco-evolutionary model of a prey-predator community. *J Math Biol* 72(3):573–622
- Crona K (2020) Rank orders and signed interactions in evolutionary biology. *Elife* 9:e51004
- Crona K, Gavryushkin A, Greene D, Beerenwinkel N (2017) Inferring genetic interactions from comparative fitness data. *Elife* 6:e28629
- Deutekom HV, Wijnker G, Boer RD (2013) The rate of immune escape vanishes when multiple immune responses control an HIV infection. *J Immunol* 191:3277–3286
- Eble H, Joswig M, Lamberti L, Ludington W (2020) Higher-order interactions in fitness landscapes are sparse. *arXiv preprint arXiv:2009.12277*
- Ferretti L, Schmiegelt B, Weinreich D, Yamauchi A, Kobayashi Y, Tajima F, Achaz G (2016) Measuring epistasis in fitness landscapes: the correlation of fitness effects of mutations. *J Theor Biol* 396:132–143
- Ganusov VV, De Boer RJ (2006) Estimating costs and benefits of CTL escape mutations in SIV/HIV infection. *PLoS Comput Biol* 2(3):e24
- Ganusov VV, Goonetilleke N, Liu MK, Ferrari G, Shaw GM, Borrow AJMP, Korber BT, Perelson AS (2011) Fitness costs and diversity of the cytotoxic t lymphocyte (CTL) response determine the rate of CTL escape during acute and chronic phases of HIV infection. *J Virol* 85(20):10518–10528
- Goh B (1978) Sector stability of a complex ecosystem model. *Math Biosci* 40(1–2):157–166
- Gould AL, Zhang V, Lamberti L, Jones EW, Obadia B, Korasidis N, Gavryushkin A, Carlson JM, Beerenwinkel N, Ludington WB (2018) Microbiome interactions shape host fitness. *Proc Natl Acad Sci* 115(51):951–960
- Gurney J, Aldakak L, Betts A, Gougat-Barbera C, Poisot T, Kaltz O, Hochberg ME (2017) Network structure and local adaptation in co-evolving bacteria-phage interactions. *Mol Ecol* 26(7):1764–1777
- Hallgrímsdóttir IB, Yuster DS (2008) A complete classification of epistatic two-locus models. *BMC Genet* 9(1):1–15
- Hofbauer J, Sigmund K (1998) *Evolutionary games and population dynamics*. Cambridge University Press, Cambridge

- Jover LF, Cortez MH, Weitz JS (2013) Mechanisms of multi-strain coexistence in host-phage systems with nested infection networks. *J Theor Biol* 332:65–77
- Kessinger TA, Perelson S, Neher RA (2015) Inferring HIV escape rates from multi-locus genotype data. *Immune system modeling and analysis* p 348
- Korytowski DA, Smith HL (2015) How nested and monogamous infection networks in host-phage communities come to be. *Thyroid Res* 8(1):111–120
- Kraut A, Bovier A (2019) From adaptive dynamics to adaptive walks. *J Math Biol* 79(5):1699–1747
- Leviyang U, Ganusov VV (2015) road CTL response in early HIV infection drives multiple concurrent CTL escapes. *PLoS Comput Biol* 11(10):e1004492
- Liu MK, Hawkins N, Ritchie AJ, Ganusov VV, Whale V, Brackenridge S, Li H, Pavlicek JW, Cai F, Rose-Abrahams M et al (2013) Vertical t cell immunodominance and epitope entropy determine HIV-1 escape. *J Clin Investig* 123(1):380–393
- Mani R, Onge RPS, Hartman JL, Giaever G, Roth FP (2008) Defining genetic interaction. *Proc Natl Acad Sci* 105(9):3461–3466
- McMichael AJ (2006) HIV vaccines. *Annu Rev Immunol* 24:227–255
- Nowak MA, Bangham CR (1996) Population dynamics of immune responses to persistent viruses. *Science* 272(5258):74–79
- Pantaleo G, Lévy Y (2013) Vaccine and immunotherapeutic interventions. *Curr Opin HIV AIDS* 8(3):236–242
- Rife Magalis, B., Autissier, P., Williams, K.C., Chen, X., Browne, C., Salemi, M.: Predator-prey dynamics of intra-host simian immunodeficiency virus evolution within the untreated host. *Front Immunol* p 4191 (2021)
- Seifert D, Di Giallonardo F, Metzner KJ, Günthard HF, Beerenwinkel N (2015) A framework for inferring fitness landscapes of patient-derived viruses using quasispecies theory. *Genetics* 199(1):191–203
- Stadler PF (2002) Fitness landscapes. *Biological Evolution and Statistical Physics*. Springer, Berlin, pp 183–204
- Walker BD, Xu GY (2013) Unravelling the mechanisms of durable control of HIV-1. *Nat Rev Immunol* 13(7):487–498
- Weinreich DM, Lan Y, Wylie CS, Heckendorn RB (2013) Should evolutionary geneticists worry about higher-order epistasis? *Curr Opin Genet Dev* 23(6):700–707
- Weitz JS, Poisot T, Meyer JR, Flores CO, Valverde S, Sullivan MB, Hochberg ME (2013) Phage-bacteria infection networks. *Trends Microbiol* 21(2):82–91
- Wolkowicz GS (1989) Successful invasion of a food web in a chemostat. *Math Biosci* 93(2):249–268

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.