Epidemic Outbreaks in Networks with Equitable or Almost-Equitable Partitions

Stefano Bonaccorsi  
*Mathematics Department, University of Trento, via Sommarive 14, 38123 Povo (Trento), Italy*

Stefania Ottaviano  
*Mathematics Department, University of Trento, via Sommarive 14, 38123 Povo (Trento), Italy and CREATE-NET, Via alla Cascata 56/d, 38123 Povo (Trento), Italy*

Delio Mugnolo  
*Lehrgebiet Analysis, Fakultät für Mathematik und Informatik, FernUniversität Hagen, 58084 Hagen, Germany*

Francesco De Pellegrini  
*CREATE-NET, Via alla Cascata 56/d, 38123, Trento, Italy*  
*(Dated: December 22, 2014)*

We study the diffusion of epidemics on networks that are partitioned into local communities. The gross structure of hierarchical networks of this kind can be described by a *quotient graph*. The rationale of this approach is that individuals infect fellows with higher probability than individuals in other communities, so that small communities may approximately reach equilibrium at a time when the network is still evolving as a whole. In community models the nodal infection probability is thus expected to depend mainly on the interaction of a few, large interconnected clusters. In this work, we describe the epidemic process as a continuous-time susceptible—infected—susceptible (SIS) model. Upon performing a first-order mean-field approximation it is possible to describe the nodal infection probability by a lower-dimensional dynamical system associated with the evolution of the process on the quotient graph. Our main result is that the spectral radius of this much smaller quotient graph (which only captures the macroscopic structure of the community network) is all we need to know in order to decide whether the overall healthy-state defines a stable or an unstable equilibrium. Our investigations are based on the graph-theoretical notion of *equitable partition* and of its recent and rather flexible generalization, that of *quasi-equitable partition*.

PACS numbers: 89.75.Hc, 89.20.Ha

Keywords: SIS model, hierarchical networks, graph spectra, equitable partitions

I. INTRODUCTION

In epidemic models population is frequently partitioned into different communities (also called in the literature households, clusters or subgraphs) where each group could e.g. be characterized by a common environment or preferred relationship. The spread of epidemics through the whole population is then often modeled in terms of the adjacency matrix of (local and global) contact network.

Models where the epidemic process within community is faster compared to the rate at which it spreads across communities, have been studied in literature [1–3].

In this work, we study the diffusion of epidemics over an undirected graph $G = (V,E)$ with edge set $E$ and node set $V$. The order of $G$, denoted here $N$, is the cardinality of $V$, whereas the size of $G$ is the cardinality of $E$, denoted here $M$. The connectivity of the graph $G$ is conveniently expressed by the $N \times N$ adjacency matrix $A$.

Our network is partitioned into $n$ communities: they are represented by a node set partition $\pi = \{V_1,\ldots,V_n\}$, i.e., a sequence of mutually disjoint nonempty subsets of $V$, called cells, whose union is $V$.

The epidemic model adopted in the rest of the paper is a continuous-time Markovian SIS model: in the SIS model a node can be repeatedly infected, recovered and yet infected again. The state of each node is thus described by a stochastic process of Bernoulli type $X_i(t)$, where at time $t$ we set $X_i(t) = 0$ if $i$ is healthy and $X_i(t) = 1$ if $i$ is infected. This process is described by the probability of being infected, $p_i(t) = P(X_i(t) = 1)$ which also implies that the probability of being healthy is $P(X_i(t) = 0) = 1 - p_i(t)$. Each node becomes infected following a Poisson process with rate $\beta$. Also, $i$ recovers following a Poisson process with rate $\delta$. We further assume that infection and curing processes occurring at different nodes are independent [4]. The ratio $\tau = \beta/\delta$ is called the effective spreading rate.

Recently, also non Markovian types of epidemic spread were introduced in the literature, by choosing other than the exponential interaction time for infection and/or cur-
ing (see [5, 6]), or by incorporating dependencies between the infection and curing process.

However, this seems to complicate the analysis considerably.

Compared to the homogeneous case where the infection rate is the same for all pairs of nodes, in our model we consider two infection rates: the intra-community infection rate $\beta$ for infecting individuals in the same community and the inter-community infection rate $\varepsilon \beta$, i.e., the rate at which individuals among different communities get infected.

The customary physical interpretation is that when $0 < \varepsilon < 1$ contacts across communities occur at a much smaller rate. The case $\varepsilon = 1$ corresponds to the case of a network not partitioned in communities, i.e., the homogeneous case.

As described in [4, 7], the SIS process developing on a graph with $N$ nodes is modeled as a continuous-time Markov process with $2^N$ states. The dynamics of the nodal infection probability is obtained by the corresponding Kolmogorov differential equations. But, the resulting dynamical system consists of $2^N$ linear differential equations, not a viable solution for large networks. An alternative first order mean-field approximation is the NIMFA approach proposed by Van Mieghem et al. in [4, 8, 9].

NIMFA replaces the original $2^N$ linear differential equations by $N$ non-linear differential equations; they represent the time-change of the infection probability of a node. As typical in first order approximations of SIS dynamics, the only approximation required by NIMFA is that the infectious state of two nodes in the network are uncorrelated, i.e., $E[X_i(t)X_j(t)] = E[X_i(t)]E[X_j(t)]$.

### A. Epidemic threshold

For a network with finite size $N$, the exact SIS Markov process will always converge towards the absorbing zero-state, in which the virus has disappeared from the network. This absorbing state is the unique stationary state of the SIS Markov chain.

Nevertheless, in the critical case $\tau > \tau_c$, prior to absorption the SIS process approaches what appears to be a quasi-stationary distribution that is different from the disease-free equilibrium.

This distribution can also be regarded as the limiting conditional distribution, useful in representing the long-term behavior of the process “that in some sense terminates, but appears to be stationary over any reasonable time scale” [10]. After a mean-field approximation a limiting occupancy probability appears as the second solution of the non-linear system, which exists, apart from the zero-vector solution that reflects the absorbing state.

This non-zero solution is the metastable state described in [4]. It is interpreted as the nodes’ ergodic infection probability of a modified Markovian SIS model (see [3]), where either the absorbing state is removed or where a vanishing self-infection probability is imposed.

This long-term behavior also justifies the presence of a characteristic epidemic threshold $\tau_c$, mentioned by many authors, (e.g.,[11, 12]). The literature is mainly concerned with establishing the value of this threshold, i.e., the critical ratio $\tau_c$ above which infections become endemic. It is a key parameter behind immunization strategies related to the network protection against viral infection.

The delicate aspect is that in the exact Markovian SIS model on a finite graph, the epidemic threshold is undefined, since the system is expected to reach surely the unique absorbing state. But, for large $N$, numerical simulations have revealed that, when $\tau > \tau_c$, the overall-healthy state is only reached after an unrealistically long time [14, 15]. Hence, the exact steady-state is hardly ever reached in real networks, while the metastable state reflects the observed viral behavior fairly well [8]. Below the critical threshold ($\tau < \tau_c$), the infection vanishes exponentially fast in time. Results about the convergence time towards the absorbing state can be found also in [13].

The relation of the epidemic threshold with the spectral radius $\lambda_1(A)$ of the adjacency matrix $A$ of the graph can be found in [4, 16]. In NIMFA, the SIS epidemic threshold for the effective infection rate $\tau$ is $\tau_c^{(1)}$ where $\tau_c^{(1)} = \frac{1}{\lambda_1(A)}$.

However, the state of two nodes was recently proved to be not negatively correlated in the exact Markovian SIS model [13]. Due to the assumption of independence, this proves that NIMFA always yields an upper bound for the probability of infection of each node, as well as a lower bound for the epidemic threshold, i.e., $\tau_c = \frac{\alpha_{c}^{(1)}}{1}$ with $\alpha \geq 1$.

Note that, from an applicative point of view, it is of fundamental importance to understand for which networks with $N$ nodes NIMFA is worst, in the sense that $\alpha_{c}^{(1)}$ is largest. Further efforts are being made to quantify satisfactorily the accuracy of NIMFA.

### B. Outline and main results

As already observed in [3], the presence of communities generates a strong mixing effect at local level (e.g., the rate of infection inside a community tends to be homogeneous) as opposed to the much lower speed of mixing (i.e., much larger inhomogeneity) within the whole population.

More precisely, in [1] a complete graph represents the internal structure of a community. Such assumption appears natural for small community sizes, for example, because the members of a small community usually know each other (e.g., let them be friends in a certain club, members of the same family, employees in the same department, etc.). Moreover, in [3], given two different communities, either all or none of their nodes are mutually linked.

In this work, instead, we consider sparser community
structures. First of all, we require that all nodes belonging to the same community have the same internal degree, i.e. each node has the same number of neighbors belonging to the same community: actually, we ask that the subgraph of $G$ induced by $V_i$ is regular for all $i$. Furthermore, for any two adjacent communities $V_i, V_j$ we require each node in $V_i$ has the same number of adjacent nodes belonging to $V_j$. The macroscopic structure of such a network can be described by the quotient graph $G/\pi$, an oriented graph (possibly) featuring loops and multiple edges. The nodes of the quotient graph are the communities in $G$. In Sec. III after defining the so-called quotient matrix $Q$ that is related to $G/\pi$ and that has the same spectral radius of the adjacency matrix $A$, we find a novel expression for the bound on the epidemic threshold $\tau_2^{(1)}$. Thus we have a relation between the epidemic threshold and the spectral properties of the corresponding quotient matrix. In Sec. IV we study the evolution of the epidemics inside the population, described in terms of the time-dependent infection probabilities, via the matrix $Q$, reducing the system of $N$ differential equations, where $n$ is the number of communities in the network.

In Sec. IV we finally extend our development to the case of almost equitable partitions. We consider the special case when an almost equitable partition is obtained by a perturbation of an equitable one, i.e., by adding/deleting a certain set of edges from an equitable partition. Thus, we relax the assumption that the internal structure of each community is regular. And, in this case, too, we obtain a lower bound for the epidemic threshold as a function of network metrics.

II. THE EPIDEMIC MODEL

In this section we first describe the epidemic model for graphs with equitable partitions. The original definition of equitable partition is due to Schwenk [18].

Definition 1. Let $G = (V, E)$ be a graph. The partition $\pi = \{V_1, ..., V_n\}$ of the node set $V$ is called equitable if $\forall i, j \in I$, there is an integer $d_{ij}$ such that

$$d_{ij} = \deg(v, V_j) := \# \{e \in E : e = \{v, w\}, w \in V_j\}.$$ independent of $v \in V_i$.

We shall identify the set of all nodes in $V_i$ as the $i$-th community of the whole population. In particular, each $V_i$ induces a subgraph of $G$ that is necessarily regular, but need not bear any symmetry. Further, although not every value of $d_{ij}$ is allowed, it is not required that $d_{ij} = k_j$, i.e., it is not mandatory that a node $v \in V_i$ is either connected to all or none of the nodes in $V_j$, as next remark makes precise.

Remark 1. We use the notation of lcm and gcd to denote the least common multiple and greatest common divisor, respectively. We can observe that the partition of a graph is equitable if and only if

$$d_{ij} = \alpha \frac{\text{lcm}(k_i, k_j)}{k_i}$$

where $\alpha$ is an integer satisfying $1 \leq \alpha \leq \gcd(k_i, k_j)$, $k_i$ the number of nodes in $V_i$, $\forall i = 1, ..., n$.

An equitable partition generates the quotient graph $G/\pi$, which is a multigraph with cells as vertices and $d_{ij}$ edges between $V_i$ and $V_j$. For the sake of explanation, in the following we will identify $G/\pi$ in a (simple) graph having the same cells vertex set, and where an edge

![FIG. 1. A sample graph with equitable partition $V = \{\{v_1\}, \{v_2, v_3\}, \{v_3, v_4, v_5, v_6\}, \{v_7, v_8, v_9, v_{10}, v_{11}, v_{12}, v_{13}\}\}$.](image-url)
exists between communities \( V_i \) and \( V_j \) if at least one exists in the original multigraph. \( B \) is hence denoting the adjacency matrix of the graph \( G/\pi \).

**Remark 2.** In \( \mathbb{H} \) the special case when each community has a clique structure – i.e., \( d_{ij} = k_i - 1 \) – has been considered. Moreover all nodes belonging to two linked communities are connected, \( d_{ij} = k_j \). By considering the theory of equitable partition, we generalize the cited work and consider any kind of regular graph to represent the internal structure of a community. Moreover, unlike before, if two communities \( i \) and \( j \) are connected, each node in community \( i \) is connected only with \( d_{ij} \) of the nodes in community \( j \).

### A. Example

Assume for instance that we are given the adjacency matrix \( B \) of the quotient graph and that for any \( i, j \in I \), \( b_{ij} \neq 0 \) implies \( d_{ij} = k_j \), i.e., each node in \( V_i \) is connected with every node inside \( V_j \). In this case we can explicitly write the adjacency matrix \( A \) in a block form. Let \( C_{V_i} = (c_{ij})_{k_i \times k_i} \), be the adjacency matrix of the subgraph induced by \( V_i \) and \( J_{k_i \times k_j} \) is a all ones \( k_i \times k_j \) matrix; then

\[
A = \begin{bmatrix}
C_{V_i} & \varepsilon J_{k_i \times k_j} b_{12} & \cdots & \varepsilon J_{k_i \times k_j} b_{1n} \\
\varepsilon J_{k_j \times k_i} b_{21} & C_{V_2} & \cdots & \varepsilon J_{k_j \times k_i} b_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
\varepsilon J_{k_j \times k_i} b_{n1} & \varepsilon J_{k_j \times k_i} b_{n2} & \cdots & C_{V_n}
\end{bmatrix}
\]  

### B. The mathematical model

The process of diffusion of epidemics on a graph is described by NIMFA with the time-change of probability \( p_i \) that a node \( i \) is infected. Thus, node \( i \) obeys a following differential equation \( \mathbb{H} \)

\[
\dot{p_i}(t) = (1 - p_i(t)) \beta \left( \sum_{j=1}^{N} a_{ij} p_j(t) \right) - \delta p_i(t), \quad i = 1, \ldots, N
\]

In \( \mathbb{H} \) the time-derivative of the infection probability of node \( i \) consists of two competing processes:

1. While healthy with probability \( 1 - p_i(t) \), all infected neighbors, whose average number is \( \sum_{j=1}^{N} a_{ij} p_j(t) \), try to infect node \( i \) at rate \( \beta \).
2. While node \( i \) is infected with probability \( p_i(t) \), it is cured at rate \( \delta \).

We have the following matrix representation

\[
\frac{dP(t)}{dt} = \beta AP(t) - \text{diag}(p_i(t)) (\beta AP(t) + \delta u)
\]

where \( P(t) = [p_1(t) p_2(t) \ldots p_N(t)]^T \), \( \text{diag}(p_i(t)) \) is the diagonal matrix with elements \( p_1(t), p_2(t), p_N(t) \) and \( u \) is the all-one vector. From \( \mathbb{H} \), considering \( P(t) = \text{diag}(p_i(t)) u \), we can write

\[
\frac{dP(t)}{dt} = \beta AP(t) - \delta \text{diag}(p_i(t)) (\beta AP(t) + \delta u)
\]

\[
= (\beta A - \delta I) P(t) - \beta \text{diag}(p_i(t)) AP(t)
\]

Let us denote with \( P_\infty \) an equilibrium point of \( \mathbb{H} \), that is, a solution of the equation

\[
P_\infty = \frac{\beta}{\delta} (I - \text{diag}(p_i(\infty))) AP_\infty.
\]

As recalled previously, from equation \( \mathbb{H} \) the epidemic threshold can be expressed in terms of the effective infection rate \( \tau = \beta / \delta \)

\[
\tau_c^{(1)} = \frac{1}{\lambda_1(A)}
\]

where the superscript \( (1) \) in \( \tau_c^{(1)} \) refers to the first-order mean-field approximation (see \( \mathbb{H} \)).

Hence, the value \( \tau_c^{(1)} = \frac{1}{\lambda_1(A)} \) is a critical value separating the absorbing phase from the endemic phase. This can be seen by studying stability of the zero steady-state solution.

**Theorem 1 (Epidemic Threshold).** Let \( P_0 = 0 \) be the vector of all zero components.

1. If \( \tau < 1 / \lambda_1(A) \), then \( P_0 \) is a globally asymptotically stable equilibrium point of the system \( \mathbb{H} \).
2. If \( \tau > 1 / \lambda_1(A) \), then \( P_0 \) is an unstable equilibrium point of the system \( \mathbb{H} \).

**Proof.** Let us denote by \( f \) the right hand side of \( \mathbb{H} \), i.e., \( \mathbb{H} \) can be re-written as a vector-valued differential equation

\[
\dot{P} = f(P).
\]

where \( f : [0, 1]^N \rightarrow [0, 1]^N \) is a smooth \( C^\infty \) function. One can easily check that \( P_0 \) is an equilibrium point of the system \( \mathbb{H} \). By the principle of linearized stability, we know that topologically the local behavior of the non-linear system \( \mathbb{H} \), near an equilibrium point, is determined by the behavior near the origin \( P_0 \) of the system of linear ODEs \( \mathbb{H} \):

\[
\dot{P} = D_f P,
\]

where \( D_f = Df(P_0) \) is the Jacobian matrix of \( f \) at \( P_0 \). From \( \mathbb{H} \) we have

\[
(Df(P_0))_{ij} = \begin{cases}
\beta a_{ij} & i \neq j \\
-\delta & i = j
\end{cases}
\]

We have that \( D_f = \beta A - \delta I \). Since the adjacency matrix \( A \) is real and symmetric its eigenvalues are real, hence also the eigenvalues of \( H \) are real and of the form

\[
\lambda_i(D_f) = \beta \lambda_i(A) - \delta.
\]
in particular we have
\[ \lambda_1(D_f) = \beta \lambda_1(A) - \delta. \]

We consider now the case 1). From the Perron–Frobenius Theorem \[21\] \( \lambda_1(A) \) is positive, hence \( \beta/\delta < 1/\lambda_1(A) \) implies \( \lambda_1(D_f) < 0 \) as well as all the other eigenvalues \( \lambda_i(D_f) \). Since all eigenvalues of \( J_f \) are negative, then \( P_0 \) is a globally asymptotically stable equilibrium point for the Lyapunov’s Linearization (or First) Method.

On other hands, in the case 2), when \( \beta/\delta > 1/\lambda_1(A) \) we have \( \lambda_1(D_f) > 0 \) and \( P_0 \) is unstable. \( \square \)

Theorem 1 states essentially that \( \tau\epsilon^{(1)} = 1/\lambda_1(A) \) is a critical value below which, the infection probabilities of each nodes will go even closer to zero over time and the epidemics will die out. Conversely above the epidemic threshold the infection probabilities diverges from zero and another nonzero solution takes over as the actual steady-state—the metastable solution.

III. THE QUOTIENT MATRIX

We search for a smaller matrix \( Q \) that contains the relevant information for the evolution of the system, in order to express the dynamics of the SIS model via an equivalent reduced dynamics based on such matrix \( Q \).

A such matrix is the quotient matrix of the equitable partition. In Theorem 2 we will show that matrix \( Q \) shares with \( A \) some properties of the spectrum: remarkably they have the same spectral radius. As a consequence, we can compute its spectral radius in order to estimate the epidemic threshold.

The quotient matrix \( Q \) can be defined for any equitable partition: in view of the internal structure of a graph with an equitable partition, it is natural to consider the cell-wise average value of a function on the node set, that is to say the projection of the node space onto the subspace of cell-wise constant functions. This projector is the matrix \( S \) defined in the following.

**Definition 2.** Let \( \pi = \{V_i, i = 1, \ldots, n\} \) be an equitable partition, let us consider the \( n \times N \) matrix \( S = (s_{iv}) \), where

\[
   s_{iv} = \begin{cases} 
   \frac{1}{\sqrt{|V_i|}} & v \in V_i \\
   0 & \text{otherwise}
   \end{cases}
\]

The quotient matrix of \( G \) (with respect to the given partition) is

\[
   Q := SAS^T.
\]

Observe that by definition \( SS^T = I \) and the equivalently introduced in an algebraic way via \( Q \) \[22\].

In the case of the example in Section II.A the form of \( Q \) is rather simple:

\[
   q_{ii} = \sum_{h=1}^{k_i} \frac{1}{\sqrt{k_i}} \sum_{k=1}^{k_i} (C_{Vi})_{kh} = \frac{1}{k_i} \sum_{h,k=1}^{k_i} (C_{Vi})_{kh}
\]

and

\[
   q_{ij} = \sum_{h=1}^{k_j} \frac{1}{\sqrt{k_j}} \sqrt{k_i k_j} \epsilon b_{ij} = \sqrt{k_i k_j} \epsilon b_{ij}.
\]

Hence we obtain that

\[
   Q = \text{diag}(d_{ii}) + (\sqrt{k_i k_j} \epsilon b_{ij})_{i,j=1,\ldots,n}, \quad (7)
\]

where \( d_{ii} = \frac{1}{k_i} \sum_{h,k=1}^{k_i} (C_{Vi})_{kh} \) is the internal degree of the subgraph induced by \( V_i \).

In the general case, the expression for \( Q \) is the following

\[
   Q = \text{diag}(d_{ii}) + (\sqrt{d_{ii} d_{jj} \epsilon b_{ij}})_{i,j=1,\ldots,n}. \quad (8)
\]

The following theorem shows that there exists a close relationship between the spectral properties of \( Q \) and that of \( A \).

**Theorem 2.** Let \( \pi = \{V_i, i = 1, \ldots, n\} \) be an equitable partition. Each eigenvalue of the reduced matrix \( Q \) is an eigenvalue of \( A \), too, and furthermore their spectral radii coincide, i.e.,

\[
   \sigma(Q) \subseteq \sigma(A) \quad \text{and} \quad \lambda_1(Q) = \lambda_1(A)
\]

**Proof.** Being the order of \( Q \) smaller of that of \( A \), the fact that \( \sigma(Q) \subseteq \sigma(A) \) is immediate from the following result, that has been proved in \[22\]: for any vector \( x \) of size \( n \) and scalar \( \lambda \), \( Qx = \lambda x \) if and only if \( A(S^T x) = \lambda(S^T x) \).

The second statement follows from the Perron–Frobenius Theorem for irreducible and non negative matrices. First, note that \( A, Q \geq 0 \) are irreducible, i.e., they possess a unique communicating class of nodes \[23\] in fact, since \( G \) is by assumption a connected graph, the corresponding adjacency matrix \( A \) is irreducible. Now, \( Q \) is irreducible as well, since it has the same zero/nonzero pattern as the adjacency matrix \( A \).

From the Perron–Frobenius Theorem we know that the (unique, up to re-normalization) strictly positive eigenvector of \( Q \) is the Perron vector \[21\]. Now let \( x \) be the Perron vector of \( A \) corresponding to \( \lambda_1(Q) \), by the result in \[22\] cited above, \( S^T x \) is an eigenvector of \( A \) corresponding to the eigenvalue \( \lambda = \lambda_1(Q) \), and it is strictly positive too.

Since \( A \) is a non negative and irreducible matrix, \( S^T x \) is the unique Perron vector of \( A \), hence \( \lambda = \lambda_1(Q) \) must be also the spectral radius of \( A \). \( \square \)

In the framework of this work, Theorem 2 shows that the macroscopic network structure is encoded in the connectivity of a quotient graph and the epidemic threshold \( \tau\epsilon^{(1)} \) can be expressed in terms of the spectral radius of \( Q \). Since the order of \( Q \) is smaller than the order of \( A \), this can provide a computational advantage.
Remark 3. Formula \( \mathbb{E} \) suggests that the internal structure of each community is not relevant for the calculus of the epidemic threshold \( \tau^{(1)}_c \) which seems to be determined by the internal degree \( d_i \) of each community instead.

### A. Lower bound for \( \tau^{(1)}_c \)

We can write \( Q = D + \tilde{B} \), where \( D = \text{diag}(d_i) \) and \( \tilde{B} = (\sqrt{d_i d_j} \varepsilon b_{ij})_{i,j=1,...,n} \). Now let \( z \) be the unique normalized eigenvector associated with the maximal eigenvalue of \( \tilde{B} \); by the Collatz–Wielandt formula \( \mathbb{E} \) we have, for any \( i = 1, \ldots, n \)

\[
\lambda_1(Q) \leq \frac{(Qz)_i}{z_i} \leq \max_{1 \leq i \leq n} \lambda_1(C_{V_i}) + \lambda_1(\tilde{B}) \quad (9)
\]

Equality can be attained in \( \mathbb{E} \). Consider for instance the graph \( G = (V; E) \) described by the adjacency matrix \( A \) in Fig. 1 where, further, that we require all communities \( V_i \) have the same number of elements \( k \) and same internal degree \( d_i = d \). In this case \( Q = \text{Id}_n + \tilde{B} \), where \( \tilde{B} := (k \varepsilon b_{ij})_{i,j=1,...,n} \), and hence

\[
\lambda_1(Q) = d + k \varepsilon \lambda_1(B),
\]

which gives the exact value of \( \lambda_1(A) \) as well, since we know from Corollary \( \mathbb{G} \) that \( \lambda_1(Q) = \lambda_1(A) \).

In general, from estimate \( \mathbb{F} \) we have a lower bound for the epidemic threshold \( \mathbb{F} \), as

\[
\tau^{(1)}_c = 1/\lambda_1(A) = 1/\lambda_1(Q)
\]

and hence

\[
\tau^{(1)}_c \geq \tau^* = \min_{i} \frac{1}{\lambda_1(C_{V_i}) + \lambda_1(B)}.
\]

In applications, when designing or controlling a network, this value can be adopted to determine a safety region \{\( \tau \leq \tau^* \} \) for the effective spreading rate that guarantees the extinction of epidemics.

### IV. INFECTION DYNAMICS FOR EQUITABLE PARTITIONS

In this section we show that the matrix \( Q \) can be used in order to express the epidemic dynamics introduced in \( \mathbb{F} \). The advantage is that while \( A \) is an \( N \times N \) matrix, \( Q \) is an \( n \times n \) matrix where \( n \leq N \); this allows us to describe the dynamics of the infection probability by a system of \( n \) differential equations instead of \( N \) involving the matrix \( Q \).

By internal symmetries, all nodes inside community \( j \) have the same infection probability \( p_j(t) \) at time \( t \) (in fact the internal structure of a community is a regular graph). Thus, we can rewrite \( \mathbb{H} \) to express the infection probability of a node in the \( j \)-th community:

\[
\frac{dp_j(t)}{dt} = \beta(1 - p_j(t)) \sum_{m=1}^{n} \varepsilon b_{jm} d_{jm} p_m(t)
\]

\[
+ \beta d_j(1 - p_j(t)) p_j(t) - \delta p_j(t) \quad j = 1, \ldots, n (10)
\]

Hence, based on Thm. 2.1 in \( \mathbb{F} \), we observe that

\[
q_{ij} = ((k_j / k_i)^{1/2} d_{ji}, \quad (11)
\]

This relation in our case brings

\[
d_{jm} = \left( \frac{k_j}{k_m} \right)^{-1/2} q_{mj} = \left( \frac{k_j}{k_m} \right)^{-1/2} q_{jm} / \varepsilon,
\]

where the last inequality holds because \( Q \) is symmetric. We can so rewrite \( \mathbb{F} \) as

\[
\frac{dp_j(t)}{dt} = \beta(1 - p_j(t)) \sum_{m=1, m \neq j}^{n} \left( \frac{k_j}{k_m} \right)^{-1/2} q_{jm} p_m(t)
\]

\[
+ \beta q_{jj}(1 - p_j(t)) p_j(t) - \delta p_j(t) \quad j = 1, \ldots, n (11)
\]

where \( q_{jj} = d_j = \lambda_1(C_{V_j}) \).

Now, by comparing \( \mathbb{H} \) with \( \mathbb{F} \), we obtain the expression of the dynamics in matrix form

\[
\frac{dP(t)}{dt} = \beta (\text{Id}_n - \text{diag}(P(t)) \tilde{Q}) (\tilde{Q}P(t) - \delta P(t)) \quad (12)
\]

where \( \tilde{Q} = \text{diag} \left( \frac{1}{\sqrt{k_j}} \right) Q \text{diag}(\sqrt{k_j}) \). It is immediate to observe that \( \sigma(Q) = \sigma(\tilde{Q}) \): the whole dynamics of infection probabilities is expressed using a reduced ODE system involving only \( n \) differential equations, instead of \( N \). The reduced system is written using the quotient matrix \( Q \). This allow to have some computational advantage in order to calculate e.g. the steady-state infection probability.

Remark 4. The framework of quotient graphs is useful for generalizing the NIMFA setting to the case of community network structure and reduces to NIMFA when \( k_j = 1 \) for all \( j = 1, \ldots, n \), in this case the matrix \( A \) in \( \mathbb{F} \) becomes the usual adjacency matrix with \( \varepsilon = 1 \).

### A. Example

In Fig. 1 we provide an example of a graph which has an equitable partition with respect to the communities \( V_1 = \{v_1\}, V_2 = \{v_2, v_3\}, V_3 = \{v_4, v_5, v_6\}, V_4 = \{v_7, v_8, v_9, v_{10}, v_{11}, v_{12}, v_{13}\} \).

The corresponding quotient matrix reads

\[
Q = \begin{bmatrix}
0 & \varepsilon \sqrt{2} & \varepsilon 2 & 0 \\
\varepsilon \sqrt{2} & 1 & \varepsilon \sqrt{3} & \varepsilon \sqrt{3} \\
\varepsilon 2 & \varepsilon \sqrt{3} & 2 & 0 \\
0 & \varepsilon \sqrt{3} & 3 & 0
\end{bmatrix}
\]
so that the dynamical system in (12) can be explicitly written

$$\frac{dP(t)}{dt} = (\text{Id}_n - \text{diag}(P)) \text{diag}(s_j)Q \text{diag}(1/s_j)P - \delta P,$$

where $s_j$ are the entries of the vector $s = (1, \sqrt{2}, 2, \sqrt{6})$.

In Figures 2 and 3 we provide graphs for the dynamics of the reduced ODE system (12) for the equitable graph described in Fig. 1 versus the averaged sample paths resulting from a discrete event simulation. The discrete events simulator is based on the generation of independent Poisson processes for both the infection of healthy nodes and the recovery of infected ones. We observe in the reported figures that the model predicts well the dynamics of the infection probabilities over time. Moreover, as already stated in the Introduction, NIMFA provides an upper bound to the dynamics of the nodal infection probabilities. Also, in Fig. 4 we observe that the dynamics for the communities that are initially healthy is characterized by a unique maximum for the nodal infection probability, which decreases afterwards. The communities initially infected, conversely, do show a monotone decrease of the infection probability.

B. Meta-stable state

We focus now on the steady-state $P_\infty = (p_{j\infty})_{j=1,...,n}$. From (11) we obtain

$$\beta(1 - p_{j\infty}) \sum_{m=1}^n \left( \frac{k_j}{k_m} \right)^{-1/2} q_{jm} p_{m\infty} - \delta p_{j\infty} = 0,$$

from where it follows

$$p_{j\infty} = \frac{\beta \sum_{m=1}^n \left( \frac{k_j}{k_m} \right)^{-1/2} q_{jm} p_{m\infty}}{\beta \sum_{m=1}^n \left( \frac{k_j}{k_m} \right)^{-1/2} q_{jm} p_{m\infty} + \delta} = 1 - \frac{1}{1 + \tau \sum_{m=1}^n \left( \frac{k_j}{k_m} \right)^{-1/2} q_{jm} p_{m\infty}}. \quad (13)$$

In the special case where each community $j$ has the same number of elements $k_j = k$ for all $1,...,n$ and all nodes of two connected community are linked (in this case the matrix $Q$ has the form in (14)), we can rewrite (13) as

$$p_{j\infty} = 1 - \frac{1}{1 + \tau \sum_{m=1}^n q_{jm} p_{m\infty}} = 1 - \frac{1}{1 + \tau (d_{jj} + c_j \varepsilon k) - \tau \sum_{m=1}^n q_{jm} (1 - p_{m\infty})} \quad (14)$$

which implies that the steady-state infection probability of any node $j$ is bounded by

$$0 \leq p_{j\infty} \leq 1 - \frac{1}{1 + \tau (d_{jj} + c_j \varepsilon k)}. \quad (15)$$
where \( c_j \) is the number of communities that are linked with community \( j \); the inequality holds because \( p_{j\infty} \in [0,1] \) for all \( j \).

By introducing \( 1 - p_{j\infty} = \frac{1}{1 + \tau \sum_{i=1}^{m} q_{ij} p_{i\infty}} \) in (14), we can express \( p_{j\infty} \) as a continued fraction iterating the formula
\[
x_{j,s+1} = f(x_{j,s}, \ldots, x_{n,s}) \\
= 1 - \frac{1}{1 + \tau (d_{jj} + c_j \epsilon k) - \tau \sum_{m=1}^{n} q_{jm} (1 - x_{m,s})},
\]
see also (13).

Even after limited number of iterations we obtain a good approximation of \( p_{j\infty} \), even if the accuracy seems worst around \( \tau = t \). The numerical estimation can be used to improve the bound in (14).

In the general case (14) writes
\[
p_{j\infty} = 1 - \frac{1}{1 + \tau g(P)},
\]
where we define
\[
g(P) := \left ( d_{jj} + \epsilon \sum_{m=1}^{n} \left ( \frac{k_j}{k_m} \right )^{-1/2} \sqrt{d_{jm} d_{mj}} \right ) \\
- \sum_{m=1}^{n} \left ( \frac{k_j}{k_m} \right )^{-1/2} q_{jm} (1 - p_{m\infty}).
\]

Since \( p_{j\infty} \approx 1 - (1/\tau) \left ( d_{jj} + \epsilon \sum_{m=1}^{n} \sqrt{d_{jm} d_{mj}} \right ) \), we estimate the steady-state fraction of infected nodes \( y_{\infty}(\tau) = (1/n) \sum_{i=1}^{n} p_{i\infty}(\tau) \) by
\[
y_{\infty}(\tau) \approx 1 - \frac{1}{\tau n} \sum_{j=1}^{n} d_{jj} + \epsilon \sum_{m=1}^{n} \sqrt{d_{jm} d_{mj}}.
\]

V. ALMOST EQUITABLE PARTITIONS

In this section we consider a graph whose partition of node set is almost equitable: this assumption let us relax the quest for regularity of the internal community structure, as implied by the definition of an equitable partition.

Ideally we can think of a network \( \tilde{G} \) whose node set has an almost equitable partition as a network \( G \) with equitable partition to which have been added or removed links between nodes in one or more communities. However, in order to obtain a lower bound on the epidemic threshold, useful in determining a safety region for the extinction of epidemics, it is appropriate to assume that links are added only. In fact, by link addition, the spectral radius increases, and this let us improve the lower bound on the epidemic threshold based on the NIMFA approximation.

**Definition 3.** The partition \( \pi = \{V_1, \ldots, V_n\} \) is called almost equitable if \( \forall i, j \in I \) with \( i \neq j \), there is an integer \( d_{ij} \) such that for all \( v \in V_i \), it holds
\[
d_{ij} = \text{deg}(v, V_j) := \# \{ e \in E : e = \{v, w\}, w \in V_j \}
\]

independently of \( v \in V_i \).

It is evident that the difference between equitable and almost equitable partitions is that in the former case, we shall have a regular structure of the subgraph \( G_i \) of \( G \) induced by \( V_i \), while the latter does not require any structural condition inside the community itself.

Let us consider two graphs \( G = (V,E) \) and \( \tilde{G} = (V, \tilde{E}) \) with the same partition \( \{V_1, \ldots, V_n\} \), but different edge sets \( E \subsetneq \tilde{E} \), and assume \( G \) to have an equitable partition but \( \tilde{G} \) to have merely an almost equitable partition. Then if \( \tilde{A} \) and \( A \) are the adjacency matrices of \( \tilde{G} \) and \( G \) respectively it holds
\[
\tilde{A} = A + R,
\]
where \( R = \text{diag}(R_1, \ldots, R_n) \); the dimension of \( R_i \) is \( k_i \times k_i \) for \( i = 1, \ldots, n \), as before \( k_i \) is the order of \( G_i \) and \( n \) is the number of the communities.

The formula of Collatz–Wielandt can be applied to \( \tilde{A} = A + R \) and then yields
\[
\lambda_1(\tilde{A}) \leq \lambda_1(A) + \lambda_1(R).
\]

In the sequel, we try to give a more explicit formulation of the right hand side of (16) involving the number of edges added.

![Comparison of the bound and the spectral radius for a 40-communities network. Each community has \( k = 25 \) nodes, whose internal structure is a initially a ring; the perurbation graph is obtained by adding in each of them the same increasing number of links. The spectral radius of the adjacency matrix \( \tilde{A} \) is compared to the upper bound as a function of the links added in each community.](image)
equitable for $G$ and almost equitable for $\tilde{G}$. Let $E \subset \tilde{E}$ with

$$\tilde{E} \setminus E = \bigcup_{i=1}^{n}(\tilde{E}_i \setminus E_i)$$

(i.e., the edge sets can only differ within cells) and denote by $R$ the adjacency matrix corresponding to $\tilde{E} \setminus E$. Finally, let us denote by $G^C_i$ the graph with edge set $\tilde{E}_i \setminus E_i$ and whose node set is simply the set of endpoints of its edges (i.e., no further isolated nodes).

(1) If $\Delta(G^C_i)$ denotes the maximal degree in $G^C_i$, $i = 1, \ldots, n$, then

$$\lambda_1(R) \leq \max_{1 \leq i \leq n} \min \left\{ \sqrt{\frac{2e_i(k_i - 1)}{k_i}}, \Delta(G^C_i) \right\}, \quad (17)$$

where $e_i$ is the number of edges added to $G_i$, i.e., $e_i = (|\tilde{E}_i| - |E_i|)$, and $k_i$ is the number of nodes in $V_i$.

(2) If additionally $G^C_i$ is connected for each $i = 1, \ldots, n$, then

$$\lambda_1(R) \leq \max_{1 \leq i \leq n} \min \left\{ \sqrt{2e_i - k_i' + 1}, \Delta(G^C_i) \right\}, \quad (18)$$

where $k_i'$ is the number of nodes of $G^C_i$.

Proof. (1) By assumption, $R$ is a diagonal block matrix whose blocks $R_i$ are the adjacency matrices of the induced subgraphs $G^C_i$. Thus, $\lambda_1(R)$ is the maximum of all spectral radii $\lambda_1(R_i)$. On the other hand, one has by [24] (3.45) that

$$\lambda_1(R_i) \leq \min \left\{ \sqrt{\frac{2e_i(k_i - 1)}{k_i}}, \Delta(G^C_i) \right\}.$$

and the claim follows.

(2) By Gershgorin’s theorem, the spectral radius of an adjacency matrix is never larger than the graph’s maximal degree, i.e., $\lambda_1(R_i) \leq \Delta(G^C_i)$. By assumption, there exists a permutation of the vertices in $V_i$ such that the matrix $R_i$ has the form

$$R_i = \begin{bmatrix} R_i' & 0 \\ 0 & 0 \end{bmatrix}$$

where $R_i'$ is the adjacency matrix of a connected graph with $k_i'$ nodes (i.e., the block $R_i'$ has dimension $k_i' \times k_i'$). Now, we deduce from [24, art. 50] that

$$\lambda_1(R_i) \leq \sqrt{2e_i - k_i' - 1},$$

and since $\lambda(R_i) = \lambda(R_i')$, the assertion follows.

A. Lower bound for $\tau_c^{(1)}$

By using estimate (9) and Proposition [1] in the first and the second term on the right hand side of (15), respectively, we deduce

$$\lambda_1(\hat{A}) \leq \max_{1 \leq i \leq n} \lambda_1(C_{V_i}) + \lambda_1(\hat{B}) + \max_{1 \leq i \leq n} \min \left\{ \sqrt{\frac{2e_i(k_i - 1)}{k_i}}, \Delta(R_i) \right\}. \quad (19)$$

The inequality in (19) gives us a lower bound for the epidemic threshold in the case of a graph whose partition of nodes set is almost equitable. Actually

$$\tau_c^{(1)} = \frac{1}{\lambda_1(A)} \geq \tau^* = \frac{1}{\max_{1 \leq i \leq n} \lambda_1(C_{V_i}) + \lambda_1(\hat{B}) + \max_{1 \leq i \leq n} \min \left\{ \sqrt{\frac{2e_i(k_i - 1)}{k_i}}, \Delta(R_i) \right\}}.$$

VI. CONCLUSION

The main goal of this work is to find a relation between the epidemic threshold of a given graph and the spectral properties of the corresponding quotient matrix.

Because the quotient matrix $Q$ has the same spectral radius of $A$, one motivation is the computational advantage in the calculus of $\lambda_1(A)$ and consequently of $\tau_c^{(1)}$, since the order of $Q$ is smaller than that of $A$.

We have derived a novel expression for the lower bound on $\tau_c^{(1)}$ as function of network metrics, e.g., the maximum among the internal degrees of the nodes over all compartments. In practice this value can be adopted to determine a safety region for the extinction of epidemic by forcing the effective infection rate below the lower bound; also, it can be useful for designing new network architectures to safeguard them from long-term, massive infection.

Moreover, we have showed that it is possible to reduce the number of equations representing the time evolution of infection probabilities using the quotient matrix $Q$. This is also convenient in order to calculate numerically the solution and the steady-state infection probability.

Finally we have also considered the case when the partition is almost equitable. Starting from a graph whose partition is equitable, we choose to perturb it by adding links inside some/every communities. Doing so, we allow for graphs whose communities’ internal structure is not that of a regular graph. In this case we have also provided a lower bound for the epidemic threshold and we have explored how it changes depending on the modifications introduced in the communities.
Acknowledgments

We thank Prof. Piet Van Mieghem for providing interesting and useful comments.

This work has been partially supported by the European Commission within the framework of the CONGAS project FP7-ICT-2011-8-317672 (see http://www.congas-project.eu).

[1] S. Bonaccorsi, S. Ottaviano, F. De Pellegrini, A. Socievole and P. Van Mieghem. Epidemic outbreaks in two-scale community networks, Physical Review E, 90(1): 012810, 2014.
[2] F. Ball, D. Mollison, and G. Scalia-Tomba. Epidemics with two levels of mixing. The Annals of Applied Probability, 7(1):46–89, 1997.
[3] F. Ball and P. J. Neal. Network epidemic models with two levels of mixing. Math. Biosci, 212:69-87, 2008.
[4] P. Van Mieghem, J. Omic and R. E. Kooij. Virus Spread in Networks, IEEE/ACM Transaction on Networking, 17(1):1-14, 2009.
[5] P. Van Mieghem and R. van de Bovenkamp. Non-Markovian infection spread dramatically alters the SIS epidemic threshold in networks, Physical Review Letters, 110(10):108701, 2013.
[6] E. Cator, R. van de Bovenkamp, and P. Van Mieghem. Susceptible-Infected-Susceptible epidemics on networks with general infection and curing times, Physical Review E, 87(6):062816, 2013.
[7] P. Van Mieghem and E. Cator. Epidemics in networks with nodal self-infection and the epidemic threshold. Physical Review E, 86(1):016116, 2012.
[8] P. Van Mieghem. The viral conductance of a network. Computer Communications, 35(12):1494–1506, 2012.
[9] P. Van Mieghem. The N-Intertwined SIS epidemic network model. Springer Computing, 93(2):147–169, 2011.
[10] P. K. Pollett. Quasi-stationary distributions: a bibliography. http://www.maths.uq.edu.au/~pkp/papers/qsds/qsds.pdf, 2008.
[11] N.T.J. Bailey. The mathematical theory of infectious diseases and its applications. Charles Griffin, London, 1975.
[12] R. Pastor-Satorras, and A. Vespignani. Epidemic spreading in scale free networks Physical Review Letters, 86(14): 3200, 2001.
[13] D.J. Daley, and J. Gani. Epidemic Modelling: An Introduction Cambridge Univ. Press, Cambridge, 1999.
[14] P. Van Mieghem. Decay towards the overall-healthy state in SIS epidemics on networks. arXiv:1310.3980, 2014.
[15] M. Draief and L. Massoulié. Epidemics and Rumours in Complex Networks, London Math. Society Lecture Note Ser. 369, Cambridge University Press, Cambridge, 2010.
[16] Y. Wang, D. Chakrabarti, C. Wang, C. Faloutsos. Epidemic spreading in real networks: an eigenvalue viewpoint, Proc. 22nd Int. Symposium on Reliable Distributed Systems, IEEE, 25-33, 2003.
[17] E. Cator and P. Van Mieghem. Nodal infection in Markovian SIS and SIR epidemics on networks are non-negatively correlated, Physical Review E, 89(5): 052802, 2014.
[18] A.J. Schwenk. Computing the characteristic polynomial of a graph. in: Graphs and Combinatorics (Proc. Washington D.C. 1973). Springer, New York, 153-172, 1974.
[19] P. Van Mieghem. Exact Markovian SIR and SIS epidemics on networks and an upper bound for the epidemic threshold, arXiv:1402.1731, 2014.
[20] L. Perko. Differential Equations and Dynamical Systems, Springer-Verlag, New York, 2001.
[21] C.D. Meyer. Matrix Analysis and Applied Linear Algebra. SIAM, 2000.
[22] C.D. Godsil and B.D. McKay. Feasibility conditions for the existence of walk-regular graphs. Linear Algebra and its Applications, 30: 15-61, 1980.
[23] E. Seneta. Non negative matrix and Markov chains. Springer-Verlag, New York, 2006.
[24] P. Van Mieghem. Graph Spectra for Complex Networks. Cambridge University Press, Cambridge, 2011.