

A fixed point theorem for a general epidemic model

Adam R. Lucas*

*Department of Mathematics and Computer Science, Saint Mary's College of California,
1928 Saint Mary's Rd., Moraga, CA 94575, U.S.A.*

(Dated: August 20, 2010)

We provide a rigorous axiomatic framework to study the critical behavior of disease spreading on top of a complex network. A necessary and sufficient condition for our general epidemic model to undergo a phase transition is proved. It is known that an *epidemic state* undergoes a phase transition when the infection rate surpasses the epidemic threshold. However, for networks having degree-degree correlations, the epidemic threshold has never formally been defined. We define the epidemic threshold as, $\lambda_c := 1/\lambda'$ with λ' denoting the largest positive eigenvalue of an operator T given in the axioms of our model. When the *epidemic state* is a strictly positive solution to a fixed point equation our model is guaranteed to have a single phase transition. Percolation as well as SIS/SIR epidemic models on complex correlated networks satisfy the axioms of our model. A benefit of our axiomatic framework is that it highlights commonalities in a variety of interacting particle systems.

PACS numbers: 82.39.Rt, 89.75.-k, 87.15.Zg, 02.30.Sa

I. INTRODUCTION

There is considerable interest today in understanding the mechanism of disease spreading. This is not only because of the prevalence of news headlines about the Swine Flu, or the impending extinction of the Tasmanian Devil, or the mystery computer worm Conficker. Epidemic models offers scientists in such far ranging fields as biology, physics, social science and mathematics, an important example of the phenomenon of critical behavior on the top of complex networks. A complex network, is a network or graph with non-trivial topological features-features that do not occur in simple networks such as lattices or random graphs. Critical behavior is characterized by a dramatic and sudden change, or *phase transition*, in a macroscopic state. On a complex network, a macroscopic behavior can't be simply reduced to the microscopic behavior of its constituents. One reason for this is because in a complex network the distribution of the number of neighbors of a vertex (degree) is typically broad, with a tail that often follows a power law [1]. Real world examples of complex networks are social networks of acquaintances, protein-protein interactions, food webs, and the World Wide Web.

From the *Susceptible-Infected-Susceptible* (abbrev. SIS) model in epidemiology a rich mathematical theory is developing which provides insight into how a disease can spread across complex networks [2, 3]. In the SIS model on a connected undirected graph, the vertices represent individuals who are in one of two states: infected (those carrying the disease) or susceptible (those who do not have the disease yet but can catch it). The edges of the graph correspond to the contacts between individuals. Only susceptible individuals in contact with one or more infected individual may become infected. Infected

individuals can spontaneously become susceptible again. When the infection rate exceed the network's *epidemic threshold* there is a phase transition and a strictly positive fraction of the population is infected in the long term. Similar critical behavior phenomena are observed in many other physical systems including percolation [4], Ising-Potts models [5, 6], synchronization [7], reaction-diffusion processes [8], sandpiles [9] and avalanches [10].

Real world networks are correlated (i.e. they cannot be completely defined by a degree distribution, $P(k)$, alone), and so it is important to study critical phenomena on complex networks with degree-degree correlations [11, 12]. For an uncorrelated network, the epidemic threshold for the SIS model is the quotient of the expected degree of the network and the expected degree squared [3]. How does one define the epidemic threshold when the degree of neighboring vertices in the network are correlated? In [13], they relate the presence or absence of an epidemic threshold to the eigenvalue spectra of certain connectivity matrices of the network. One of the goals of this paper is to relate this observation to a general fixed point theorem. The importance of doing this is that one can begin to see commonalities between different physical systems. For example, in the Bond Percolation model a strictly positive state satisfies a fixed point equation if and only if the probability that a bond is retained in a randomly damaged network exceeds the percolation threshold. Similarly in the SIS model, a strictly positive epidemic states satisfies a fixed point equation if and only if the infection rate exceeds the epidemic threshold.

In this paper we define a general epidemic model guaranteed to exhibit a unique phase transition. The model is expressed in the language of Banach space theory. The model consists of four primitive notions related to one another by six axioms, given in Section II B 1. The function, F , involved in the fixed point equation mentioned above, is an operator on the states of the system. The

* alucas@alum.mit.edu; <http://www.dillgroup.ucsf.edu/~alucas/>

states of the system are elements of a real Banach space of functions, $X(\Omega)$, on a sample space Ω . There is also an integral operator T , on $X(\Omega)$, and an ‘infection rate’, $\lambda > 0$. Our main result is a fixed point theorem which provides a necessary and sufficient condition for a general epidemic model, $(X(\Omega), \lambda, T, F)$, to have a strictly positive epidemic state.

Theorem I.1 (*Theorem III.1*) *Let $(X(\Omega), \lambda, T, F)$ be a general epidemic model satisfying the axioms of Section II B 1. There exists a unique strictly positive solution $x : \Omega \rightarrow (0, 1]$ to the fixed point equation $x = F(x)$ if and only if $\lambda > \lambda_c$ where $\lambda_c := 1/\lambda'$ with λ' denoting the largest positive eigenvalue of the operator T . The number λ_c is thus referred to as the **epidemic threshold**.*

One may wonder why the author chose Banach space theory to describe the model. Indeed, one could argue that the set of degrees Ω is finite for a real world network and so a Banach space, $X(\Omega)$, can be identified with Euclidean space, as is discussed in Section VII A. The answer is that Banach space theory provides a convenient language to discuss many of the ideas in this paper. The theory and application of fixed point theorems on Banach spaces have a rich history and is the subject of current investigation [14–17]. Using the language of Banach spaces, only a little extra effort is required to prove Theorem III.1 in its full generality.

The organization of the paper is as follows. In Section II we introduce the primitive notions of a general epidemic model and give the axioms of our model. In Section III we state and prove the fixed point theorem, Theorem III.1, stated above. The next two sections contain examples of our general epidemic model. In Section IV we introduce a Hybrid SIS/SIR model. In section V we introduce the Bond and Site Percolation model. Section VII provides mathematical prerequisites. The Hybrid SIS/SIR model and Percolation model provide an example of three different fixed point formulas. The techniques used in this paper to verify the axioms should extend to other interacting particle systems.

II. AXIOMS FOR A GENERAL EPIDEMIC MODEL

A. Preliminaries

Our general epidemic model shall be described in terms of four primitive notions:

$$X(\Omega), \lambda, T \text{ and } F.$$

$X(\Omega)$ is a Banach space of real valued functions on a sample space Ω , containing the function $\mathbf{1}$, defined by $\mathbf{1}(k) = 1$, for all $k \in \Omega$. A Banach space is a vector space with a topology defined by some norm $\|\bullet\|$ that is complete (see section VII for background material). The choice of Banach space will depend on the specific

application at hand. For ease of notation we henceforth write X instead of $X(\Omega)$. The infection rate, λ , is a nonnegative number. $T : X \rightarrow X$, is a linear operator on X . The function $F : X \rightarrow X$ satisfies a fixed point equation $x = F(x)$.

To make these notions more intuitive here is an example of their role in the Susceptible-Infected-Susceptible (SIS) model on a complex network with nearest neighbor degree correlations. Given a finite undirected connected network (graph) the *degree* of a node is the number of edges connecting to that node. Let Ω denote the set of distinct degrees present in the network (for example $\Omega = \{1, 2, 5\}$ if each vertex of the network has degree either 1, 2 or 5). Let $P(k)$ be the distribution of degrees, and (Ω, P) a probability space with the probability measure given by

$$P(\omega) = \sum_{k \in \omega} P(k) \quad \text{for all } \omega \subseteq \Omega; \quad (1)$$

here we follow the standard notation of identifying the probability measure P (on the left in (1)) with its distribution $P(k)$ (on the right in (1)). We omit from our notation for the probability space (P, Ω) the collection of events which can be assigned probabilities, known as the σ -**algebra** of Ω . The σ -algebra over a set Ω is a nonempty collection of subsets of Ω (including Ω itself) that is closed under complements and countable unions.

Our Banach space, X , is the space of square integrable functions $L^2(\Omega)$, which because Ω is finite of size n , can be identified with Euclidean space \mathbb{R}^n (see Section VII A). The operator $T : X \rightarrow X$ is defined by

$$T(x)(k) = \sum_{k' \in \Omega} k P(k'|k) x(k'),$$

where $P(k'|k)$ is the probability a degree k node leads to a degree k' node. Identifying X with \mathbb{R}^n , the operator T is a nonnegative diagonalizable matrix with real eigenvalues related to the correlation matrix for the network. We allow for our network to have degree-degree correlations but we will ignore non nearest neighbor degree correlations. Hence the network is Markovian in the sense that it can be fully described by $P(k)$ and $P(k'|k)$.

In the SIS model, at a given time, an individual has one of two possible states, healthy (susceptible) or infected. Let $\rho(k, t)$ be the density of infected degree k individuals in the network at time t .

The *epidemic state* of the model

$$\rho : \Omega \times [0, \infty) \rightarrow [0, 1]$$

satisfies the system of nonlinear differential equations

$$\boxed{\begin{cases} \partial_t \rho(k, t) = -\rho(k, t) + \lambda(1 - \rho(k, t))(T\rho)(k, t), \\ \rho(k, 0) = \rho_0(k), \end{cases}} \quad (2)$$

where $\rho_0 : \Omega \rightarrow [0, 1]$ is the initial state of the system, and where $\lambda > 0$ is called the (*effective*) *rate of infection* [13]. The first term $-\rho(k, t)$ on the right of the

differential equation (2) says that infected individuals become healthy again with a decay rate proportional to the density of infected individuals. The factor $1 - \rho(k, t)$ represents the probability that a degree k node is susceptible and $(T\rho)(k, t)$ is a measure of the possibility that a degree k susceptible node can become infected.

Remark Here, $T\rho$ is the operator T acting on the function $\rho_t : \Omega \rightarrow \mathbb{R}$, defined by $\rho_t(k) := \rho(k, t)$, for all $k \in \Omega$.

The infection rate, λ is a proportionality constant.

Setting $\partial_t \rho(k, t) = 0$ in the differential equation (2) gives rise to a fixed point equation $x = F(x)$, where $x(k)$ is the steady state density of degree k nodes. This fixed point equation always has a zero solution. As we shall prove, it has a strictly positive solution $x : \Omega \rightarrow (0, 1]$ exactly when the infection rate λ is greater than the reciprocal of the largest eigenvalue of the operator T .

Below we give prerequisite mathematical background to understand the axioms in Section IIB1 and establish notation.

Definition A non-empty closed convex subset K of a real Banach space X is a *cone* if it satisfies

$$x \in K, \gamma \geq 0 \Rightarrow \gamma x \in K$$

$$x, -x \in K \Rightarrow x = \theta,$$

where θ is the zero element of X (i.e. $\theta(k) = 0$ for all $k \in \Omega$) and γ is a scalar.

A cone K in X induces a *partial ordering* \leq by the rule: $u \leq v$ if and only if $v - u \in K$. In this paper, K is the cone of nonnegative functions, $\Omega \rightarrow [0, \infty)$. It follows that $u \leq v$ is equivalent to $u(k) \leq v(k)$, for all $k \in \Omega$. In this paper, we reserve the letters k and k' for members of the set Ω . If the context is clear, we will henceforth omit saying, “for all $k \in \Omega$ ”.

If $\overline{K - K} = X$, (i.e., the set $\{u - v : u, v \in K\}$ is dense in X), then K is called a *total cone*. A cone, K is called *normal* if there exists a constant $N > 0$ such that, for all $x, y \in X$, $\theta \leq x \leq y$ implies $\|x\| \leq N\|y\|$. This condition ensures that the norm is compatible with the partial order. In this case N is called the *normality* constant of K .

We denote by M_n the set of real $n \times n$ square matrices. A matrix, $A \in M_n$ is *irreducible* if there is no permutation matrix P such that

$$P^T A P = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

for some square matrices A_{ij} . An n vertex digraph, $\Gamma(A)$, is defined to have a directed arc from vertices v_i to v_j if and only if the matrix element $a_{i,j}$ in A is nonzero. A digraph $\Gamma(A)$ is strongly connected if between every pair of distinct nodes v_i and v_j there is a directed path of finite length that begins at v_i and ends at v_j . It is known

that A , is irreducible if and only if its associated directed graph $\Gamma(A)$ is strongly connected [18].

The *spectral radius* of a matrix or a bounded linear operator is the supremum among the absolute values of the elements in its spectrum. The spectrum of an operator on a finite dimensional space is its set of eigenvalues. On infinite dimensional spaces it can contain more than just eigenvalues [19]. We will write the spectrum of a matrix A , as $\rho(A)$ and more generally of an operator T as $\rho(T)$.

We are now ready to state the classical theorem of O. Perron and F.G. Frobenius for nonnegative matrices.

Theorem II.1 (Perron Frobenius Theorem [18]) *Let $A \in M_n$ and suppose that A is irreducible and nonnegative. Then*

$$(a) \rho(A) > 0$$

$$(b) \rho(A) \text{ is an eigenvalue of } A$$

$$(c) \text{ There is a vector } \nu, \text{ with strictly positive entries, such that } A\nu = \rho(A)\nu$$

$$(d) \rho(A) \text{ is simple (i.e. its corresponding eigenspace is one-dimensional)}$$

Below we will use the Krein-Rutman theorem which generalizes the Perron Frobenius theorem of nonnegative matrices to positive operators on a Banach space:

Theorem II.2 (Krein-Rutman[20]) *Let X be a Banach space, $K \subset X$ a total cone and $T : X \rightarrow X$ a compact linear operator that is positive (i.e. $T(K) \subset K$) with positive spectral radius $\rho(T)$. Then $\rho(T)$ is an eigenvalue with an eigenvector $\nu \in K - \theta : T\nu = \rho(T)\nu$.*

Here, an operator T is called *compact* if T maps bounded subsets of X to subsets of X whose closure is compact.

We consider only Banach spaces that contain the function $\mathbf{1}$, defined by $\mathbf{1}(k) = 1$. This assumption allows us to make the following definition:

Definition The set $K_{(0,1]} := \{x \in K \mid \theta < x \leq \mathbf{1}\}$ are the *epidemic states* of X .

To prove our main result, Theorem III.1, we will require our function F to be α -concave on $K_{(0,1]}$ [15, 16].

Definition An operator $F : K_{(0,1]} \rightarrow K_{(0,1]}$ is α -concave if for all $0 < t < 1$ and $x \in K_{(0,1]}$ there exists $0 < \alpha = \alpha(t) < 1$ such that

$$F(tx) \geq t^\alpha F(x)$$

We will need the following reformulation of the definition of α -concave.

Lemma II.3 *The following are equivalent:*

(a): F is α -concave.

(b): *There exists a real number $\eta = \eta(t) > 0$ such that*

$$F(tx) \geq t(1 + \eta)F(x), \quad \forall x \in K_{(0,1]} \text{ and } 0 < t < 1.$$

Proof To show (a) implies (b), take $\eta(t) = t^{\alpha(t)-1} - 1$. Conversely to show (b) implies (a), take $\alpha(t) = \frac{\ln(t(1+\eta(t)))}{\ln t}$. ■

B. Axioms for a general epidemic model

The axioms of our general epidemic model ensure that Theorem III.1 hold. Any physical system satisfying these axioms has an “epidemic threshold”.

1. Axioms

Using the four primitive notions $X(\Omega)$, λ , T , and F introduced above, we describe our model. A system $(X(\Omega), \lambda, T, F)$ shall be called a *general epidemic model* if the following six axioms are satisfied.

Axiom 1:: Ω is the sample space of a probability space (Ω, P) , where P is a probability measure on a σ -algebra of subsets of Ω .

Axiom 2:: X is a Banach space of real valued functions on Ω , containing the element $\mathbf{1}$, defined by $\mathbf{1}(k) = 1$ for all $k \in \Omega$. X has norm $\|\bullet\|$, partially ordered by a normal total cone K of nonnegative functions.

Axiom 3:: The operator $T : X \rightarrow X$ is a compact integral operator that is positive (i.e. $T(K) \subset K$) with positive spectral radius $\lambda' = \rho(T)$. Here, integral operator means there is a measurable function, called the *kernel* of T ,

$$\tau : \Omega \times \Omega \rightarrow [0, \infty)$$

such that for all $f \in X$, we have

$$T(f)(k) = \int_{\Omega} \tau(k, k') f(k') dP(k').$$

By the Krein Rutman theorem, λ' is an eigenvalue of T with eigenvector $\nu \in K - \theta$.

Axiom 4:: The eigenvector ν , given in axiom 3, has strictly positive entries.

Axiom 5:: $\|T(x)\| < \lambda' \|x\|$ ($x \in K$)

Axiom 6:: Let

$$K_{(0,1]} = \{x \in K \mid \theta < x \leq \mathbf{1}\},$$

and for $\epsilon > 0$, let

$$C_{\nu, \epsilon} = \{x \in K \mid \epsilon \nu \leq x \leq \mathbf{1}\}.$$

There is a fixed point equation $x = F(x)$ where $F : X \rightarrow X$ is a function satisfying the following conditions:

- a.: F is increasing on $K_{(0,1]}$ and $F(K_{(0,1]}) \subseteq K_{(0,1]}$
- b.: $\|F(x)\| \leq \lambda \|T(x)\|$ for some $\lambda > 0$ (called the *infection rate*)
- c.: For $\lambda > 1/\lambda'$ there exists an $\epsilon > 0$ such that $F(\epsilon \nu) \in C_{\nu, \epsilon}$
- d.: F is α -concave

2. Remarks

Axiom 1 says that our epidemic model is on top of a “network” with probability distribution $P(k)$, where $k \in \Omega$ is a (possibly infinite) set. Axiom 2 says that the model contains epidemic states or probabilities that a node of degree k is infected. Axiom 3 provides a linear operator, T , on the space of positive functions on Ω . The operator T is a measure of the possibility that a degree k node can be infected. Axiom 4 is necessary when the underlying network isn’t connected or when Ω isn’t finite. The inequality in Axiom 5 is used in Theorem III.1 to derive a necessary condition for our fixed point equation to have a solution.

Axiom 6 says that our general epidemic model has a fixed point equation with certain properties. It is helpful to consider a simple example arising from the SIS model on a homogeneous network. Let

$$F(x) = \frac{\lambda x}{1 + \lambda x},$$

where x and λ are real numbers. This F satisfies axioms 6a-d (here the set Ω consists of a single point, $T(x) = x$, and $\|\bullet\|$ is the absolute value). Axioms 6a, c, d are required for the existence and uniqueness part of the proof of Theorem III.1. Axiom 6b is needed to derive a necessary condition for our fixed point equation to have a solution. Note that Axiom 6b implies

$$F(\theta) = \theta, \quad (3)$$

since $T(\theta) = \theta$ and by the nonnegative and positive properties of a vector norm (see section VII B).

III. THE EPIDEMIC THRESHOLD FOR A GENERAL EPIDEMIC MODEL

The following theorem defines the epidemic threshold for a general epidemic model.

Theorem III.1 *Let $(X(\Omega), \lambda, T, F)$ be a general epidemic model satisfying the axioms of Section II B 1. There exists a unique strictly positive solution $x : \Omega \rightarrow (0, 1]$ to the fixed point equation $x = F(x)$ if and only if $\lambda > \lambda_c$ where $\lambda_c := 1/\lambda'$, with λ' denoting the largest positive eigenvalue of the operator T . The number λ_c is thus referred to as the **epidemic threshold**.*

Proof First we show that for the fixed point equation $x = F(x)$ in axiom 6 to have a strictly positive solution it is necessarily that $\lambda > \lambda_c$, where $\lambda_c = 1/\lambda'$. Lets assume that $x \in K_{(0,1]}$ is a strictly positive solution. By axiom 6b, we have the inequality $\|x\| = \|F(x)\| < \lambda \|T(x)\|$.

Furthermore, by axiom 5, $\|T(x)\| \leq \lambda' \|x\|$. We conclude that

$$\|x\| < \lambda \lambda' \|x\|.$$

Dividing by $\|x\| \neq 0$ we see that $\lambda \lambda' > 1$.

Next we prove a sufficient condition for our fixed point equation to have a solution $x \in K_{(0,1]}$.

Step 1 (existence): We assume that $\lambda \lambda' > 1$. By axiom 6d (which requires that $\lambda \lambda' > 1$), there exists an $\epsilon > 0$ such that $F(\epsilon\nu) \in C_{\nu,\epsilon}$. We will show that there exists a strictly positive $x \in C_{\nu,\epsilon}$ satisfying the fixed point equation $F(x) = x$ by constructing two sequences, $\{u_n\}, \{w_n\}$, one increasing and one decreasing, that converge to x . Notice that x is strictly positive since by Axiom 4, ν is strictly positive.

We define the following sequences:

$$u_0 = \epsilon\nu, \quad w_0 = \mathbf{1}, \quad u_n = F(u_{n-1}), \quad w_n = F(w_{n-1}),$$

for $n = 1, 2, \dots$. Then $u_0, w_0 \in C_{\nu,\epsilon}$, $u_0 \leq w_0$, and

$$u_1 = F(\epsilon\nu) \geq \epsilon\nu = u_0. \quad (4)$$

$$w_1 = F(\mathbf{1}) \leq \mathbf{1} = w_0. \quad (5)$$

Here, equation (4) holds because $F(\epsilon\nu) \in C_{\nu,\epsilon}$. Equation (5) holds because $\mathbf{1} \in K_{(0,1]}$ and so by axiom 6a, $F(\mathbf{1}) \in K_{(0,1]}$.

By equations (4),(5), we see that $u_0 \leq u_1 \leq w_1 \leq w_0$. It follows by induction and the monotonicity of F (axiom 6a) that

$$u_0 \leq u_1 \leq \dots u_n \leq \dots \leq w_n \leq \dots w_1 \leq w_0.$$

Notice that $u_n, w_n \in C_{\nu,\epsilon}$ because $\epsilon\nu \leq u_0, w_0 \leq \mathbf{1}$.

In what follows we prove that $\{u_n\}$ and $\{w_n\}$ are Cauchy sequences. For any $n \in \mathbb{N}$, there exists $\mu > 0$ such that $\mu w_n \leq u_n \leq w_n$ since u_n and w_n are interior points of K . Let

$$A_n = \{\mu > 0 | u_n \geq \mu w_n\},$$

and set

$$r_n = \sup A_n.$$

We see that $0 < r_n \leq 1$ and r_n is non-decreasing. So $\lim_{n \rightarrow \infty} r_n = r$ for some $0 < r \leq 1$. We show that $r = 1$ by assuming that $0 < r < 1$ and deriving a contradiction.

Since K is a closed set, $r_n \in A_n$. Furthermore we have,

$$\begin{aligned} u_{n+1} &= F(u_n) \\ &\geq F(r_n w_n) \quad (\text{since } F \text{ is increasing}) \\ &= F\left(\frac{r_n}{r}(r w_n)\right) \\ &\geq \frac{r_n}{r} \left(1 + \eta\left(\frac{r_n}{r}\right)\right) F(r w_n), \quad \text{for} \\ &\quad \eta\left(\frac{r_n}{r}\right) > 0 \quad (\text{by } \frac{r_n}{r} < 1, \text{ Lemma II.3, axiom 6d}) \\ &\geq \frac{r_n}{r} F(r w_n) \\ &\geq r_n (1 + \eta(r)) F(w_n), \quad \text{for} \\ &\quad \eta(r) > 0 \quad (\text{since } 0 < r < 1) \\ &= r_n (1 + \eta(r)) w_{n+1}. \end{aligned}$$

Hence $r_n(1 + \eta(r)) \in A_{n+1}$ and since r_{n+1} is the least upperbound of A_{n+1} we have $r_{n+1} \geq r_n(1 + \eta(r))$. Letting $n \rightarrow \infty$ we obtain $r \geq r(1 + \eta(r))$. But this leads to a contradiction since $\eta(r) > 0$. So we conclude that $r = 1$ as claimed.

For any natural number p we have

$$0 \leq u_{n+p} - u_n \leq w_n - u_n \leq w_n - r_n w_n \leq (1 - r_n) w_0.$$

From the normality of K (axiom 2) it follows that $\|u_{n+p} - u_n\| \leq N(1 - r_n)\|w_0\| \rightarrow 0$ as $n \rightarrow \infty$. So u_n is a Cauchy sequence; the same reasoning shows that $\{w_n\}$ is also a Cauchy sequence.

Now we prove that F has a fixed point $x \in C_{\nu,\epsilon}$. By the completeness property of Banach spaces, there exists $u^*, w^* \in X$ such that $u_n \rightarrow u^*$ and $w_n \rightarrow w^*$. We have,

$$\begin{aligned} u_n \leq w_n &\Leftrightarrow w_n - u_n \in K \\ &\Rightarrow w^* - u^* \in K \quad (\text{since } K \text{ is closed}) \\ &\Leftrightarrow u^* \leq w^*. \end{aligned}$$

It follows that $u_n \leq u^* \leq w^* \leq w_n$ for $n = 0, 1, 2, \dots$, and so $u^*, w^* \in C_{\nu,\epsilon}$. Furthermore,

$$w^* - u^* \leq w_n - u_n \leq (1 - r_n) w_0.$$

Hence by the normality of K , $\|w^* - u^*\| \rightarrow 0$. Let $x \equiv u^* = v^*$. To complete the proof it remains to show that x is a fixed point of F . Indeed,

$$\begin{aligned} u_{n+1} &= F(u_n) \leq F(u^*) = F(x) \\ &= F(w^*) \leq F(w_n) = w_{n+1}, \end{aligned}$$

and letting $n \rightarrow \infty$ yields $F(x) = x$. It follows that F has a fixed point $x \in K_{(0,1]}$ as claimed.

step 2 (uniqueness): We show that F has a unique fixed point $x \in K_{(0,1]}$.

Suppose that x and y are fixed points of F in $K_{(0,1]}$. Because x and y are interior points of K there exists an $a > 0$ such that $ay \leq x \leq \frac{y}{a}$. Let

$$A = \{a > 0 | ay \leq x \leq \frac{y}{a}\},$$

and let

$$a_0 = \sup A.$$

Note that $a_0 \in A$ since K is closed. We claim that $a_0 = 1$ (clearly $0 < a_0 \leq 1$). Suppose not; if $0 < a_0 < 1$ then $x = F(x) \geq F(a_0 y) \geq a_0^\alpha F(y) = a_0^\alpha y$ (by axioms 6a and 6d). Similarly, we have $y = F(y) \geq F(a_0 x) \geq a_0^\alpha F(x) = a_0^\alpha x$, so $a_0^\alpha \in A$. But $a_0^\alpha > a_0$ (since $0 < \alpha < 1$ and $0 < a_0 < 1$) contradicting the fact that a_0 is the least upper bound of A . We conclude that $x = y$, as required. ■

Theorem III.1 guarantees a unique strictly positive solution to the fixed point equation, $x = F(x)$, when the infection rate, λ , is above the epidemic threshold. Here, strictly positive means that the density of infected individuals ($\rho(k)$ in the SIS language of equation (2)) is positive for all $k \in \Omega$. The theorem doesn't rule out the possibility of a solution in which $\rho(k)$ is zero for some but not all k . This situation however doesn't happen in the SIS model. For the SIS model, when the infection rate is above the epidemic threshold there is a unique nonzero fixed point [21]. It follows by Theorem III.1 that the only nonzero solution in the SIS model is strictly positive.

IV. A HYBRID SIS/SIR MODEL ON A COMPLEX NETWORK

On our complex network we consider a three state hybrid model

$$S \xrightleftharpoons[\lambda_1]{\lambda_2} I \xrightarrow{\lambda} R, \quad (6)$$

where each node is one of three states: S =healthy (susceptible to infection), I =infective, or R =immune (recovered). λ_1 , and λ_2 , is the rate that an infected individual becomes susceptible, and recovered, respectively, and 1 is the rate that a susceptible individual becomes infected. The dynamics of our model will be discussed in section IV C.

We are in particular interested in probability functions $f : \Omega \rightarrow [0, 1]$ that assign to each degree $k \in \Omega$ the probability that that a degree k vertex is susceptible, infected or removed. Lets write the density of susceptible, infected and recovered nodes of degree k at time t as $S(k, t)$, $I(k, t)$ and $R(k, t)$ respectively with $S(k, t) + I(k, t) + R(k, t) = 1$. These are the states of the system at time t . We will write $S(k)$, $I(k)$ and $R(k)$ for the steady state values of the epidemic states.

A. The SIS case

We call our model the Hybrid SIS/SIR model because it includes the classical *Susceptible-Infected-Susceptible* (SIS) and *Susceptible-Infected-Recovered* (SIR) epidemic

models as special cases depending on the values of λ_2 and λ_1 . In the first case, (called the SIS case) $\lambda_2 = 0$, we get the classical two state SIS model,

$$S \xrightleftharpoons[1]{\lambda} I.$$

Here the infection rate is $\lambda = \frac{1}{\lambda_1}$ after dividing both the rates by λ_1 and rescaling time.

The operator

$$T : L^2(\Omega) \rightarrow L^2(\Omega)$$

in case 1 (SIS case) of our hybrid model, is defined by

$$T(f)(k) = \sum_{k' \in \Omega} k P(k'|k) f(k') \quad (7)$$

for all $f \in L^2(\Omega)$.

This definition of T is motivated by the following:

Proposition IV.1

$$T(I)(k) = k \sum_{k'} P(k'|k) I(k') \quad (8)$$

is a measure of the possibility that a degree k node can be infected.

Proof Consider a undirected connected graph, and the set of all possible orientations of its edges. Let E be the subset of directed edges emanating from a vertex of degree k . Also, let E' be the set of directed edges leading to a vertex of degree k' . The number of edges originating at k and ending at k' , $n(E \cap E')$, multiplied with the probability a degree k' vertex is infected, $I(k')$, gives the number of links originating from a degree k vertex and leading to an infected degree k' vertex. Dividing $n(E \cap E')I(k')$ by the size of E ,

$$\frac{n(E \cap E)I(k)}{n(E)} = P(E'|E)I(k'),$$

gives the probability that an edge originating at a degree k vertex leads to infected degree k' vertex. We rewrite $P(E'|E)$ as $P(k'|k)$. Summing over all the degrees k' in our graph we find that the probability that a half edge originating at a node of degree k leads to an infected node is

$$\sum_{k'} P(k'|k) I(k').$$

The likelihood that a healthy node of degree k makes contact with a degree k' infected node is k times more likely then if the healthy node only has a single edge, proving the claim. ■

B. The SIR case

The second case (called the SIR case) is where $\lambda_2 \neq 0$,

$$S \xrightleftharpoons[\frac{\lambda_1}{\lambda_2}]^{\lambda} I \xrightarrow{1} R.$$

The infection rate is $\lambda = \frac{1}{\lambda_2}$ after dividing the rates by λ_2 and renormalizing time. In the case when $\lambda_1 = 0$ we get the classical three state SIR model.

The operator

$$T : L^2(\Omega) \rightarrow L^2(\Omega)$$

if $\lambda_1 = 0$ (i.e. classical SIR model) is defined by

$$T(f)(k) = \sum_{k' \in \Omega} k \frac{(k' - 1)}{k'} P(k'|k) f(k') \quad (9)$$

for all $f \in L^2(\Omega)$.

We now multiply $P(k'|k)I(k', t)$ by the fraction, $\frac{k'-1}{k'}$, since not all of the edges leading to an infected node can transmit the disease. In the SIR model, we assume that if a degree k' node is infective that one of its edges cannot transmit disease since it leads to the infected (or removed) node that originally infected it. Hence we have,

Proposition IV.2

$$T(I)(k) = k \sum_{k'} \frac{k' - 1}{k'} P(k'|k) I(k')$$

is a measure of the possibility that a degree k node can be infected in the SIR case.

When $\lambda_1 \neq 0$, there is a nonzero possibility that an infective node will become susceptible again. Hence T will be the same as the SIS case when $\lambda_2 = 0$ but $\lambda_1 \neq 0$.

C. Fixed point formula for the SIS/SIR model

Our fixed point formula arises as the steady state solution of the system of differential equations,

$$\begin{cases} \partial_t S(k, t) = \lambda_1 I(k, t) - S(k, t)T(I)(k, t), \\ \partial_t I(k, t) = -(\lambda_1 + \lambda_2)I(k, t) + S(k, t)T(I)(k, t), \\ \partial_t R(k, t) = \lambda_2 I(k, t) \\ S(k, 0) = 1 - I_0(k), \\ I(k, 0) = I_0(k), \\ R(k, 0) = 0 \end{cases} \quad (10)$$

where $I_0 : \Omega \rightarrow [0, 1]$ is the initial state of the system.

Remark Here, $T(I)$ is the operator T acting on the function $I_t : \Omega \rightarrow \mathbb{R}$, defined by $I_t(k) := I(k, t)$.

These nonlinear differential equations describes the dynamics of (6). Our assumptions about the transmission of the infection are as follows:

- i:** The gain of infected degree k nodes each time step is $S(k, t)T(I)(k, t)$
- ii:** The loss of infected degree k nodes each time step is proportional to $I(k, t)$ with proportionality constant $(\lambda_1 + \lambda_2)$.
- iii:** The gain of recovered degree k nodes each time step is proportional to $I(k, t)$ with proportionality constant λ_2

SIS case ($\lambda_2 = 0$):

In this case $R(k, t) = 0$ so $S(k, t) = 1 - I(k, t)$. Setting $\frac{\partial I(k, t)}{\partial t} = 0$ for the differential equation of $I(k, t)$ in (10) gives

$$0 = -(\lambda_1 + \lambda_2)I(k) + (1 - I(k))T(I)(k). \quad (11)$$

Solving equation (11) for $I(k)$ gives the fixed point formula

$$I(k) = \frac{\lambda T(I)(k)}{1 + \lambda T(I)(k)}, \quad (12)$$

where $\lambda = \frac{1}{\lambda_1}$ is the infection rate.

SIR case ($\lambda_2 \neq 0$):

To simplify our fixed point equation we assume $I_0 \approx 0$. We also assume $S(k, t) \neq 0$ and $\lim_{t \rightarrow \infty} \frac{I(k, t)}{S(k, t)} = 0$. This is justified since we are primarily interested in infection rates $\lambda < 1$ so the rate of conversion of S to I is slower than the rate of conversion of I to S or R . Furthermore, we necessarily have $\lim_{t \rightarrow \infty} I(k, t) = 0$ since the antiderivative of I is R and R is bounded above by 1.

The differential equation for $S(k, t)$ in (10) gives us

$$\frac{dS(k, t)}{S(k, t)} = (\lambda_1 \frac{I(k, t)}{S(k, t)} - T(I)(k, t))dt. \quad (13)$$

Integrating both sides of equation (13) we have

$$\ln S(k, t) = \lambda_1 \int \frac{I(k, t)}{S(k, t)} dt - \frac{1}{\lambda_2} T(R)(k, t). \quad (14)$$

Here we solve the differential equation for R in (10) giving us

$$\int I(k, t) dt = \frac{1}{\lambda_2} R(k, t),$$

and hence

$$\int T(I)(k, t) dt = \frac{1}{\lambda_2} T(R)(k, t).$$

Taking $t \rightarrow \infty$ on both sides of (14) we have

$$\ln S(k) = -\frac{1}{\lambda_2} T(R)(k),$$

using our assumption that $\lim_{t \rightarrow \infty} \frac{I(k,t)}{S(k,t)} = 0$. Solving for S and using our assumption that $I_0 \approx 0$ we have

$$S(k) = e^{-\lambda T(R)(k)}, \quad (15)$$

where $\lambda = \frac{1}{\lambda_2}$ is the infection rate. Because $R(k, t)$ is the antiderivative of $I(k, t)$ and $R(k, t)$ is bounded above by 1, it follows that $I(k) = 0$ and $R(k) = 1 - S(k)$. Hence from equation (15) we get the fixed point equation

$$R(k) = 1 - e^{-\lambda T(R)(k)}. \quad (16)$$

D. Checking the axioms for the Hybrid SIS/SIR model

We confirm that our hybrid SIS/SIR model satisfy the 6 axioms given in Section II B 1.

Axiom 1:: Let $\Omega \subseteq \mathbb{N}$ be the finite set of degrees of the network. For $A \subseteq \Omega$ we define the probability measure $P(A) = \sum_{k \in A} P(k)$ where $P(k)$ is a probability mass function on Ω .

Axiom 2:: Our Banach space is the space of square integral functions $X = L^2(\Omega)$ (see Appendix VII A). Since Ω is finite, all functions $f : \Omega \rightarrow \mathbb{R}$ are in $L^2(\Omega)$. In particular the element $\mathbf{1} \in L^2(\Omega)$.

Let $K \subset L^2(\Omega)$ be the cone consisting of nonnegative functions $x : \Omega \rightarrow [0, \infty)$. K is a total cone since any real function can be written as the difference of two nonnegative functions. For the SIS case we give $L^2(\Omega)$ the norm $\|x\|_2 = \sqrt{\langle x, x \rangle}$ given by the inner product (32). From the partial ordering on $L^2(\Omega)$ it is immediately clear that $x \leq y \Rightarrow \|x\|_2 \leq \|y\|_2$. Hence K is a normal cone with respect to this norm. In fact K is a normal cone with respect to any norm on $L^2(\Omega)$ since all norms on a finite dimensional vector space are equivalent [22]. For the SIR case, we use the vector norm

$$\|x\|_{S^{-1}} \equiv \|S^{-1}x\|_2, \quad (17)$$

where S is a nonsingular matrix such that $S^{-1}TS$ is a diagonal matrix..

Axiom 3:: For the SIS case or the SIR case with $\lambda_1 \neq 0$ writing equation (7) as

$$T(f)(k) = \sum_{k' \in \Omega} \frac{kP(k'|k)}{P(k')} f(k')P(k'),$$

we see that the kernel of T is

$$\tau(k, k') = \frac{kP(k'|k)}{P(k')}.$$

Notice that we have replaced the integral in (3) by a summation since Ω is a discrete set. By definition, for $\tau(k, k')$ to be a measurable function, the preimage of any interval in $[0, \infty)$ must be an element of the product sigma algebra of $\Omega \times \Omega$. This is trivially satisfied here since the sigma algebra is the power set of $\Omega \times \Omega$. That T is a compact operator follows from that fact that $L^2(\Omega) \cong \mathbb{R}^N$ and any linear map on \mathbb{R}^N is compact by the Bolzano-Weierstrass theorem.

For the SIR case with $\lambda_1 = 0$ writing equation (9) as

$$T(f)(k) = \sum_{k' \in \Omega} \frac{k(k'-1)P(k'|k)}{k'P(k')} f(k')P(k'),$$

we see that the kernel of T is

$$\tau(k, k') = \frac{k(k'-1)P(k'|k)}{k'P(k')}.$$

We can analogously argue that $\tau(k, k')$ is measurable and T is compact.

Axiom 4:: Identifying $L^2(\Omega)$ with Euclidean space, and the operator T with a matrix, we have $T = [T_{k,k'}]$ where

$$T_{k,k'} = \frac{kP(k'|k)}{P(k')}$$

or

$$T_{kk'} = \frac{k(k'-1)}{k'P(k')} P(k'|k),$$

depending on what case we are in (compare with matrices given in [13]). Since the network is connected, the associated directed graph $\Gamma(T)$ is strongly connected and hence T is irreducible. Furthermore, T is nonnegative since $\tau(k, k') \geq 0$. We apply the Perron-Frobenius theorem to show that the largest eigenvalue λ' of T has an associated strictly positive eigenvector $v : \Omega \rightarrow (0, \infty)$.

Axiom 5:: In the SIS case or the SIR case with $\lambda_1 \neq 0$, we have $T_{k,k'} = \frac{kP(k'|k)}{P(k')}$. Since an edge in the graph links a node to another node, one can derive the following degree detailed balance condition [13]:

$$kP(k'|k)P(k) = k'P(k|k')P(k'). \quad (18)$$

It follows immediately that $T_{k,k'} = T_{k',k}$ so T is a symmetric matrix. By the Spectral theorem for symmetric matrices, there is a real orthogonal matrix $P \in M_n$ and a real diagonal matrix $\Lambda \in M_n$ such that $T = P\Lambda P^t$. By Theorem VII.8 we have $\|Ax\|_2 \leq \lambda_1 \|x\|_2$ as required.

For the SIR case with $\lambda_1 = 0$, the operator T isn't symmetric. Nevertheless, we show below that it has a basis of eigenvectors and hence is diagonalizable.

Proposition IV.3 *The nonsymmetric matrix, $[T_{k,k'}]$, where $T_{k,k'} = \frac{k(k'-1)}{k'P(k')}P(k'|k)$, has the same eigenstructure as the symmetric matrix, $[S_{k,k'}]$ where $S_{k,k'} = \sqrt{T_{k,k'}T_{k',k}}$, and hence is diagonalizable.*

Proof Using the detailed balance condition in equation (18) we have the symmetric equation

$$\frac{(k-1)}{k}T_{k,k'} = \frac{(k'-1)}{k'}T_{k',k}. \quad (19)$$

Let $\{V^i\}$ be a basis of eigenvectors of the symmetric operator $[S_{k,k'}]$ where $S_{k,k'} = \sqrt{T_{k,k'}T_{k',k}}$. In other words we have

$$\sum_{k'} \sqrt{T_{k,k'}T_{k',k}}V_{k'}^i = \lambda^i V_k^i. \quad (20)$$

We show that

$$W_k^i = \sqrt{\frac{(k-1)}{k}}V_k^i \quad (21)$$

is an eigenvector of $T_{k',k}$ with eigenvalue λ^i . Indeed, from equation (20) we have

$$\sum_{k'} \sqrt{T_{k,k'}T_{k',k}} \frac{(k'-1)}{k'} W_{k'}^i = \lambda^i \sqrt{\frac{(k-1)}{k}} W_k^i.$$

Equation (19) then implies

$$\sum_{k'} T_{k,k'} \sqrt{\frac{(k-1)}{k}} W_{k'}^i = \lambda^i \sqrt{\frac{(k-1)}{k}} W_k^i. \quad (22)$$

Cancelling $\sqrt{\frac{(k-1)}{k}}$ from both sides of equation (22), gives

$$\sum_{k'} T_{k,k'} W_{k'}^i = \lambda W_k^i,$$

so $\{W^i\}$ are eigenvectors of $[T_{k,k'}]$ with eigenvalues $\{\lambda^i\}$. The independence of $\{V^i\}$ implies that $\{W^i\}$ form a basis. Hence $T_{k,k'}$ is diagonalizable. \blacksquare

Next applying Theorem VII.9 we conclude that Axiom 5 holds.

Axiom 6:: SIS case: From equation (12) let

$$F(x)(k) = \frac{\lambda T(x)(k)}{1 + \lambda T(x)(k)},$$

where

$$T(f)(k) = \sum_{k' \in \Omega} kP(k'|k)f(k').$$

SIR case: From equation (16) let

$$F(x)(k) = 1 - e^{-\lambda T(x)(k)},$$

where

$$T(f)(k) = \sum_{k' \in \Omega} \frac{kP(k'|k)}{P(k')} f(k')P(k') \quad (\lambda_1 = 0)$$

$$T(f)(k) = \sum_{k' \in \Omega} kP(k'|k)f(k') \quad (\lambda_1 \neq 0).$$

(a): SIS case: Since $T(x)(k) > 0$ for $x \in K_{(0,1]}$ it is clear that $F(x)(k) \in K_{(0,1]}$. Next we show that the function $F(x)(k)$ is increasing on $K_{(0,1]}$. Indeed, suppose $x \leq y$. We have $\lambda T(x) \leq \lambda T(y)$ because the function $\tau(k, k') \geq 0$ for all $k, k' \in \Omega$ and $\lambda > 0$. Then,

$$\begin{aligned} x \leq y &\Rightarrow \lambda T(x) \leq \lambda T(y) \\ &\Rightarrow \frac{\lambda T(x)(k)}{1 + \lambda T(x)(k)} \leq \frac{\lambda T(y)(k)}{1 + \lambda T(y)(k)}, \end{aligned}$$

for all $k \in \Omega$.

SIR case: Analogous to the SIS case.

(b): SIS case: We have

$$F(x)(k) = \frac{\lambda T(x)(k)}{1 + \lambda T(x)(k)} \leq \lambda T(x)(k).$$

Our norm in the SIS case is the Euclidean norm $\|\bullet\|_2$ (which has normality constant 1 by Proposition VII.2), so

$$F(x) \leq \lambda T(x) \Rightarrow \|F(x)\|_2 \leq \lambda \|T(x)\|_2.$$

SIR case: By the inequality $1 - e^{-x} \leq -x$, for $x \leq 0$, we have

$$F(x) = 1 - e^{-\lambda T(x)(k)} \leq \lambda T(x).$$

For the SIR case with $\lambda_1 \neq 0$ we have norm $\|\bullet\|_{S^{-1}}$ given in equation (17). By Proposition VII.2, the norm $\|\bullet\|_{S^{-1}}$ has normality constant 1. Hence $\|F(x)\|_{S^{-1}} \leq \lambda \|T(x)\|_{S^{-1}}$ as required. For the SIR case with $\lambda_1 = 0$ we have the Euclidean norm $\|\bullet\|_2$ (which has normality constant 1 by Proposition VII.2). Hence $\|F(x)\|_2 \leq \lambda \|T(x)\|_2$ as required.

(c): Assuming $\lambda > \lambda'$, we show that $\epsilon\nu \leq F(\epsilon\nu)$.

SIS case: We have the following equivalences,

$$\begin{aligned} \epsilon\nu < F(\epsilon\nu) &\iff \\ \epsilon(k)\nu(k) &< \frac{\lambda\epsilon(k)T(\nu)(k)}{1 + \lambda\epsilon(k)T(\nu)(k)} \iff \\ \nu(k) &< \frac{\lambda\lambda'\nu(k)}{1 + \lambda\lambda'\epsilon(k)\nu(k)} \iff \\ &1 < \frac{\lambda\lambda'}{1 + \lambda\lambda'\epsilon(k)\nu(k)}. \end{aligned}$$

The last inequality is true if $\epsilon(k) = 0$, because we are assuming that $\lambda\lambda' > 1$. By the continuity of $\frac{\lambda\lambda'}{1+\lambda\lambda'\epsilon(k)\nu(k)}$ as a function of $\epsilon(k)$, the last inequality holds for a small number $\epsilon(k) > 0$. We also require $F(\epsilon\nu) \leq \mathbf{1}$. This holds if $\nu\epsilon \leq \mathbf{1}$. Hence $F(\epsilon\nu) \in C_{\nu,\epsilon}$ for $\epsilon := \min\{\epsilon(k), \frac{1}{\nu(k)}\}$.

SIR case: We have the following equivalences,

$$\begin{aligned} \epsilon\nu < F(\epsilon\nu) &\iff \\ \epsilon(k)\nu(k) < 1 - e^{-\lambda\lambda'\epsilon(k)\nu(k)} &\iff \\ \nu(k) < \lambda\lambda'\nu(k)(1 - \epsilon(k)G(\epsilon(k))), & \end{aligned}$$

where $\lim_{\epsilon(k) \rightarrow 0} G(\epsilon(k)) = \frac{\lambda\lambda'\nu(k)}{2}$. The last equivalence holds by writing the Taylor series of $1 - e^{-\lambda\lambda'\epsilon(k)\nu(k)}$. Preceding as in the SIS case, $F(\epsilon\nu) \in C_{\nu,\epsilon}$ for $\epsilon := \min\{\epsilon(k), \frac{1}{\nu(k)}\}$.

(d): To show that F is α -concave, by Lemma II.3, we show that for $0 < t < 1$, and $x \in K_{(0,1]}$, there exists $\eta > 0$ such that $F(tx)(k) \geq t(1 + \eta)F(x)(k)$.

SIS case: For $0 < t < 1$ we have

$$\begin{aligned} F(tx)(k) &= \frac{\lambda t T(x)(k)}{1 + \lambda t T(x)(k)} \\ &> \frac{\lambda t T(x)(k)}{1 + \lambda T(x)(k)} = tF(x)(k). \end{aligned}$$

The inequality above is strict since $x \in K_{(0,1]}$ implies that $T(x)(k) > 0$. We conclude that there is a small positive number η such that $F(tx) \geq t(1 + \eta)F(x)$ as required.

SIR case: We show next that

$$\begin{aligned} F(tx)(k) - tF(x)(k) &= 1 - e^{-t\lambda T(x)(k)} \\ &\quad - t + te^{-\lambda T(x)(k)}, \end{aligned}$$

is strictly positive for $x \in K_{(0,1]}$ and $0 < t < 1$. Define

$$f(t, y) = 1 - e^{-ty} - t + te^{-y},$$

where $0 < t < 1$ and $y = \lambda T(x)(k) > 0$. Then

$$\frac{\partial f(t, y)}{\partial y} = t(e^{-ty} - e^{-y}) > 0,$$

so $f(t, y)$ is strictly increasing in y for every $0 < t < 1$. Because $f(0, t) = 0$ we conclude that $f(y, t) > 0$, for $0 < t < 1$ and $y > 0$. Equivalently, $F(tx) > tF(x)$ for $0 < t < 1$ and $x \in K_{(0,1]}$. It then follows that there is an $\eta > 0$ such that $F(tx) \geq t(1 + \eta)F(x)$ as required.

E. Remarks

(1) The uncorrelated situation is really a special case of the correlated situation, with $P(k'|k) = \frac{k'P(k')}{\langle k \rangle}$, where average degree of a node is $\langle k \rangle = \sum_i iP(i)$. To see this, in Proposition IV.1 we have $P(k'|k) = P(E'|E) = P(E')$ when E' and E are independent. If the network is uncorrelated, E' is the subset of half edges connected to a vertex of degree k' . Let S be set of all half edges. The size of these sets are,

$$n(E') = (\text{number vertices degree } k')k' \text{ and,}$$

$$n(S) = \sum_i (\text{number vertices degree } i)i.$$

Then $P(E') = \frac{n(E')}{n(S)}$. Dividing numerator and denominator by N we get

$$P(k'|k) = P(E') = \frac{k'P(k')}{\sum_i iP(i)}.$$

(2) The operator T in the uncorrelated SIS/SIR hybrid model is the projection operator onto the eigenvector $\nu(k) = k$. For the SIS case or the SIR case with $\lambda_1 \neq 0$, we have

$$T(f)(k) = \frac{k}{\langle k \rangle} \sum_{k'} f(k')k'P(k') = \frac{1}{\langle k \rangle} \langle f, \nu \rangle \nu(k),$$

where $\langle f, \nu \rangle$ is defined in equation (32), and $\nu(k) = k$. Hence,

$$T = \frac{1}{\langle k \rangle} \langle \cdot, \nu \rangle \nu.$$

It follows that the eigenvalue with eigenvector ν , is $\lambda' = \frac{1}{\langle k \rangle} \langle \nu, \nu \rangle = \frac{\langle k^2 \rangle}{\langle k \rangle}$. Analogously, for the SIR case with $\lambda_1 = 0$ we have

$$T(f)(k) = \frac{k}{\langle k \rangle} \sum_{k'} f(k')(k' - 1)P(k'),$$

and

$$T = \frac{1}{\langle k \rangle} \langle \cdot, \nu - 1 \rangle \nu,$$

where the eigenvector ν has eigenvalue $\lambda' = \frac{1}{\langle k \rangle} \langle \nu, \nu - 1 \rangle = \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle}$. These results agree with the epidemic threshold for the SIS and SIR uncorrelated case given in [13].

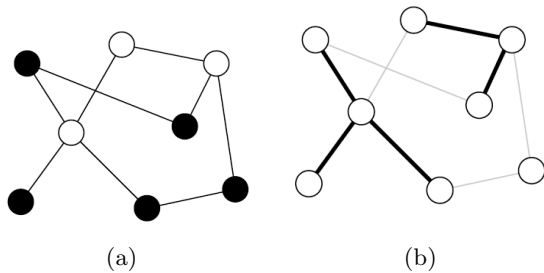


FIG. 1. Site and bond percolation on a network. (a) In site percolation, vertices (sites) are either “occupied” (solid circles) or “unoccupied” (open circles). There are three components of occupied sites (b) In bond percolation, it is the edges (bonds) that are occupied or not (black or gray lines) and the vertices are connected together by occupied edges. There are two components in this case.

V. THE PERCOLATION MODEL ON A COMPLEX NETWORK

In 1959 mathematicians Paul Erdős and Alfred Rényi showed that the largest component of a network, formed by randomly connecting two existing vertices per time step, grows rapidly when the average number of connections per vertex exceeds one [23]. By *component* here we mean a group of vertices that are all connected to each other, directly or indirectly. When the average number of connections per vertex reaches one, a phase transition occurs. The network changes from having many small, fragmented components, to having a single *giant component* of size $\theta(n)$.

Random tree like networks with degree-degree correlations undergo a similar phase transition phenomena [4]. In a *site* or *bond percolation* process, vertices or edges respectively, on a network are randomly “occupied” (with probability λ) or “unoccupied” (with probability $1 - \lambda$), and its components are studied (as shown in Figure 1 adapted from [24]). There is a critical point $0 \leq \lambda_c \leq 1$, called the *percolation threshold*, when a giant connected component forms. We show below that the percolation threshold is an epidemic threshold of our general epidemic model.

A. Fixed point formula for the percolation model

Just as in the SIS/SIR hybrid model, the percolation model has a fixed point equation, $x = F(x)$. In the SIS/SIR model the fixed point equation consists of the stationary solution of a differential equation. In the percolation model we derive the fixed point equation by writing a self consistent equation using first step analysis techniques [25, 26].

We consider a network, with no closed loops and a locally tree-like structure. Let Ω be the degrees of the vertices of the network. Let $y(k)$ be the average prob-

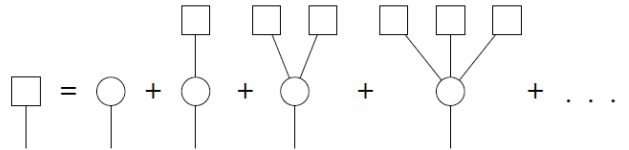


FIG. 2. Schematic representation of the sum rule for the connected component of vertices reached by following a randomly chosen edge. The probability of each such component (left hand side) can be represented as the sum of the probabilities (right hand side) of having only a single vertex, having a single vertex connected to one other component, or two other components, and so forth.

ability that an edge connected to a vertex of degree k , leads to another vertex that doesn’t belong to the giant component [27]. If a pair of vertices aren’t connected to the giant component, then because of the tree like structure of our network, the vertices along the branches going outwards through the second vertex aren’t connected to the giant component, and are independent of one another. It follows that $y(k)$ can be written according to the following probability sum rule (see Figure 2 amended from [26]). We can write the event that a branch through a degree k vertex and a second vertex don’t connect to the giant component as the disjoint union of events that the second vertex has $0, 1, 2, \dots$ branches not connecting to the giant component. If the second vertex has degree k' then the probability that a degree k vertex through this degree k' vertex doesn’t go through the giant component is $P(k'|k)y(k')^{k'-1}$, since the k' vertex has $k' - 1$ branches not connected to the giant component. We can then write $y(k)$ using the following self consistent formula,

$$y(k) = \sum_{k' \in \Omega} P(k'|k)y(k')^{k'-1}.$$

Now suppose we damage the network by removing edges (i.e. bond percolation) with probability $1 - \lambda$. The probability the edge connecting our degree k vertex and the second vertex is unoccupied is $1 - \lambda$. Furthermore, each term of the sum in Figure 2 is multiplied by the probability that it is not damaged, λ . Hence we can write $y(k)$ for the damaged network using the following self consistent formula,

$$y(k) = 1 - \lambda + \lambda \sum_{k' \in \Omega} P(k'|k)y(k')^{k'-1}. \quad (23)$$

It isn’t difficult to see that the same formula results if we remove vertices instead of edges (i.e. site percolation).

Equation (23) is a fixed point equation, however $F(\theta) \neq \theta$ so equation (3) doesn’t hold and hence axiom 6b can’t be true. Nevertheless, the probability $y(k) = 1$ satisfies equation (23) and so we rewrite the equation in terms of $x(k) = 1 - y(k)$. By the binomial theorem we

have

$$\begin{aligned} (1-x(k'))^{k'-1} &= \sum_{n=0}^{k'-1} \binom{k'-1}{n} (-x(k'))^n \\ &= 1 - (k'-1)x(k') + \\ &\quad \sum_{n=2}^{k'-1} \binom{k'-1}{n} (-x(k'))^n \end{aligned} \quad (24)$$

Letting $y(k') = 1 - x(k')$ in equation (23) and using equation (24) we have,

$$\begin{aligned} x(k) &= \lambda - \lambda \sum_{k' \in \Omega} P(k'|k) \\ &\times \left(1 - (k'-1)x(k') + \sum_{n=2}^{k'-1} \binom{k'-1}{n} (-x(k'))^n \right), \end{aligned}$$

or equivalently,

$$\begin{aligned} x(k) &= \lambda \sum_{k' \in \Omega} (k'-1)P(k'|k)x(k') \\ &- \lambda \sum_{k' \geq 3} \sum_{n=2}^{k'-1} \binom{k'-1}{n} (-x(k'))^n P(k'|k). \end{aligned}$$

Writing

$$T(f)(k) = \sum_{k' \in \Omega} (k'-1)P(k'|k)f(k') \quad (25)$$

we have the fixed point equation for bond or site percolation,

$$\begin{aligned} x(k) &= \\ \lambda T(x)(k) &- \lambda \sum_{k' \geq 3} \sum_{n=2}^{k'-1} \binom{k'-1}{n} (-x(k'))^n P(k'|k). \end{aligned} \quad (26)$$

B. Checking the axioms for the Percolation model

Axiom 1:: Same as in the Hybrid SIS/SIR model.

Axiom 2:: Same as in the Hybrid SIS/SIR model. We use the vector norm

$$\|x\|_{S^{-1}} \equiv \|S^{-1}x\|_2, \quad (27)$$

where S is a nonsingular matrix such that $S^{-1}TS$ is a diagonal matrix.

Axiom 3:: From equation 25 we have

$$T(x)(k) = \sum_{k' \in \Omega} (k'-1)P(k'|k)x(k').$$

As before, the kernel, $\tau(k, k') = \frac{(k'-1)P(k'|k)}{P(k')}$, is a measurable function. Under the identification of $L^2(\Omega)$ with Euclidean space given in Appendix VII A, we can write $T = [T_{kk'}]$, where $T_{kk'} = \frac{(k'-1)P(k'|k)}{P(k')}$. This is related to the branching matrix of [4].

Axiom 4:: Same as in SIS/SIR model.

Axiom 5:: We have the following proposition whose proof is analogous to the proof of Proposition IV.3,

Proposition V.1 *The nonsymmetric matrix $[T_{k,k'}]$, where $T_{k,k'} = \frac{k'-1}{P(k')}P(k'|k)$, has the same eigenstructure as the symmetric matrix $[S_{k,k'}]$, where $S_{k,k'} = \sqrt{T_{k,k'}T_{k',k}}$, and hence is diagonalizable.*

The axiom then follows by applying Theorem VII.9.

Axiom 6:: From equation (26) let

$$\begin{aligned} F(x)(k) &= \\ \lambda T(x)(k) &- \lambda \sum_{k' \geq 3} \sum_{n=2}^{k'-1} \binom{k'-1}{n} (-x(k'))^n P(k'|k), \end{aligned} \quad (28)$$

where $T(x)(k) = \sum_{k' \in \Omega} (k'-1)P(k'|k)x(k')$.

(a): Adapting equation (23), we see that equation (28) can be written as

$$F(x)(k) = \lambda \left(1 - \sum_{k' \in \Omega} P(k'|k)(1-x(k'))^{k'-1} \right). \quad (29)$$

If $x \in K_{(0,1]}$, we have $0 \leq (1-x(k'))^{k'-1} < 1$. It follows that

$$0 \leq \sum_{k' \in \Omega} P(k'|k)(1-x(k'))^{k'-1} < 1.$$

Since $0 \leq \lambda \leq 1$, we conclude that $F(x) \in K_{(0,1]}$, as required.

Next we verify that for $x, w \in K_{(0,1]}$, if $x \leq w$ then $F(x) \leq F(w)$. Indeed,

$$\begin{aligned} x(k) \leq w(k) &\Rightarrow \\ -(1-x(k'))^{k'-1} &\leq -(1-w(k'))^{k'-1}. \end{aligned}$$

By equation (29), we then see that $F(x)(k) \leq F(w)(k)$, because λ and $P(k'|k)$ are nonnegative.

(b): For the percolation model we have norm $\|\bullet\|_{S^{-1}}$ defined in equation (27). By Proposition VII.2, this norm has normality constant 1. Hence to show $\|F(x)\|_{S^{-1}} \leq \lambda \|T(x)\|_{S^{-1}}$, it suffices to show that $F(x) \leq \lambda T(x)$. We achieve this by proving that

$$\lambda \sum_{k' \geq 3} P(k'|k) \sum_{n=2}^{k'-1} \binom{k'-1}{n} (-x(k'))^n \geq 0,$$

from equation (28). For $k, k' \in \Omega$, since λ and $P(k'|k)$ are nonnegative, it suffices to show that

$$\sum_{n=2}^{k'-1} \binom{k'-1}{n} (-x(k))^n \geq 0,$$

or equivalently (by equation (24)),

$$-1 + (k' - 1)x(k) + (1 - x(k))^{k'-1} \geq 0. \quad (30)$$

Notice that the derivative of the right hand side of equation (30) with respect to $x(k)$ is

$$(k' - 1) + (k' - 1)(1 - x(k))^{k'-2} > 0,$$

for $0 \leq x(k) \leq 1$ and $k' \geq 3$. It follows that since equation (30) is true for $x(k) = 0$, it is true for $0 \leq x(k) \leq 1$.

(c): We have the following equivalences,

$$\begin{aligned} \epsilon\nu < F(\epsilon\nu) &\Leftrightarrow \\ \epsilon\nu(k) < \lambda T(\epsilon\nu)(k) - \lambda \sum_{k' \geq 3} \sum_{n=2}^{k'-1} & \\ \times \binom{k'-1}{n} (-\epsilon\nu(k))^n P(k'|k) &\Leftrightarrow \\ \nu(k) < \lambda \lambda' \nu(k) (1 - \epsilon(G(\epsilon))) & \end{aligned}$$

where

$$\lim_{\epsilon \rightarrow 0} G(\epsilon) = \lambda \sum_{k' \geq 3} \binom{k'-1}{2} (\nu(k))^2 P(k'|k).$$

The last equivalence is true for $\epsilon = 0$ since $\lambda \lambda' > 1$ by assumption. The function $G(\epsilon)$ is a continuous function of ϵ and so the last equivalence holds for a small positive $\epsilon(k)$ for each $k \in \Omega$. We also require that $F(\epsilon\nu) \leq 1$. This holds if $\nu\epsilon \leq 1$. Hence, $F(\epsilon\nu) \in C_{\nu, \epsilon}$ for $\epsilon := \min\{\epsilon(k), \frac{1}{\nu(k)}\}$.

(d): For $0 < t < 1$ and $x \in K_{(0,1]}$ we first show that $F(tx)(k) - tF(x)(k) > 0$. Using equation (28), it suffices to show that, for $k' \geq 3$,

$$\sum_{n=2} \binom{k'-1}{2} [t(-x(k))^n - (-tx(k))^n] > 0,$$

or equivalently, (by equation (30)),

$$1 - t + t(1 - x(k))^{k'-1} - (1 - tx(k))^{k'-1} > 0.$$

Define

$$f(t, x(k)) = 1 - t + t(1 - x(k))^{k'-1} - (1 - tx(k))^{k'-1}.$$

Then

$$\begin{aligned} \frac{\partial f(t, x(k))}{\partial x(k)} &= t(k' - 1)[(1 - tx(k))^{k'-2} \\ &\quad - (1 - x(k))^{k'-2}] > 0. \end{aligned} \quad (31)$$

The inequality in equation (31) is strict for $0 < t < 1$ and $0 < x(k) \leq 1$ since

$$\begin{aligned} (1 - tx(k)) &> (1 - x(k)) > 0 \Rightarrow \\ (1 - tx(k))^{k-2} &> (1 - x(k))^{k-2}. \end{aligned}$$

Hence, $f(t, x(k))$ is a strictly increasing function for $0 < t < 1$ and $0 < x(k) \leq 1$. Because, $f(t, 0) = 0$, we have $f(t, x(k)) > 0$, or equivalently, $F(tx) > tF(x)$. It follows that there exists an $\eta > 0$ such that $F(tx) \geq t(1 + \eta)F(x)$ as required.

C. Remarks

- (1) As with the hybrid SIS/SIR model the uncorrelated situation is really a special case of the correlated situation, with $P(k'|k) = \frac{k'P(k')}{\langle k \rangle}$, where average degree of a node is $\langle k \rangle = \sum_i iP(i)$.
- (2) The operator T in the uncorrelated percolation model is the projection operator onto the eigenvector $\mathbf{1}(k) = 1$. From equation 25 we have

$$\begin{aligned} T(f)(k) &= \frac{1}{\langle k \rangle} \sum_{k'} f(k') k' (k' - 1) P(k') \\ &= \frac{1}{\langle k \rangle} \langle f, \nu(\nu - 1) \rangle \mathbf{1}(k), \end{aligned}$$

where $\langle f, \nu \rangle$ is defined in equation (32), and $\nu(k) = k$. Hence,

$$T = \frac{1}{\langle k \rangle} \langle \cdot, \nu(\nu - 1) \rangle \mathbf{1}.$$

It follows that the eigenvalue with eigenvector $\mathbf{1}$, is $\lambda' = \frac{1}{\langle k \rangle} \langle \nu, \nu - 1 \rangle = \frac{\langle k^2 \rangle - \langle k \rangle^2}{\langle k \rangle}$. This result agrees with the epidemic threshold for the uncorrelated percolation model in [4].

VI. CONCLUSIONS

In this paper we provide a rigorous axiomatic foundation to the study of epidemiology and ‘‘epidemic-like’’ physical systems. Our general epidemic model consists of four primitive notions $(X(\Omega), \lambda, T, F)$ that satisfy the axioms given in Section IIB1. It has a unique strictly positive fixed point if and only if the infection rate λ exceeds the epidemic threshold, equal to the reciprocal of the largest eigenvalue of the operator T . The fixed point represents the steady state fraction of infected or recovered individuals in the Hybrid SIS/SIR model. In the Percolation model it represents the probability that a vertex connects to the networks’ giant component. Our axioms should apply to a broad range of interacting particle systems.

VII. APPENDIX

A. Banach Spaces

Many of the infinite-dimensional spaces studied in mathematics are **Banach spaces**, including continuous functions on the unit interval $[0, 1]$ (or more generally on a compact Hausdorff space), spaces of Lebesgue integrable functions known as L^p spaces, and spaces of holomorphic functions. Banach spaces are defined as complete normed vector spaces. This means that a Banach space is a vector space X over the real or complex numbers with a norm $\|\cdot\|$ such that every Cauchy sequence (with respect to the metric $d(x, y) = \|x - y\|$) in X has a limit in X . In a metric space (X, d) , a sequence x_n is a **Cauchy sequence** if for any $\epsilon > 0$, there is some integer N such that $d(x_m, x_n) < \epsilon$ whenever m and n are $\geq N$. In other words, a Cauchy sequence is a sequence whose elements become arbitrarily close to each other as the sequence progresses. The property of a space X that every Cauchy sequence in X converges is known as **completeness**. The line \mathbb{R} for example is complete.

Some examples of Banach spaces include:

1. Euclidean space \mathbb{R}^n . The space consists of n -tuples of scalars $a = (a_1, \dots, a_n)$ with $\|a\| = (\sum_{k=1}^n |a_k|)^{1/2}$. \mathbb{R}^n has a **cone** K consisting of points in the first quadrant (i.e. $a_i \geq 0$ for all i). This cone induces a **partial ordering** \leq on \mathbb{R}^n given by $x \leq y$ if $x - y \in K$.
2. Continuous functions on the unit interval $[0, 1]$, with values in the real or complex numbers. This space, denoted by $C([0, 1])$ is a vector space with respect to the pointwise addition of functions and scalar multiplication by constants. It is a normed space with norm defined by the uniform norm $\|f\| := \sup_{x \in [0, 1]} |f(x)|$.
3. The space of square integrable functions L^2 . Let (Ω, P) be a probability space. The Banach space $L^2(\Omega)$ consists of all functions $f : \Omega \rightarrow \mathbb{R}$ such that

$$\int f^2 dP < \infty.$$

Technically, L^2 is a Banach space only if you consider equivalence classes of functions up to a set of measure zero. However, throughout this paper, for brevity, we shall ignore sets of probability zero and write equality between elements of L^2 without prefacing the equality with ‘‘almost everywhere’’. On $L^2(\Omega)$, for all $f, g \in L^2(\Omega)$, we define the inner product by

$$\langle f, g \rangle = \int_{k' \in \Omega} f(k') g(k') dP(k'),$$

or in the case where Ω is a discrete set,

$$\langle f, g \rangle = \sum_{k' \in \Omega} f(k') g(k') P(k'). \quad (32)$$

$L^2(\Omega)$ has a **cone** K consisting of nonnegative functions on Ω . This cone induces a **partial ordering** \leq on $L^2(\Omega)$ given by $x \leq y$ if $x - y \in K$.

If n is the number of distinct degrees of the network then $L^2(\Omega) \cong \mathbb{R}^n$. Here we make the identification of a function $f : \Omega \rightarrow \mathbb{R}$ with the vector $(f(i_1)\sqrt{P(i_1)}, f(i_2)\sqrt{P(i_2)}, \dots, f(i_n)\sqrt{P(i_n)}) \in \mathbb{R}^n$, where $\Omega = \{i_1, \dots, i_n\}$. This identification maps the orthonormal basis $\{\frac{e_i}{\sqrt{P(i)}}\}$, where $e_i(j) = \delta_{i,j}$ is the delta function in $L^2(\Omega)$ with respect to inner product (32), to the orthonormal basis in \mathbb{R}^n with respect to the Euclidean inner product.

B. Vector and Matrix Norms

To make this paper more self contained we give the results about matrix and vector norms used in this paper. All the results not proved below can be found in [18]. We denote by M_n the set of $n \times n$ matrices of real numbers. A matrix $S \in M_n$ is *nonsingular* if it has nonzero determinant. For $A \in M_n$, A^* will denote its transpose. A matrix $A \in M_n$ is *unitary* if $A^*A = I$, where $I \in M_n$ is the identity matrix. Let θ denote the zero vector in \mathbb{R}^n . $\rho(A)$ is the spectral radius of A defined in Section II A.

Definition Let V be a vector space over \mathbb{C} . A function $\|\bullet\| : V \rightarrow \mathbb{R}$ is a *vector norm* if for all $x, y \in V$,

- (1): $\|x\| \geq 0$ Nonnegative
- (2): $\|x\| = 0$ if and only if $x = 0$ Positive
- (3): $\|cx\| = |c|\|x\|$ for all scalars $c \in \mathbb{R}$. Homogeneous
- (4): $\|x + y\| \leq \|x\| + \|y\|$ Triangle inequality

A frequently encountered vector norm is the Euclidean norm (or l_2 norm) on \mathbb{R}^n :

$$\|x\|_2 \equiv (|x_1|^2 + \dots + |x_n|^2)^{\frac{1}{2}}.$$

Given a vector norm such as the Euclidean norm, the following result is easy to verify and of great use in tailoring a vector norm for a specific purpose.

Proposition VII.1 *If $\|\bullet\|$ is a vector norm on \mathbb{R}^n and if $T \in M_n$ is nonsingular, then $\|\bullet\|_T$ defined by $\|x\|_T \equiv \|Tx\|$, $x \in \mathbb{R}^n$, is also a vector norm in \mathbb{R}^n*

Denote a vector $x \in \mathbb{R}^n$ by $(x(1), x(2), \dots, x(n))$. Let \leq be a partial ordering on \mathbb{R}^n defined by $x \leq y$ if and only if $y(k) - x(k) \geq 0$. We show next that the vector norm $\|\bullet\|_T$ defined above has normality constant 1.

Proposition VII.2 Let $T \in M_n$ be nonsingular and $x, y \in \mathbb{R}^n$. If $\theta \leq x \leq y$ then $\|x\|_T \leq \|y\|_T$

Proof By the linearity of T it suffices to prove the claim when only a single component of x, y is nonzero. Hence, assume $0 \leq x(i) \leq y(i)$, for some $i \in \{1, 2, \dots, n\}$, and $x(k) = y(k) = 0$ for $k \neq i$. Let $e_i \in \mathbb{R}^n$ be the unit column vector with 1 in the i^{th} position. Then $\|Tx\|_2 = |x(i)|\|Te_i\|_2$ and $\|Ty\|_2 = |y(i)|\|Te_i\|_2$. Note that $\|Te_i\|_2 \neq 0$ since T is nonsingular. It follows that $\|Tx\|_2 \leq \|Ty\|_2$. ■

We call a function $\|\bullet\| : M_n \rightarrow \mathbb{R}$ a *matrix norm* if for all $A, B \in M_n$ it satisfies the following five axioms:

- (1): $\|A\| \geq 0$ Nonnegative
- (2): $\|A\| = 0$ if and only if $A = 0$ Positive
- (3): $\|cA\| = |c| \|A\|$ for all scalars c Homogeneous
- (4): $\|A + B\| \leq \|A\| + \|B\|$ Triangle Inequality
- (5): $\|AB\| \leq \|A\| \|B\|$ Submultiplicative

Notice that properties 1–4 are identical to the properties of vector norms.

We will need the following result, analogous to Proposition VII.1 for vector norms, that shows that one matrix norm may be transformed into another by a fixed similarity.

Proposition VII.3 If $\|\bullet\|$ is a matrix norm on M_n and if $S \in M_n$ is nonsingular, then $\|A\|_S \equiv \|S^{-1}AS\|$ is a matrix norm.

Associated with each vector norm $\|\bullet\|$ on \mathbb{R}^n is a natural matrix norm $\|\bullet\|$ that is *induced* by $\|\bullet\|$ on M_n .

Definition Let $\|\bullet\|$ be a vector norm on \mathbb{R}^n . Define $\|\bullet\|$ on M_n by

$$\|A\| = \max_{\|x\|=1} \|Ax\|$$

A particularly important matrix norm is the spectral norm.

Definition The spectral norm $\|\bullet\|_2$ is defined on M_n by

$$\|A\|_2 \equiv \max\{\sqrt{\lambda} : \lambda \text{ is an eigenvalue of } A^*A\}.$$

Proposition VII.4 The spectral norm $\|\bullet\|_2$ is induced by the Euclidean vector norm $\|\bullet\|_2$.

Proof It isn't difficult to verify that $\|Ax\|_2^2 = |x^*A^*Ax|$. Since A^*A is a normal matrix we can write $A^*A = U\Lambda U^*$ where U is a unitary matrix and Λ is a diagonal matrix of eigenvalues of A^*A . We can write

$$\begin{aligned} |x^*A^*Ax| &= |y^*\Lambda y| \text{ where } y = Ux \\ &\leq \rho(A^*A)\|y\|_2^2 \\ &= \rho(A^*A)\|x\|_2^2. \end{aligned}$$

It follows that

$$\max_{\|x\|_2=1} \|Ax\| \leq \sqrt{\rho(A^*A)}.$$

Next we show that equality holds. Reordering the eigenvalues in Λ if necessary lets assume that $e_1^*\Lambda e_1 = \rho(A^*A)$, where $e_1 = [1, 0, \dots, 0]^* \in \mathbb{R}^n$. Let $x = Ue_1$. Then $|x^*A^*Ax| = \sqrt{\rho(A^*A)}$. It follows that $\max_{\|x\|_2=1} \|Ax\| = \sqrt{\rho(A^*A)}$ as required. ■

The next result says that the spectral norm is a unitary invariant matrix norm.

Proposition VII.5 $\|UAV\|_2 = \|A\|_2$ for any $A \in M_n$ and any unitary matrices $U, V \in M_n$.

Proof For U and V unitary matrices we have $(UAV)^*UAV = V^*A^*U^*UAV = V^*A^*AV = V^{-1}A^*AV$. The claim then follows since A^*A and $V^{-1}A^*AV$ have the same eigenvalues. ■

An important example of a matrix norm induced by a vector norm is given below.

Proposition VII.6 If $\|\bullet\|$ is an induced matrix norm on M_n and $S \in M_n$ is nonsingular, $\|\bullet\|_S$ is induced by the vector norm $\|\bullet\|_{S^{-1}} \equiv \|S^{-1}x\|$

Proof It isn't difficult to show that

$$\max_{\|x\|=1} \|Ax\|_{S^{-1}} = \max_{x \neq 0} \frac{\|Ax\|_{S^{-1}}}{\|x\|_{S^{-1}}}.$$

Hence we prove that

$$\|A\|_S = \max_{x \neq 0} \frac{\|Ax\|_{S^{-1}}}{\|x\|_{S^{-1}}}.$$

Let $y = S^{-1}x$. We have

$$\begin{aligned} \max_{x \neq 0} \frac{\|Ax\|_{S^{-1}}}{\|x\|_{S^{-1}}} &= \max_{Sy \neq 0} \frac{\|S^{-1}ASy\|}{\|y\|} \\ &= \max_{y \neq 0} \frac{\|S^{-1}ASy\|}{\|y\|} \\ &= \|A\|_S \end{aligned}$$

as required. ■

The following inequality, says that a vector norm is compatible with its induced matrix norm.

Proposition VII.7 If the matrix norm $\|\bullet\|$ is induced by the vector norm $\|\bullet\|$, then

$$\|Ax\| \leq \|A\| \|x\|$$

for all $A \in M_n$ and all $x \in \mathbb{R}^n$

Theorem VII.8 Let $A \in M_n$ be given. If there is a unitary matrix U such that U^*AU is diagonal, then $\|Ax\|_2 \leq \rho(A)\|x\|_2$, where $\rho(A)$ is the spectral radius of A .

Proof By Proposition VII.4 the spectral norm, $\|\bullet\|_2$, is induced from the Euclidean norm $\|\bullet\|_2$. Hence, we can apply Proposition VII.7. By Proposition VII.5, the spectral norm is unitarily invariant. The result follows immediately. ■

Theorem VII.9 *Let $A \in M_n$ be given. If there is a nonsingular matrix S such that $S^{-1}AS$ is diagonal, then $\|Ax\|_{S^{-1}} \leq \rho(A)\|x\|_{S^{-1}}$ where $\rho(A)$ is the spectral radius of A , and the vector norm $\|x\|_{S^{-1}} \equiv \|S^{-1}x\|_2$.*

Proof The vector norm $\|\bullet\|_{S^{-1}}$ induces the matrix norm $\|A\|_S$ by Proposition VII.6. By Proposition VII.7 we have the inequality

$$\|Ax\|_{S^{-1}} \leq \|A\|_S \|x\|_{S^{-1}}.$$

Since $\Lambda = S^{-1}AS$ is a diagonal matrix of eigenvalues of A , $\|A\|_S = \|\Lambda\|_2 = \sqrt{\rho(\Lambda\Lambda^*)} = \rho(A)$ proving the claim. ■

ACKNOWLEDGMENTS

The author would like to thank Paul Loya, Steve Pressé and Eric Bahuaud for helpful discussions.

-
- [1] R. Albert, H. Jeong, and A.-L. Barabási, *Nature*, **401**, 130 (1999).
 - [2] M. Newman, *Phys. Rev. E*, **66**, 016128 (2002).
 - [3] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, *Journal of Theoretical Biology*, **235**, 275 (2005).
 - [4] A. Goltsev, S. Dorogovtsev, and J. Mendes, *Phys. Rev. E*, **78**, 051105 (2008).
 - [5] F. Alcaraz and M. Barber, *Journal of Statistical Physics*, **46**, 435 (1987).
 - [6] J. Yeomans, *Statistical Mechanics of Phase Transitions* (Oxford University Press, New York, 2002).
 - [7] H. Hong and M. Choi, *Phys. Rev. E*, **65**, 026139 (2002).
 - [8] G. Ódor, *Brazilian Journal of Physics*, **33**, 431 (2003).
 - [9] E. Ivashkevich, *Phys. Rev. Lett.*, **76**, 3368 (1996).
 - [10] Y. Liu and A. Dahmen, *Phys. Rev. E*, **79**, 061124 (2009).
 - [11] S. Boccaletti, Y. Latora, M. Moreno, M. Chavez, and U. Hwang, *Phys. Rep.*, **424**, 175 (2006).
 - [12] R. Pastor-Satorras, A. Vázquez, and A. Vespignani, *Phys. Rev. Lett.*, **87**, 258701 (2001).
 - [13] M. Boguna, R. Pastor-Satorras, and A. Vespignani, *Lecture Notes in Physics*, **625**, 127 (2003).
 - [14] R. Leggett and L. Williams, *Journal of mathematical analysis and application*, **76**, 91 (1980).
 - [15] L. Zhandong, Z. Lingling, and L. Shengjia, *Journal for Analysis and its Applications*, **22**, 529 (2003).
 - [16] Z. Zengqin and D. Xinsheng, *Journal for Analysis and its Applications*, **334**, 1426 (2007).
 - [17] M. Krasnosel'skii, *Positive solutions of operator equations* (Noordhoff, Groningen, 1964).
 - [18] R. Horn and C. Johnson, *Matrix Analysis* (Cambridge University Press, New York, 1999).
 - [19] F. Riesz and B. Sz.-Nagy, *Functional analysis* (Dover Publications Inc., New York, 1990) translated from the second French edition by Leo F. Boron, Reprint of the 1955 original.
 - [20] M. G. Kreĭn and M. A. Rutman, *Amer. Math. Soc. Translation*, **1950**, 128 (1950).
 - [21] A. R. Lucas, "A one-parameter family of stationary solutions in the Susceptible-Infected-Susceptible epidemic model," (2010), submitted to *Journal of Mathematical Analysis and Applications*.
 - [22] L. Debnath and P. Mikusinski, *Introduction to Hilbert Spaces with Applications* (Academic Press, New York, 1990).
 - [23] P. Erdős and A. Rényi, *Publicationes Mathematicae*, **6**, 290 (1959).
 - [24] M. Newman, *SIAM Review*, **45**, 167 (2003).
 - [25] A. Vázquez and Y. Moreno, *Phys. Rev. E*, **67**, 015101(R) (2003).
 - [26] M.E.J. Newman, S.H. Strogatz, and D.J. Watts, *Phys. Rev. E*, **64**, 026118 (2001).
 - [27] M.E.J. Newman, *Phys. Rev. Lett.*, **89**, 208701 (2002).