Diffusion and Consensus on Weakly Connected Directed Graphs

J. J. P. Veerman[*]
Ewan Kummel[*]

September 2, 2018

Abstract

Let G be a weakly connected directed graph with asymmetric graph Laplacian $\mathcal{L}$. Consensus and diffusion are dual dynamical processes defined on G by $\dot{x} = -\mathcal{L}x$ for consensus and $\dot{p} = -p\mathcal{L}$ for diffusion. We consider both these processes as well their discrete time analogues. We define a basis of row vectors $\{\bar{\gamma}_i\}_{i=1}^k$ of the left null-space of $\mathcal{L}$ and a basis of column vectors $\{\gamma_i\}_{i=1}^k$ of the right null-space of $\mathcal{L}$ in terms of the partition of G into strongly connected components. This allows for complete characterization of the asymptotic behavior of both diffusion and consensus — discrete and continuous — in terms of these eigenvectors.

As an application of these ideas, we present a treatment of the pagerank algorithm that is dual to the usual one. We further show that the *teleporting* (see below) feature usually included in the algorithm is not strictly necessary.

Together with [13], this is a complete and self-contained treatment of the asymptotics of consensus and diffusion on digraphs. Many of the ideas presented here can be found scattered in the literature, though mostly outside mainstream mathematics and not always with complete proofs. This paper seeks to remedy this by providing a compact and accessible survey.

1 Introduction

Directed graphs are an important generalization of undirected graphs, because they have wide-ranging applications. Examples include models of the internet [11] and social networks [12], food webs [28], epidemics [23], chemical reaction networks [34], databases [4], communication networks [3], and networks of autonomous agents in control theory [18] to name but a few. In each of these examples, the relation defining adjacency between neighboring vertices is often not symmetric. Such

---

[*]Maseeh Dept. of Math. and Stat., Portland State Univ., Portland, OR, USA

networks are most naturally represented by directed graphs. Still, in many standard references, directed graphs rate no more than one section or chapter out of many. A notable exception to this is [5], but even here, little mention is made of the *algebraic* theory for digraphs. One reason for this is perhaps that adjacency matrices (and hence Laplacians) of undirected graphs are symmetric and thus there is a complete basis of orthonormal eigenvectors, which is not generally true for digraphs.

One aim of this article is to collect and explain some important results in algebraic digraph theory in a unified context. We will see that the lack of symmetry, referred to above, can actually be used to reveal a great deal of structure in general digraphs that has *no counterpart* in undirected graphs. Many of these results are known or published in varied disciplines outside the mainstream of mathematics, but frequently are only stated in the case of strongly connected graphs and not always with proofs that apply more generally.

In considering the many applications, such as those mentioned in the first paragraph, we found it a useful heuristic to fix once and for all *the direction of the edges in the digraph as the direction of the flow of information.* In most cases, this is a convenient way to determine the (ultimately arbitrary) direction of the arrows of the graph, and unify the treatment of these different applications. In our article, the direction of the flow of information is exemplified by the consensus problem (see below): if i's opinion is influenced by j, then there is an edge from j to i. We will see that a random walk on the same digraph naturally is defined to go in the *opposite* direction, i.e. from i to j. One of our main conclusions is that these processes in opposite directions can be considered as *duals* of one another.

Finally, we apply some of these insights to the pagerank algorithm most famously associated with Google founders Brin and Page. The usual treatment of this algorithm involves a random walk in a graph defined by web links, together with some random jumps. We give a dual interpretation in terms of how information added to one page influences the collective network as a whole.

The outline of this article is as follows. In the next section, we give important definitions and some background relating to our topic. In Section 3, we describe the structure of the right kernel of the random walk Laplacian $\mathcal{L}$ (with or without teleporting). In particular, we show how it relates to the way strong components and reaches are linked. Most of this section is based on [13]. In Section 4, we give the structure of the left kernel of $\mathcal{L}$ and similarly relate it to how the different components of G are linked. Furthermore, we show that the left kernel encodes the asymptotic behavior of a random walk defined on G. In Section 5, we use both left and right kernels, and show that the asymptotic behavior of the random walk on the one hand and consensus on the other, are natural duals. In Section 6, we give the dual interpretation of the classical pagerank algorithm, and show that the common practice of adding "teleporting" to so-called "dangling nodes" adds no new information to the ranking order. Finally, in Section 7, we illustrate our definitions and algorithms in one (small) graph.

2   Definitions and Background

We now give an overview of the most relevant definitions concerning digraphs.

Let G be a directed graph on $V = \{1, \ldots, n\}$ with (directed) edge-set $E \subseteq V \times V$. A directed edge $j \to i$ will be referred to as ji. We also write $j \rightsquigarrow i$ if there exists a directed path in G from vertex j to vertex i. To each (directed) edge $ji \in E$, we assign a positive weight $w_{ji} > 0$. The *adjacency matrix* Q is the $n \times n$ matrix whose ij entry equals $w_{ji} > 0$ if $ji \in E$ and zero otherwise.

The notion of "connectedness" in undirected graphs has several natural extensions to directed graphs. See [5] for more information. A digraph G with vertex set V is *strongly connected* if for all i and j in V, there is a path $i \rightsquigarrow j$ (and thus also a path $j \rightsquigarrow i$). Such components are called *strong components*. A much weaker form of connectedness is the following. G is *weakly connected* if for all i and j in V, there is an undirected path in G from i to j. Equivalently, G is weakly connected if its underlying undirected graph (all edges replaced by undirected edges) is connected. There is an intermediate form of connectedness, called *unilateral connectedness* in [5]. G is *unilaterally connected* if for all i and j in V, there is a undirected path $i \rightsquigarrow j$ or $j \rightsquigarrow i$. A digraph that is not weakly connected is disconnected. To study its properties, it is sufficient to study the properties of its weakly connected components. Thus in this article, we assume without loss of generality that our digraphs are weakly connected, but not necessarily unilaterally connected. From now on, we will abbreviate *(weighted) directed graph* to *graph* unless misunderstanding is possible.

Definition 2.1 *[13] For any vertex j, we define the* reachable set $R(j)$ *of vertex j to be the set containing j and all vertices i such that $j \rightsquigarrow i$.*

Definition 2.2 *[13] A set R of vertices in a graph will be called a* reach *if it is a maximal reachable set; in other words, R is a reach if $R = R(i)$ for some i and there is no j such that $R(i) \subsetneq R(j)$. Since our graphs all have finite vertex sets, such maximal sets exist and are uniquely determined by the graph. For each reach R_i of a graph, we define the* exclusive part *of R_i to be the set $H_i = R_i \backslash \cup_{j \neq i} R_j$. Likewise, we define the* common part *of R_i to be the set $C_i = R_i \backslash H_i$.*

Thus, a reach is a maximal unilaterally connected set. A graph will typically contain more than one reach. Note that, by definition, the pairwise intersection of two exclusive sets is empty: $H_i \cap H_j = \emptyset$ if $i \neq j$. The common sets can, however, intersect. Note further that each reach R contains at least one vertex r such that its reachable set $R(r)$ equals the entire reach. Such a vertex is called a *root* of R. By definition, any root must be contained in the exclusive part of its reach.

Definition 2.3 *Given a digraph G. Then each reach R_i contains a set of roots B_i. The set B_i is called the* cabal *of R_i and is contained in H_i. (A cabal consisting of 1 vertex is usually called a leader [36].)*

Dynamical processes on G are often defined using a version of the Laplacian operator. In this paper we find it convenient to define these operators in terms of a normalized adjacency matrix Q obtained by dividing each row by its sum. This is straightforward, unless Q contains a row $\{Q_{ij}\}_{j=1}^n$ that consists of entirely of zeroes. One possible solution is to replace Q_{ii} with a 1. We will denote the resulting matrix also by Q. Another possibility is to replace each entry of the row with $1/n$. The resulting matrix will also be called Q. This variation is often called *teleporting*. The reason for that name will be made clear in Section 4. Define the *in-degree* matrix D is the $n \times n$ diagonal matrix

3

whose i-th diagonal entry equals the sum of the i-th row of Q. Then, $S \equiv D^{-1}Q$. If we used "teleporting", we will denote the resulting matrix by S_t. Both S and S_t are *row stochastic* matrices, that is: they are non-negative and have row sum one. We can now list three commonly used Laplacians together with their names:

$$\begin{aligned} &\text{the } \textit{combinatorial Laplacian} & L &= D - Q, \\ &\text{the } \textit{random walk (rw) Laplacian} & \mathcal{L} &= I - S, \\ &\text{the } \textit{rw Laplacian with teleporting} & \mathcal{L} &= I - S_t. \end{aligned}$$

All these Laplacians share the fact that they can be written as follows.

Definition 2.4 *[13] We define the following (non-symmetric) Laplacian matrix:*

$$M \equiv D - DS ,$$

where D is non-negative on the diagonal and zero elsewhere, and S is a row stochastic matrix. M has non-negative diagonal entries and non-positive off-diagonal entries and any row-sum is zero.

There are many other conventions that may give different Laplacians for a given graph. For example, if Q is *irreducible*, there is a leading all-positive eigenvector r with (real) eigenvalue λ. Let R be the diagonal matrix whose diagonal entries equal r_i. Then $S = \lambda^{-1}R^{-1}QR\lambda$ is the *stochasticization* [8] of Q. If Q is undirected (symmetric), then the *normalized* adjacency is $S = D^{-\frac{1}{2}}QD^{-\frac{1}{2}}$. This normalized matrix has the advantage that it is symmetric if and only if Q is symmetric. More sophisticated definitions that give symmetric Laplacians for digraphs can be found in [14] and [15]. We will, however, not pursue these in this article.

The statement that i is in the same strong component as j determines an equivalence relation on the vertices of the graph G. Thus the vertices can be partitioned into strong components. In turn, there is a natural partial order on the strong components $S_i \subset V$. Namely, $S_1 < S_2$ if S_2 is reachable from S_1. The fact that any partial order can be extended to a total order implies [5] that the strong components themselves can be ordered in such a way that the adjacency matrix becomes block upper triangular with each diagonal block corresponding to the adjacency matrix of a subgraph of G induced by one of its strong components. Thus, the spectrum of Q in is one-to-one correspondence with the union of the spectra of the strong components of G and gives us no information about how these strong components are linked together. This is not true for the various Laplacian matrices given in definition 2.4. We will see that the multiplicity of the eigenvalue 0 in such a matrix gives precisely this kind of information; it counts the number of reaches of G. Further, the structure of the associated eigenvectors are intimately related to how the strong components (and, indeed, the reaches) are connected to one another.

To describe this relationship, we define two well-known dynamical processes on digraphs, namely diffusion and consensus. We consider them from the start as dual to each other. [1]

Definition 2.5 *Let $\mathcal{L}$ be the rw Laplacian (with or without teleporting).*
i) The consensus problem *on G is the differential equation $\dot{x} = -\mathcal{L}x$.*
ii) The diffusion problem *on G is the differential equation $\dot{p} = -p\mathcal{L}$.*
From hereon, x is a column vector and p is a row vector.

---

[1]For an introduction to dynamical systems defined on graphs, see [32].

If we discretize both these problems with time-steps of size 1, we get (noting that $-\mathcal{L} = S - I$)

$$\dot{x} = (S - I)x \qquad \longrightarrow \qquad x^{(n+1)} = Sx^{(n)} ,$$
$$\dot{p} = p(S - I) \qquad \longrightarrow \qquad p^{(n+1)} = p^{(n)}S .$$

This inspires our next definition.

Definition 2.6 *Let $S = I - \mathcal{L}$ where $\mathcal{L}$ is the rw Laplacian (with or without teleporting).*
i) The discrete consensus problem *on G is the difference equation $x^{(n+1)} = Sx^{(n)}$.*
ii) The discrete diffusion problem *or* random walk *on G is the difference equation $p^{(n+1)} = p^{(n)}S$.*

Note that the second process is also known as a discrete-time finite-state Markov chain [25].

In the above definition, S can be either the matrix $D^{-1}Q$ or $D^{-1}Q_t$ that we defined earlier. The difference is that in the former case, if the walker reaches a vertex with in-degree zero, it will stay there forever, while in the latter case, the walker will be sent to an arbitrary vertex (with uniform distribution) in the graph. Hence the annotation "with teleporting" for the corresponding Laplacian. Note that the random walker moves in the direction *opposite* the direction of the flow of information.

The (continuous) consensus problem can be characterized by two requirements. First, we require that $\mathbf{1}$, the vector of all ones, is an equilibrium. Second, if $\mathbf{1}_{\{i\}}$ is the vector that is 1 on the ith vertex and 0 everywhere else, then $-\mathcal{L}\mathbf{1}_{\{i\}}$ must be non-negative except on the ith vertex. This requirement indicates that i pulls other vertices in *its* direction. It is easy to see that the first requirement is equivalent to saying that $\mathcal{L}$ has row sum zero, while the second indicates that off-diagonal components of $\mathcal{L}$ are non-positive.

A similar characterization can be given for the diffusion problem. First, we need this definition.

Definition 2.7 *A* probability vector *or a (discrete)* measure *on G is a row-vector in $\mathbb{R}^V$ such that for all i, $p(i) \geq 0$ and $\sum_i p(i) = 1$. The* support, $\text{supp}(p)$, *of the measure p, is the set of vertices on which p takes a positive value.*

The characterization of the diffusion problem is as follows. First, total probability is conserved, or $\sum_i \dot{p}_i = 0$. Second, all components except the ith of $-\mathbf{1}_{\{i\}}^T\mathcal{L}$ are non-negative, because probability streams from vertex i to other vertices. One immediately sees that this gives the same requirements on $\mathcal{L}$. Similar characterizations can be given for the discrete equivalent of these problems.

Thus the discrete versions of each process can be treated in pretty much the same way. In the interest of brevity, we will limit the remainder of the exposition to considering only the discrete diffusion problem and the continuous consensus problem. The other two problems can be treated in the same way and with the same conclusions. A more rigorous treatment of the discrete/continuous distinction is deferred to the appendix.

3   The Right Kernel of $D - DS$

In this section, we analyze the right kernel of the Laplacian.

Lemma 3.1 *(i): There are no edges from the complement H_i^c of an exclusive set to H_i.*
(ii): There are no edges from the complement B_i^c of a cabal to B_i.
(iii): The graph induced by a cabal B_i is strongly connected.

Proof: (i): Suppose that for $j \neq i$, $v \in R_j$ and $vu \in E$. Then by the definition of reach, $u \in R_j$. Thus, by the definition of exclusive set, $u \notin H_i$.
(ii): Suppose $u \in B_i$ and $vu \in E$. By the definition of cabal, the entire reachable from u and therefore from v. So $v \in B_i$.
(iii): Suppose v and w in B_i and there is no path $v \rightsquigarrow w$ in B_i. Then v cannot be a root, which is a contradiction. ∎

Theorem 3.2 *[13] Suppose $M = D - DS$, where D is a nonnegative $n \times n$ diagonal matrix and S is (row) stochastic. Suppose G has k reaches, denoted R_1 through R_k, where we denote the exclusive and common parts of each R_i by H_i, C_i respectively. Then the eigenvalue 0 has algebraic and geometric multiplicity k and the kernel of M has a basis $\gamma_1, \gamma_2, \ldots \gamma_k$ in $\mathbb{R}^n$ whose elements satisfy:*
(i) $\gamma_i(v) = 1$ for $v \in H_i$;
(ii) $\gamma_i(v) \in (0,1)$ for $v \in C_i$;
(iii) $\gamma_i(v) = 0$ for $v \notin R_i$;
(iv) $\sum_i \gamma_i = \mathbf{1}_n$ (the all ones vector).

Theorem 3.3 *[2], [13] Any nonzero eigenvalue of a Laplacian matrix of the form $D - DS$, where D is nonnegative diagonal and S is stochastic, has (strictly) positive real part.*

Theorem 3.2 states, among other things, that the null space contains no non-trivial Jordan blocks. Since we are dealing with digraphs, this was by no means a priori obvious. Indeed, other eigenspaces may contain Jordan blocks, *even for strongly connected digraphs.* An example is the (rw) Laplacian

$$\mathcal{L} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 & 1 & -\frac{1}{2} \\ 0 & 0 & -1 & 1 \end{pmatrix}.$$

One checks that $(\mathcal{L} - I)(1, 0, 0, -1)^T = (0, 0, 0, 0)^T$, and that $(\mathcal{L} - I)(0, -1, 1, 0)^T = (1, 0, 0, -1)^T$, producing a Jordan block of dimension 2. Similar examples can be found for the combinatorial Laplacian.

The definitions of reach, cabal, exclusive part, and common part are illustrated in Section 7. Here we wish to note that the reaches of even a large graph G are relatively easy to compute by first computing the condensation digraph $SC(G)$. This is the graph whose vertices are the strong components S_i of G and $S_i \to S_j$ if there is an edge from S_i to S_j. Let V_0 be the vertices of $SC(G)$ with in-degree 0. Since $SC(G)$ has no directed cycles, there must be at least one such vertex. Any such vertex v represents a cabal in G by Lemma 3.1. We can then use breadth first search from v to find the reach of v in $SC(G)$ and finally recover a reach in G.

Theorem 3.2 also shows that the underlying unweighted graph of G gives a great deal of information about the basis of the kernel. In particular, it determines γ_i on H_i and on the complement of R_i. However, the exact value of $\gamma_i(v)$ if v is a vertex in a common part is only determined once the weights on the edges are fixed.

6

4 Random Walks on Directed Graphs

In this section, we analyze random walks on G and analyze its invariant measure in terms of the left kernel of the Laplacian on G.

Definition 4.1 *The probability vector p is an* invariant probability measure *(or a* stationary distribution*) for T if $pS = p$. $K \subseteq V(G)$ is a forward invariant set under T if $\mathrm{supp}(p) \subseteq K$ implies $\mathrm{supp}(pS) \subseteq K$.*

Lemma 4.2 *Given a random walk random walk T, every exclusive set H_i and its cabal B_i are forward invariant sets under T.*

Proof: A walker leaving C_i means that the graph must have an edge into C_i. This contradicts Lemma 3.1. The same holds for B_i. ∎

Theorem 4.3 *Let G be a weighted digraph with Laplacian $\mathcal{L}$ and k reaches. The probability that a random walker under T starting at v is absorbed into the cabal B_r equals $\gamma_r(v)$ (defined in Theorem 3.2).*

Proof: Let $q(j)$ be the probability that a random walker starting at $j \in V$ reaches B_m for some fixed m. (Note that q is not a probability vector.) Then $q : V \to [0,1]$ is well-defined and is constant in time. Since, by Lemma 4.2, B_m is forward invariant, $q(j)$ is also equal to the probability that the walker starting at j ends up and stays in B_m.

The probability $q(j)$ concerns the future (under T) of the walker on j. Therefore it is equal to the appropriately weighted average of $q(i)$ of j's successors under T. Thus from Definition 2.6,

$$q(j) = \sum_i \mathrm{prob}(j \to i) q(i) = \sum_i S_{ji} q(i) \ .$$

From this we conclude $q = Sq$, which is equivalent to $\mathcal{L}q = 0$. Thus q is in the kernel of $\mathcal{L}$, and therefore

$$q(j) = \sum_m \alpha_m \gamma_m(j) \ .$$

By Lemma 4.2, if j is a vertex in B_r, then $q(j) = 1$. Also if j is in any B_m with $m \neq r$, then $q(j) = 0$. Since the cabals are disjoint, we must have that $\alpha_r = 1$ and $\alpha_m = 0$ if $m \neq r$. ∎

We remark that it follows from Theorem 3.2(iv) that the probability is one that a walker will be absorbed in B_r for some r.

Lemma 4.4 *Let G be a digraph that has a reach R, consisting of an exclusive part H which contains a cabal B, and a common part C. Under the random walk T on G, there is a unique invariant measure p with support in R. Furthermore, $\mathrm{supp}(p)$ equals B.*

Proof: Consider a reach R with its cabal B and denote the vertex set $R \backslash B$ by Y and the vertex set $V \backslash R$ by Z. Since directed paths in G cannot leave the reach R, we have $S_{ZB} = S_{ZY} = 0$. By Lemma 3.1, $S_{BY} = S_{BZ} = 0$. So

$$S = \begin{pmatrix} S_{BB} & \mathbf{0} & \mathbf{0} \\ S_{YB} & S_{YY} & S_{YZ} \\ \mathbf{0} & \mathbf{0} & S_{ZZ} \end{pmatrix} \; .$$

We solve for p in $pS = p$, where $p = (a_B, b_Y, c_Z)$ and assume $c_Z = 0$. This gives

$$a_B S_{BB} + b_Y S_{YB} = a_B \; , \quad b_Y S_{YY} = b_Y \quad \text{and} \quad b_Y S_{ZY} + c_Z S_{ZZ} = c_Z \; .$$

The proof of Theorem 2.7 in [13] establishes that the spectral radius of S_{YY} is strictly less than 1. Thus the middle equation can only be satisfied if $b_Y = 0$. Since B is a strong component, S_{BB} is irreducible. It follows directly from the Perron-Frobenius theorem (see [8, 35]) that the eigenvalue 1 is simple and its associated eigenvector is strictly positive. ∎

Theorem 4.5 *Let G be a graph with Laplacian $\mathcal{L} = I - S$ with k reaches. Then the eigenvalue 0 of $\mathcal{L}$ has algebraic and geometric multiplicity k. The left kernel of $\mathcal{L}$ has a basis $\bar{\gamma}_1, \bar{\gamma}_2, \ldots \bar{\gamma}_k$ in $\mathbb{R}^n$ whose elements satisfy:*
(i) For all $i \in \{1, \cdots k\}$ and all $v \in \{1, \cdots n\}$: $\bar{\gamma}_i(v) \geq 0$;
(ii) $supp(\bar{\gamma}_i) = B_i$;
(iii) $\sum_v \bar{\gamma}_i(v) = 1$;
(iv) The vectors $\{\bar{\gamma}_i\}_{i=1}^k$ are orthogonal.

Proof: The first statement (the multiplicity of 0) is in fact part of Theorem 3.2. For each reach R_i, Lemma 4.4 gives an invariant measure $\bar{\gamma}_i$ satisfying (i), (ii), and (iii). Since $\bar{\gamma}_i S = \bar{\gamma}_i$, we have that $\bar{\gamma}_i$ is a left null vector of $\mathcal{L}$. These k vectors are orthogonal, because the sets B_i are mutually disjoint. ∎

Corollary 4.6 *Let G be a graph with Laplacian $L = D - DS$ with k reaches. Then the left kernel of L has a basis $\{\bar{\gamma}_i D^{-1}\}_{i=1}^k$, where the $\{\bar{\gamma}_i\}_{i=1}^k$ are as in Theorem 4.5.*

Proof: Construct the basis $\bar{\gamma}_1, \bar{\gamma}_2, \ldots \bar{\gamma}_k$ of the left kernel of $I - S$. Then it is easy to see that the $\bar{\gamma}_i D^{-1}$ form a basis of the left kernel of $D - DS$. ∎

5 The Left and Right Kernels of $\mathcal{L}$

We start this section with a general lemma and a definition that will be used at the end of the section. Subsequently, we establish the dual relationship between the equilibria of the random walk and those of the consensus problem. In this context, the duality is very natural. The matrix is multiplied from the *left* by probability vectors (measures) and from the right by vectors whose Laplacian is zero. The latter are also known as *harmonic vectors or functions* and have many other

applications in graph theory. Two notable examples focused on undirected graphs are [7] and [10]. The latter reference also discusses the duality mentioned above in the context of electrical networks. At the end of the section, we briefly note that the duality is independent of the discrete or continuous nature of the processes discussed. This has also been observed by other authors, for example [27] and [31].

Lemma 5.1 *Let A be an $n \times n$ matrix. There are bases of (generalized) right eigenvectors (columns), $\{\eta_i\}_{i=1}^n$, and of (generalized) left eigenvectors (rows), $\{\bar{\eta}_i\}_{i=1}^n$, such that the matrices:*

$$\Gamma = \begin{pmatrix} \eta_1 & \eta_2 & \cdots & \eta_n \end{pmatrix} \quad \text{and} \quad \bar{\Gamma} = \begin{pmatrix} \bar{\eta}_1 \\ \bar{\eta}_2 \\ \vdots \\ \bar{\eta}_n \end{pmatrix}$$

are inverses of one another. We will assume the basis-vectors ordered according to ascending eigenvalue.

Proof: This follows directly from the Jordan Decomposition Theorem [21]. Let J be the Jordan normal form of A. Then that theorem tells us that there is an invertible matrix Γ such that $A\Gamma = \Gamma J$ or $\Gamma^{-1}A = J\Gamma^{-1}$. Right multiply the first equation by the standard column basis vector $\mathbf{1}_{\{i\}}$ to show that the ith column of Γ is a generalized right eigenvector. Left multiply by $\mathbf{1}_{\{i\}}^T$ to see that the ith row of Γ^{-1} is a generalized left eigenvector. ∎

Definition 5.2 *Let G have k reaches. Let $\{\gamma_i\}_{i=1}^k$ be the (column) vectors of Theorem 3.2 and $\{\bar{\gamma}_i\}_{i=1}^k$ the (row) vectors of Theorem 4.5. Define*

$$\Gamma^0 = \begin{pmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_k \end{pmatrix} \quad \text{and} \quad \bar{\Gamma}^0 = \begin{pmatrix} \bar{\gamma}_1 \\ \bar{\gamma}_2 \\ \vdots \\ \bar{\gamma}_k \end{pmatrix}.$$

Theorem 5.3 *Let G be a graph with random walk laplacian $\mathcal{L} = I - S$ with k reaches $\{R_i\}_{i=1}^k$. Let $p^{(0)}$ be an initial measure (at $t = 0$). Then*

$$\lim_{\ell \to \infty} \frac{1}{\ell} \sum_{i=0}^{\ell-1} T_*^i p^{(0)} = \sum_{m=1}^k \left(p^{(0)}\gamma_m\right) \bar{\gamma}_m .$$

Note that $(p^{(0)}\gamma_m)$ is a scalar. (Recall that p is a row vector and γ is a column vector.)

Proof: Given an initial measure $p^{(0)}$, then Theorem 4.3 implies that $\alpha_r \equiv p^{(0)}\gamma_r$ is the probability to be absorbed in the cabal B_r. By Lemma 3.1, the graph induced by B_i is strongly connected. This is equivalent with the matrix $S_{B_iB_i}$ being *irreducible*. The Perron-Frobenius theorem (see [8, 35]) implies that $S_{B_iB_i}$ satisfies:

9

1. eigenvalue 1 has algebraic and geometric multiplicity 1, and

2. there are possibly other eigenvalues that are roots of 1, all with multiplicity 1, and

3. all eigenvalues not listed under cases 1 or 2, have absolute value less than 1.

The eigenvector corresponding to case 1 is the unique invariant measure in B_r. For the eigenvector p of case 2, we have that $T_*^i p = \lambda^i p$ with λ a root of unity. This implies that

$$\lim_{\ell \to \infty} \frac{1}{\ell} \sum_{i=0}^{\ell-1} T_*^i p = 0 .$$

The same holds for case 3.

Thus for each r, the probability is α_r that a random walker is absorbed in B_r. Thus for large i, $T_*^i p$ is a combination of invariant measures in B_r, a finite number of periodic measures, and a remainder that decays exponentially. The averaging cancels the periodic (case 2) and decaying (case 3) parts. Hence,

$$\lim_{\ell \to \infty} \frac{1}{\ell} \sum_{i=0}^{\ell-1} T_*^i p^{(0)} = \sum_{m=1}^{k} \alpha_r \, \bar{\gamma}_m = \sum_{m=1}^{k} \left(p^{(0)} \gamma_m \right) \bar{\gamma}_m .$$

∎

Corollary 5.4 *Assume the hypotheses of Theorem 5.3. Suppose that in addition the following is true: for each cabal B_i, there is a k_i so that the k_i-th power of $S_{B_i B_i}$ is strictly positive, then*

$$\lim_{\ell \to \infty} T_*^\ell p^{(0)} = \sum_{m=1}^{k} \left(p^{(0)} \gamma_m \right) \bar{\gamma}_m .$$

Proof: The proof of this statement is exactly the same as the proof of Theorem 5.3, except that now each matrix $S_{B_i B_i}$ is *primitive*. Primitivity guarantees that case 2 of that proof does not occur. For a (generalized) eigenvector p as in case 3, $T_*^i p$ converges to 0. Hence the averaging is not necessary. ∎

Theorem 5.5 *Let G be a graph with Laplacian $\mathcal{L} = I - S$ with k reaches $\{R_i\}_{i=1}^k$. Then the solution of the consensus problem of Definition 2.5 with initial condition (at $t = 0$) $x^{(0)}$ satisfies*

$$\lim_{t \to \infty} x^{(t)} = \sum_{m=1}^{k} \left(\bar{\gamma}_m x^{(0)} \right) \gamma_m .$$

Note that $(\bar{\gamma}_m x^{(0)})$ is a scalar.

Proof: The right and left eigenvectors of $\mathcal{L}$ defined in Theorems 3.2 and 4.5 satisfy $\bar{\gamma}_i \gamma_j = \delta_{ij}$ for $i, j \in \{1, \cdots k\}$. According to Lemma 5.1, one can extend theses sets of vectors to dual bases of right and left (generalized) eigenvectors $\{\gamma_\ell\}_{\ell=1}^n$ and $\{\bar{\gamma}_\ell\}_{\ell=1}^n$ such that for $i \leq k$ and $\ell > k$

$$\bar{\gamma}_i \gamma_\ell = \bar{\gamma}_\ell \gamma_i = 0 .$$

10

Let λ_i be the ith eigenvalue associated with the (generalized) eigenvector γ_i (or $\bar{\gamma}_i$). Now we consider the consensus problem with initial condition $x^{(0)} = y^{(0)} + z^{(0)}$ where

$$x^{(0)} = y^{(0)} + z^{(0)} \quad \text{where} \quad y^{(0)} = \sum_{i=1}^{k} \alpha_i \gamma_i \quad \text{and} \quad z^{(0)} = \sum_{i=k+1}^{n} \alpha_i \gamma_i \ . \tag{5.1}$$

From the standard theory of linear differential equations (see, for example, [1]), one easily derives that the general solution of the consensus problem is given by

$$x^{(t)} = e^{-\mathcal{L}t} x^{(0)} = \sum_{i=1}^{n} \beta_i \gamma_i e^{-\lambda_i t} p_i(t) \ , \tag{5.2}$$

where $beta_i$ are constants. Here $p_i(t)$ are polynomials whose degree is less than the size of the corresponding Jordan block. Furthermore, if γ_i is an actual eigenvector (not a generalized one), then $\alpha_i = \beta_i$.

By Theorem 3.2, we have that in equation (5.2), $\lambda_i = 0$ for $i \in \{1, \cdots k\}$. Also, the eigenvalue zero has only trivial Jordan blocks and so for $i \in \{1, \cdots k\}$, $p_i = 1$ and $\beta_i = \alpha_i$. By Theorem 3.3, we have that all the other terms (for $i > k$) converge to zero. Therefore, using equation (5.1),

$$\lim_{t \to \infty} x^{(t)} = \sum_{m=1}^{k} \alpha_m \gamma_m = y^{(0)} \ .$$

Next, we determine the α_i. Let Γ be the matrix whose ith column equals γ_i. The previous equation implies that (γ_i being he columns of Γ)

$$\Gamma \left(\sum_{i=1}^{k} \alpha_i \mathbf{1}_{\{i\}} \right) = y^{(0)} \quad \Longrightarrow \quad \sum_{m=1}^{k} \alpha_m \mathbf{1}_m = \bar{\Gamma} y^{(0)} \ .$$

With the definition of $x^{(0)}$ and first paragraph of this proof, this implies that

$$\alpha_m = \bar{\gamma}_m y^{(0)} = \bar{\gamma}_m x^{(0)} \ ,$$

from which the result follows. ∎

An alternate proof of this result can be found in [31].

There is a striking way to express the duality between consensus and diffusion. In Theorem 5.5, the factor within the parentheses in $\left(\bar{\gamma}_m x^{(0)} \right) \gamma_m$ is a real number. Since multiplication of a real and vector is commutative, it can be written as $\gamma_m \left(\bar{\gamma}_m x^{(0)} \right)$. But multiplication is associative, and thus this is equal to $(\gamma_m \bar{\gamma}_m) x^{(0)}$. The same reasoning works for the corresponding expression in Theorem 5.3. The notation becomes more compact upon observing that $\left[\sum_{m=1}^{k} \gamma_m \bar{\gamma}_m \right] = \Gamma^0 \bar{\Gamma}^0$. Thus we obtain the following.

Corollary 5.6 *Let G be a graph with Laplacian $\mathcal{L} = I - S$ with k reaches $\{R_i\}_{i=1}^k$. Then the solution of the random walker with initial condition $p^{(0)}$ and of the consensus problem with initial condition $x^{(0)}$, respectively, satisfy*

$$\lim_{\ell \to \infty} \frac{1}{\ell} \sum_{i=0}^{\ell-1} T_*^i p^{(0)} = p^{(0)} \, \Gamma^0 \bar{\Gamma}^0 \qquad \text{and} \qquad \lim_{t \to \infty} x^{(t)} = \Gamma^0 \bar{\Gamma}^0 \, x^{(0)} \; .$$

Equivalently, $\lim_{\ell \to \infty} \frac{1}{\ell} \sum_{i=0}^{\ell-1} S^j = \lim_{t \to \infty} e^{-\mathcal{L}t} = \Gamma^0 \bar{\Gamma}^0$.

Of course, the same result also holds for the discrete version of the consensus and the continuous version of diffusion mentioned in Definitions 2.5 and 2.6. Note that the matrix $e^{-\mathcal{L}t}$ is well studied in the context of a continuous-time Markov chains where it is called the *heat kernel*. We refer the reader to [25] for more details.

6   Application to Ranking Algorithms

In this section, we derive the pagerank algorithm using the consensus problem. This explanation is the "dual" of the usual one, which is given in terms of diffusion. Our analysis does not require use of the teleportation matrix Q_t.

For a graph with vertex set V and $U \subseteq V$, let $\mathbf{1}_U$ denote the (column) vector that has the value 1 on U and 0 elsewhere. We set $\mathbf{1}_V \equiv \mathbf{1}$ (as before). The second part of Corollary 5.6 tells us that in the consensus problem the final displacements when the initial condition is $\mathbf{1}_{\{v\}}$ are equal to $[\sum_m \gamma_m \bar{\gamma}_m)] \, \mathbf{1}_{\{v\}}$. As first noted in [26], the *mean* of these displacements over all vertices is a good measure of the influence that vertex v exerts over the long term behavior of the solution to the consensus problem. This naturally leads to the following definition.

Definition 6.1 *For a graph G with n vertices, the* influence *of a vertex v is given by*

$$I(v) = \frac{\mathbf{1}^T}{n} \, \Gamma^0 \bar{\Gamma}^0 \, \mathbf{1}_{\{v\}} \; .$$

The influence vector I *is the row vector whose jth component equals $I(j)$ or*

$$I = \frac{\mathbf{1}^T}{n} \, \Gamma^0 \bar{\Gamma}^0 \; .$$

From this we see that the influence of v equals the average of the vth column of the matrix $\Gamma^0 \bar{\Gamma}^0$. It is nonnegative. It is positive if and only if some $\bar{\gamma}_m$ is non-zero at v, that is: if v is part of a cabal. The sum of all influences equals 1. Thus the influence vector is a probability measure. It is straightforward to see that the influence of a subset U of vertices equals: $\frac{\mathbf{1}^T}{n} \Gamma^0 \bar{\Gamma}^0 \, \mathbf{1}_U$.

The fact that $I(v) > 0$ only if v is part of a cabal is not entirely realistic for networks such as the internet. A page may contain information even though it has both in- and out-links. That information, of course, had to be put there by some other entity in the first place. Following [26], one could say that, in fact, each page v has an associated (meta) page b_v together with an edge $b_v \to v$.

Definition 6.2 *Let G be a graph with n vertices and stochastic weighted adjacency matrix S. The extended graph $E_\alpha[G]$ is defined as the graph G together with one new vertex b_v and one new edge $b_v \to v$ of weight $\alpha > 0$ for each $v \in V$.*

The new graph $E_\alpha[G]$ has exactly n reaches $\{\tilde{R}_i\}_{i=1}^n$ and R_i the new vertex b_i as leader. The natural choice for the weight α would appear to be 1, but in actual applications it tends to have different values. So we keep it as a (positive) parameter. In what follows we need to distinguish vectors in $\mathbb{R}^{2n}$ from those in $\mathbb{R}^n$. To do that we mark the former with a tilde. The same device is used to distinguish $2n \times 2n$ matrices from $n \times n$ matrices. We label the vertices of $E_\alpha[G]$ in the following order: $\{b_1, \cdots b_n, 1, \cdots n\}$. The first n vertices are the leaders and this set will be denoted by B. The second set of n vertices we continue to refer to as V.

The Laplacian and the random walk Laplacian associated with $E_\alpha[G]$ are

$$\tilde{L} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ -\alpha I & (1+\alpha)I - S \end{pmatrix} \implies \tilde{\mathcal{L}} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ -\frac{\alpha}{1+\alpha}I & I - \frac{1}{1+\alpha}S \end{pmatrix} .$$

The normalized Laplacian with teleportation is given in the same way, except that we use S_t instead of S. We are now ready to discuss pagerank. We begin with an alternate formulation. In Corollary 6.5, we will show how this is equivalent to the standard treatment (as found, for example, in [35]).

Definition 6.3 *For a graph G with n vertices, the* pagerank *of a vertex v is given by*

$$\wp(v) = 2\tilde{I}(b_v) - \frac{1}{n} .$$

Here, $\tilde{I}$ is the influence of b_v in the extended graph $E_\alpha[G]$. The pagerank *vector $\wp \in \mathbb{R}^n$ is the row vector whose jth component equals $\wp(j)$.*

Theorem 6.4 *Let G be a graph with vertex set V, $|V| = n$, and Laplacian $\mathcal{L} = I - S$ or $\mathcal{L} = I - S_t$. The pagerank vector of Definition 6.3 is the unique probability measure on $\mathbb{R}^n$ that satisfies*

$$\wp = \frac{\alpha}{n}\mathbf{1}^T(\alpha I + \mathcal{L})^{-1} .$$

Proof: From the comments after Definition 6.1, we know that the sum of the influences equals 1. We know that the influence of non-leaders is zero and so $\sum_v \left(2\tilde{I}(b_v) - \frac{1}{n} \right) = 2 - 1 = 1$. Also $\wp(v) \geq 0$, because the displacement of b_v in the consensus problem is 1 and the others are non-negative (Theorem 5.5).

Sections 3 and 4 provides us with appropriate bases $\{\tilde{\tilde{\gamma}}_i\}_{i=1}^n$ and $\{\tilde{\gamma}_i\}_{i=1}^n$ of the left and right kernels of $\tilde{\mathcal{L}}$. From Theorems 4.5 and 3.2 we know that for $1 \leq m \leq n$,

$$\tilde{\tilde{\gamma}}_m = (\mathbf{1}_{\{m\}}^T, \mathbf{0}^T) \quad \text{and} \quad \tilde{\gamma}_m = \begin{pmatrix} \mathbf{1}_{\{m\}} \\ \eta_m \end{pmatrix} .$$

With these constraints, $\tilde{\gamma}_m$ is entirely determined by $\tilde{L}\tilde{\gamma}_m = 0$. This is equivalent to

$$(\alpha I + \mathcal{L})\eta_m = \alpha \mathbf{1}_{\{m\}} .$$

Since $\alpha > 0$ and $\mathcal{L}$ has no negative eigenvalues, the matrix $(\alpha I + \mathcal{L})$ can be inverted. Thus $\wp(m)$ is the average over V of the displacements η_m:

$$\wp(m) = \frac{\mathbf{1}^T}{\alpha}\eta_m = \frac{\alpha}{n}\mathbf{1}^T(\alpha I + \mathcal{L})^{-1}\mathbf{1}_{\{m\}} \ .$$

The result follows immediately from this. ∎

We use the vectors $\tilde{\tilde{\gamma}}_m$ and $\tilde{\gamma}_m$ to define the matrices $\tilde{\Gamma}^0$ and $\tilde{\tilde{\Gamma}}^0$ analogous to Definition 5.2. From Definition 6.1, we obtain that

$$\tilde{I}(b_v) = \frac{\tilde{\mathbf{1}}^T}{2n}\tilde{\Gamma}^0\tilde{\tilde{\Gamma}}^0\,\tilde{\mathbf{1}}_{\{b_v\}} = \frac{\tilde{\mathbf{1}}_B^T + \tilde{\mathbf{1}}_V^T}{2n}\tilde{\Gamma}^0\tilde{\tilde{\Gamma}}^0\,\tilde{\mathbf{1}}_{\{b_v\}} \ .$$

The initial condition $\tilde{\mathbf{1}}_{\{b_v\}}$ in the consensus problem for $E_\alpha[G]$ displaces no leaders, except b_v itself by one unit. Thus $\frac{\tilde{\mathbf{1}}_V^T}{2n}\tilde{\Gamma}^0\tilde{\tilde{\Gamma}}^0\,\tilde{\mathbf{1}}_{\{b_v\}} = \dfrac{1}{2n}$. Putting this together, we see that

$$\wp(v) = \frac{\tilde{\mathbf{1}}_B^T}{n}\tilde{\Gamma}^0\tilde{\tilde{\Gamma}}^0\,\tilde{\mathbf{1}}_{\{b_v\}} \ .$$

In other words, the pagerank of the vertex v is in fact the mean over *only* the vertices in V of the "old" graph G of the displacements in the consensus problem on $E_\alpha[G]$ caused by initial condition $\mathbf{1}_{\{b_v\}}$.

We now turn to a different interpretation of the pagerank.

Corollary 6.5 *Let G be a graph with vertex set V, $|V| = n$, and Laplacian $\mathcal{L} = I - S$. The pagerank vector of Definition 6.3 is the probability measure on $\mathbb{R}^n$ that satisfies*

$$\wp = \wp\left(\beta S + \frac{1-\beta}{n}J\right) = \beta\wp S + \frac{1-\beta}{n}\mathbf{1}^T = 0 \ ,$$

where J is the all ones matrix and $\beta \equiv \dfrac{1}{1+\alpha} \in (0,1)$.

Proof: Since $\wp$ is a probability measure, we have that $\mathbf{1}^T = \wp J$, where J is the all ones matrix. Substituting this in Theorem 6.4 and right multiplying by $(\alpha I + \mathcal{L})$ gives

$$\wp\left(\alpha I + \mathcal{L} - \frac{\alpha}{n}J\right) = 0 \ .$$

Substitute $\mathcal{L} = I - S$ and divide by $1 + \alpha$ to get

$$\wp\left(I - \frac{1}{1+\alpha}S - \frac{\alpha}{n(1+\alpha)}J\right) = 0 \ .$$

Substituting β for α gives the required statement. ∎

The traditional interpretation of the pagerank (as discussed in [35]) is evident in this result. The pagerank vector is the unique invariant probability measure that results from applying the random walk with probability β and uniform teleporting with probability $1 - \beta$. It is clear that the resulting matrix $S_{pagerank} \equiv \beta S + \frac{1-\beta}{n} J$ is row stochastic and primitive. Thus standard arguments using the Perron-Frobenius theorem show that the eigenvalue 1 is simple and all other eigenvalues are strictly smaller. In fact, since (generalized) eigenvectors of S are eigenvectors of J, it follows that all other eigenvalues have modulus less than β. The leading eigenvector is strictly positive. Google's pagerank algorithm takes $\beta = 0.85$ [35], and so after roughly 57 iterates, convergence of $S_{pagerank}^n$ to the pagerank is already accurate to 4 decimal places. The quick convergence plus the fact that $S_{pagerank}^n$ can be cheaply computed without using the full matrix, guarantees efficient algorithms for the computation of the pagerank vector.

The last result describes essentially two ranking algorithms, depending on how S is defined. In the introduction, we gave the definition of S and of S_t with teleporting. These give rise to pageranks $\wp$ and $\wp_t$. These two ranks have an interesting relation to one another, essentially that of deck of cards before and after one shuffle.

To see this, denote the set of vertices that are leaders by L and the rest by R. Thus

$$S = \begin{pmatrix} S_{LL} & S_{LR} \\ S_{RL} & S_{RR} \end{pmatrix} \; , \quad \wp = (\wp_L, \wp_R) \, , \quad \mathbf{1}^T = (\mathbf{1}_L^T, \mathbf{1}_R^T) \, , \quad \text{etc.}$$

Proposition 6.6 *With the above notation, we have*

$$\begin{aligned} \wp_{t,L} &= (\beta \pi_t + (1 - \beta)) \, \wp_L \, , \\ \wp_{t,R} &= \left(\frac{\beta}{1-\beta} \pi_t + 1 \right) \wp_R \, , \end{aligned}$$

where $\pi_t = \sum_{j \in L} \wp_t(j)$.

Proof: Corollary 6.5 says:

$$\begin{aligned} \wp_L (I - \beta S_{LL}) &= \beta \wp_R S_{RL} + \frac{1-\beta}{n} \mathbf{1}_L^T \, , \\ \wp_R (I - \beta S_{RR}) &= \beta \wp_L S_{LR} + \frac{1-\beta}{n} \mathbf{1}_R^T \, . \end{aligned}$$

In the case without teleporting, we get:

$$S_{LL} = I_{LL} \quad \text{and} \quad S_{LR} = 0 \, .$$

With teleporting, this becomes:

$$S_{LL} = \frac{1}{n} J_{LL} \quad \text{and} \quad S_{LR} = \frac{1}{n} J_{LR} \, .$$

15

The former will not receive a subscript, the latter will be denoted with the subscript "t". Denote $\pi = \sum_{j \in L} \wp(j)$ and $\pi_t = \sum_{j \in L} \wp_t(j)$. Thus, without teleporting, the equations of Corollary 6.5 become:

$$\wp_L = \frac{\beta}{1-\beta} \wp_R S_{RL} + \frac{1}{n} \mathbf{1}_L^T ,$$

$$\wp_R = \frac{1-\beta}{n} \mathbf{1}_R^T (I - \beta S_{RR})^{-1} .$$

Since S has spectral radius 1 and $\beta \in (0,1)$, the inverse is well-defined. Noting that $\wp_{t,L} J_{LR} = \pi_t \mathbf{1}_R$ and $\wp_{t,L} J_{LL} = \pi_t \mathbf{1}_L$, the equations *with* teleporting become

$$\wp_{t,L} = \beta \wp_{t,R} S_{RL} + \frac{1}{n} \left(\beta \pi_t + (1-\beta) \right) \mathbf{1}_L^T ,$$

$$\wp_{t,R} = \frac{1-\beta}{n} \left(\frac{\beta}{1-\beta} \pi_t + 1 \right) \mathbf{1}_R^T (I - \beta S_{RR})^{-1} .$$

This gives the required relation between $\wp_{t,R}$ and $\wp_R$. Substituting these expressions into $\wp_L$ and $\wp_{t,L}$, respectively, gives:

$$\wp_L = \frac{\beta}{n} \mathbf{1}_R^T (I - \beta S_{RR})^{-1} S_{RL} + \frac{1}{n} \mathbf{1}_L^T .$$

$$\wp_{t,L} = \frac{\beta}{n} \left(\beta \pi_t + (1-\beta) \right) \mathbf{1}_R^T \mathbf{1}_R^T (I - \beta S_{RR})^{-1} S_{RL} + \frac{1}{n} \left(\beta \pi_t + (1-\beta) \right) \mathbf{1}_L^T .$$

And this yields the relation between $\wp_{t,L}$ and $\wp_L$. ∎

To express $\wp_t$ completely in terms of $\wp$, we add the following result.

Corollary 6.7 *With the above notation, we also have*

$$\pi_t = \frac{(1-\beta)\pi}{1-\beta\pi} .$$

Proof: Summing the coefficients on the left and right hand sides of the equations of Proposition 6.6, and taking into account that $\sum_{j \in R} \wp(j) = 1 - \pi$ (and similar for π_t), we get

$$\pi_t = \left(\beta \pi_t + (1-\beta) \right) \pi$$

$$1 - \pi_t = \frac{1}{1-\beta} \left(\beta \pi_t + (1-\beta) \right) (1 - \pi) .$$

Dividing the first equation by the second gives $\dfrac{\pi_t}{1 - \pi_t} = (1 - \beta) \dfrac{\pi}{1 - \pi}$, which is equivalent to the statement of the Corollary. ∎

16

7 A Simple Example

In this section, we illustrate the definitions and results with the graph given in Figure 7.1. The unweighted Laplacian matrix equals (row and column i correspond to vertex i)

$$M = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & -1 & 0 & 0 & -1 & 2 \end{pmatrix}.$$

By ordering the vertices so that we list the exclusive part first, we see that the matrix M is lower block-triangular. In the exclusive blocks, we list the vertices in the cabal first, and then the others. The exclusive block itself is then also lower block-triangular.
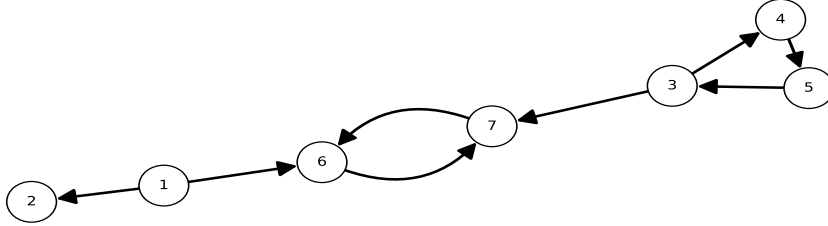

Figure 7.1: *A graph (vertices as labeled) with two reaches $R_1 = \{1, 2, 6, 7\}$ and $R_2 = \{3, 4, 5, 6, 7\}$. Each reach R_i has cabal B_i, exclusive part H_i, and common part C_i. According to the definitions in the text: $B_1 = \{1\}$, $H_1 = \{1, 2\}$, $C_1 = \{6, 7\}$, $B_2 = \{3, 4, 5\}$, $H_2 = \{3, 4, 5\}$, $C_2 = \{6, 7\}$.*

Normalizing, as discussed in Section 2, gives the stochastic matrix S as

$$S = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 & 0 & 1/2 \\ 0 & 0 & 1/2 & 0 & 0 & 1/2 & 0 \end{pmatrix} \quad \text{and} \quad \mathcal{L} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ -1/2 & 0 & 0 & 0 & 0 & 1 & -1/2 \\ 0 & 0 & -1/2 & 0 & 0 & -1/2 & 1 \end{pmatrix}.$$

The stochastic matrix with teleporting S_t is nearly the same: the first row entries are replaced by $1/7$.

The bases of the left and right kernels of the Laplacian $\mathcal{L} = I - S$ of Theorems 3.2 and 4.5 are, respectively,

$$\gamma_1^T = \left(1, \quad 1, \quad 0, \quad 0, \quad 0, \quad \tfrac{2}{3}, \quad \tfrac{1}{3} \right) \quad \text{and} \quad \gamma_2^T = \left(0, \quad 0, \quad 1, \quad 1, \quad , 1, \quad \tfrac{1}{3}, \quad \tfrac{2}{3} \right),$$

$$\bar{\gamma}_1 = \left(1, \quad 0, \quad 0, \quad 0, \quad 0, \quad 0, \quad 0 \right) \quad \text{and} \quad \bar{\gamma}_2 = \left(0, \quad 0, \quad \tfrac{1}{3}, \quad \tfrac{1}{3}, \quad \tfrac{1}{3} \quad 0, \quad 0 \right).$$

The general solutions to both the random walk and the consensus problems are given in terms of these vectors as in Theorems 5.3 and 5.5. Note that for the former, Corollary 5.4 does not hold. In this example, the matrix $\left[\frac{1}{n}\sum_m \gamma_m\bar{\gamma}_m(v)\right]$ is:

$$
\frac{1}{7\cdot 9}
\begin{pmatrix}
9 & 0 & 0 & 0 & 0 & 0 & 0 \\
9 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 3 & 3 & 3 & 0 & 0 \\
0 & 0 & 3 & 3 & 3 & 0 & 0 \\
0 & 0 & 3 & 3 & 3 & 0 & 0 \\
6 & 0 & 1 & 1 & 1 & 0 & 0 \\
3 & 0 & 2 & 2 & 2 & 0 & 0
\end{pmatrix}.
$$

Thus the influence vector is

$$
\left(\ \tfrac{3}{7},\ \ 0,\ \ \tfrac{4}{21},\ \ \tfrac{4}{21},\ \ \tfrac{4}{21},\ \ 0,\ \ 0\ \right)
$$

Next, we solve for the pageranks with $\alpha = 1$ (or $\beta = \frac{1}{2}$). Without teleporting we get

$$
(I+\mathcal{L})^{-1} = \frac{1}{14\cdot 15}
\begin{pmatrix}
210 & 0 & 0 & 0 & 0 & 0 & 0 \\
105 & 105 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 120 & 30 & 60 & 0 & 0 \\
0 & 0 & 60 & 120 & 30 & 0 & 0 \\
0 & 0 & 30 & 60 & 120 & 0 & 0 \\
56 & 0 & 8 & 2 & 4 & 112 & 28 \\
14 & 0 & 32 & 8 & 16 & 28 & 112
\end{pmatrix}.
$$

From this the pagerank vector follows from Theorem 6.4:

$$
\wp = \frac{1}{294}\left(\ 77,\ \ 21,\ \ 50,\ \ 44,\ \ 46,\ \ 28,\ \ 28\ \right).
$$

The same calculation using $\mathcal{L}_t = I - S_t$ gives:

$$
\wp_t = \frac{1}{273}\left(\ 56,\ \ 21,\ \ 50,\ \ 44,\ \ 46,\ \ 28,\ \ 28\ \right).
$$

Notice that the teleporting decreases the rank of vertex 1 and re-distributes the remaining measure over the other vertices. It is easy to check that $L = \{1\}$ and $\wp_t(1) = \dfrac{(1-\beta)\wp(1)}{1-\beta\wp(1)}$ (Corollary 6.7).

8 Appendix: Discrete versus Continuous Consensus

We briefly discuss the way discrete and continuous consensus relate to one another. The case of diffusion is exactly the same.

If we start with the flow defined by $\dot{x} = -\mathcal{L}x$ on a digraph G then, we can easily determine its time 1 map, namely $x^{(n+1)} = e^{-\mathcal{L}}x^{(n)}$, as well as its properties. A graph $\tilde{G}$ is a *transitive closure* of the graph G if for every possible (directed) path in $i \rightsquigarrow j$ in G there is a (directed) edge $i \to j$ in $\tilde{G}$.

Proposition 8.1 *Consider a non-negative row-stochastic adjacency matrix S associated with the graph G and a (rw) Laplacian $\mathcal{L} = I - S$. Then*
(i): $e^{-\mathcal{L}}$ is row-stochastic,
(ii): $e^{-\mathcal{L}}$ is non-negative, and
(iii): $e^{-\mathcal{L}}$ is the adjacency matrix of a transitive closure of G.

Proof: (i): From the expansion $e^{-\mathcal{L}}$ in powers of $\mathcal{L}$, we see that $e^{-\mathcal{L}}$ must have row sum 1.
(ii): If we expand $e^{-\mathcal{L}} = e^{S-I} = e^{-1}e^{S}$ in powers of S, we see that $e^{-\mathcal{L}}$ is non-negative.
(iii): By (i) and (ii), we may consider $e^{-\mathcal{L}}$ as a (weighted) adjacency matrix $\tilde{S}$ of a graph $\tilde{G}$. The ji entry of S^k positive if and only if there is a path $i \rightsquigarrow j$ of length k in G. The expansion in e^{S} shows that every power of S occurs with a positive coefficient. Thus the ji entry of $\tilde{S}$ is positive if and only if there is a path $i \rightsquigarrow j$. ∎

Corollary 8.2 *The right kernels of $\mathcal{L}$ an $\tilde{\mathcal{L}} = I - e^{-\mathcal{L}}$ are equal. Similar for the left kernels.*

Proof: The structure of the reaches in a graph is entirely determined by the paths in that graph. Thus, the reaches and in particular their subdivision in cabals, exclusive parts, and common parts are the same for G and $\tilde{G}$. Thus the dimension of the kernels is the same. Since kernels are linear spaces, it is sufficient to show that $\ker \mathcal{L} \subseteq \ker \tilde{\mathcal{L}}$. So, let $v \in \mathcal{L}$, then

$$\tilde{\mathcal{L}}v = \left(I - e^{-\mathcal{L}}\right) v = 0 \; ,$$

where the last equality follows upon expanding the exponential. ∎

Remark: There is one substantial difference between S and $\tilde{S} = e^{-\mathcal{L}}$. While the strongly connected components of the two are equal, the restriction of $\tilde{S}$ to such a component is strictly positive. That means that the matrix restricted to such a component is *primitive*, and no periodic behavior occurs. Hence the averaging we see in Theorem 5.3 for a discrete system is replaced by a straightforward limit in Theorem 5.5 for a continuous system.

Now let us start with a time 1 map $x^{(n+1)} = Sx^{(n)}$ and consider the much more challenging problem of constructing (if possible) a flow $\dot{x} = -\mathcal{L}x$ that generates it. The matrix $-\mathcal{L}$ is sometimes referred as a *logarithm* of S [20]. A stochastic matrix S (or its corresponding Markov chain) for which such a flow exists is called *embeddable*. The problem of characterizing the embeddable stochastic matrices is known as the *embedding problem for finite Markov chains* and is an area of current research interest. The general problem of determining if a given stochastic matrix has such a flow has been recently shown to be NP-complete [16]. Recent characterizations for special classes of stochastic matrices can be found in [22] and [37].

There is an obvious obstruction to the construction of a continuous flow $\dot{x} = -\mathcal{L}x$ that generates a given discrete system $x^{(n+1)} = Sx^{(n)}$ as its time 1 map $x^{(n+1)} = e^{-\mathcal{L}}x^{(n)}$. Namely, if S has an eigenvalue 0, then there is there is no flow whose time 1 map generates it. Intuitively, this is because $e^{-\mathcal{L}}v = 0$ would imply that $\mathcal{L}v$ diverges, which, as we have seen, is impossible for a laplacian $\mathcal{L}$. For more details we refer the reader to [20] (under *logarithms of matrices*).

A more interesting obstruction follows from Proposition 8.1. If the matrix S does not correspond to a transitively closed graph, then any logarithm of S (if one exists) cannot have the structure

of the negative of a random walk Laplacian. Here is a simple example.

$$
S = \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 3/5 & 2/5 \end{pmatrix} \quad \text{and} \quad \ln(S) = \begin{pmatrix} 0 & 0 & 0 \\ \ln(2) & -\ln(2) & 0 \\ -\ln(5^5/2^{11}) & \ln(5^6/2^{12}) & -\ln(5/2) \end{pmatrix} .
$$

Perhaps surprisingly, transitive closure of the graph associated with a stochastic matrix S is not sufficient to ensure that S is embeddable. Examples can be found in [22]. These examples give real logarithms of S which are not Laplacians. One can show that there are no other real logarithms of S (c.f. Theorem 2 in [17]).

References

[1] V. I. Arnold, *Ordinary Differential Equations*, 3rd Edn, Springer, Berlin, 1992.

[2] R. Agaev, P. Chebotarev, *On the Spectra of Nonsymmetric Laplacian Matrices*, Lin. Alg. and its Appl., 399, pp. 157-168, 2005.

[3] R. Ahlswede *et al.*, *Network Information Flow*, IEEE Transactions on Information Theory, Vol. 46, No. 4, pp. 1204-1216, 2000.

[4] R. Angles, C. Guiterrez, *Survey of Graph Database Models*, ACM Computing Surveys, Vol. 40, No. 1, pp. 1-39, 2008.

[5] J. Bang-Jensen, G. Gutin, *Digraphs: Theory, Algorithms, and Applications*, Springer, Berlin, 2007.

[6] F. Bauer, *Normalized Graph Laplacians for Directed Graphs*, Lin. Alg. and its Appl.,436, pp. 4193-4222, 2012.

[7] I. Benjamini, L. Lovász, *Harmonic and Analytic Functions on Graphs*, J. Geom. 76, pp. 3-15, 2003.

[8] M. Boyle, *Notes on Perron-Frobenius Theory of Nonnegative Matrices*, www.math.umd.edu/~mboyle/courses/475sp05/spec.pdf

[9] D. Boley, G. Ranjan, Z. Zhang, *Commute Times for a Directed Graph using an Asymmetric Laplacian*, Lin. Alg. and its Appl., 435, pp. 224-242, 2011.

[10] B. Bollobás, *Modern Graph Theory*, Springer, Berlin, 1998.

[11] A. Broder *et al.*, *Graph Structure of the Web*, Computer Networks, 33, pp. 309-322, 2000.

[12] P. Carrington, J. Scott, S. Wasserman, *Models and Methods in Social Network Analysis*, Cambridge University Press, 2005.

[13] J. S. Caughman, J. J. P. Veerman, *Kernels of Directed Graph Laplacians*, Electronic Journal of Combinatorics 13, No 1, R39, 2006.

[14] F. Chung, *Laplacians and the Cheeger Inequality for Directed Graphs*, Annals of Combinatorics 9, pp. 1-19, 2005.

[15] F. Chung, W. Zhao, *Pagerank and Random Walks of Graphs*, Fete of Combinatorics and Computer Science, (G. O. H. Katona, A. Schrijver and T. Szonyi, Eds.), Springer, Berlin, pp. 43-63, 2010.

[16] J. Bausch, T. Cubitt, *The complexity of divisibility*, Lin. Alg. and its Appl. 504, 64-107, 2016.

[17] W. Culver, *On the existence and uniqueness of the real logarithm of a matrix*, Proceedings of the American Mathematical Society 17, No 5, 1146-1151, 1966.

[18] J. Fax, R Murray, *Information Flow and Cooperative Control of Vehicle Formations*, IEEE Transactions on Automatic Control, Vol. 49, No. 9, 2004.

[19] J. L. Gross, J. Zellen, P. Zhang, *Handbook of Graph Theory*, 2nd Edn, CRC Press, Boca Raton FL, 2014.

[20] N. Higham, *Functions of Matrices. Theory and Computation*, SIAM, 2008.

[21] R. A. Horn, C. R. Johnson, *Matrix Analysis*, 2nd Edn, Corrected Reprint, Cambridge University Press, New York, 2017.

[22] C. Jia, *A solution to the reversible embedding problem for finite Markov chains*, Statistics and Probability Letters 116, 122-130, 2016.

[23] T. Jombert *et al.*, *Reconstructing disease outbreaks from genetic data: a graph approach*,Heredity 106, 383-390, 2011.

[24] A. Langville, C. Meyer, *Google's PageRank and Beyond*, Princeton University Press, 2011.

[25] D. A. Levin, Yuval Peres, *Markov Chains and Mixing Times*, 2nd Edn, American Mathematical Society, Providence RI, 2017

[26] N. Masuda, H. Kori, *Dynamics Based Centrality for Directed Networks*, Phys. Review E 82, 2010.

[27] N. Masuda, *Random Walks and Diffusion on Networks*, Physics Reports, Vols. 716-717, pp. 1-59, 2017.

[28] Robert M. May, *Qualitative Stability in Model Ecosystems*, Ecology, Vol. 54, No. 3. (May, 1973), pp. 638-641.

[29] P. McGraw, M. Menzinger, *Laplacian Spectra as a Diagnostic Tool for Network Structure and Dynamics*, Physical Review E 77, 2008.

[30] M. Mesbahi, M. Egerstedt, *Graph Theoretic Methods in Multiagent Networks*, Princeton University Press, 2010.

[31] I. Mirzaev, J. Gunawardena, *Laplacian Dynamics on General Graphs*, Bull. Math. Biol., Vol. 75(11), pp. 2118-2149, 2013.

[32] M. Newman, *Networks: An Introduction*, Oxford University Press, 2010.

[33] G. Punzo *et al.*, *Using Network Dynamical Influence to Drive Consensus*, Sci. Rep. 6, 26318, 2016

[34] S. Rao, A. van der Schaft, B. Jayawardhana, *A graph-theoretical approach for the analysis and model reduction of complex-balanced chemical reaction networks*, J. Math. Chem., Vol. 51, No. 9, pp. 2401-2422, 2013.

[35] S. Sternberg, *Dynamical Systems*, Dover Publications, Mineola NY, 2010.

[36] J. J. P. Veerman, G. Lafferriere, John S. Caughman, A. Williams, *Flocks and Formations*, J. Stat.Phys. 121, Vol 5-6, pp. 901-936, 2005.

[37] A. Van-Brunt, *Infinitely divisible nonnegative matrices, M-matrices, and the embedding problem for finite state stationary Markov chains*, Lin. Alg. and its Appl. 541, 163-176, 2018.

[38] D. Wagner, *The Critical Group of a Graph*, 2000, https://arxiv.org/abs/math/0010241 Retrieved 2018.

[39] D. B. West, *Introduction to Graph Theory*, 2nd Edn, Prentice-Hall, Upper Saddle River NJ, 2001.