An Extension of InfoMap to Absorbing Random Walks

Esteban Vargas, Mason A. Porter, Joseph H. Tien

December 22, 2021

Abstract

InfoMap is a popular approach for detecting densely connected ‘communities’ of nodes in networks. To detect such communities, it builds on the standard type of Markov chain and ideas from information theory. Motivated by the dynamics of disease spread on networks, whose nodes may have heterogeneous disease-removal rates, we extend InfoMap to absorbing random walks. To do this, we use absorption-scaled graphs, in which the edge weights are scaled according to the absorption rates, along with Markov time sweeping. One of our extensions of InfoMap converges to the standard version of InfoMap in the limit in which the absorption rates approach 0. We find that the community structure that one detects using our extensions of InfoMap can differ markedly from the community structure that one detects using methods that do not take node-absorption rates into account. Additionally, we demonstrate that the community structure that is induced by local dynamics can have important implications for susceptible-infected–recovered (SIR) dynamics on ring-lattice networks. For example, we find situations in which the outbreak duration is maximized when a moderate number of nodes have large node-absorption rates. We also use our extensions of InfoMap to study community structure in a sexual-contact network. We consider the community structure that corresponds to different absorption rates for homeless individuals in the network and the associated impact on syphilis dynamics on the network. We observe that the final outbreak size can be smaller when treatment rates are lower in the homeless population than in other populations than when they are the same in all populations.

1 Introduction

In the study of dynamical processes on networks, such as the spread of infectious diseases [15, 24] and the spread of ideas [18], a key issue is how the structural features of a network affect dynamical processes on it [27]. One important structural feature is community structure, in which tightly-knit sets of nodes called ‘communities’ are connected relatively sparsely to other tightly-knit sets of nodes [9, 28]. Community structure in a network can exert much influence on dynamical processes on a network. For example, Salathé and Jones [33] showed that making a network more modular can lead to a longer outbreak duration of a disease. In the present paper, we examine how dynamical processes can influence community structure. One can interpret some objective functions (e.g., various types of modularity) for studying community structure in terms of random walks [16, 21], and random walks also underlie many other community-detection algorithms, such as InfoMap [32]. To explore how dynamical processes can affect community structure, we extend InfoMap to absorbing random walks and explore how that affects communities in a network.

To illustrate the implicit impact of a dynamical process on community structure, consider an absorbing random walk on an undirected and unweighted line network with four nodes (see Figure 1). Suppose that the absorption rate$^1$ of node 2 is much larger than those of the other nodes. The large absorption rate of node 2 is a barrier to the absorbing random walk, as transitions from nodes in $\{3, 4\}$ to nodes in $\{1\}$ or vice versa are not likely. The local dynamics (specifically, absorption at the nodes) introduces a partition of the set of nodes that we can interpret as an effective community structure of the network. Such effective community structure can be rather different from the community structure that one detects based on network structure alone.

$^1$The absorption rate of a node is the sum of the transition rates from that node to the absorbing states.
There are a large variety of community-detection algorithms [9, 23, 28]. Well-known approaches to community detection including the maximization of a modularity objective function [23] (e.g., by using the locally greedy Louvain algorithm [2]), methods that are based on statistical inference (e.g., via stochastic block models [25]), methods that are based on dynamical processes [12], and methods that are based on both dynamical processes and information theory (e.g., InfoMap [31, 32]). A key rationale for using random-walk-based approaches such as InfoMap is that a random walker tends to become trapped in sets of nodes with dense internal connections and sparse connections to other dense sets of nodes. InfoMap seeks a partition of the set of nodes of a network that minimizes what a so-called ‘map function’, which is a weighted sum of entropies that are associated with one-step transitions within communities and one-step transitions between communities [31, 32].

In the present paper, we extend InfoMap to absorbing random walks. Absorbing random walks arise naturally in biological contagion processes on networks, where recovery or removal of infectious nodes corresponds to ‘absorption’. This interpretation of recovered or removed individuals as ‘absorbed’ underlies the next-generation-matrix approach [38] for calculating the basic reproduction number \( R_0 \), a staple quantity of mathematical epidemiology that indicates the mean number of secondary infections that arise when a single infected individual enters a population of susceptible individuals [4]. Recent work described how to use an ‘absorption-scaled graph’ and a generalized inverse (the so-called ‘absorption inverse’) of the unnormalized graph Laplacian to derive analytical results for \( R_0 \) on ‘movement networks’ (where edge weights correspond to per capita movement rates between locations) [11, 36]. References [11, 36] also discussed how different absorption rates can shape effective community structure and how such community structure can impact the spread of infectious diseases. In particular, Tien et al. [36] showed that clustering of disease hot spots on balanced2 networks is associated with larger values of \( R_0 \). Jacobsen and Tien [11] discussed a spectral partitioning method that is based on the absorption inverse and absorption-scaled graphs. Benzi et al. [1] introduced centrality measures based on absorption inverses.

Our paper proceeds as follows. In Section 2, we present the original InfoMap algorithm and the Markov time-sweeping technique that we use in our extensions of InfoMap. In Section 3, we present extensions of InfoMap to absorption-scaled graphs. In Section 4, we introduce a definition of a map function \( L^{(a)} \) for absorbing random walks and we relate this new map function to the extension of InfoMap that is associated with a particular absorption-scaled graph. We also relate this particular absorption-scaled graph to the absorption-scaled graph in [11] that has a direct relationship to the corresponding absorbing random walk. In Section 5, we discuss three toy examples to illustrate how our extensions of InfoMap yield effective community structures that are driven by different absorption rates. In Section 6.1, we examine one of our extensions of InfoMap on a network that we construct from a ring-lattice network and we relate epidemiological quantities to the effective community structure that we obtain from that extension of InfoMap. In Section 6.2, we examine the same extension of InfoMap on a sexual-contact network that we extract from [20] and we examine the effect of different absorption rates of a specific set of nodes on the effective community structure. We also examine epidemiological quantities that we obtain from a model of syphilis transmission on this sexual-contact network.3 In Section 7, we summarize and discuss our main conclusions. In Section 8, we present the proofs of the main propositions in Section 4, discuss adjusted mutual information, and give the details for the model of syphilis transmission.

---

2 We call a network balanced if the in-degree equals the out-degree for each node in the network.

3 We have posted the code that generates the numerical results in Section 6 at https://gitlab.com/esteban_vargas_bernal/extending-infomap-to-absorbing-random-walks.
2 Background: InfoMap

We now present background material on InfoMap and an extension of it to continuous-time Markov chains using Markov time sweeping [34]. In Section 2.1, we define a ‘map’ objective function, which the InfoMap algorithm attempts to minimize [3, 31, 32]. In Section 2.2, we present Markov time sweeping for a version of the map function that includes a resolution parameter for community detection [34]. This resolution parameter, which amounts to a Markov time, allows us to tune the sizes of the communities that we obtain using InfoMap. In Section 3, we use Markov time sweeping for the map function to extend InfoMap to absorbing random walks. Before proceeding with these discussions, we summarize our notation in Table 1.

| Symbol | Expression | Description |
|--------|------------|-------------|
| $G$    | directed and weighted graph | |
| $A$    | adjacency matrix of a directed and weighted graph | |
| $W$    | diagonal out-degree matrix | |
| $\mathcal{L}$ | $W - A$ | unnormalized graph Laplacian |
| $P$    | $AW^{-1}$ | transition matrix of the Markov chain that is associated with $A$ |
| $M$    | $\{M_1, \ldots, M_m\}$ | graph partition |
| $Q$    | $\{q_1, \ldots, q_n\}$ | distribution that is associated with inter-community transitions |
| $\mathcal{P}^*$ | $\{p_{i,j}\}$ | distributions that are associated with intra-community transitions in $M_i$ |
| $\mathcal{H}$ | $\mathcal{H}(P), \mathcal{H}(Q)$ | entropy of a distribution |
| $L(M, P, \pi)$ | $q_{i,j} \mathcal{H}(Q) + \sum_{i=1}^n p_{i,j} \mathcal{H}(P^*_{i,j})$ | map function |
| $H$    | diagonal weight matrix in Algorithms 2a and 2b | |
| $d$    | generic absorption-rate vector | |
| $\delta$ | node-absorption-rate vector | |
| $D_\delta$ | $\text{diag}\{\delta\}$ | diagonal matrix of node-absorption rates |
| $d_\delta(D_\delta, H)$ | $(d_\delta)_i = h_i w_i + \delta_i$ | scaled rate vector, which is defined as the diagonal of $HW + D_\delta$ |
| $G(D_\delta, H)$ | absorption-scaled graph that is associated with the pair $(G, d_\delta(D_\delta, H))$ | |
| $\mathcal{L}(D_\delta, H)$ | $(W - A)(HW + D_\delta)^{-1}$ | unnormalized graph Laplacian of an absorption-scaled graph |
| $\tilde{u}$ | vector in Ker $\mathcal{L}(D_\delta, 0)$ with non-negative entries that satisfy $\sum u_i = 1$ | |
| $P_t(D_\delta, H)$ | $e^{-t \mathcal{L}(D_\delta, H)}$ | transition matrix for a Markov chain on an absorption-scaled graph |
| $P_l(D_\delta, H)$ | $I - t \mathcal{L}(D_\delta, H)$ | linearized transition matrix for a Markov chain on an absorption-scaled graph |
| $A$    | adjacency matrix of a graph with an absorbing state | |
| $P$    | transition matrix that is induced by $A$ | |
| $N$    | $(I - Q)^{-1}$ | fundamental matrix of a discrete-time absorbing Markov chain |
| $t$    | $N^{T}1$ | the scalar $t_i$ is the expected number of transitions before absorption when the initial state is $i$ |
| $N$    | $ND_t^{-1}$ | normalized fundamental matrix of a discrete-time absorbing Markov chain |
| $P_t$  | $P_l(D_\delta, I, 1)$ | special case of $P_l(D_\delta, H)$ for $H = I$ and $t = 1$ |
| $\pi_\delta$ | $N \pi_0$ | the quantity $(\pi_\delta_{i,j})$ is the probability that node $i$ is the last node before absorption if the initial distribution is $\pi_0$ |
| $L^{(a)}(M, A, \delta, \pi_0)$ | $L(M, P_\delta, \pi_\delta^{(a)})$ | map function for a Markov chain with an absorbing state |

Table 1: Summary of our notation.
2.1 The map function and the standard InfoMap algorithm

Let $A = (a_{ij})_{i,j \in \{1, \ldots, n\}}$ be the adjacency matrix of a directed and weighted graph with the set $\{1, \ldots, n\}$ of nodes. The adjacency-matrix element $a_{ij}$ encodes the weight of the edge from node $j$ to $i$, and $a_{ij} = 0$ implies that there is no edge from $j$ to $i$. The map function $L(M)$ measures the strength of the community structure of a partition $M = \{M_1, \ldots, M_m\}$ of the set of nodes. The map function evaluated at $M$ is [32]

$$L(M) = q_\rightarrow \mathcal{H}(Q) + \sum_{i=1}^m p_i^1 \mathcal{H}(P^i). \quad (1)$$

The quantity $\mathcal{H}(Q)$ is the entropy of a distribution $Q$ that is associated with transitions between communities, and the $\mathcal{H}(P^i)$ terms are entropies of the distributions $P^i$ that are associated with transitions within communities\(^4\). One can interpret these entropies as optimal mean encoding lengths in the sense of Shannon’s source-coding theorem [6]. Specifically, Shannon’s source-coding theorem states that if $X$ is a random variable with finitely many states and $p$ is a probability mass function with entropy $\mathcal{H}(p)$, then the mean length of a code that describes the states of $X$ cannot be smaller than $\mathcal{H}(p)$. Additionally, as the size of the set of states of the random variable becomes infinite, one approaches the lower bound arbitrarily closely. Therefore, we call $\mathcal{H}(p)$ the ‘optimal mean encoding length’ (see Theorem 6 in [35]), and we regard the map function (1) as a weighted sum of optimal mean encoding lengths for one-step transitions between and within communities.

Intuitively, if the term $q_\rightarrow \mathcal{H}(Q)$ in (1) is small, then the connections between communities are sparse. If the terms $p_i^1 \mathcal{H}(P^i)$ are small, then the intra-community connections are dense [32, 34]. Therefore, we expect that minimizing $L(M)$ yields a partition with dense connections within communities and sparse connections between communities.

The following definition describes the terms of $L(M)$ in equation (1).

**Definition 1. (Map function)** Let $P = (p_{ij})_{i,j \in \{1, \ldots, n\}}$ be a stochastic matrix, and let $\pi = (\pi_1, \ldots, \pi_n)^T$ be a distribution. For a partition $M = \{M_1, \ldots, M_m\}$ of the set of nodes of a network, we define the probability of a transition out of community $M_i$ with an initial state from the distribution $\pi$ by $q_\rightarrow := \sum_{j \in M_i, k \notin M_i} \pi_j p_{kj}$ and the probability of a transition into community $M_i$ with an initial state from the distribution $\pi$ by $q_\leftarrow := \sum_{k \in M_i, j \notin M_i} \pi_j p_{kj}$. We also define $q_\rightarrow := \sum_{i \in \{1, \ldots, m\}} q_{\rightarrow} Q_i$ and $p_i^1 := q_\rightarrow + \sum_{j \in M_i} \pi_j$. The optimal mean encoding length that is associated with inter-community transitions is the entropy $\mathcal{H}(Q)$, where the distribution $Q$ is

$$Q := \{q_{\rightarrow}/q_{\rightarrow}, \ldots, q_{n_{\rightarrow}}/q_{\rightarrow}\}.$$ 

The optimal mean encoding length that is associated with intra-community transitions for community $M_i$ is the entropy $\mathcal{H}(P^i)$, where the distribution $P^i$ is

$$P^i := \{q_{\rightarrow}/p_i^1, \ldots, \pi_k/p_i^1|k \in M_i\}.$$ 

The map function that is associated with $P$ and $\pi$ evaluated at the partition $M$ is

$$L(M, P, \pi) := q_\rightarrow \mathcal{H}(Q) + \sum_{i=1}^m p_i^1 \mathcal{H}(P^i).$$

**Definition 2. (Standard map function)** Let $P$ be a regular matrix (i.e., some power of $P$ has only positive entries), and let $\pi$ be its corresponding stationary distribution. We define $L(M, P)$ by $L(M, P, \pi)$, and we refer to $L(M, P)$ as a standard map function.

The map function $L(M)$ in the following definition corresponds to the case in which the Markov chain that is induced by $A$ is regular\(^5\). If $A$ is not regular, then the definition of $L(M)$ follows from Algorithm 1.

**Definition 3. (Map function associated with an adjacency matrix $A$)** Let $A$ be an adjacency matrix such that $AW^{-1}$ is regular, where $W := \text{diag}\{\omega_1, \ldots, \omega_n\}$ and $\omega_j := \sum_i a_{ij} \neq 0$ for $j \in \{1, \ldots, n\}$. The map function $L(M)$ that is associated with $A$ is

$$L(M) := L(M, AW^{-1}).$$

\(^4\)Recall that the entropy of a distribution with strictly positive probabilities $p_1, \ldots, p_r$ is $\mathcal{H}(\{p_1, \ldots, p_r\}) := -\sum_i p_i \log_2(p_i)$.

\(^5\)A Markov chain is regular if its associated transition-probability matrix is regular.
Algorithm 1 InfoMap algorithm [3]

Input: An adjacency matrix $A = (a_{ij})_{i,j \in \{1, \ldots, n\}}$ of a directed and weighted graph.

Output: A partition $M$ of the set of nodes that minimizes $L(M)$.

Let $M = \{M_1, \ldots, M_n\}$ be a partition of the set $\{1, \ldots, n\}$ of nodes. With steps (1)–(6), we compute the map function

$$L(M) = q_{\bowtie} H(Q) + \sum_{i=1}^{m} p_i^{(i)} H(P^i).$$

The greedy algorithm InfoMap attempts to minimize $L(M)$.

1: Define the preference vector: Let $\omega_j := \sum_i a_{ij}$ and $\omega := \sum_i a_{ij}$. Define $v_i := \frac{\omega_j}{\omega}$ and $v := (v_1, \ldots, v_n)^T$.

2: Define the transition matrix $T_j = (b_{ij})_{i,j \in \{1, \ldots, n\}}$, where

$$b_{ij} = \begin{cases} a_{ij}/\omega_j, & \text{if } \omega_j \neq 0; \\ 0, & \text{if } \omega_j = 0. \end{cases} \quad (2)$$

3: Let $\tau$ be a teleportation probability. We use the value $\tau = 0.15$. Define the transition matrix $P = (p_{ij})_{i,j \in \{1, \ldots, n\}}$, where

$$p_{ij} = \begin{cases} (1-\tau)b_{ij} + \tau v_i, & \text{if } \omega_j \neq 0; \\ v_i, & \text{if } \omega_j = 0. \end{cases}$$

The definition of $P$ is inspired by the PageRank algorithm [5]. It guarantees the existence of a stationary distribution in which all probabilities are strictly positive. The value $\tau = 0.15$ is the most common value of PageRank’s teleportation parameter.

4: Find the stationary distribution $p^* = (p_1^*, \ldots, p_n^*)^T$ that is associated with $P$ by solving $Pp = p^*$ (for example, by using the power method).

5: Define the matrix $Q := (q_{ij})_{i,j \in \{1, \ldots, n\}}$, where $q_{ij} := q_{ij}/(\sum_k q_{kj})$ and $q_{ij} = b_{ij}p_{ij}^*$. Although $p_{ij}^*$ depends on the teleportation probability $\tau$, the parameter $b_{ij}$ does not. Additionally, define $p_i := \sum_j q_{ij}$. (In [17], this step is called ‘unrecorded teleportation’.)

6: Calculate the following ingredients of the map function $L(M)$:

$$q_{i\bowtie} := \sum_{k \in M_i} \sum_{l \notin M_i} q_{kl},$$

$$q_{i\bowtie} := \sum_{k \in M_i} \sum_{l \notin M_i} q_{ik},$$

$$q_{\bowtie} := \sum_{i=1}^{m} q_{i\bowtie},$$

$$q_{\bowtie} := \sum_{i=1}^{m} q_{i\bowtie},$$

$$H(Q) := -\sum_{i=1}^{m} q_{i\bowtie} \log_2(q_{i\bowtie}/q_{\bowtie}),$$

$$p_{i\bowtie} := q_{i\bowtie} + \sum_{k \in M_i} p_k,$$

$$H(P^i) := -q_{i\bowtie}/p_{i\bowtie} \log_2(q_{i\bowtie}/p_{i\bowtie}) - \sum_{k \in M_i} p_k/p_{i\bowtie} \log_2(p_k/p_{i\bowtie}).$$

7: Using a computational heuristic, minimize $L(M)$. We do this using an iterative process. We start with a partition of a network that consists of communities that each have a single node. At step $k+1$, we obtain a new partition $M_{k+1}$ such that $L(M_{k+1}) \leq L(M_k)$ and $M_{k+1} = (M_k \setminus \{A, B\}) \cup \{A \cup B\}$ for some $A, B \in M_k$. 

5
2.2 Markov time sweeping

The map function (1) is associated with one-step transitions of a random walk. Schaub et al. [34] incorporated Markov time sweeping into InfoMap to tune the time scales of transitions by encoding not only one-step transitions but also transitions with steps\(^6\) of any length \(t > 0\). We think of the Markov process that is determined by \(T_j\) as having time steps with size 1, where \(T_j\) is defined as in Algorithm 1. Markov time sweeping uses the transition matrix of a continuous-time Markov chain in which the step in time is \(T\) determined by \(T_j\) if we assume that \(a_{ii} = 0\) for \(i \in \{1, \ldots, n\}\). These diagonal elements correspond to self-edges with weight \(1 - t\). If we use InfoMap with \((1 - t)I + tT_j\) instead of \(T_j\), we recover the version of InfoMap in Algorithm 1 by setting \(t = 1\).

As discussed in [14], for \(t < 1\), one can also consider the linearization
\[
e^{-t(I-T_j)} \approx I - t(I-T_j) = (1-t)I + tT_j
\]
as an input of InfoMap. In (3), \((1-t)I + tT_j\) has diagonal elements that are all equal to \(1 - t\) if we assume that \(a_{ii} = 0\) for \(i \in \{1, \ldots, n\}\). These diagonal elements correspond to self-edges with weight \(1 - t\). If we use InfoMap with \((1-t)I + tT_j\) instead of \(T_j\), we thereby obtain large communities. See [14] for further discussion of encoding and Markov times.

As discussed in [14], for \(t < 1\), one can also consider the linearization
\[
e^{-t(I-T_j)} \approx I - t(I-T_j) = (1-t)I + tT_j
\]
as an input of InfoMap. In (3), \((1-t)I + tT_j\) has diagonal elements that are all equal to \(1 - t\) if we assume that \(a_{ii} = 0\) for \(i \in \{1, \ldots, n\}\). These diagonal elements correspond to self-edges with weight \(1 - t\). If we use InfoMap with \((1-t)I + tT_j\) instead of \(T_j\), we thereby obtain large communities. See [14] for further discussion of encoding and Markov times.

As discussed in [14], for \(t < 1\), one can also consider the linearization
\[
e^{-t(I-T_j)} \approx I - t(I-T_j) = (1-t)I + tT_j
\]
as an input of InfoMap. In (3), \((1-t)I + tT_j\) has diagonal elements that are all equal to \(1 - t\) if we assume that \(a_{ii} = 0\) for \(i \in \{1, \ldots, n\}\). These diagonal elements correspond to self-edges with weight \(1 - t\). If we use InfoMap with \((1-t)I + tT_j\) instead of \(T_j\), we thereby obtain large communities. See [14] for further discussion of encoding and Markov times.

As discussed in [14], for \(t < 1\), one can also consider the linearization
\[
e^{-t(I-T_j)} \approx I - t(I-T_j) = (1-t)I + tT_j
\]
as an input of InfoMap. In (3), \((1-t)I + tT_j\) has diagonal elements that are all equal to \(1 - t\) if we assume that \(a_{ii} = 0\) for \(i \in \{1, \ldots, n\}\). These diagonal elements correspond to self-edges with weight \(1 - t\). If we use InfoMap with \((1-t)I + tT_j\) instead of \(T_j\), we thereby obtain large communities. See [14] for further discussion of encoding and Markov times.

3 Markov time sweeping and extensions of InfoMap to absorbing random walks

We now introduce extensions of InfoMap that take the absorption rates of the nodes of a network into account. Our approach uses absorption-scaled graphs, which arise naturally in the context of absorbing random walks [11].

Definition 4. (Absorption-scaled graph) Let \(G\) be a directed and weighted graph with adjacency matrix \(A = (a_{ij})_{i,j \in \{1, \ldots, n\}}\), where \(a_{ij}\) encodes the weight of the edge from node \(j\) to node \(i\). Let \(\vec{d} = (d_1, \ldots, d_n)^T\) be a vector (which we call an ‘absorption-rate vector’) with positive entries that we call the ‘absorption rates’. We define the absorption-scaled graph that is associated with the pair \((G, \vec{d})\) as the graph \(\tilde{G}\) with adjacency matrix \(\tilde{A} := AD^{-1}\), where \(D := \text{diag}\{d_1, \ldots, d_n\}\).

![Figure 2: An absorption-scaled graph. (a) A graph \(G\) with absorption-rate vector \(\vec{d}\). The red node has a large absorption rate. (b) The associated absorption-scaled graph \(\tilde{G}\), where the arrow length is proportional to the corresponding edge weight.](image)

We now define a couple of important mathematical objects that we use in our extensions of InfoMap.

Definition 5. Let \(A = (a_{ij})_{i,j \in \{1, \ldots, n\}}\) be the adjacency matrix of a graph \(G\). Let \(\vec{\delta} = (\delta_1, \ldots, \delta_n)^T\) be a vector with positive entries, where \(\delta_i\) is specified independently from the matrix \(A\) and is the node-absorption rate of node \(i\). We refer to the vector \(\vec{\delta}\) as the node-absorption-rate vector. Let \(D_\delta := \text{diag}\{\vec{\delta}\}\) and consider a diagonal matrix \(H = \text{diag}\{h_1, \ldots, h_n\}\), where \(h_i \geq 0\). We define the scaled rate vector \(\vec{d}_s(D_\delta, H)\) as the diagonal of \(HW + D_\delta\).

The node-absorption-rate vector \(\vec{\delta}\) in Definition 5 does not depend on the adjacency matrix \(A\), whereas the scaled rate vector \(\vec{d}_s(D_\delta, H)\) depends on the out-degrees \(\omega_i\) through \(H\). In particular, \(\vec{\delta} = \vec{d}_s(D_\delta, 0)\) when

\(^6\)Markov sweeping time had been used previously in contexts other than InfoMap [8, 16].
Consider a family of absorption-scaled graphs with absorption-rate vectors that are equal to scaled rate vectors $\mathbf{d}$. In the expression for $\mathbf{d}$, we have $\theta = (\delta_1, \ldots, \delta_n)^T$. We consider a family of absorption-scaled graphs with absorption-rate vectors that are equal to scaled rate vectors $\tilde{\mathbf{d}}(D_\delta, H)$, where $H = \text{diag}(h_1, \ldots, h_n)$ and $h_i \geq 0$.

We denote the absorption-scaled graph that is associated with the pair $(G, \tilde{\mathbf{d}}(D_\delta, H))$ by $\tilde{G}(D_\delta, H)$. The unnormalized graph Laplacian of $\tilde{G}(D_\delta, H)$ is

$$\tilde{\mathbf{L}}(D_\delta, H) := (W - A)(HW + D_\delta)^{-1}.$$  

For any Markov time $t > 0$, the transition matrix that is associated with the infinitesimal generator $-\tilde{\mathbf{L}}(D_\delta, H)$ is

$$P_t(D_\delta, H, t) := e^{-t\tilde{\mathbf{L}}(D_\delta, H)}.$$  

For a Markov time $t$, the linearization of $P_t(D_\delta, H, t)$ is

$$P_t(D_\delta, H, t) := I - t\tilde{\mathbf{L}}(D_\delta, H) = (I - tW(HW + D_\delta)^{-1}) + (tAHW + D_\delta)^{-1}.$$  

For this linearization, we require that $0 < t < 1/\max_i \{\omega_i/(h_i\omega_i + \delta_i)\}$ to ensure that $P_t(D_\delta, H, t)$ is a transition-probability matrix.

The reason for considering the matrix $H$ in Definition 5 is that the choice of the matrix $H$ allows us to tune the relative effects of the edge weights and the node-absorption rates on the communities that we detect using our extensions of InfoMap. We give our extensions of InfoMap in Algorithms 2a and 2b.

**Algorithm 2a** InfoMap for absorbing random walks with a linear input.

Input: An adjacency matrix $A = (a_{ij})_{i,j \in \{1, \ldots, n\}}$ of a directed and weighted graph, a node-absorption-rate vector $\tilde{\mathbf{d}} = (\delta_1, \ldots, \delta_n)^T$ with strictly positive entries, and a diagonal matrix $H$ with non-negative entries.

Output: A partition $M$ of the set of nodes that minimizes $L(M, P_t(D_\delta, H, t))$ for a Markov time $t$.

1. Construct the unnormalized graph Laplacian $\tilde{\mathbf{L}}(D_\delta, H) = (W - A)(HW + D_\delta)^{-1}$ for the absorption-scaled graph $\tilde{G}(D_\delta, H)$.
2. Choose a Markov time such that

$$0 < t < 1/\max_i \{\omega_i/(h_i\omega_i + \delta_i)\}.$$  

3. Apply Algorithm 1 (i.e., the standard InfoMap algorithm) with $P_t(D_\delta, H, t)$ as input.

**Algorithm 2b** InfoMap for absorbing random walks with an exponential input.

Input: An adjacency matrix $A = (a_{ij})_{i,j \in \{1, \ldots, n\}}$ of a directed and weighted graph, a node-absorption-rate vector $\tilde{\mathbf{d}} = (\delta_1, \ldots, \delta_n)^T$ with strictly positive entries, and a diagonal matrix $H$ with non-negative entries.

Output: A partition $M$ of the set of nodes that minimizes $L(M, P_t(D_\delta, H, t))$ for a Markov time $t$.

1. Construct the unnormalized graph Laplacian $\tilde{\mathbf{L}}(D_\delta, H) = (W - A)(HW + D_\delta)^{-1}$ for the absorption-scaled graph $\tilde{G}(D_\delta, H)$.
2. Choose any Markov time $t > 0$.
3. Apply Algorithm 1 (i.e., the standard InfoMap algorithm) with $P_t(D_\delta, H, t)$ as input.

The idea of our extensions of InfoMap to absorbing random walks is to introduce a family of associated absorption-scaled graphs and then apply Markov time sweeping to these absorption-scaled graphs. To illustrate how the node-absorption rates impact the communities that we detect, consider the matrix $P_t$ in Algorithm 2a. In the expression for $P_t(D_\delta, H, t)$ in (5), the term

$$(I - tW(HW + D_\delta)^{-1}) = \text{diag} \left\{ 1 - \frac{t\omega_i}{h_i\omega_i + \delta_i} \right\}$$

creates self-edges that are positively correlated with the node-absorption rates $\delta_i$. This correlation reflects the idea that the flow of a random walk gets stuck longer in nodes with larger node-absorption rates.
If we assume that \( D_\delta = \delta I, \omega_i \geq 1 \) (with \( i \in \{1, \ldots, n\} \)) and set \( h_i = \delta_i (\omega_i - 1)/\omega_i \) and \( t = \delta_n \), then \( P_l(\delta_i I, H, \delta_n) = AW^{-1} \). We thereby recover the input of the standard version of InfoMap when all the absorption rates are the same. If we assume that \( H = hI \) and \( t = h \), we obtain
\[
\lim_{\|\delta\|\to 0} P_l(D_\delta, hI, h) = \lim_{\|\delta\|\to 0} (I - h W(hW + D_\delta)^{-1}) + h A(hW + D_\delta)^{-1} = AW^{-1},
\]
so we again recover the input of the standard version of InfoMap in the limit.

4 A map function \( L^{(a)} \) for Markov chains with an absorbing state

4.1 A map function for an absorbing random walk

The extensions of InfoMap in Section 3 are associated with Markov chains that do not have an absorbing state. They account for absorption through absorption-scaled graphs, which have associated absorbing random walks [11]. It is natural to ask whether we can interpret map functions that are associated with the extensions of InfoMap in Section 3 in terms of corresponding Markov chains that have an absorbing state. We answer this question by directly defining a map function \( L^{(a)} \) for Markov chains with an absorbing state that coincides with \( L(M, P_\delta) \) for \( P_\delta = P_l(D_\delta, I, 1) \). The map function \( L(M, P_\delta) \) is associated with the extension in Algorithm 2a. We also show that the map function \( L^{(a)} \) converges to the map function \( L(M) \) that is associated with the Markov chain without absorption when the absorption rates approach 0. In Section 4.1.1, we present the building blocks of \( L^{(a)} \), the definition of \( L^{(a)} \), and a related example. In Section 4.1.2, we present our two main results about \( L^{(a)} \).

4.1.1 Construction and main results for a map function for an absorbing random walk

We assume that the Markov chain that is associated with the adjacency matrix \( A \) is regular, and we add an absorbing state and node-absorption rates \( \delta_1, \ldots, \delta_n \), which are the transition rates from states that are associated with \( A \) to the absorbing state. The adjacency matrix of the absorbing Markov chain is
\[
\tilde{A} = \begin{pmatrix} A & \tilde{0} \\ \tilde{0}^T & 0 \end{pmatrix},
\]
where \( \tilde{\delta} = (\delta_1, \ldots, \delta_n)^T \). From \( \tilde{A} \), we obtain the transition-probability matrix
\[
\tilde{P} = \begin{pmatrix} Q & \tilde{0} \\ \tilde{0}^T & 1 \end{pmatrix},
\]
where \( \tilde{r} = (\delta_1/\omega_1 + \delta_1), \ldots, \delta_n/\omega_n + \delta_n)^T \) gives transition probabilities to absorption, with \( \omega_j = \sum_i a_{ij} \) (for \( j \in \{1, \ldots, n\} \)), and \( Q = A(W + D_\delta)^{-1} \), with \( D_\delta = \text{diag}(\tilde{\delta}) \) and \( W = \text{diag}(\omega_1, \ldots, \omega_n) \). The Markov chain with the transition probability matrix \( \tilde{P} \) is the absorbing Markov chain that is associated with \( A \) and \( \delta_1, \ldots, \delta_n \).

By Definition 1, we know that if \( P \) is a regular transition matrix and \( \pi \) is a probability mass function, then \( L(M, P, \pi) \) depends only on \( M \), \( P \), and \( \pi \). We write \( L(M, P, \pi) \) as \( L(M, P) \) if \( \pi \) is the unique stationary distribution of \( P \). The quantity \( L(M, P, \pi) \) is a weighted sum of optimal mean encoding lengths for one-step transitions (with probabilities in \( P \) and starting from the distribution \( \pi \)) between and within communities. We define a map function \( L^{(a)} \) for the absorbing random walk that is associated with (9) by \( L(M, P_\delta, \pi^{(a)}_\delta) \) for an appropriate distribution \( \pi^{(a)}_\delta \) (see Section 4.1.1.1) and appropriate transition-probability matrix \( P_\delta \) (see Section 4.1.1.2).

4.1.1.1 The distribution \( \pi^{(a)}_\delta \)

We define a distribution \( \pi^{(a)}_\delta \) that has the desirable property of recovering the stationary distribution \( \pi \) of a Markov chain without absorption (with adjacency matrix \( AW^{-1} \)) when the absorption rates approach 0. That is,
\[
\lim_{\|\delta\|\to 0} \pi^{(a)}_\delta = \pi.
\]

Consider the normalized fundamental matrix \( \tilde{N} = ND_\delta^{-1} \), where \( N = (I - Q)^{-1} = \sum_{k=0}^\infty Q^k \) is the fundamental matrix of the absorbing Markov chain that is given by \( \tilde{P} \) and \( \tilde{T} = N^T \tilde{T} = (t_1, \ldots, t_n)^T \) (with \( \tilde{T} := (1, \ldots, 1)^T \))
is the vector whose entries are the expected numbers of steps after starting from the non-absorbing states [13]. For each node pair \((i, j)\), the entry \(n_{ij}/t_j\) of \(N\) gives the probability that node \(i\) is the last node before absorption if we start at node \(j\). This probability depends on the node-absorption rates. For an initial distribution \(\pi_0\), we obtain the distribution

\[
\pi^{(a)}_\delta = N\pi_0.
\]  

(11)

The following proposition states that property (10) holds.

**Proposition 1.** Suppose that the Markov chain with transition matrix \(AW^{-1}\) is regular. Let \(\bar{\delta}\) be a node-absorption-rate vector in which all entries are strictly positive, and let \(D_\delta := \text{diag}\{\bar{\delta}\}\). Let \(N = (I - A(W + D_\delta)^{-1})^{-1}\) be the fundamental matrix of the absorbing Markov chain that is associated with \(A\) and \(\delta_1, \ldots, \delta_n\), let \(D_\bar{\delta}\) be the diagonal matrix with the column sums of \(N\) in its diagonal, and let \(\pi\) be the stationary-distribution vector that is associated with \(AW^{-1}\).

It then follows that

\[
\lim_{||\bar{\delta}|| \to 0} ND_{\bar{\delta}^{-1}} = \pi T^T.
\]

**Proof.** Consider the vector 1-norm \(\|\bar{\delta}\|_1 = \sum_i |\bar{\delta}_i|\) and its induced matrix norm \(\|B\|_1 = \max_j \sum_i |b_{ij}|\). Fix \(\epsilon > 0\).

From \(\lim_{n \to \infty} (AW^{-1})^n = \pi T^T\), it follows that there is a positive integer \(N_1\) such that

\[
\|(AW^{-1})^n - \pi T^T\|_1 < \epsilon \quad \text{for} \quad n \geq N_1.
\]

(12)

In particular,

\[
(AW^{-1})^N_1 = \pi T^T + \Lambda,
\]

(13)

where \(\|\Lambda\|_1 < \epsilon\). Let \(Q := A(W + D_\delta)^{-1}\). From \(\|Q\|_1 \leq \|Q\|_2 \leq \|AW^{-1}\|_2 = 1\) and \(t_i \to \infty\) as \(\|\bar{\delta}\|_1 \to 0\), it follows that

\[
\lim_{||\bar{\delta}|| \to 0} \|\sum_{j<N_1} Q^j D_{\bar{\delta}^{-1}}\|_1 = 0.
\]

(14)

Additionally, \(\lim_{||\bar{\delta}|| \to 0} Q^N_1 = (AW^{-1})^N_1\). Therefore, there is a \(\eta_0 > 0\) such that

\[
\left\|\sum_{j<N_1} Q^j D_{\bar{\delta}^{-1}}\right\|_1 < \epsilon \quad \text{and} \quad Q^N_1 = (AW^{-1})^N_1 + \Delta
\]

(15)

whenever \(0 < \|\bar{\delta}\|_1 < \eta_0\), where \(\|\Delta\|_1 < \epsilon\). Let \(\Delta' := \Lambda + \Delta\). From (13) and (15), we have

\[
\sum_{j \geq N_1} Q^j D_{\bar{\delta}^{-1}} = Q^N_1 \sum_{j \geq 0} Q^j D_{\bar{\delta}^{-1}}
\]

\[
= Q^N_1 ND_{\bar{\delta}^{-1}}
\]

\[
= (\pi T^T + \Delta')ND_{\bar{\delta}^{-1}}
\]

\[
= \pi T^T ND_{\bar{\delta}^{-1}} + \Delta'ND_{\bar{\delta}^{-1}}
\]

\[
= \pi T^T D_{\bar{\delta}^{-1}} + \Delta'ND_{\bar{\delta}^{-1}}
\]

\[
= \pi T^T + \Delta'ND_{\bar{\delta}^{-1}},
\]

(16)

where \(\|\Delta'ND_{\bar{\delta}^{-1}}\|_1 \leq \|\Delta'\|_1 \|ND_{\bar{\delta}^{-1}}\|_1 < 2\epsilon\). From (15) and (16), it follows for \(0 < \|\bar{\delta}\|_1 < \eta_0\) that

\[
\|ND_{\bar{\delta}^{-1}} - \pi T^T\|_1 \leq \|\sum_{j<N_1} Q^j D_{\bar{\delta}^{-1}}\|_1 + \|\Delta'ND_{\bar{\delta}^{-1}}\|_1 < 3\epsilon.
\]

\[
\square
\]

4.1.1.2 The probability-transition matrix \(P_\delta\)

Let \(P_\delta\) denote the linearization (6) with \(H = I\) and \(t = 1\). That is,

\[
P_\delta := P_1(D_\delta, I, 1) = (I - W(W + D_\delta)^{-1}) + A(W + D_\delta)^{-1} = D_\bar{\delta} + Q,
\]

(17)
where \( D_T = \text{diag}(\{r_i\}) \). Our choice of \( P_\delta \) is motivated by property
\[
\lim_{\|\delta\| \to 0} P_\delta = P = AW^{-1}. \tag{18}
\]
From (18), it follows that we recover the transition-probability matrix of the regular Markov chain that is induced by \( A \) in the limit in which there is no absorption. Furthermore, for all pairs \((i, j)\), the off-diagonal entries \( a_{ij}/(\omega_i + \delta_i) \) of \( P_\delta \) give the one-step probabilities for the absorbing Markov chain that is associated with \( \hat{P} \) to go from node \( i \) to node \( j \). The one-step absorption probabilities \( \delta_i/(\omega_i + \delta_i) \) of the absorbing Markov chain that is associated with \( P \) are on the diagonal of \( P_\delta \); these are self-edges in the Markov chain that is associated with \( P_\delta \).

The following proposition states that the time to self-transitions of the Markov chain that is associated with \( P_\delta \) is equal to the time to absorption in the absorbing Markov chain that is associated with \( \hat{P} \).

**Proposition 2.** Let \{\(X_n\)\}_{n \in \mathbb{N}} be the Markov chain with transition-probability matrix \( P_\delta = P_\delta(D_\delta, I, 1) = (p^{(1)}_{ij})_{i,j \in \{1,\ldots,n\}} \). Define the random variable
\[
T_j := \min\{n : X_n = X_{n-1} \text{ and } X_0 = j\}.
\]
Let \( \theta_j = \mathbb{E}(T_j) \) be the expectation of \( T_j \), and let \( \bar{\theta} = (\theta_1, \ldots, \theta_n)^T \). If \( N = (I - Q)^{-1} \) is the fundamental matrix of the absorbing Markov chain that is associated with \( A \) and \( \delta_1, \ldots, \delta_n \), it is then the case that
\[
\bar{\theta}^T = \bar{1}^T N.
\]

**Proof.** From the law of total expectation,
\[
\theta_j = \sum_{i \neq j} (\mathbb{E}(T_i) + 1)p^{(1)}_{ij} + p^{(1)}_{jj} = \sum_{i \neq j} \theta_i p^{(1)}_{ij} + 1.
\]
Let \((P_\delta)_{dg}\) denote the diagonal matrix with the same diagonal as \( P_\delta \). We write (21) as
\[
\bar{\theta}^T[I - (P_\delta - (P_\delta)_{dg})] = \bar{1}^T.
\]
Because \( P_\delta = D_T + Q \), it follows from (22) that \( Q = P_\delta - (P_\delta)_{dg} \) and \( \bar{\theta}^T = \bar{1}^T N \).

Using Definition 1, we obtain that the map function \( L(M, P_\delta, \hat{\pi}_0) \) is a weighted sum of the optimal mean encoding lengths for one-step transitions (that are associated with the transition matrix \( P_\delta \) and the distribution \( \pi_0^{(1)} \)) between and within communities.

In the following definition, we define a map function for the absorbing random walk that is associated with \( A \) and \( \delta_1, \ldots, \delta_n \).

**Definition 6.** (Map function for absorbing random walks) Let \( A \) be the adjacency matrix of an absorbing random walk, and let \( \delta \) be the walk’s node-absorption-rate vector. Let \( M \) be a partition of the set of nodes that is associated with \( A \), and let \( \pi_0 \) be an initial distribution. We define the map function \( L^{(a)}(M, A, \delta, \pi_0) := L(M, P_\delta, \hat{\pi}_0) \).

As an instructive example, we show the map function \( L^{(a)} \) for all the possible partitions \( M \) of the set of nodes of a three-node network with adjacency matrix (23). The set of nodes of the network is \{1, 2, 3\}, and the adjacency matrix is
\[
A = \begin{pmatrix}
0 & 1 & 1 \\
0 & 0 & 0 \\
1 & 1 & 0
\end{pmatrix}.
\tag{23}
\]
Intuitively, if the node-absorption rate of node 2 is larger than the node-absorption rates of nodes 1 and 3, then node 2 is in a different community than nodes 1 and 3 in the effective community structure. We want to check whether the partition with the minimum value of \( L^{(a)} \) captures this intuition.

Let \( \delta = (\delta_1, \delta_2, \delta_3)^T \) be a node-absorption-rate vector. We fix \( \delta_1 = \delta_3 = 0.1 \) and vary \( \delta_2 \) in the interval \([0.1, 10]\). In Figure 3, we show the values of \( L^{(a)}(M, A, \delta, \pi_0) \) for every partition \( M \) of the set \{1, 2, 3\} of nodes, where \( A \) is defined in (23) and \( \pi_0 \) is the uniform distribution over \{1, 2, 3\}. We always attain the minimum value of \( L^{(a)}(M, A, \delta, \pi_0) \) for the partition \{\{2\}, \{1, 3\}\}, so we obtain this partition if we select the optimal encoding \( L^{(a)}(M, A, \delta, \pi_0) \). This is consistent with the intuition that \{\{2\}, \{1, 3\}\} is the effective community structure because \( \delta_2 > \delta_1 \) and \( \delta_2 > \delta_3 \).
Figure 3: The encoding length $L^{(a)}(M,A,\tilde{\delta},\pi_0)$ for all five possible partitions $M$ of the set of nodes of the three-node network with adjacency matrix (23). The node-absorption-rate vector is $\tilde{\delta} = (\delta_1, \delta_2, \delta_3)^T$, with $\delta_1 = 0.1$ and $0.1 \leq \delta_2 \leq 10$. The initial distribution is $\pi_0 = (1/3,1/3,1/3)^T$.

### 4.1.2 Primary results for the map function $L^{(a)}$

We first show that if we choose $P_3$ as in (17), then the standard map function $L(M,P_3)$ (i.e., the objective function that we minimize using Algorithm 2a with input $P_3(D_3, I, 1)$) coincides with the encoding $L^{(a)}(M,A,\tilde{\delta},\pi_0)$ that corresponds to the absorbing random walk that is associated with the adjacency matrix $A$ and the node-absorption-rate vector $\tilde{\delta}$ for an appropriate distribution $\pi_0$.

**Proposition 3.** Suppose that $P_3 = P_3(D_3, I, 1) = (I - W(W + D_3)^{-1}) + A(W + D_3)^{-1} = D + Q$, let $\pi^{(na)}_\delta$ be the stationary distribution of $P_3$, and let $\pi_0 := D_1 D_2 \pi^{(na)}_\delta = (\pi^{(na)}_\delta)\tilde{\delta}/(\omega_1 + \delta_1))_i$. We have that

$$L(M,P_3) = L^{(a)}(M,A,\tilde{\delta},\pi_0).$$

**Proof.** Because $\pi^{(na)}_\delta$ is the stationary distribution of $P_3$, it follows that $(D + Q)\pi^{(na)}_\delta = \pi^{(na)}_\delta$, which implies that

$$\pi^{(na)}_\delta = N D_2 D_3 \pi^{(na)}_\delta = N \pi_0.$$

Because the entries of $\pi^{(na)}_\delta$ sum to 1 and the columns of $N$ each sum to 1, we know that $\pi_0$ is a probability distribution. Therefore, for this choice of $\pi_0$, it follows that

$$L(M,P_3) = L(M,P_3,\pi^{(na)}_\delta) = L(M,P_3,N \pi_0) = L^{(a)}(M,A,\tilde{\delta},\pi_0),$$

where $L(M,P_3)$ is the standard map function with input $P_3$ and $L^{(a)}(M,A,\tilde{\delta},\pi_0)$ is the map function in Definition 6.

The second main result of the present subsubsection follows from equations (10) and (18), which imply that

$$L^{(a)}(M,A,\tilde{\delta},\pi_0) \rightarrow L(M,P) \quad \text{as} \quad \|\tilde{\delta}\| \rightarrow 0$$

for any $\pi_0$. Therefore, the map function $L^{(a)}(M,A,\tilde{\delta},\pi_0)$ that is associated with the absorbing random walk converges to the map function $L(M,P)$ that is associated with the (non-absorbing) Markov chain in the limit $\|\tilde{\delta}\| \rightarrow 0$.

### 4.2 Relating $\tilde{G}(D_\delta, I)$ to $\tilde{G}(D_\delta, 0)$

In Section 4.1, we used $L^{(a)}$ to relate the extension of InfoMap that is associated with the absorption-scaled graph $\tilde{G}(D_\delta, I)$ and the absorbing random walk that is associated with $A$ and $\delta_1, \ldots, \delta_n$. There is also a relationship between the absorption-scaled graph $\tilde{G}(D_\delta, 0)$ and the absorbing random walk that is associated with $A$ and $\delta_1, \ldots, \delta_n$. This relationship arises by using the generalized inverse of $\mathcal{L} = W - A$ that is known as the “absorption inverse” [11].
It is natural to ask how the absorption-scaled graphs $\tilde{G}(D_\delta, 0)$ and $\tilde{G}(D_\delta, I)$ are related. In this subsection, we find relationships between the Markov processes that are associated with the graph Laplacians $\tilde{L}(D_\delta, 0)$ and $\tilde{L}(D_\delta, I)$ through the fundamental matrices and absorption inverses. We describe these relationships, which are the main results of this subsection, in Propositions 4, 8, and 9. These results also yield connections between $\tilde{G}(D_\delta, 0)$, $\tilde{G}(D_\delta, I)$, and the fundamental matrix $(L + D_\delta)^{-1}$ through Propositions 5 and 6.

We first look at the fundamental matrices of the discrete-time Markov chains that are associated with $\tilde{L}(D_\delta, 0)$, and $\tilde{L}(D_\delta, I)$. Definition 7 describes the fundamental matrix of a regular Markov chain.

**Definition 7. (Fundamental matrix)** Let $P$ be a regular transition-probability matrix, and let $\bar{p}$ be its corresponding stationary distribution. The fundamental matrix $Z$ of the Markov chain that is associated with $P$ is

$$Z = (I - P + \bar{p}I)^{-1}.$$  \hspace{1cm} (28)

The matrix $Z - \bar{p}I$ measures the deviation of the expected number of visits between the Markov chain that is associated with $P$ and the Markov chain that is associated with $\bar{p}I$. (See Theorem 4.3.4 of [13].) The fundamental matrix $Z$ of a regular Markov chain satisfies

$$Z\bar{p} = \bar{p} \quad \text{and} \quad \bar{p}^T Z = \bar{p}^T.$$ \hspace{1cm} (29)

See Chapter IV of [13].

The following proposition gives the relationships between the fundamental matrices of the discrete-time Markov chains that are associated with the graph Laplacians $\tilde{L}(D_\delta, 0)$ and $\tilde{L}(D_\delta, I)$. We prove this proposition in Appendix 8.1.

**Proposition 4.** Let $P_0 = AW^{-1}$ be the transition-probability matrix of the discrete-time Markov chain that is associated with $\tilde{L}(D_\delta, 0)$, and let $P_1 = (A + D_\delta)(W + D_\delta)^{-1}$ be the transition-probability matrix of the discrete-time Markov chain that is associated with $\tilde{L}(D_\delta, I)$. Let $\pi$ and $\pi'$ be the stationary distributions that are associated with $P_0$ and $P_1$, respectively. Let $Z_1$ be the fundamental matrix that is associated with $P_i$ (for $i \in \{1, 2\}$). Let $U := (1/\bar{\delta}^T \bar{u})\bar{u}^T$ and $\alpha := \bar{\delta}^T \bar{u}/(\bar{w}^T \bar{u} + \bar{\delta}^T \bar{u})$, where $\bar{u} = (u_1, \ldots, u_n)^T$ is a vector in $\text{Ker} \tilde{L}(D_\delta, 0)$ with non-negative entries such that \(\sum_{i=1}^{n} u_i = 1\). We have that

$$Z_1 = W^{-1}(W + D_\delta) \left[ Z_0 + \alpha(1 - \alpha) \pi I^T - \alpha \left[ Z_0 D_\delta U + W \frac{\bar{u}^T \bar{u}}{\bar{\delta}^T \bar{u}} W^{-1} Z_0 (I - \alpha D_\delta U) \right] \right].$$ \hspace{1cm} (30)

We seek a connection between the unnormalized graph Laplacians $\tilde{L}(D_\delta, 0)$ and $\tilde{L}(D_\delta, I)$ through absorption inverses. In Definition 8, we describe an absorption inverse of an unnormalized graph Laplacian.

**Definition 8. (Absorption inverse)** Let $\tilde{L}$ be the unnormalized graph Laplacian of a strongly connected graph, and let $\bar{d}$ be an absorption-rate vector. Let $D := \text{diag}\{\bar{d}\}$, $N_{1,0} := \{\bar{x} \in \mathbb{R}^n : D\bar{x} \in \text{Range} \tilde{L}\}$, and $R_{1,0} := \{D\bar{x} : \bar{x} \in \text{Ker} \tilde{L}\}$. An absorption inverse $\tilde{L}^\delta$ of $\tilde{L}$ with respect to $\bar{d}$ is defined by the following properties:

$$\tilde{L}^\delta \tilde{L} \bar{y} = \bar{y} \quad \text{for} \quad \bar{y} \in N_{1,0},$$

$$\tilde{L}^\delta \bar{y} = 0 \quad \text{for} \quad \bar{y} \in R_{1,0}.$$ \hspace{1cm} (31)

The absorption inverse of a graph exists and is unique if the graph is strongly connected and the absorption rates are all positive. (See Theorem 2 in [11].) In Definition 9, we describe the group inverse of a matrix that is related to absorption inverses via Proposition 5.

**Definition 9. (Group inverse)** Let $X$ be a matrix such that $\text{rank}(X) = \text{rank}(X^2)$. The group inverse of $X$ is the unique matrix $X^#$ that satisfies

$$XX^# X = X,$$

$$X^# XX^# = X^#,$$

$$XX^# = X^# X.$$ \hspace{1cm} (32)

The following proposition relates the absorption inverse $\tilde{L}^\delta$ to the group inverse of $\tilde{L} D^{-1}$, where $D = \text{diag}\{\bar{d}\}$. 

12
Proposition 5 ([11], Proposition 2). Let \( \tilde{L} \) be the unnormalized graph Laplacian of a strongly connected graph, let \( \tilde{d} \) be an absorption-rate vector, and let \( D = \text{diag}(\tilde{d}) \). The following relationship holds:

\[
\left( \tilde{L}D^{-1} \right)^\# = D\tilde{L}^\#.
\]

The following proposition relates the absorption inverse \( \tilde{L}^\# \) and the fundamental matrix \( (\tilde{L} + D)^{-1} \) of the continuous-time absorbing random walk that associated with \( \tilde{L} \) and \( d_1, \ldots, d_n \).

Proposition 6 ([11], Proposition 4). Let \( \tilde{L}^\# \) be the absorption inverse of \( \tilde{L} \) with respect to \( \tilde{d} \). Let \( \tilde{u} \) be a vector in \( \text{Ker} \tilde{L} \) with positive entries such that \( \sum_i u_i = 1 \). Let \( D := \text{diag}(\tilde{d}) \) and \( U := (1/\tilde{u})\tilde{u}^T \), where \( d = \sum_i (u_id_i) \). The following relationship holds:

\[
(\tilde{L} + D)^{-1} = U + (1 + \tilde{L}^\#D)^{-1}\tilde{L}^\#.
\] (33)

In particular, if \( \tilde{d} = \tilde{\delta} \) and \( \tilde{L} = L \), then Propositions 5 and 6 relate the graph Laplacian \( \tilde{L}(D_\delta, 0) = L D_\delta^{-1} \) that corresponds to the absorption-scaled graph \( \tilde{G}(D_\delta, 0) \) and the continuous-time absorbing random walk that is associated with \( L = W - A \) and \( \delta_1, \ldots, \delta_n \). We suppose that the spectral radius \( \rho(L^\#D_\delta) \) satisfies \( \rho(L^\#D_\delta) < 1 \) and use the series expansion \( (I + L^\#D_\delta)^{-1} = \sum_{k=0}^\infty (-L^\#D_\delta)^k \) in (33) to obtain

\[
(\tilde{L} + D_\delta)^{-1} = U + \tilde{L}^\# + \sum_{k=1}^\infty (-\tilde{L}^\#D_\delta)^k \tilde{L}^\#.
\] (34)

For \( \epsilon := ||D_\delta||/||L|| \ll 1 \), it follows from (34) that

\[
\tilde{L}^\# = (\tilde{L} + D_\delta)^{-1} - (1/\tilde{\delta})u^T + O(\epsilon).
\] (35)

This approximation indicates that when \( \epsilon \ll 1 \), the entries of \( \tilde{L}^\# \) approximate the deviation in expected time to absorption between the continuous-time absorbing random walk that is associated with \( L \) and \( \delta_1, \ldots, \delta_n \) and the Markov chain without absorption that is associated with \( (1/\tilde{\delta})u^T \).

To compute the absorption inverse, we use the following proposition.

Proposition 7 ([11], Lemma 3 and Theorem 3). Let \( \tilde{d} \) be an absorption-rate vector, and let \( \tilde{u} \in \text{Ker}(W - A) \). Let \( \tilde{\omega} := \text{diag}(\tilde{W}) \), \( \pi := \tilde{\omega}/(\tilde{\omega}^T\tilde{u}) \), \( D := \text{diag}(\tilde{\delta}) \), \( U := (1/\tilde{u}^T\tilde{u})\tilde{u}^T \), and \( Z := W^{-1}Z_0 \), where \( Z_0 = \left( I - AW^{-1} + \pi\tilde{\omega}^T \right)^{-1} \) is the fundamental matrix that is associated with \( AW^{-1} \). We have that

\[
\tilde{L}^\# = (I - UD)Z(I - DU).
\] (36)

If \( \tilde{d} = \tilde{d}_\delta(D_\delta, 0) \), then Proposition 5 implies that

\[
\left[ \tilde{L}(D_\delta, 0) \right]^\# = D_\delta L^\#,
\]

where \( L = W - A \).

(37)

The following proposition relates the absorption inverse \( \tilde{L}^\# \) [which is related to \( L(D_\delta, 0) \) through (37)] and the absorption inverse of \( \tilde{L}(D_\delta, I) \) with respect to the diagonal of \( D_\delta(W + D_\delta)^{-1} \). We prove this proposition in Appendix 8.1.

Proposition 8. Denote \( \tilde{L}_1 := \tilde{L}(D_\delta, I) = (W - A)(W + D_\delta)^{-1} \), and let \( \tilde{d}_1 \) be the diagonal of \( D_\delta(W + D_\delta)^{-1} \). Let \( U := \tilde{\omega}\tilde{\omega}^T/(\tilde{\delta}^T\tilde{u}) \), \( U_1 := (W + D_\delta)U \) and \( D_1 := D_\delta(W + D_\delta)^{-1} \). We have that

\[
\tilde{L}^\#_1 = (W + D_\delta)\tilde{L}^\#
\]

and

\[
(\tilde{L} + D_\delta)^{-1} = (W + D_\delta)^{-1} \left( U_1 + (1 + \tilde{L}^\#_1 D_1)^{-1}\tilde{L}^\#_1 \right).
\] (39)

If \( \tilde{d} = \tilde{d}^\# := \tilde{d}_\delta(D_\delta, I) \), then Proposition 5 implies that

\[
\left[ \tilde{L}(D_\delta, I) \right]^\# = (W + D_\delta)L^\#, \quad \text{where} \quad L = W - A.
\] (40)

The following proposition relates the absorption inverses \( \tilde{L}^\# \) [which is related to \( \tilde{L}(D_\delta, 0) \) through (37)] and \( \tilde{L}^\# \) [which is related to \( \tilde{L}(D_\delta, I) \) through (40)]. We prove this proposition in Appendix 8.1.
Proposition 9. Let $\mathbf{\bar{d}} := \mathbf{d}(D, I) = \mathbf{w} + \mathbf{\delta} = (\omega_1 + \delta_1, \ldots, \omega_n + \delta_n)^T$ be the absorption-rate vector that is associated with the absorption-scaled graph $\hat{G}(D, I)$. With $L = W - A$, $\alpha := \mathbf{\delta}^T \mathbf{\bar{u}} / (\mathbf{\bar{w}}^T \mathbf{\bar{u}} + \mathbf{\delta}^T \mathbf{\bar{u}})$, $\pi = W \mathbf{\bar{u}} / (\mathbf{\bar{w}}^T \mathbf{\bar{u}})$, $Z_0 = (I - AW^{-1} + \pi \mathbf{\bar{u}})^{-1}$, and $Z_* = W^{-1} (Z_0 - \pi \mathbf{\bar{u}})^{-1}$, it follows that

$$L^\delta = \alpha^2 L^\delta + \alpha (1 - \alpha) \left( L^\delta Z_0 + Z_* L^\bar{\delta} \right) + (1 - \alpha)^2 Z_* .$$  \hspace{1cm} (41)

Because $Z_0$ is the fundamental matrix of the regular Markov chain with transition-probability matrix $AW^{-1}$, the entries of $Z_* = W^{-1} (Z_0 - \pi \mathbf{\bar{u}})^{-1}$ measure the deviation in time to absorption between the Markov chain that is associated with $AW^{-1}$ and the Markov chain that is associated with $\pi \mathbf{\bar{u}}$. Therefore, if $\| D \| / \| L \| \ll 1$, then the matrix $Z_*$ is the analog of $L^\delta$. Note from (41) that $L^\delta$ is a linear combination of terms that include $L^\delta$ and $Z_*$, where the corresponding coefficients depend on $\alpha = \mathbf{\delta}^T \mathbf{\bar{u}} / (\mathbf{\bar{w}}^T \mathbf{\bar{u}} + \mathbf{\delta}^T \mathbf{\bar{u}})$. Additionally, $\alpha \approx 1$ implies that $L^\delta \approx L^\bar{\delta}$ and $\alpha \approx 0$ implies that $L^\delta \approx Z_*$.

5 Examples

We apply Algorithms 2a and 2b to three networks. In Section 5.1, we examine a three-node network and examine the map function in Algorithm 2a with input $P_t(D, 0, t)$ for all possible partitions of the network. In Section 5.2, we study a network that consists of four cliques that are connected as in Figure 6. In this example, we examine the output of Algorithms 2a and 2b using different input matrices and discuss how Algorithms 2a and 2b yield effective community structures that are driven by node-absorption rates or are driven by edge weights, depending on the Markov times and the matrix $H$ in the input. In Section 5.3, we examine a grid of nodes with four different node-absorption rates and examine how Algorithms 2a and 2b can yield effective community structures that are driven by the four node-absorption rates.

5.1 A three-node network

We consider the three-node network (see Figure 4) with adjacency matrix $A$ in (23) and node-absorption-rate vector $\mathbf{\delta} = (\delta_1, \delta_2, \delta_3)^T$.

![Figure 4: An example three-node network.](image)

Consider the absorption-scaled graph $\hat{G}(D, H)$ with $H = 0$ and Algorithm 2a with input $P_t(D, 0, t)$ for a fixed $t$. In Figure 5, we show the values of $L(M) = L(M, P_t(D, 0, 1/20))$ for all five possible partitions $M$ of the three nodes, where we fix $\delta_1 = \delta_3 = 0.1$ and vary $\delta_2$ in the interval [0.1, 1]. When all the absorption rates are equal (and hence $\delta_2 = 0.1$), we see that $L(M)$ is minimized by the partitions $M = \{\{2\}, \{1, 3\}\}$ and $M = \{\{1, 2, 3\}\}$. (See the dashed orange curve.) However, when $\delta_2 > 0.1$, the partition $M = \{\{2\}, \{1, 3\}\}$ produces a smaller value of the map function (see the solid blue curve); this partition is the output of Algorithm 2a.
all of the edge weights are equal to 1. The node-absorption rates of the nodes in cliques $C_1 := \{1, 2, 3, 4\}$ and $C_3 := \{9, 10, 11, 12\}$ are $\delta_i = 7$ (with $i \in C_1 \cup C_3$), and the node-absorption rates of the nodes in cliques $C_2 := \{5, 6, 7, 8\}$ and $C_4 := \{13, 14, 15, 16\}$ are $\delta_i = 1$ (with $i \in C_2 \cup C_4$).

Consider the node-absorption-rate vector $\bar{\delta} = (\delta_1, \ldots, \delta_{16})^T$ and the absorption-scaled graphs: $\tilde{G}(D_D, 0)$ and $\tilde{G}(D_D, (3/2)I)$. We use Algorithm 2a with inputs $P_l(D_D, 0, t)$ and $P_l(D_D, (3/2)I, t)$ and Algorithm 2b with inputs $P_r(D_D, 0, t)$ and $P_r(D_D, (3/2)I, t)$ for different Markov times $t$. Arguably, the network structure (including both network topology and edge weights) on its own favors the partition $M^* := \{C_1, C_2, C_3, C_4\}$ in Figure 7(a). However, the larger absorption rates in cliques $C_1$ and $C_3$ lead to the partition $M^{**} := \{\{1\}, \{2\}, \{3\}, \{4\}, C_2, \{9\}, \{10\}, \{11\}, \{12\}, C_4\}$ in Figure 7(b).
Figure 7: Two partitions of the set of nodes of the network in Figure 6. Each color indicates a different community. The partition $M^*$ in (a) is the four-clique planted partition, which arises from the network structure (i.e., the network topology and edge weights). The partition $M^{**}$ in (b) arises from a combination of the network structure and the node-absorption rates. In (b), the nodes with node-absorption rate $\delta_1 = 7$ (the largest nodes) belong to single-node communities. (The node sizes are arbitrary; they are not proportional to the node-absorption rates of the associated nodes.)

We observe that Algorithms 2a and 2b with $H = 0$ produce the partition $M^*$ for a smaller range of Markov times than Algorithms 2a and 2b with $H = (3/2)I$. Specifically, Algorithm 2a with input $P_l(D_\delta, 0, t)$ does not produce the partition $M^*$ for any Markov time $t$ that satisfies (7) [see Figure 8(a)], whereas Algorithm 2a with input $P_l(D_\delta, (3/2)I, t)$ produces the partition $M^*$ when $1.47 \lesssim t \lesssim 1.75$ [see Figure 8(c)]. Additionally, Algorithm 2b with input $P_l(D_\delta, 0, t)$ produces $M^*$ when $1.28 \lesssim t \lesssim 2.67$ [see Figure 8(b)], whereas Algorithm 2b with input $P_l(D_\delta, (3/2)I, t)$ produces $M^*$ when $2.01 \lesssim t \lesssim 13.73$ [see Figure 8(d)]. These results are consistent with the fact that the choice $H = (3/2)I$ gives more importance to the edge weights than the choice $H = 0$.

Figure 8: The numbers of communities in the partitions that we obtain using Algorithms 2a and 2b with four different inputs. The partitions that consist of four communities coincide with $M^*$ in Figure 7(a), and the partitions that consist of ten communities coincide with $M^{**}$ in Figure 7(b).
5.3 A square grid network with four different node-absorption rates

We consider a square grid network with 36 nodes (see Figure 9). We divide the set of nodes into four subgrids (which we label as $B_1$, $B_2$, $B_3$, and $B_4$), which we illustrate using nodes of different sizes in Figure 9. We endow the nodes in subgrid $B_1$ with a node-absorption rate of 0.2, those in subgrid $B_2$ with a node-absorption rate of 0.7, those in subgrid $B_3$ with a node-absorption rate of 1.5, and those in subgrid $B_4$ with a node-absorption rate of 1.7.

Figure 9: A grid network with nodes that have node-absorption rates in the set $\{0.2, 0.7, 1.2, 1.7\}$. The node sizes are positively correlated with (but are not linearly related to) the node-absorption rates of the associated nodes.

We first look at the community structure that we obtain using Algorithm 2a with input $P_l(D, \delta, 0, t)$ for a Markov time $t$ that satisfies (7). In Figure 10(a), we see that we obtain a partition with 28 communities when $t \geq 0.0345$. This partition has $B_1$ as a community, and all of the other communities are single-node communities. Indeed, Algorithm 2a yields a partition in which all of the nodes with $\delta_i \in \{0.7, 1.5, 1.7\}$ belong to single-node communities [see Figures 10(a) and 10(c)]. By contrast, it is possible for Algorithm 2b to produce alternative communities, depending upon the choice of $\mathcal{H}$ and the Markov time $t$. For example, Algorithm 2b with input $P_t(D, I, 5.25)$ produces the partition $\{B_1, B_2, B_3, B_4\}$ [see Figure 10(d)].
Figure 10: (a,b) The number of communities in partitions that we obtain using (a) Algorithm 2a with input \( P_l(D_8, 0, t) \) and (b) Algorithm 2b with input \( P_e(D_8, I, t) \). In (c) and (d), we show the partitions that we obtain from Algorithm 2a with input \( P_l(D_8, 0, 0.04) \) and Algorithm 2b with input \( P_e(D_8, I, 5.25) \), respectively. Each color in panels (c) and (d) indicates a different community; the node sizes are positively correlated with (but are not linearly related to) the node-absorption rates of the associated nodes.

6 Epidemiological implications and empirical data

6.1 Effective community structure and susceptible–infected–recovered dynamics on networks with ring-lattice communities

We explore the effects of the node-absorption configurations of a contact network on quantities such as a disease’s outbreak duration, final size, and outbreak peak in simulations of a susceptible–infected–recovered (SIR) model of disease spread on the network. We use a Gillespie algorithm [15] to simulate the disease dynamics, and we refer to these effects as epidemiological implications. We also examine the effective community structure that we obtain using Algorithm 2b with input \( P_e(D_8, 0, t) \) for specific node-absorption configurations and explore the connection of the resulting effective community structure with the dynamics of the above epidemiological quantities.

Salathé and Jones [33] explored the effect of changes in community structure on the above epidemiological quantities. They considered a network that consists of Watts–Strogatz small-world graphs [26] that are connected to each other by ‘community bridges’ in the form of edges that are assigned uniformly at random from all possible node pairs. They then rewired the community bridges, one at a time, into edges within communities in the following manner. First, they selected a community bridge \( \{i_1, i_2\} \) uniformly at random from all of the community bridges and removed it from the set of edges. They then picked \( i_k \) uniformly at random from \( \{i_1, i_2\} \). Finally, they selected a node \( i_3 \neq i_k \) uniformly at random from the Watts–Strogatz small-world subgraph that is associated with \( i_k \) and added \( \{i_k, i_3\} \) to the set of edges. They ran SIR simulations on their network after each rewiring step (for fixed contact rates and fixed node-absorption rates) and recorded the means of the outbreak durations, final sizes, and outbreak peaks in these simulations.

We use a similar process as in [33], but we consider directed (and unweighted) ring-lattice graphs instead of Watts–Strogatz networks and we modify the node-absorption rates of nodes in community bridges instead of
rewiring community bridges. We use Algorithm 3 to obtain Figure 11. In step 2 of Algorithm 3, increasing the node-absorption rate of the bridging nodes is analogous to removing community bridges in [33] and increasing the transmission rate of other nodes (which we call ‘balancing nodes’) is analogous to adding edges within a community in [33]. We refer to each configuration of node-absorption rates and transmission rates in step 2 of Algorithm 3 as a stage.

Algorithm 3 SIR simulations on networks for different node-absorption configurations. Each $n_{WS}$-node ring-lattice graph in the input of this algorithm is isomorphic to a network with the set \{0, \ldots, n_{WS} - 1\} of nodes, where the neighbors of node $i$ are of the form $i \pm j \pmod{n_{WS}}$, with $j \in \{1, \ldots, k_{WS}/2\}$.

**Input:** Size $n_{WS}$ of each of the ring-lattice graphs, number $N_{WS}$ of disconnected $n_{WS}$-node ring-lattice graphs (these are our planted communities), even number $k_{WS}$ of neighbors of each node in its planted community (where each edge is reciprocated by the edge in the other direction), transmission rates $\beta_s$ and $\beta_{ss}$ (with $\beta_{ss} > \beta_s$), and node-absorption rates $\delta_s$ and $\delta_{ss}$ (with $\delta_{ss} > \delta_s$).

**Output:** Means of the outbreak durations, final sizes, and outbreak peaks over $N_{sim}$ SIR simulations for each parameter stage (where a parameter stage is determined by the node-absorption and transmission rates of the nodes of the graph $G$ that we define in step 1 of the algorithm).

1. Define the graph $G$: select $n_{WS} \times n_{WS}$ pairs of nodes uniformly at random from all node pairs and add bidirectional bridging edges for each of these node pairs. Set all of the edge weights to 1.

2. Define the stages of parameter configurations (i.e., the ‘parameter stages’): in stage 1, the transmission rate is $\beta_s$ and the node-absorption rate is $\delta_s$ for all of the nodes of the graph $G$ in step 1. If the parameters of stage $n_s$ have been defined, select a community bridge $\{i_1, i_2\}$ (where $i_1$ and $i_2$ belong to distinct ring-lattice graphs) uniformly at random from the set of node pairs in which each node has node-absorption rate $\delta_s$. Set the node-absorption rates of $i_1$ and $i_2$ to be $\delta_{ss}$ for stage $n_s + 1$. Additionally, select a balancing node $l_k$ uniformly at random from the ring-lattice subgraph that is associated with $i_k$. (We require that $l_k$ is distinct from $i_k$ and that it is not neighbor of $i_k$ for $k \in \{1, 2\}$.) Set the transmission rates of $l_1$ and $l_2$ to $\beta_{ss}$ for stage $n_s + 1$.

3. Run SIR simulations: for each parameter stage, run $N_{sim}$ simulations of the SIR model on $G$ with the corresponding parameter configuration. Record the means of the outbreak durations, the final sizes, and the outbreak peaks over all the simulations in each stage.

We choose the transmission rate $\beta_{ss}$ to compensate for the decrease in new infections that occur because of the chosen bridging nodes. We estimate this decrease by calculating $\langle k \rangle \beta_s / \delta_s - \langle k \rangle \beta_{ss} / \delta_{ss}$ (using an approximation that is similar to one in [22]), where $\langle k \rangle$ is the mean degree of the bridging nodes. We compensate for the decrease in infections by choosing $\beta_{ss}$ such that $\langle k \rangle \beta_{ss} / \delta_s = \langle k \rangle \beta_s / \delta_s + \alpha (\langle k \rangle \beta_s / \delta_s - \langle k \rangle \beta_{ss} / \delta_{ss})$, which estimates the new infections that arise from balancing nodes, where $\alpha$ is a tuning parameter that we use to preserve the value of the basic reproduction number. This yields

$$\beta_{ss} = \beta_s + \alpha \delta_s \beta_s \left( \frac{1}{\delta_s} - \frac{1}{\delta_{ss}} \right).$$

With (42), the mean numbers of infectious individuals that arise from the first infectious individual are 2.57, 2.24, and 1.8 for stages 1, 29, and 68, respectively. In Figure 11, we show the means that we obtain for the outbreak duration, the final size, and the outbreak peak for the 68 stages of the node-absorption configurations. The qualitative behavior of these epidemiological quantities is consistent with the observations in [33]. Specifically, the final outbreak size and the outbreak peak decrease monotonically as we increase the number of stages [see Figures 11(b,c)]. By contrast, for the first 29 stages in Figure 11(a), the mean outbreak duration increases as we increase the number of stages.
Figure 11: The (a) mean outbreak duration, (b) mean final outbreak size, and (c) mean outbreak peak as a function of the stage $n_s$ that we obtain from Algorithm 3 with $N_{sim} = 1000$ simulations of SIR dynamics, $N_{WS} = 20$ ring-lattice graphs in the overall network, $n_{WS} = 12$ nodes in each ring-lattice graph, $k_{WS} = 6$ neighbors for each node in its ring-lattice graph, $N_s = 68$ stages, transmission rates of $\beta_\ast = 0.125$ and $\beta_{\ast\ast}$ [from (42)], absorption rates of $\delta_\ast = 0.2$ and $\delta_{\ast\ast} = 1$, and a tuning parameter of $\alpha = 0.1$.

We now examine the effective community structure of one planted community of the network on which we ran the SIR simulations for Figure 11. We use Algorithm 2b with input $P_t(D_\delta, 0, t)$ for Markov times $t \in (0.01, 0.05)$ at different stages of the node-absorption configurations $\delta_1, \ldots, \delta_{240}$, where $\vec{\delta} = (\delta_1, \ldots, \delta_{240})^T$ is the node-absorption-rate vector and $\delta_i \in \{\delta_\ast, \delta_{\ast\ast}\}$. We use $H = 0$ as an input of Algorithm 2b because the influence of absorption on the effective community structure is more noticeable when $H \neq 0$ [e.g., see Figures 8(b,d)]. In Figure 12(a), we show the size of the resulting partition of the network $G$ (where $G$ is defined in step 1 of Algorithm 3) by Algorithm 2b with input $P_t(D_\delta, 0, t)$ for Markov times $t \in (0.01, 0.05)$ and for the node-absorption rates in the initial stage (stage 1, which we show using the dash-dotted green curve), the stage with the peak duration (stage 29, which we show using the solid red curve), and the final stage (stage 68, which we show using the dashed blue curve). We then select the fifth planted community [which encompasses the nodes with labels 49–60 in Figures 12(c,d)]. For a partition $M$ of the set of nodes, we say that a subset of the fifth planted community is a subcommunity of this community if it is the intersection between a community in $M$ and the fifth planted community. In Figure 12(b), we show the number of subcommunities of the fifth planted community in the partition that we obtain using Algorithm 2b with input $P_t(D_\delta, 0, t)$ and Markov times $t \in (0.01, 0.05)$ for the initial stage (see the dash-dotted green curve), the stage with the peak duration (see the solid red curve), and the final stage (see the dashed blue curve). In Figures 12(c,d), we show the subcommunities of the fifth community for Markov time $t = 0.025$.

In Figures 12(c,d), we see that the above partitions have more than one community in the peak-duration and final stages. By contrast, in the initial stage, the partition has only one community. In Figure 12(c), we see that the nodes with node-absorption rate $\delta_{\ast\ast}$ (i.e., nodes 54, 56, and 59) are in different communities but that the disease can flow through the blue community and enter a different planted community (i.e., a different ring-lattice graph in the input of Algorithm 3). For example, nodes 51, 52, and 53 are neighbors that belong to the same community and have edges that connect to nodes in planted communities other than their own planted community (i.e., they are bridging nodes). Furthermore, the increase in transmission rate of some of the nodes (e.g., nodes 50, 51, and 53) implies that the disease can still spread to other planted communities, so the duration of the outbreak is longer than the outbreak duration in the initial stage [see Figure 11(a)]. At the final stage (i.e., stage 68), there are more bridging nodes with node-absorption rate $\delta_{\ast\ast}$ than in the peak-duration stage; these nodes lead to six subcommunities [see Figure 12(d)]. Specifically, in the fifth planted
community, nodes 50, 53, 54, 56, and 59 have a node-absorption rate of $\delta^{**}$ at stage 68 (see the larger nodes in Figure 12(d)), and these five nodes belong to single-node communities of the final partition. Moreover, in stage 68 (unlike in stage 29), the disease can die out more easily in nodes with larger node-absorption rates without spreading to other communities. Therefore, we expect that the outbreak duration, final outbreak size, and outbreak peak in Figure 11 in the final stage are smaller than in previous stages. We can think of the increase in the node-absorption rates of bridging nodes as analogous to cutting bridging edges in [33] and the increase in transmission rates of some nodes as analogous to rewiring bridging edges inside a planted community in [33]. Moreover, as in [33], the outbreak duration peaks at intermediate stages (see Figure 11(a)).

Figure 12: Comparison of community structures in the initial, peak-duration, and final stages of Algorithm 2b for SIR dynamics on networks with ring-lattice communities. (a) The number of communities in the network that we obtain using Algorithm 2b with input $P_t(D, 0, t)$ for the node-absorption configuration in the initial stage (dash-dotted green curve), the peak-duration stage (solid red curve), and the final stage (dashed blue curve). (b) The number of subcommunities of the fifth planted community that we obtain using Algorithm 2b in the initial stage (dash-dotted green curve), the peak-duration stage (solid red curve), and the final stage (dashed blue curve). (c,d) The subcommunities of the fifth planted community that we obtain using Algorithm 2b in (c) the peak-duration stage with $t = 0.025$ and (d) the final stage with $t = 0.025$. The set of subcommunities of the fifth planted community in the initial stage consists of a single community (so we do not include a panel for the subcommunities of the fifth planted community from the initial stage). Each color indicates a subcommunity; larger nodes have a larger node-absorption rate $\delta^{**}$.

6.2 HIV Transmission Networks

6.2.1 Community structure

We examine the effective community structure of one of the networks in the HIV Transmission Network Metastudy Project [20]. This project collected networks of social and sexual contacts from eight different studies that took place between 1988 and 2001. One goal of this project was to create a collection of data sets that people can use to analyze the effects of contacts on the transmission of diseases. Some attributes of these networks
include the mode of connection between participants of the studies (e.g., drug users or sexual contacts), the races of the participants, the sexual orientations of the participants, whether or not a participant is a prostitute, whether or not a participant is homeless, whether or not a participant has been infected with diseases (such as HIV, syphilis, chlamydia, and gonorrhea), and other metadata.

In our investigation, we examine the sexual-contact network in the Atlanta Urban Networks Project [20]. The Atlanta Urban Networks Project that produced this network took place from June 1996 until April 1999. In this project, 228 respondents were interviewed at 6-month intervals about their sexual behavior and activity. The resulting network consists of a single connected component of 2297 nodes and 7501 edges. This network includes 89 homeless individuals. The network in the Atlanta Urban Networks Project is undirected and unweighted. In our code for our extensions of InfoMap, we use a directed and weighted network with a symmetric adjacency matrix in which the weight of an edge \((i_1,i_2)\) is 1 if \(i_1\) and \(i_2\) have a sexual contact and 0 if they do not. We refer this network as the “Atlanta Urban Network”.

The homeless population in the Atlanta Urban Network has a larger mean degree (it is about 17) than the rest of the population; the mean degree of the whole network is about 2.4. Additionally, given that some individuals in the homeless population may lack health services, we assume that the homeless individuals in the network have lower treatment rates for sexual diseases (and for other infections) [29].

In our investigation, we examine the sexual-contact network in the Atlanta Urban Networks Project [20]. The Atlanta Urban Networks Project that produced this network took place from June 1996 until April 1999. In this project, 228 respondents were interviewed at 6-month intervals about their sexual behavior and activity. The resulting network consists of a single connected component of 2297 nodes and 7501 edges. This network includes 89 homeless individuals. The network in the Atlanta Urban Networks Project is undirected and unweighted. In our code for our extensions of InfoMap, we use a directed and weighted network with a symmetric adjacency matrix in which the weight of an edge \((i_1,i_2)\) is 1 if \(i_1\) and \(i_2\) have a sexual contact and 0 if they do not. We refer this network as the “Atlanta Urban Network”.

We seek to compare the effective community structure of the Atlanta Urban Network in two scenarios: (1) the node-absorption rate \(\delta_\ast\) of the homeless nodes is smaller than the node-absorption rate \(\delta\) of the non-homeless nodes and (2) \(\delta_\ast = \delta\). The node-absorption configuration of the first scenario has a node-absorption-rate vector \(\vec{\delta}(1)\) with node-absorption rates \(\delta_j = \delta_\ast = 0.2\) for non-homeless nodes \(j\) and \(\delta_i = \delta_\ast = 0.04 < \delta\) for homeless nodes\(^8\) \(i\). The node-absorption configuration of the second scenario has a node-absorption-rate vector \(\vec{\delta}(2)\) with node-absorption rates \(\delta_i = \delta_\ast\), where \(i\) is any node.

We use Algorithm 2b with input \(P_e(D_\delta, 0, t)\) because the choice of \(H = 0\) makes the effect of absorption on the effective community structure more conspicuous than the choice \(H \neq 0\) when we compare the produced partitions for different Markov times. In Figure 13, we show the number of communities that we obtain using Algorithm 2b with input \(P_e(D_\delta, 0, t)\) for \(D_\delta = \text{diag}\{\vec{\delta}(1)\}\) (see the solid red curve) and \(D_\delta = \text{diag}\{\vec{\delta}(2)\}\) (see the dashed blue curve).

![Figure 13: The number of communities that we obtain using Algorithm 2b with input \(P_e(D_\delta, 0, t)\) for a node-absorption configuration with a node-absorption rate of \(\delta_\ast = \delta_\ast = 0.2\) for all of the nodes (dashed blue curve) and a node-absorption configuration with a node-absorption rate of \(\delta_\ast = 0.04\) for homeless nodes and a node-absorption rate of \(\delta_\ast = 0.2\) for non-homeless nodes (solid red curve). We consider Markov times \(t \in (0, 0.05)\).](image)

In Figure 13, we examine Markov times \(t \in (0, 0.05)\) and observe for \(t \in (0, 0.02)\) that the number of communities for the node-absorption configuration with node-absorption-rate vector \(\vec{\delta}(2)\) is larger than the number of communities for the node-absorption configuration with node-absorption-rate vector \(\vec{\delta}(1)\). For example, if we look at the subgraph \(G_0\) that consists of homeless node 92 and the neighbors of node 92, we observe [see Figure 14(a)] that all of the homeless nodes (the smaller nodes) belong to the same community (as indicated by the fact that they have the same color) when the node-absorption rate for the homeless nodes is smaller than that of the non-homeless nodes. By contrast, we see in Figure 14(b) that the homeless nodes are in different communities if the node-absorption rate is the same for all nodes. Therefore, the disease (which we interpret as

\(^8\)The factor of 5 difference in the values of \(\delta_\ast\) and \(\delta\) is plausible for sexually transmitted diseases such as HIV. For example, in one study, Robertson et al. [30] found that approximately one fifth of the HIV-positive-homeless individuals reported using medication. Additionally, if we interpret node-absorption rates as recovery rates, then the use of medication increases the node-absorption rate of a node.
a random walker on the network) may remain longer among the homeless nodes of $G_0$ in Figure 14(a) because these nodes have smaller node-absorption rates but are connected densely to each other.

Figure 14: Intersections of the subgraph $G_0$ (which consists of node 92 and the neighbors of node 92) with the communities of the partition that we obtain using Algorithm 2b with input $P(t|D,\vec{\delta},0)$. In (a), we use the node-absorption-rate vector $\vec{\delta}(1)$; in (b), we use the node-absorption-rate vector $\vec{\delta}(2)$. The small nodes represent the homeless individuals, the numbers 1, \ldots, 2297 are the labels of the nodes, and each color indicates a community. (The node sizes are arbitrary; they are not proportional to the node-absorption rates of the associated nodes.)

The homeless nodes’ large mean degree and smaller node-absorption rate $\delta_{**}$ (in comparison to the node-absorption rate of the non-homeless nodes) produce partitions with fewer communities than in the node-absorption configuration that is associated with $\vec{\delta}(2)$ [see Figure 14(a)]. However, we obtain a rather different community structure in a third scenario, in which 89 non-homeless nodes have a node-absorption rate that is smaller than the node-absorption rate of the rest of the nodes. More specifically, consider a node-absorption configuration with a node-absorption vector of $\vec{\delta}(3)$ with a node-absorption-rate of $\delta_i = \delta_* = 0.2$ for nodes $i$ that include the homeless nodes and a node-absorption rate of $\delta_j = \delta_{**} = 0.04$ for 89 nodes $j$ that we select uniformly at random from the non-homeless nodes. In Figure 15(a), we show that the number of communities that we obtain using Algorithm 2b with input $P(t|D,\vec{\delta},0)$ for the node-absorption configuration with the same node-absorption rates for all nodes (see the dashed blue curve) is larger for $t \in (0,0.05)$ than the number of communities that we obtain for the node-absorption configuration with node-absorption-rate vector $\vec{\delta}(3)$ (see the solid red curve).

We use adjusted mutual information (AMI) [39] to measure the similarity of the partitions in Figure 13(a) and Figure 15(a). See Appendix 8.2 for the definition of AMI. If the AMI between two partitions of the set of nodes of a network is close to 1, then the partitions are similar; if the AMI between two partitions is close to 0, then the partitions are very different from each other. The solid black curve in Figure 15(b) gives the AMI between the partitions that we obtain for node-absorption configurations with node-absorption-rate vectors $\vec{\delta}(1)$ and $\vec{\delta}(2)$ [see Figure 13(a)], and the dashed green curve gives the AMI between the partitions that we obtain for node-absorption configurations with node-absorption-rate vectors $\vec{\delta}(2)$ and $\vec{\delta}(3)$ [see Figure 15(a)]. The AMI is approximately 0 when $0.0034 \approx t \approx 0.012$ for the solid black curve and is approximately 0 when $0.0067 \approx t \approx 0.012$ for the dashed green curve. Therefore, we see that the smaller node-absorption rates of the homeless nodes have a noticeable effect on a wider range of Markov times than the smaller node-absorption rates of non-homeless nodes.
Figure 15: (a) The number of communities that we obtain using Algorithm 2b with input $P_t(D_\delta, 0, t)$ for a node-absorption configuration with a node-absorption vector of $\vec{\delta}^{(2)}$ (dashed blue curve) and a node-absorption configuration with a node-absorption vector $\vec{\delta}^{(3)}$ (solid red curve). The node-absorption rates of $\vec{\delta}^{(2)}$ are $\delta_* = \delta_{**} = 0.2$ for all nodes. The node-absorption rates of $\vec{\delta}^{(3)}$ are $\delta_* = 0.2$ for a set of nodes that includes the homeless nodes and $\delta_{**} = 0.04$ for 89 non-homeless nodes that we select uniformly at random from the set of non-homeless nodes. (b) The solid black curve gives the adjusted mutual information (AMI) between the partitions that we obtain from node-absorption configurations with node-absorption-rate vectors of $\vec{\delta}^{(1)}$ and $\vec{\delta}^{(2)}$. The dashed green curve gives the AMI between the partitions that we obtain from node-absorption configurations with node-absorption-rate vectors of $\vec{\delta}^{(2)}$ and $\vec{\delta}^{(3)}$.

6.2.2 Syphilis dynamics

We now compare the effects of the node-absorption configurations that are associated with $\vec{\delta}^{(1)}$, $\vec{\delta}^{(2)}$, and $\vec{\delta}^{(3)}$ on the number of syphilis infections that we obtain from a stochastic model of syphilis on the Atlanta Urban Network. The possible states of the nodes in the model are susceptible ($S$); exposed ($E$), which is a non-infectious, incubating state; primary and secondary syphilis ($I_1$ and $I_2$, respectively), which are infectious; early latent and late latent syphilis ($L_1$ and $L_2$, respectively), which are non-infectious; and treated individuals ($T_1$, $T_2$, and $T_3$). In Figure 16, we show the possible transitions between the states (i.e., compartments) of the model. See Algorithm 4 in Appendix 8.3 for more details.

Figure 16: Possible state transitions of the nodes in the stochastic process in Algorithm 4.

We explore three scenarios that have different treatment rates and thus have different associated node-absorption rates. In the first scenario (‘scenario 1’), the homeless nodes have a treatment rate of $\alpha_1/5$ if they are in state $I_1$, a treatment rate of $\alpha_2/5$ if they are in state $I_2$, a treatment rate of $\alpha_3/5$ if they are in state $L_1$, and $\alpha_4/5$ if they are in state $L_2$. The non-homeless nodes have a treatment rate of $\alpha_5/5$ if they are in state $E$. In scenario 2, the homeless nodes have a treatment rate of $\alpha_6/5$ if they are in state $I_1$, $\alpha_7/5$ if they are in state $I_2$, $\alpha_8/5$ if they are in state $L_1$, and $\alpha_9/5$ if they are in state $L_2$. The non-homeless nodes have a treatment rate of $\alpha_{10}/5$ if they are in state $E$. In scenario 3, the homeless nodes have a treatment rate of $\alpha_{11}/5$ if they are in state $I_1$, $\alpha_{12}/5$ if they are in state $I_2$, $\alpha_{13}/5$ if they are in state $L_1$, and $\alpha_{14}/5$ if they are in state $L_2$. The non-homeless nodes have a treatment rate of $\alpha_{15}/5$ if they are in state $E$. The treatment rates for $T_1$, $T_2$, and $T_3$ are not specified in the scenario descriptions.
and a treatment rate of $\alpha_4/5$ if they are in state $L_2$. Additionally, the non-homeless nodes have a treatment rate of $\alpha_1$ if they are in state $I_1$, a treatment rate of $\alpha_2$ if they are in state $I_2$, a treatment rate of $\alpha_3$ if they are in state $L_1$, and a treatment rate of $\alpha_4$ if they are in state $L_2$. (See the values of $\alpha_i$ in Table 2 of Appendix 8.3.)

In other words, the treatment rates of node $k$ are $\alpha_i \delta_k^{(1)}/\delta_*$ (with $i = \{1, 2, 3, 4\}$), so this scenario corresponds to the node-absorption configuration that is associated with $\delta^{(1)}$. In the second scenario (‘scenario 2’), the treatment rates of all nodes $k$ is $\alpha_i \delta_k^{(2)}/\delta_*$ with $i \in \{1, 2, 3, 4\}$. In the third scenario (‘scenario 3’), the treatment rates of 89 non-homeless nodes $k$ (which we choose uniformly at random) are $\alpha_i \delta_k^{(3)}/\delta_*$ = $\alpha_i/5$ with $i \in \{1, 2, 3, 4\}$, and the treatment rates for all other nodes $j$ are $\alpha_i \delta_j^{(3)}/\delta_*$ = $\alpha_i$ (with $i \in \{1, 2, 3, 4\}$). We run 500 simulations of the syphilis model in each of scenarios 1, 2, and 3 for 120 months and two different types of initial conditions (ICs). In the first type of IC, we select a homeless node uniformly at random and start it in the exposed state. In the second type of IC, we select a non-homeless node uniformly at random and start it in the exposed state. In both types of ICs, all other nodes are in the susceptible state.

In Figure 17(a,b), we show histograms of the total number of new infections (i.e., the newly exposed nodes) for the three different scenarios with each of the two types of IC. If the initial exposed node is non-homeless, we observe larger outbreaks than if it is homeless. In Figure 17(a), we see in all scenarios that less than 40% of the simulations have fewer than 300 infections when the initial exposed individual is homeless. We say that an outbreak is ‘small’ if it has fewer than 300 infections. By contrast, in Figure 17(b), we see in all scenarios that more than 70% of the simulations have fewer than 300 infections when the initial exposed node is non-homeless.

The above observations occur because the large mean degree of the homeless nodes that are initially exposed leads to fewer small outbreaks. We also see in Figures 17(a,b) that scenario 1 has fewer new infections on average than the other two scenarios. This is the case because more nodes with a small treatment rate and large mean degree enter the latent compartments in scenario 1 [see Figure 17(c)], so the spread of the infection is hindered because new infections arise only from contacts with nodes in the primary and secondary syphilis states. The smaller mean outbreak size in scenario 1 than in scenario 2, despite the fact that the overall treatment is better in scenario 2, agrees with results in studies such as the one by Tuite and Fisman [37], who observed numerically that syphilis incidence at steady state increases with the treatment rate unless the treatment rate is large enough.

In Figures 17(a,b), we observe that scenarios 2 and 3 have similar distributions of the number of infections. This is the case because the nodes with the smallest treatment rate in scenario 3 (these are non-homeless nodes) have a smaller mean degree than that of the homeless nodes and the homeless nodes in scenario 2 have the same treatment rates as the homeless nodes in scenario 3.
Figure 17: Histograms of (a,b) the total number of infections and (c) the total number of latent infections from 500 runs of the syphilis model in Algorithm 4 on the Atlanta Urban Network. In the input of Algorithm 4, we use the rates $\beta, \gamma_0, \gamma_1, \gamma_2, \gamma_3, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \lambda_1, \lambda_2, \lambda_3$ from Table 2 and we simulate for $T = 10$ years. (See Appendix 8.3 for a detailed discussion of the model.) In each of our simulations, a single node starts in the exposed state and all other nodes start in the susceptible state. In panels (a) and (b), we show the histograms of the total number of new infections. In scenario 1, homeless nodes have a smaller treatment rate than the other nodes. In scenario 2, all nodes have the same treatment rate. In scenario three, the 89 non-homeless nodes that we choose uniformly at random have a smaller treatment rate than that of the other nodes. In (c), we show the histograms of the total number of nodes with latent syphilis that are homeless in scenario 1 (light gray bars) and non-homeless with a small treatment rate in scenario 3 (black bars).

7 Conclusions and Discussion

Random walks are one of the most fundamental dynamical processes on networks, and many studies have used random walks to gain insights both into network structure and into how network structure affects dynamical processes [19]. Much research has focused on standard random walks (which are regular Markov chains), and it is important to understand the relationships between network structure and different types of random walks. In particular, absorbing random walks arise naturally in many applications (including disease dynamics), and they have been used to develop centrality measures [10], rank nodes [40], and identify social circles [7].

We extended InfoMap, which is a popular method for detecting communities in networks, to use absorbing random walks (which are not regular) instead of standard random walks. We based our extension on the idea of looking at associated regular random walks on a family of absorption-scaled graphs [11]. Our results demonstrate that heterogeneous node-absorption rates of the nodes in a network lead to very different effective community structures than when one does not incorporate the node-absorption rates of nodes.

Part of our extension of InfoMap involves a weight matrix $H$ that allows one to tune the relative importances of the edge weights and the node-absorption rates when detecting communities. Each choice of $H$ is associated with a corresponding absorption-scaled graph. For example, $H = 0$ corresponds to the absorption-scaled graph that arises in connection with the fundamental matrix of the absorbing Markov chain [11]. The choice $H = I$ is also important because it yields connections (1) between standard InfoMap and our extension of InfoMap (see
Propositions 1 and 3) and (2) between the associated absorption-scaled graph and the absorbing random walk. (See Propositions 2, 8, and 9.) We also related the absorption inverses of the unnormalized graph Laplacians for \( H = 0 \) and \( H = I \) to each other through a fundamental matrix for the regular Markov chain on the graph without absorption (see Proposition 9). In future work, it is worth further exploring the relationships between the absorption-scaled graphs for \( H = 0 \) and \( H = I \). It will also be worthwhile to develop systematic criteria for for choosing \( H \) for community detection in practice.

In the present paper, we defined a a map \( L^{(a)} \) for absorbing random walks. The map \( L^{(a)} \) is the map that is associated with Algorithm 2a for \( H = I \) under the assumptions of Proposition 3. Unlike the map function of Algorithm 1 (which is associated with the stationary distribution of a regular Markov chain), the map function \( L^{(a)} \) is not associated with the stationary distribution of an absorbing random walk. (Such a stationary distribution is trivial; the probability of the absorbing state is 1, and all other states have probability 0.) To define \( L^{(a)} \), we chose a distribution \( \pi^{(a)}_i \) from which we determine the initial state. We suggest that \( \pi^{(a)}_i \) is a reasonable choice because it consists of the probabilities of visiting a state before absorption and it converges to the stationary distribution of the associated regular random walk with no absorption as the node-absorption rates tend to 0.

Our extensions of InfoMap in Algorithm 2a and Algorithm 2b use a Markov-time parameter like the Markov time sweeping that was used in [34]. We observed that varying the Markov-time parameter reveals differences in node-absorption rates that uncover effective community structure. It is desirable to develop criteria for choosing Markov times. In this paper, we used an ad hoc approach. Specifically, we plotted the number of communities versus Markov time and sought intervals in which the plots are flat with respect to the number of communities for Markov times that are close to 0. However, as one can see in Section 6, it may not be easy to identify intervals in which the plots are flat (see Figure 12). We do not have a generic criterion for choosing a single partition, and it is desirable to try to develop one.

When studying disease dynamics, our extensions of InfoMap are relevant for examining relationships between epidemiological quantities and effective community structure. Heterogeneous node-absorption rates of nodes can help shape effective community structure and impact disease dynamics. For example, we observed that simultaneously increasing the node-absorption rates of nodes that connect tightly-knit sets (specifically, the graphs with ring-lattice communities in Algorithm 3) and increasing transmission rates in other nodes (in order to preserve the basic reproduction number) leads to partitions with more communities and that these increases have a similar effect on epidemiological quantities as rewiring edges [33]. One example of this similarity is a peak in the outbreak duration when a moderate number (specifically, 29 of 240) of the nodes have larger node-absorption rates than the other nodes [see Stage 29 in Figure 11(a)]. A particular concern is situations in which vulnerable populations (e.g., people who are homeless) have lower treatment rates (e.g., due to barriers to healthcare access). By examining the spread of a model disease on the sexual-contact network data from the HIV Transmission Network Metastudy Project [20], we illustrated that lower treatment rates in homeless individuals can result in larger effective communities than in the communities that one obtains from considering only network structure. These low treatment rates in homeless individuals and associated effective community structure correspond to a smaller total outbreak size than in other treatment-rate scenarios. This finding may seem surprising, but it is consistent with the results of [37]. It is also supported by our computations that there are more latent infections when homeless nodes have smaller treatment rates than the other nodes than is the case in our other scenarios (see Figure 17). This observation suggests that the large mean degree of homeless nodes combines with the small node-absorption rates to impact effective community structure and the number of infections.

8 Appendix

8.1 Proofs of the propositions in Section 4.2

We now prove Propositions 4, 8, and 9 from Section 4.2.

We start with Proposition 4.

**Proposition 4.** Let \( P_0 = AW^{-1} \) be the transition-probability matrix of the discrete-time Markov chain that is associated with \( \tilde{L}(D_0, 0) \), and let \( P_1 = (A + D)(W + D)^{-1} \) be the transition-probability matrix of the discrete-time Markov chain that is associated with \( \tilde{L}(D_1, 1) \). Let \( \pi \) and \( \pi' \) be the stationary distributions that are associated with \( P_0 \) and \( P_1 \), respectively. Let \( Z_i \) be the fundamental matrix that is associated with \( P_i \) (for \( i \in \{1, 2\} \)). Let \( U := (1/(\delta^T \bar{u}))\bar{u}I^T \) and \( \alpha := \delta^T \bar{u}/(\bar{w}^T \bar{u} + \delta^T \bar{u}), \) where \( \bar{u} = (u_1, \ldots, u_n)^T \) is a vector in \( \text{Ker} \, \tilde{L}(D_0, 0) \) with non-negative entries such that \( \sum_{i=1}^n u_i = 1 \). We have that
\[ Z_1 = W^{-1}(W + D_\delta) \left[ \sum_{i} + \alpha(1 - \alpha)\pi \bar{T} \right] - \alpha \left[ Z_0D_\delta U + W \frac{\bar{u} \bar{u}^T}{\bar{u}^T \bar{u}} W^{-1}Z_0(I - \alpha D_\delta U) \right] . \] (30)

**Proof.** Let \( \bar{u} \) be a column vector in \( \ker L \) whose entries \( u_i \) are all positive and sum to 1. It then follows that \( \pi = W\bar{u} / (u^T \bar{u}) \), \( \pi' = (W + D_\delta)\bar{u} / (u^T \bar{u} + \bar{u}^T \bar{u}) \), and
\[ \pi' = (1 - \alpha)\pi + \frac{D_\delta \bar{u}}{\bar{u}^T \bar{u}} . \] (43)

Additionally,
\[ Z_1 = \left( I - P_1 + \pi' \bar{T} \right)^{-1} , \]
which we can write as
\[ Z_1 = W^{-1}(W + D_\delta) \left[ I - P_0 + \pi \bar{T} - \alpha \pi \bar{T} + \frac{D_\delta \bar{u}}{\bar{u}^T \bar{u}} \bar{u}^T + \left( (1 - \alpha)\pi + \frac{D_\delta \bar{u}}{\bar{u}^T \bar{u}} \right) \bar{T}D_\delta W^{-1} \right]^{-1} . \] (44)

We now use the Sherman–Morrison formula to compute the inverse in (44). The Sherman–Morrison formula states that if \( B \) is a non-singular matrix and \( \bar{v}_1 \) and \( \bar{v}_2 \) are column vectors such that \( 1 + \bar{v}_1^T B \bar{v}_2 \neq 0 \), then
\[ \left( B + \bar{v}_1 \bar{v}_2^T \right)^{-1} = B^{-1} - \frac{B^{-1} \bar{v}_1 \bar{v}_2^T B^{-1}}{1 + \bar{v}_2^T B^{-1} \bar{v}_1} . \] (45)

Define
\[
\begin{align*}
B_0 &:= I - P_0 + \pi \bar{T} , \\
B_1 &:= I - P_0 + \pi \bar{T} - \alpha \pi \bar{T} , \\
B_2 &:= I - P_0 + \pi \bar{T} - \alpha \pi \bar{T} + \frac{D_\delta \bar{u}}{\bar{u}^T \bar{u}} \bar{T} , \\
B_3 &:= I - P_0 + \pi \bar{T} - \alpha \pi \bar{T} + \frac{D_\delta \bar{u}}{\bar{u}^T \bar{u}} \bar{T} + \left( (1 - \alpha)\pi + \frac{D_\delta \bar{u}}{\bar{u}^T \bar{u}} \right) \bar{T}D_\delta W^{-1} .
\end{align*}
\]

Using the properties in (29) and the formula (45), we obtain
\[
\begin{align*}
B_1^{-1} &= \left( B_0 - \alpha \pi \bar{T} \right)^{-1} = Z_0 + \frac{\alpha}{1 - \alpha} \pi \bar{T} , \\
B_2^{-1} &= \left( B_1 + \alpha \frac{D_\delta \bar{u}}{\bar{u}^T \bar{u}} \bar{T} \right)^{-1} = Z_0 + \alpha \pi \bar{T} - \alpha Z_0 D_\delta U , \\
B_3^{-1} &= \left( B_2 + \left( (1 - \alpha)\pi + \alpha \frac{D_\delta \bar{u}}{\bar{u}^T \bar{u}} \right) \bar{T}D_\delta W^{-1} \right)^{-1} = Z_0 + \alpha(1 - \alpha)\pi \bar{T} - \alpha \left[ Z_0D_\delta U + W \frac{\bar{u} \bar{u}^T}{\bar{u}^T \bar{u}} W^{-1}Z_0(I - \alpha D_\delta U) \right] .
\end{align*}
\] (47)

From equation (44), it follows that \( Z_1 = W^{-1}(W + D_\delta)B_3^{-1} \). Combining this relation with (47) yields (30). \( \square \)

We now prove Proposition 8.

**Proposition 8.** Let \( \tilde{L}_1 := \tilde{L}(D_\delta, I) = (W - A)(W + D_\delta)^{-1} \), and let \( \tilde{d}_1 \) be the diagonal of \( D_\delta(W + D_\delta)^{-1} \). Let \( U := \bar{u} \bar{T} / (\bar{u}^T \bar{u}) \), \( U_1 := (W + D_\delta)U \), and \( D_1 := D_\delta(W + D_\delta)^{-1} \). We have that
\[ \tilde{L}_1 \tilde{d}_1 = (W + D_\delta)\tilde{L} , \] (38)
and
\[ (\mathcal{L} + D_\delta)^{-1} = (W + D_\delta)^{-1} \left( U_1 + (I + (L_1)^{-1} D_1)^{-1} \tilde{L}_1 \tilde{d}_1 \right) . \] (39)
Proof. The adjacency matrix $A_1 = P_1 = (A + D_b)(W + D_b)$ has the associated unnormalized graph Laplacian

$$\tilde{L}_1 = (W - A)(W + D_b)^{-1}.$$  

From Proposition 4, it follows that

$$Z_1 = W^{-1}(W + D_b)[Z_0 + \alpha R],$$  

where

$$R = (1 - \alpha)\pi \tilde{L}^T - Z_0 D_b U - WUD_b W^{-1}Z_0(I - \alpha D_b U).$$

Proposition 7 then implies that

$$\tilde{L}^d_1 = (I - U_1 D_1)Z_1(I - D_1 U_1).$$

Additionally,

$$I - U_1 D_1 = (W + D_b)(I - UD_b)(W + D_b)^{-1},$$

$$I - D_1 U_1 = I - UD_b.$$  

Substituting (50) into (49) yields

$$\tilde{L}^d_1 = (W + D_b)\tilde{L}^d + \alpha(W + D_b)(I - UD_b)W^{-1}R(I - D_b U).$$

Using the relations $\pi \tilde{L}^T(I - D_b U) = 0$, $D_b U(I - D_b U) = 0$, and $(I - \alpha D_b U)(I - D_b U) = I - D_b U$ yields $R(I - D_b U) = -WUD_b W^{-1}Z_0(I - D_b U)$. We then use the fact that $(I - UD_b)UD_b = 0$ to obtain

$$(I - UD_b)W^{-1}R(I - D_b U) = 0.$$  

Substituting (52) into the right-hand side of (51) yields (38).

We express the fundamental matrix $(\mathcal{L} + D_b)^{-1}$ as

$$(\mathcal{L} + D_b)^{-1} = (W + D_b)^{-1}((W - A)(W + D_b)^{-1} + D_b(W + D_b)^{-1})^{-1} = (W + D_b)(\tilde{L}_1 + D_1)^{-1}.$$  

By Proposition 6 and (38), we have

$$(\tilde{L}_1 + D_1)^{-1} = U_1 + (I + \tilde{L}^d_1 D_1)^{-1} \tilde{L}^d_1.$$  

From (53) and (54), we obtain (39).

We now prove Proposition 9.

**Proposition 9.** Let $\tilde{d} := \tilde{d}_b(D_b, I) = \tilde{w} + \tilde{\delta} = (\omega_1 + \delta_1, \ldots, \omega_n + \delta_n)^T$ be the scaled rate vector that is associated with the absorption-scaled graph $\tilde{G}(D_b, I)$. With $\mathcal{L} = W - A$, $\alpha := \tilde{\delta}^T \tilde{u}/(\tilde{w}^T \tilde{u} + \tilde{\delta}^T \tilde{u})$, $\pi = W\tilde{u}/(\tilde{w}^T \tilde{u})$, $P_0 = AW^{-1}$, $Z_0 = (I - P_0 + \pi \tilde{L}^T)^{-1}$, and $Z_* = W^{-1}(Z_0 - \pi \tilde{L}^T)$, it follows that

$$\mathcal{L}^{\tilde{d}} = \alpha^2 \mathcal{L}^\delta + \alpha(1 - \alpha)(\mathcal{L}^\delta Z_* + Z_* \mathcal{L}^\delta) + (1 - \alpha)^2 Z_*.$$  

**Proof.** By Proposition 7, we have

$$\mathcal{L}^{\tilde{d}} = (I - \frac{\tilde{u}^T}{\tilde{w}^T \tilde{u} + \tilde{\delta}^T \tilde{u}}(W + D_b))W^{-1}Z_0(I - \frac{1}{\tilde{w}^T \tilde{u} + \tilde{\delta}^T \tilde{u}}(W + D_b)\tilde{u}^T).$$

Furthermore,

$$I - \frac{\tilde{u}^T}{\tilde{w}^T \tilde{u} + \tilde{\delta}^T \tilde{u}}(W + D_b) = \alpha(I - \frac{\tilde{u}^T}{\tilde{\delta}^T \tilde{u}} D_b) + (1 - \alpha)W^{-1}(I - \pi \tilde{L}^T)W,$$

$$I - \frac{1}{\tilde{w}^T \tilde{u} + \tilde{\delta}^T \tilde{u}}(W + D_b)\tilde{u}^T = \alpha(I - D_b \frac{\tilde{u}^T}{\tilde{\delta}^T \tilde{u}}) + (1 - \alpha)(I - \pi \tilde{L}^T).$$
Substituting (56) into (55) yields
\[
\mathcal{L}^\delta = \alpha^2 \left( I - \frac{\mathbf{u}^T}{\delta^T} \mathbf{D}_\delta \right) W^{-1} Z_0 \left( I - D_\delta \frac{\mathbf{u}^T}{\delta^T} \right) + \alpha(1-\alpha) \left( I - \frac{\mathbf{u}^T}{\delta^T} \mathbf{D}_\delta \right) W^{-1} Z_0 (I - \pi \mathbf{I})
\]
\[
+ \alpha(1-\alpha)W^{-1}(I - \pi \mathbf{I})W^{-1} Z_0 \left( I - D_\delta \frac{\mathbf{u}^T}{\delta^T} \right) + (1-\alpha)^2 W^{-1}(I - \pi \mathbf{I})W^{-1} Z_0 (I - \pi \mathbf{I}).
\]

(57)

By Proposition 6, the first term of the right-hand side of (57) is \(\mathcal{L}^\delta\). Note that \(\pi\) is the stationary distribution of the Markov chain with transition-probability matrix \(P_0\). Additionally, \(Z_0\) is the fundamental matrix of this Markov chain. Therefore, from (29), it follows that \(W^{-1} Z_0 (I - \pi \mathbf{I}) = W^{-1}(I - \pi \mathbf{I}) Z_0 = W^{-1}(Z_0 - \pi \mathbf{I}) = Z_0\) and \(W^{-1}(I - \pi \mathbf{I})Z_0 (I - \pi \mathbf{I}) = W^{-1}(Z_0 - \pi \mathbf{I}) = Z_0\). Consequently, from (57), we obtain
\[
\mathcal{L}^\delta = \alpha^2 \mathcal{L}^\delta + \alpha(1-\alpha) \left[ (I - \frac{\mathbf{u}^T}{\delta^T} \mathbf{D}_\delta \right) Z_0 + Z_0 \left( I - D_\delta \frac{\mathbf{u}^T}{\delta^T} \right) + (1-\alpha)^2 Z_0.
\]
\[
(58)
\]

Additionally, \(\mathcal{L}^\delta \mathcal{L} = I - \frac{\mathbf{u}^T}{\delta^T} \mathbf{D}_\delta\) and \(\mathcal{L}^\delta \mathcal{L} = I - D_\delta \frac{\mathbf{u}^T}{\delta^T}\) (see Theorem 1 and Lemma 1 in [11]), so (41) follows from (58).

8.2 Adjusted Mutual Information

To compare the partitions in Section 6.2, we calculated adjusted mutual information (AMI). We present AMI in Definition 10.

**Definition 10. (Adjusted Mutual Information (AMI))** Let \(U = \{U_1, \ldots, U_{n_1}\}\) and \(V = \{V_1, \ldots, V_{n_2}\}\) be partitions of a set of \(n\) elements. For any set \(S\), let \(|S|\) denote the size of \(S\). We define the following quantities:

- **Entropy:** \(H(U) = -\sum_{i=1}^{n_1} (|U_i|/n) \log_2 (|U_i|/n)\).

- **Conditional entropy:** \(H(U|V) = -\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} (|U_i \cap V_j|/n) \log_2 (|U_i \cap V_j|/n)\).

- **Mutual information (MI):** \(I(U, V) = H(U) - H(U|V)\).

We then define the AMI by
\[
\text{AMI}(U, V) = \frac{I(U, V) - \mathbb{E}(I(U, V))}{(H(U) + H(V))/2 - \mathbb{E}(I(U, V))},
\]
where \(\mathbb{E}(I(U, V))\) is the expected mutual information between partitions \(U\) and \(V\) that we choose uniformly at random, given a fixed number of communities in each partition and fixed community sizes. Specifically,
\[
\mathbb{E}(I(U, V)) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left( \min\{a_i, b_j\} \sum_{n_{ij} = \max\{a_i + b_j - n, 0\}}^{\min\{a_i, b_j\}} \frac{n_{ij}}{n} \log_2 \frac{n \cdot n_{ij}}{a_i b_j} \frac{a_i b_j! (n - a_i)! (n - b_j)!}{n! n_{ij}! (a_i - n_{ij})! (b_j - n_{ij})! (n - a_i - b_j + n_{ij})!} \right),
\]
where \(a_i := \sum_{j=1}^{n_2} |U_i \cap V_j|\) and \(b_j := \sum_{i=1}^{n_1} |U_i \cap V_j|\).

As discussed in [39], the following properties are desirable for a similarity measure:

- **Metric property:** The similarity measure satisfies positive definiteness, symmetry, and the triangle inequality.

- **Normalization:** The similarity measure has values between 0 and 1, where 1 signifies a perfect match.

- **Constant-baseline property:** The expected value of the similarity measure between pairs of partitions that are sampled independently and uniformly at random is equal to 0.

In contrast to some similarity measures, such as normalized mutual information (NMI) [39], AMI does not satisfy the metric property. However, AMI satisfies the normalization property (which is also satisfied by NMI) and the constant-baseline property (which is not satisfied by NMI) [39].

We use AMI because it satisfies the constant-baseline property, which is relevant in our applications because it avoids bias in partitions that are sampled independently. (A similarity measure is “biased” if its expected value between pairs of partitions that are sampled independently and uniformly at random is nonzero.)
8.3 A Model for the Spread of Syphilis

We now present the compartmental network model for syphilis that we employed in Section 6.2. We simulate this model using the Gillespie algorithm in Algorithm 4, where the input graph $G$ is the Atlanta Urban Network. The possible states of the nodes in the model are susceptible ($S$); exposed ($E$), which is a non-infectious, incubating state; primary syphilis ($I_1$), which is infectious; secondary syphilis ($I_2$), which is infectious; early latent syphilis ($L_1$), which is non-infectious; late latent syphilis ($L_2$), which is non-infectious; treated $I_1$ and $I_2$ individuals ($T_3$); treated $L_1$ individuals ($T_1$); and treated $L_2$ individuals ($T_2$). Let $|Y|$ denote the number of nodes in state $Y \in \{S, E, I_1, I_2, L_1, L_2, T_1, T_2, T_3\}$, and let $|SI|$ denote the number of unordered node pairs with one node in state $S$ and the other node in state $I_1$ or $I_2$. In the first column of Table 2 (also see Figure 16), we show the possible state transitions for the syphilis model in Algorithm 4. The second column of Table 2 gives the rates of the Gillespie algorithm. (See step 5 of Algorithm 4.) The third column of Table 2 gives the values of the rate parameters in Algorithm 4. (We take these values from [37].) The treatment rates for each node are either $\alpha_i$ or $\alpha_i/5$ for $i = \{1, 2, 3, 4\}$, depending on the scenario that we examine.

| Transition | Rate | Parameters (per month) |
|------------|------|------------------------|
| $S \rightarrow E$ | $\beta|SI|$ | $\beta = 0.6 \times (2.4/12)$ |
| $E \rightarrow I_1$ | $\gamma_0|E|$ | $\gamma_0 = 1/0.9$ |
| $I_1 \rightarrow I_2$ | $\gamma_1|I_1|$ | $\gamma_1 = (1/1.5) \times 0.85$ |
| $I_2 \rightarrow L_1$ | $\gamma_2|I_2|$ | $\gamma_2 = (1/3.6) \times 0.75$ |
| $L_1 \rightarrow L_2$ | $\gamma_3|L_1|$ | $\gamma_3 = (1/6.9) \times 0.75$ |
| $I_1 \rightarrow T_3$ | $\alpha_1|I_1|$ | $\alpha_1 = [(1/1.5) \times 0.15] + 0.001/12 + 0.0578$ |
| $I_2 \rightarrow T_3$ | $\alpha_2|I_2|$ | $\alpha_2 = [(1/3.6) \times 0.25] + 0.001/12 + 0.0578$ |
| $L_1 \rightarrow T_1$ | $\alpha_3|L_1|$ | $\alpha_3 = [(1/6.9) \times 0.25] + 0.001/12 + 0.0578$ |
| $L_2 \rightarrow T_2$ | $\alpha_4|L_2|$ | $\alpha_4 = 0.001/12 + 0.0578$ |
| $T_3 \rightarrow S$ | $\lambda_3|T_3|$ | $\lambda_3 = 1/0.25$ |
| $T_1 \rightarrow S$ | $\lambda_1|T_1|$ | $\lambda_1 = 1/0.25$ |
| $T_2 \rightarrow S$ | $\lambda_2|T_2|$ | $\lambda_2 = 1/60$ |

Table 2: Transition rates in the model for the spread of syphilis. The treatment rates are $\alpha_i$ or $\alpha_i/5$, depending on the parameter configuration. We consider different parameter configurations in scenarios 1, 2, and 3. We take the parameter values from [37]. The parameter values give monthly rates.
Algorithm 4 Gillespie algorithm for simulating the Syphilis model

Input: Rates $\beta$, $\gamma_0$, $\gamma_1$, $\gamma_2$, $\gamma_3$, $\alpha_1$, $\alpha_2$, $\alpha_3$, $\alpha_4$, $\lambda_1$, $\lambda_2$, and $\lambda_3$, an undirected and unweighted graph, a subset $Y_0$ of the set of nodes of $G$, a final time value $T$.
Output: A matrix $N_{\text{state}}$.

1: Denote the state of a node $i$ of $G$ by $s(i)$. The possible states of the nodes are $S$, $E$, $I_1$, $I_2$, $L_1$, $L_2$, $T_1$, $T_2$, and $T_3$. Set $s(i) = S$ for each node $i$ in $G$.
2: Select a node $i_0$ uniformly at random from the nodes in $Y_0$ and set $s(i_0) = E$. Set the initial time to be $t = 0$, and set $N_{\text{state}}$ to be an empty array.
3: while $t < T$ do
4: For each state $Y \in \{S, E, I_1, I_2, L_1, L_2, T_1, T_2, T_3\}$, define $Y$ as the set of nodes of $G$ with state $Y$. Let $|Y|$ denote the size of $Y$. Define $\text{SI}$ as the set of node pairs $\{j_1, j_2\}$ of $G$ such that (1) $s(j_1) = S$ and (2) either $s(j_2) = I_1$ or $s(j_2) = I_2$. Let $|\text{SI}|$ denote the number of nodes in $\text{SI}$.
5: Define the rates

\[
\begin{align*}
\lambda_{SE} &:= \beta|\text{SI}|, \\
\lambda_{L_1L_2} &:= \gamma_3|L_1|, \\
\lambda_{L_2T_2} &:= \alpha_4|L_2|, \\
\lambda_{I_1I_2} &:= \gamma_1|I_1|, \\
\lambda_{I_1T_3} &:= \alpha_1|I_1|, \\
\lambda_{I_2T_3} &:= \alpha_2|I_2|, \\
\lambda_{T_1S} &:= \alpha_1|T_1|, \\
\lambda_{T_2S} &:= \alpha_2|T_2|, \\
\lambda_{I_2L_1} &:= \gamma_2|I_2|, \\
\lambda_{L_1T_1} &:= \alpha_3|L_1|, \\
\lambda_{T_1S} &:= \lambda_3|T_3|. \\
\end{align*}
\]

We then define

\[
\lambda_{\text{tot}} := \lambda_{SE} + \lambda_{E1} + \lambda_{I_1I_2} + \lambda_{I_2L_1} + \lambda_{I_1L_2} + \lambda_{I_1T_3} + \lambda_{I_2T_3} + \lambda_{L_1T_1} + \lambda_{L_2T_2} + \lambda_{T_1S} + \lambda_{T_2S} + \lambda_{T_3S}.
\]

6: Select a value $\Delta t > 0$ from an exponential distribution with rate $\lambda_{\text{tot}}$. Update $t$ to $t + \Delta t$.
7: Select a value $u$ uniformly at random from the interval $(0, 1)$.
8: Define the vector

\[
\Lambda := \frac{1}{\lambda_{\text{tot}}} (\lambda_{SE}, \lambda_{E1}, \lambda_{I_1I_2}, \lambda_{I_2L_1}, \lambda_{I_1L_2}, \lambda_{I_1T_3}, \lambda_{I_2T_3}, \lambda_{L_1T_1}, \lambda_{L_2T_2}, \lambda_{T_1S}, \lambda_{T_2S}, \lambda_{T_3S})^T.
\]

Define the following ordered lists of states:

\[
Y^{\text{old}} := (E, I_1, I_2, L_1, I_1, I_2, L_1, L_2, T_1, T_2, T_3), \\
Y^{\text{new}} := (I_1, I_2, L_1, L_2, T_3, T_3, T_1, T_2, S, S, S).
\]

9: if $0 < u < \Lambda_1 = \lambda_{SE}/\lambda_{\text{tot}}$ then
10: Select a pair $(j_1, j_2)$ uniformly at random from $\text{SI}$ such that $s(j_1) = S$. Set $s(j_1) = E$.
11: else
12: Find the integer $k \in \{1, \ldots, 11\}$ such that $\sum_{i=1}^k \Lambda_i < u < \sum_{i=1}^{k+1} \Lambda_i$. Select a node $l$ uniformly at random from $Y^{\text{old}}_k$ and set $s(l) = Y^{\text{new}}_k$.
13: end if
14: Append the vector

\[
(t, |S|, |E|, |I_1|, |I_2|, |L_1|, |L_2|, |T_1|, |T_2|, |T_3|)
\]

as a new row of $N_{\text{state}}$.
15: end while

Data and code availability

We have posted the code that yields the numerical results in Figures 11, 12, 13, 14, 15, and 17 of Section 6 at https://gitlab.com/esteban_vargas_bernal/extending-infomap-to-absorbing-random-walks.

Acknowledgements

This work was supported by the National Science Foundation (through DMS 1814737 and DMS 1440386) and by the Fulbright International Program.
References

[1] M. Benzi, P. Fika, and M. Mitrouli. Graphs with absorption: Numerical methods for the absorption inverse and the computation of centrality measures. Linear Algebra and its Applications, 574:123–152, 2019.

[2] V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment, 2008(10):P10008, 2008.

[3] L. Bohlin, D. Edler, A. Lancichinetti, and M. Rosvall. Community detection and visualization of networks with the Map Equation framework. In Measuring Scholarly Impact, pages 3–34. Springer-Verlag, Heidelberg, Germany, 2014.

[4] F. Brauer, C. Castillo-Chavez, and Z. Feng. Mathematical Models in Epidemiology. Springer-Verlag, Heidelberg, Germany, 2019.

[5] S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. Computer Networks and ISDN Systems, 30(1-7):107–117, 1998.

[6] T. M. Cover and J. A. Thomas. Elements of Information Theory. John Wiley & Sons, Hoboken, United States, 2012.

[7] J. De, X. Zhang, F. Lin, and L. Cheng. Transduction on directed graphs via absorbing random walks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(7):1770–1784, 2017.

[8] J. C. Delvenne, S. N. Yaliraki, and M. Barahona. Stability of graph communities across time scales. Proceedings of the National Academy of Sciences, 107(29):12755–12760, 2010.

[9] S. Fortunato and D. Hric. Community detection in networks: A user guide. Physics Reports, 659:1–44, 2016.

[10] A. J. Gurfinkel and P. A. Rikvold. Absorbing random walks interpolating between centrality measures on complex networks. Physical Review E, 101(1):012302, 2020.

[11] K. A. Jacobsen and J. H. Tien. A generalized inverse for graphs with absorption. Linear Algebra and its Applications, 537:118–147, 2018.

[12] L. G. S. Jeub, P. Balachandran, M. A. Porter, P. J. Mucha, and M. W. Mahoney. Think locally, act locally: Detection of small, medium-sized, and large communities in large networks. Physical Review E, 91(1):012821, 2015.

[13] J. G. Kemeny and J. L. Snell. Finite Markov Chains: With a New Appendix “Generalization of a Fundamental Matrix”. Springer-Verlag, Heidelberg, Germany, 1983.

[14] M. Kheirkhahzadeh, A. Lancichinetti, and M. Rosvall. Efficient community detection of network flows for varying Markov times and bipartite networks. Physical Review E, 93(3):032309, 2016.

[15] István Z. Kiss, Joel C. Miller, and Péter L. Simon. Mathematics of Epidemics on Networks: From Exact to Approximate Models. Springer International Publishing, Cham, Switzerland, 2017.

[16] R. Lambiotte, J. C. Delvenne, and M. Barahona. Laplacian dynamics and multiscale modular structure in networks. arXiv preprint arXiv:0812.1770, 2008.

[17] R. Lambiotte and M. Rosvall. Ranking and clustering of nodes in networks with smart teleportation. Physical Review E, 85(5):056107, 2012.

[18] Sune Lehmann and Yong-Yeol Ahn. Complex Spreading Phenomena in Social Systems: Influence and Contagion in Real-World Social Networks. Springer International Publishing, Cham, Switzerland, 2018.

[19] N. Masuda, M. A Porter, and R. Lambiotte. Random walks and diffusion on networks. Physics Reports, 716:1–58, 2017.

[20] M. Morris and R. Rothenberg. HIV Transmission Network Metastudy Project: An Archive of Data From Eight Network Studies, 1988–2001. Inter-university Consortium for Political and Social Research, Ann Arbor, United States, 2011.
[21] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J. P. Onnela. Community structure in time-dependent, multiscale, and multiplex networks. *Science*, 328(5980):876–878, 2010.

[22] M. E. J. Newman. Spread of epidemic disease on networks. *Physical Review E*, 66(1):016128, 2002.

[23] M. E. J. Newman. *Networks*. Oxford University Press, Oxford, UK, second edition, 2018.

[24] Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. Epidemic processes in complex networks. *Reviews of Modern Physics*, 87(3):925, 2015.

[25] T. P. Peixoto. Bayesian Stochastic Blockmodeling. *Advances in Network Clustering and Blockmodeling*, pages 289–332, 2019.

[26] M. A. Porter. Small-world network. *Scholarpedia*, 7(2):1739, 2012.

[27] M. A. Porter and J. P. Gleeson. Dynamical systems on networks: A tutorial. *Frontiers in Applied Dynamical Systems: Reviews and Tutorials*, 4, 2016.

[28] M. A. Porter, J.-P. Onnela, and P. J. Mucha. Communities in networks. *Notices of the American Mathematical Society*, 56(9):1082–1097, 1164–1166, 2009.

[29] D. Raoult, C. Foucault, and P. Brouqui. Infections in the homeless. *The Lancet Infectious Diseases*, 1(2):77–84, 2001.

[30] M. J. Robertson, R. A. Clark, E. D. Charlebois, J. Tulsky, H. L. Long, D. R. Bangsberg, and A. R. Moss. HIV seroprevalence among homeless and marginally housed adults in San Francisco. *American Journal of Public Health*, 94(7):1207–1217, 2004.

[31] M. Rosvall, D. Axelsson, and C. T. Bergstrom. The map equation. *The European Physical Journal — Special Topics*, 178(1):13–23, 2009.

[32] M. Rosvall and C. T. Bergstrom. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences of the United States of America*, 105(4):1118–1123, 2008.

[33] M. Salathé and J. H. Jones. Dynamics and control of diseases in networks with community structure. *PLoS Computational Biology*, 6(4):e1000736, 2010.

[34] M. T. Schaub, R. Lambiotte, and M. Barahona. Encoding dynamics for multiscale community detection: Markov time sweeping for the map equation. *Physical Review E*, 86(2):026112, 2012.

[35] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.

[36] J. H. Tien, Z. Shuai, M. C. Eisenberg, and P. van den Driessche. Disease invasion on community networks with environmental pathogen movement. *Journal of Mathematical Biology*, 70:1065–1092, 2015.

[37] A. Tuite and D. Fisman. Go big or go home: Impact of screening coverage on syphilis infection dynamics. *Sexually Transmitted Infections*, 92(1):49–54, 2016.

[38] P. van den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180:29–48, 2002.

[39] N. X. Vinh, J. Epps, and J. Bailey. Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance. *Journal of Machine Learning Research*, 11(Oct):2837–2854, 2010.

[40] X. Zhu, A. B. Goldberg, J. Van Gael, and D. Andrzejewski. Improving diversity in ranking using absorbing random walks. In *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Proceedings of the Main Conference*, pages 97–104, 2007.