WALK NETWORK RECOVERABILITY AND THE CARD CONJECTURE

DAVID DIAMONDSTONE

Abstract. This paper continues the work done by Tim DeVries in 2003 on random walk network recoverability. DeVries defined a random walk network as a special type of graph on which a random walk is performed, and then posed the question of whether the edge probabilities in the graph could be recovered from the graph as well as walk-endpoint probabilities from walks beginning at all vertices in the graph. He presented a conjecture hypothesized by Ryan Card (which bore a striking similarity to a conjecture posed by Kurt Krenz in 1992) on a necessary and sufficient condition for recoverability. Furthermore, he proved that the condition was necessary. In this paper it will be shown that the condition of Card's Conjecture, while necessary, is not sufficient to ensure network recoverability. The question of network recoverability is further discussed in both algebraic and geometric terms. Finally, a refined conjecture similar to (but more restrictive than) Card's Conjecture is made.

1. Introduction

In his 2003 paper [1], Tim DeVries quoted the following conjecture, which he named "Card's Conjecture":

Conjecture 1.1. (Card's Conjecture) *A random walk network is recoverable if and only if all of the edges*[1] *leaving any interior vertex can be simultaneously extended to vertex-disjoint paths to the boundary.*

DeVries proved that it was a necessary condition for recoverability, but was unable to find a proof for the other direction. It turns out that while the condition that it must be possible to extend all edges leaving an interior vertex to disjoint boundary paths is necessary, it is not sufficient for recoverability. (As an aside, we note that DeVries proof required something that DeVries stated as the "Choke Lemma" but did not provide a proof for. This lemma is equivalent to Menger's n-Arc Theorem, which was proved in 1927. [3])

In the next section, necessary background on random walk networks is given, including a brief discussion of Card's Conjecture. In section three, a simple network that is a counterexample to Card's Conjecture is discussed, and shown to be a counterexample. In sections four and five, the question of when a network is recoverable is discussed, first in algebraic terms, and then in geometric ones. Finally, in section six, a new conjecture is made, which is meant to replace Card's Conjecture.

[1] Even though he dealt entirely with directed graphs, DeVries used the term edge rather than arc. The same convention will be followed in this paper. Please note that every time the word edge is used in this paper, a directed edge is meant.

2. Background

For the following definitions, let $G = (V, \partial V, E)$ be a directed graph with boundary (where V is the set of vertices, $\partial V \subset V$ is the subset of boundary vertices, and E is the set of edges.) Let $Int = V - \partial V$ be the set of interior vertices.

Definition 2.1. A *walk* is a finite sequence of vertices $w = (v_1, v_2, ..., v_k)$ for some $v_i \in V$ with the property that for all $1 \leq i < k$, there is an edge from v_i to v_{i+1}. (The *initial vertex* of the walk is v_1, the *ith vertex* of the walk is v_i, the *final vertex* of the walk is v_k, and the number k is the *length* of w). An *infinite walk* is an infinite sequence of vertices $(v_1, v_2, ...)$ for some $v_i \in V$ with the property that for all $i \in \mathbb{N}$, there is an edge from v_i to v_{i+1}. (For infinite walks, the initial and ith vertices are still defined, but not the final vertex or length.) A *uv-walk* is a walk with initial vertex u and final vertex v.

Definition 2.2. We say a walk or infinite walk *visits* a vertex v if $v = v_i$ for some i. We say a walk or infinite walk visits a vertex v n times if $v = v_i$ for exactly n distinct i. We say a walk or infinite walk visits an edge e if e is the edge from v_i to v_{i+1} for some value of i, and we say a walk or infinite walk visits an edge e n times if e is the edge from v_i to v_{i+1} for exactly n distinct values of i.

Definition 2.3. A *path* is a walk where all vertices visited by the walk are visited once. A uv-path is a uv-walk that is a path. A *boundary path* is a path where the final vertex is a boundary vertex.

Definition 2.4. Let $u, v \in V$ be (not necessarily distinct) vertices. We write $u \sim v$ if there is an edge to u from v. Note that $u \sim v$ and $v \sim u$ are not equivalent. When $u \sim v$ we say u is adjacent to v, but again we mean this in a directional sense, so that just because u is adjacent to v (in the sense that will be used in this paper) does not mean that v is adjacent to u.

Definition 2.5. *(random walk network)*
Let G satisfy the following four properties:

(1) If $u, v \in V$ are any two (not necessarily distinct) vertices with $v \sim u$, there is exactly one uv-edge.
(2) If u, v are interior vertices, $u \sim v$ if and only if $v \sim u$.
(3) If u is any vertex and v is a boundary vertex, $u \sim v$ if and only if $u = v$.
(4) For all interior vertices v, there is a boundary path with initial vertex v.

Let ρ be a function from $V \times V$ to $[0, 1]$ which satisfies the following two equations:

$$\text{For all } u \in V, \sum_{v:v\sim u} \rho(u, v) = 1 \tag{2.1}$$

$$\rho(u, v) > 0 \text{ if and only if } v \sim u \tag{2.2}$$

(Since $\rho(u, v)$ is always zero if there is no (directed) edge from u to v, and if there is such an edge it is unique, ρ can also be thought of as a function on edges. Such a function ρ - whether defined on edges or on ordered pairs of vertices - is called a transition probability.) The pair $\Gamma = \{G, \rho\}$ is a *random walk network.*

Let Γ be a random walk network, and let S_f be the set of all walks and S_∞ the set of all infinite walks on Γ. Given a walk $w \in S_f$ with length n, let B_w be the set of all infinite walks $w' \in S_\infty$ such that the initial n terms of w' is the sequence

w. Let $A \subset \mathcal{P}(S_\infty)$ be the set defined by the properties: (i) for all walks w on Γ, $B_w \in A$; and (ii) A is closed under complement and finite and countable union. Note that $B_u = \bigcup_{v:v\sim u} B_{(e_{uv})}$ where (e_{uv}) is the walk consisting of the single edge from u to v.

Let $f : S_f \to (0, 1]$ be the function:

$$f(w) = \begin{cases} \prod_{i=1}^{k-1} \rho(v_i, v_{i+1}) & \text{if } k > 1 \\ 1 & \text{if } k = 1 \end{cases} \tag{2.3}$$

for $w = (v_1, ..., v_k) \in S_f$. For each $v \in V$, define the measure P_v on A by:

$$P_v(B_w) = \begin{cases} f(w) & \text{if } v \text{ is the initial vertex of } w \\ 0 & \text{otherwise} \end{cases} \tag{2.4a}$$

$$P_v\left(\bigcup_{i\in\mathbb{N}} B_i\right) = \sum_{i\in\mathbb{N}} m_v(B_i) \text{ for } B_i \in A \text{ pairwise disjoint sets} \tag{2.4b}$$

$$P_v(S - B) = 1 - P_v(B) \text{ for } B \in A \tag{2.4c}$$

Theorem 2.1. *A is a Borel algebra, and P_v is a probability measure on A.*

For definitions of a Borel algebra and a probability measure, as well as for a proof of this theorem, see Andrew Lewis' paper "Random Walk Networks" [2]. This makes everything we will say about the probabilities of various events rigorous as the measure (with the respect to this probability measure) of a corresponding set in A. The interperetation for $P_v(B)$ is the probability that a random walk on Γ with initial vertex v and transition probabilities ρ will be an element of the set B.

Proposition 2.2. *Given a random walk network $\Gamma = \{G, \rho\}$, and a vertex v of G, let W_v be the set of infinite walks $w = (v_1, v_2, ...)$ that visit v. Then $W_v \in A$.*

Proof. Let $W_v^{(n)}$ be the set of infinite walks $w = (v_1, ...)$ with nth vertex v, where v is not the kth vertex for $k < n$. Let $W_v'^{(n)}$ be the set of walks with length n and final vertex v that visit v once. Note that $W_v'^{(n)}$ is finite, since there are at most $|V|^n$ walks of length n. Then:

$$W_v^{(n)} = \bigcup_{w\in W_v'^{(n)}} B_w \tag{2.5}$$

and

$$W_v = \bigcup_{n\in\mathbb{N}} W_v^{(n)} \tag{2.6}$$

so $W_v \in A$. □

We can therefore talk about the probability $P_u(W_v)$ for any pair of vertices u and v that a walk with initial vertex u will visit v. We will write this as $P(u \to v)$.

Lemma 2.3. *Given a random walk network $\Gamma = \{G, \rho\}$ and vertices u and v of G,*

$$P(u \to v) = \sum_{w\in W'} f(w) \tag{2.7}$$

where W' is the set of all uv-walks that visit v once.

Proof. We have $P(u \to v) = P_u(W_v)$, where $W_v = \bigcup_{n\in\mathbb{N}} W_v^{(n)}$, and $W_v^{(n)} = \bigcup_{w\in W_v'^{(n)}} B_w$, where $W_v'^{(n)}$ is the set of walks with length n and final vertex v, that visit v once. Note that $W_v = \bigcup_{n\in\mathbb{N}} W_v^{(n)}$ is a *disjoint* union, since an infinite walk w that visits v is in the set $W_v^{(n)}$ where n is the least natural number such that v is the nth vertex, and in no other set $W_v^{(m)}$. Since $P_u(B_w) = f(w)$ for all walks w with initial vertex u, and $B_w \cap B_{w'} = \emptyset$ for $w, w' \in W_v'^{(n)}$, $w \neq w'$:

$$P_u(W_v^{(n)}) = \sum_{w\in W_{uv}'^{(n)}} f(w) \tag{2.8}$$

where $W_{uv}'^{(n)}$ is the set of uv-walks with length n that visit v once. Therefore, we have

$$\begin{aligned} P_u(W_v) = P_u\left(\bigcup_{n\in\mathbb{N}} W_v^{(n)}\right) &= \sum_{n\in\mathbb{N}} \sum_{w\in W_{uv}'^{(n)}} f(w) \\ &= \sum_{w\in W'} f(w) \end{aligned} \tag{2.9}$$

□

Proposition 2.4. *Given a random walk network $\Gamma = \{G, \rho\}$, let q be an interior vertex of G with n adjacent vertices $Q = \{q_1, ..., q_n\}$, and let s be a boundary vertex. Then:*

$$P(q \to s) = \sum_{i=1}^{n} \rho(q, q_i) P(q_i \to s) \tag{2.10}$$

Proof. From Lemma 2.3, we have:

$$P(q \to s) = \sum_{w\in W} f(w) \tag{2.11}$$

where W is the set of all qs-walks that visit s once. Any walk with initial vertex q must have second vertex q_i for some $q_i \in Q$, so any walk $w \in W$ is q followed by a $q_i s$-walk. Let $W_i \subset W$ be the set of walks $w \in W$ with second vertex q_i, and let W_i' be the set formed from W_i be removing the initial q from each walk $w \in W_i$. We must have:

$$\begin{aligned} P(q \to s) &= \sum_{i=1}^{n} \sum_{w\in W_i} f(w) \\ &= \sum_{i=1}^{n} \sum_{w\in W_i'} \rho(q, q_i) f(w) \\ &= \sum_{i=1}^{n} \rho(q, q_i) P(q_i \to s) \end{aligned}$$

□

For the following two definitions, let the vertices of V be ordered from 1 to $N = k + d$ (where k is the number of interior vertices, d is the number of boundary vertices) in such a way that the boundary vertices are labeled from 1 to d, and the interior vertices are labeled from $d+1$ to $d+k=N$. Then we can write $V=\{v_1,...,v_N\}$, $\partial V=\{v_1,...,v_d\}$, $Int=\{v_{d+1},...,v_{d+k}\}$.

Definition 2.6. Given a random walk network Γ, we define the transition matrix, $T=(t_{ij})$, for Γ by:

$$t_{ij}=\rho(v_i,v_j) \tag{2.12}$$

T can be written in block form as:

$$T=\begin{pmatrix} I & \mathbf{0} \\ X & Y \end{pmatrix} \tag{2.13}$$

where I is the $d\times d$ identity matrix, $\mathbf{0}$ is the $d\times k$ zero matrix, and X and Y are $k\times d$ and $k\times k$, respectively.

Definition 2.7. Given a random walk network Γ, we define the *absorption matrix*, $B=(b_{ij})$, for Γ as the $|Int|\times|\partial V|$ matrix:

$$b_{ij}=P(v_{i+d}\to v_j) \tag{2.14}$$

We define the *extended absorption matrix*, $\hat{B}$, to be the $|V|\times|\partial V|$ matrix:

$$\hat{b}_{ij}=P(v_i\to v_j) \tag{2.15}$$

The extended absorption matrix $\hat{B}$ gives the probabilities that a walk from an *interior or boundary* vertex i will visit a boundary vertex j, whereas the absorption matrix B gives the probabilities that a walk from an *interior* vertex i will visit a boundary vertex j. We see immediately that B and $\hat{B}$ contain the same information, since the only edges out of boundary vertices are loops, so $\hat{B}$ takes block form $\begin{pmatrix} I \\ B \end{pmatrix}$ where I is the $|\partial V|\times|\partial V|$ identity matrix.

Theorem 2.5. *If $\Gamma=\{G,\rho\}$ is a random walk network on a directed graph with boundary $G=\{V,\partial V,E\}$ (where $|\partial V|=d$ and $|V-\partial V|=k$), with transition matrix T and absorbtion matrix B, then:*

$$B=(I-Y)^{-1}X,\ \textit{where}\ T=\begin{pmatrix} I & \mathbf{0} \\ X & Y \end{pmatrix} \tag{2.16}$$

where I is the $d\times d$ identity matrix, $\mathbf{0}$ is the $d\times k$ zero matrix, and X and Y are $d\times k$ and $x\times k$, respectively. Furthermore, the probability that a random walk will reach the boundary from any vertex in the interior is 1.

For a proof, see [1].

Definition 2.8. Given a directed graph with boundary $G=(V,\partial V,E)$ which satisfies all the properties required to define a random walk network on G, Let $\mathcal{T}$ be the set of transition matrices on G. We then define the transition to absorption map for G as the function $\varphi_G:\mathcal{T}\to\mathcal{M}_{|Int|\times|\partial V|}(\mathbb{R})$ that sends a transition matrix T to the associated absorption matrix B.

Definition 2.9. A pair G,B is said to be recoverable if there is a unique transition matrix T on the graph G such that $\varphi(T)=B$. A graph G is called recoverable if φ_G is injective. Given an edge $e_{uv}\in E$ from u to v, $\rho(e_{uv})$ is said to be recoverable from G and B if $T(u;v)$ is constant on all $T\in\mathcal{T}$ such that $\varphi_G(T)=B$.

In his 2003 paper "Recoverability of Random Walk Networks" [1], Tim DeVries gave Card's Conjecture and proved one direction of it:

Theorem 2.6. (DeVries, 2003) *If a random walk network is recoverable, then all of the edges leaving any interior vertex can be simultaneously extended to vertex-disjoint paths to the boundary.*

Unfortunately, the other half of the conjecture is false.

3. A Counterexample to Card's Conjecture

Consider the random walk network in figure 3.1, which we will call C_a to denote its dependence on the parameter a. It is easy to check that all of the edges leaving any interior vertex can be simultaneously extended to vertex-disjoint paths to the boundary.

The transition matrix T_a for C_a is the following:

$$T(a) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & a & \frac{1}{2}-a & a & \frac{1}{2}-a & 0 \end{bmatrix} = \begin{pmatrix} I & \mathbf{0} \\ X & Y_a \end{pmatrix} \tag{3.1}$$

where I is the 4×4 identity matrix, $\mathbf{0}$ is the 4×5 zero matrix,

$$X = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 \end{bmatrix}, \text{ and } Y_a = \begin{bmatrix} 0 & 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & \frac{1}{3} \\ a & \frac{1}{2}-a & a & \frac{1}{2}-a & 0 \end{bmatrix}.$$

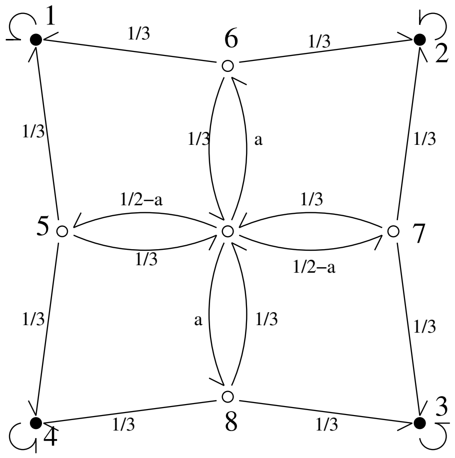

Figure 3.1. The random walk network C_a. (The center vertex is vertex 9. Black vertices are boundary, white are interior.) The parameter a may be any real number $0 < a < \frac{1}{2}$.

Now consider a walk beginning at the center vertex (9) of C_a. C_a has north-south symmetry, so the probability of such a walk visiting vertex 1 and vertex 4 are the same, and likewise for vertices 2 and 3. Likewise, C_a has east-west symmetry, so the probability of a walk visiting 1 and 2 are the same, and therefore all probabilities are the same, hence all must be $\frac{1}{4}$. A walk 'step' from any side vertex (5, 6, 7, or 8) has a $\frac{1}{3}$ probability of going to the adjacent boundary vertices, and a $\frac{1}{3}$ chance of going to the center, at which point a walk has a $\frac{1}{4}$ probability of ending up at each of the four boundary vertices, so the probabilities are $\frac{1}{3} + \frac{1}{3} \times \frac{1}{4} = \frac{5}{12}$ of ending up at the adjacent two boundary vertices, and $\frac{1}{12}$ of ending up at the other two boundary vertices. Thus our absorbtion matrix B_a must be the following, which is constant as a function of a!

$$B_a = \begin{bmatrix} \frac{5}{12} & \frac{5}{12} & \frac{1}{12} & \frac{1}{12} \\ \frac{1}{12} & \frac{5}{12} & \frac{5}{12} & \frac{1}{12} \\ \frac{1}{12} & \frac{1}{12} & \frac{5}{12} & \frac{5}{12} \\ \frac{5}{12} & \frac{1}{12} & \frac{1}{12} & \frac{5}{12} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} \tag{3.2}$$

Since the absorbtion matrix does not depend on a, the inverse problem for this graph and absorbtion matrix does not have a unique solution. Therefore Card's Conjecture must be false. (It is also possible to show algebraically that $B_a = (I - Y_a)^{-1}X$ is a constant function of a.)

4. Further Musings on Recoverability

Let $\Gamma = \{G, \rho\}$ be a random walk network with transition matrix T written in block form as

$$T = \begin{pmatrix} I & \mathbf{0} \\ X & Y \end{pmatrix}$$

where the four blocks represent boundary-boundary, boundary-interior, interior-boundary, and interior-interior connections, respectively. Let $B = (b_{ij})$ be the absorbtion matrix, $B = (I-Y)^{-1}X$, and recall that $b_{ij} = P(u \to v)$ where u is the ith interior vertex and v is the jth boundary vertex. Consider an interior vertex q with n adjacent vertices $Q = \{q_1, ..., q_n\}$, and a boundary vertex s. Recall equation 2.10:

$$P(q \to s) = \sum_{i=1}^{n} \rho(q, q_i) P(q_i \to s)$$

If we wish to solve the inverse problem for ρ from our graph G and the absorbtion matrix B, $P(q \to s)$ is an entry in B, and $P(q_i \to s)$ is an entry in the extended absorption matrix $\hat{B}$, so we may take these to be known, and $\rho(q, q_i)$ to be unknown. This is therefore a system of $|\partial V|$ equations in n variables. Because B is in fact the absorbtion matrix for some function ρ defined on pairs of vertices of G, these equations will be consistent, and there will be a unique solution if and only if n of the equations are linearly independent. Thus the system is uniquely solvable if and only if there is a set $S = \{s_1, ..., s_n\}$ of n boundary vertices such that $\det(\hat{B}(Q; S)$ is nonzero.

Theorem 4.1. *Let $\Gamma = \{G, \rho\}$ be a random walk network with absorption matrix B and extended absorption matrix $\hat{B}$. If q is an interior vertex of G with n adjacent* *vertices* $Q = \{q_1, ..., q_n\}$, *then the probabilities* $\rho(q, x)$ *on the edges leaving* q *are recoverable from* G *and* B *if and only if there exists a set* $S = \{s_1, ..., s_n\}$ *of* n *boundary vertices such that* $\det(\hat{B}(Q; S))$ *is nonzero.*

Proof. From the argument immediately prior to the statement of this theorem, we see that the probabilities are recoverable if there is such a set S. If there is no such set S, then any set S' of n boundary vertices not adjacent to q has $\det(\hat{B}(Q; S')) = 0$, so $\operatorname{rank}(\hat{B}(Q; \partial V)) \leq n - 1$, hence:

$$\dim(\ker(\hat{B}(Q; \partial V)^T)) \geq 1 \tag{4.1}$$

Let $\vec{y} = (y_1...y_n)^T$ be a nonzero (column) vector of the kernel of $\hat{B}(Q; \partial V)^T$. Since the column sums of $\hat{B}(Q; \partial V))^T$ are 1, $\vec{y} \cdot \vec{1} = 0$ where $\vec{1} \in \mathbb{R}^n$ is the vector of all ones. Let $\vec{x}$ be the (column) vector $(\rho(q, q_1)...\rho(q, q_n))^T$ of probabilities on the edges leaving vertex q. Note that:

$$\hat{B}(Q; \partial V)^T \vec{x} = (\hat{B}(q; \partial V))^T \tag{4.2}$$

Let m be the minimum entry in $\vec{x}$. Since $\rho(q, q_i)$ is positive whenever $q_i \sim q$ (which is true for all i), $m > 0$. Let $c \in (0, 1)$ be sufficiently small that $|cy_i| < m$ for all i. Then let $\vec{x}' = \vec{x} + c\vec{y}$ and let x'_i be defined by $\vec{x}' = (x'_1...x'_n)$. The sum of all entries in $\vec{x}'$ is 1, since $\vec{1} \cdot \vec{x} = 1$ and $\vec{1} \cdot \vec{y} = 0$; all entries in $\vec{x}'$ are positive, by the triangle inequality, since all entries in $\vec{x}$ are positive, and are greater in magnitude than all entries of $c\vec{y}$. From the definition of $\vec{x}'$, we obtain:

$$\hat{B}(Q; \partial V)^T \vec{x}' = \hat{B}(Q; \partial V)^T (\vec{x} + c\vec{y}) = (\hat{B}(q; \partial V))^T \tag{4.3}$$

Let ρ' be a new function on $V \times V$ defined by:

$$\rho'(u, v) = \begin{cases} \rho(u, v) & \text{if } u \neq q \text{ or } v \notin Q \\ x'_i & \text{if } u = q \text{ and } v = q_i \end{cases} \tag{4.4}$$

Note that ρ' is a transition probability. Let $\Gamma' = \{G, \rho'\}$ be a new random walk network on the original graph G with the new probability function ρ', and let B' be the absorbtion matrix of Γ'. We denote the probability that a random walk beginning at u will visit v by $P(u \to v)$ in Γ, and by $P'(u \to v)$ in Γ', and we denote the probability of an infinite walk from a vertex u being in a set $W \in A$ by $P_u(W)$ or $P'_u(W)$ similarly. Let $P_{(q)}(u \to v)$ denote the probability that a walk beginning at u will visit v before visiting q in Γ; formally:

$$P_{(q)}(u \to v) = P_u(W) \tag{4.5}$$

where W is the set of all infinite walks with initial vertex u that visit v before visiting q. We can show $W \in A$ in a manner analogous to the proof of Proposition 2.2. Again, we define $P'_{(q)}(u \to v)$ similarly.

Then, given a vertex $v \in \partial V$:

$$\begin{aligned}
P(q \to v) &= \sum_{i=1}^{n} \rho(q, q_i) P(q_i \to v) \\
&= \sum_{i=1}^{n} \rho'(q, q_i) P(q_i \to v) \\
&= \sum_{i=1}^{n} \rho'(q, q_i) \left[P_{(q)}(q_i \to v) + P(q_i \to q) P(q \to v) \right] \\
&= \left[\sum_{i=1}^{n} \rho'(q, q_i) P_{(q)}(q_i \to v) \right] + P(q \to v) \left[\sum_{i=1}^{n} \rho'(q, q_i) P(q_i \to q) \right]
\end{aligned}$$

So,

$$P(q \to v) - P(q \to v) \left[\sum_{i=1}^{n} \rho'(q, q_i) P(q_i \to q) \right] = \sum_{i=1}^{n} \rho'(q, q_i) P_{(q)}(q_i \to v)$$

$$P(q \to v) = \frac{\sum_{i=1}^{n} \rho'(q, q_i) P'_{(q)}(q_i \to v)}{1 - \sum_{i=1}^{n} \rho'(q, q_i) P'(q_i \to q)} \tag{4.6}$$

since $P'_{(q)}(q_i \to v) = P_{(q)}(q_i \to v)$ and $P'(q_i \to q) = P(q_i \to q)$ since no walk that contributes to either probability visits any of the edges where ρ' is different from ρ. We know that the denominator is nonzero, since it is the probability that a walk beginning at q will not visit q more than once, which cannot be 1 because there is a boundary path with initial vertex q, so the probability that a walk from q will reach the boundary without visiting q a second time is nonzero.

A similar chain of equations dealing with the network Γ' shows that

$$P'(q \to v) = \frac{\sum_{i=1}^{n} \rho'(q, q_i) P'_{(q)}(q_i \to v)}{1 - \sum_{i=1}^{n} \rho'(q, q_i) P'(q_i \to q)} \tag{4.7}$$

Therefore, we must have $P'(q \to v) = P(q \to v)$ for all $v \in \partial V$.

Let W be the set of infinite walks that visit v, let W' be the set of infinite walks that visit v without first having visited q, and let W'' be the set of infinite walks that visit q. Then:

$$\begin{aligned}
B'(u; v) &= P'(u \to v) = P'_u(W) = P'(W') + P'(q \to v) P'(W'') \\
&= P(W') + P(q \to v) P(W'') = P(W) = P(u \to v) = B(u; v)
\end{aligned} \tag{4.8}$$

We see that $B' = B$; there are two transition probabilities with the same associated absorbtion probability and different probabilities on edges leaving q, so the probabilities on the edges leaving q are not recoverable. □

5. From Algebra to Geometry

Now let us go back and examine the determinant of $\hat{B}(Q, S)$. From the definition of the determinant, we have:

$$
\begin{aligned}
\det(\hat{B}(Q;S)) &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \hat{b}_{i\sigma(i)} \\
&= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} P(q_i \to s_{\sigma(i)}) \\
&= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \sum_{w\in W(q_i,s_{\sigma(i)})} f(w)
\end{aligned} \tag{5.1}
$$

where $W(q_i, s)$ is the set of q_is-walks that visit s once.

Let an n-*walk* be a set $\omega = \{w_1, ..., w_n\}$ of n walks. Let $\hat{f}$ be a function from the set of n-walks to $\mathbb{R}$, defined by $\hat{f}(\omega) = \prod_{w\in\omega} f(w)$ where f is defined as in (2.7). Let $\Omega(Q, S)$ be the set of all n-walks $\omega = \{w_1, ..., w_n\}$ such that every element of Q is the initial vertex of one of the w_i, and every element of S is the final vertex of one of the w_i. Let $\sigma(\omega)$ be the permutation $\sigma \in S_n$ such that $s_{\sigma(i)}$ is the final point of the walk $w_j \in \omega$ with initial point q_i for all $i \in \{1, ..., n\}$ [2] Given $\sigma \in S_n$, let $\Omega(\sigma)$ be the set $\Omega(\sigma) = \{\omega \in \Omega(Q, S)|\sigma(\omega) = \sigma\}$. Then:

(5.2)

$$
\begin{aligned}
\det(\hat{B}(Q;S)) &= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \sum_{w\in W(q_i,s_{\sigma(i)})} f(w) \\
&= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \left(\sum_{w_1\in W(q_1,s_{\sigma(1)})} f(w_1) \right) \cdots \left(\sum_{w_n\in W(q_n,s_{\sigma(n)})} f(w_n) \right) \\
&= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \sum_{w_1\in W(q_1,s_{\sigma(1)})} \cdots \sum_{w_n\in W(q_n,s_{\sigma(n)})} f(w_1)...f(w_n) \\
&= \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma) \sum_{\omega\in\Omega(\sigma)} \hat{f}(\omega) \\
&= \sum_{\sigma\in S_n} \sum_{\omega\in\Omega(\sigma)} \hat{f}(\omega)\operatorname{sgn}(\sigma(\omega)) \\
&= \sum_{\omega\in\Omega(Q,S)} \hat{f}(\omega)\operatorname{sgn}(\sigma(\omega))
\end{aligned}
$$

Here we have gone from a purely algebraic construction (determinant) to a more geometric one (a certain sum of values of a function defined on n-walks).

6. A New Conjecture for Recoverability

Let us consider (5.2):

$$
\det(\hat{B}(Q;S)) = \sum_{\omega\in\Omega(Q,S)} \hat{f}(\omega)\operatorname{sgn}(\sigma(\omega))
$$

[2]The permutation $\sigma(\omega)$ is dependent on how we order the sets Q and S, so we must be consistent with our orderings of these sets throughout any discussion. However, we will be interested primarily in whether the two permutations $\sigma(\omega_1), \sigma(\omega_2) \in S_n$ "induced" by two n-walks ω_1, ω_2 are the same, or have the same sign, and that is independent of the orderings of Q and S.

One way of analyzing when this is non-zero is to cancel n-walks with opposite signs of $\sigma(\omega)$ on which $\hat{f}$ is the same. While it is possible for different n-walks to 'accidentally' give the same value $\hat{f}$, we will consider sets of n-walks on which $\hat{f}$ is necessarily constant - i.e. both n-walks visit each edge of the graph the same number of times.

Definition 6.1. Given an n-walk $\omega = \{w_1, ..., w_n\}$, and a vertex or edge $x \in V$ or E, we say ω visits x if w_i visits x for some i, and we say ω visits x k times if the sum from $i = 1$ to n of the number of times w_i visits x is k.

Definition 6.2. Given an n-walk $\omega = \{w_1, ..., w_n\}$, the set of initial vertices of ω, $V_I(\omega)$, is the set of vertices that are initial vertices of one of the walks w_i, and the set of final vertices of ω, $V_F(\omega)$, is the set of vertices that are final vertices of one of the walks w_i.

Definition 6.3. Given an n-walk ω, we define $R(\omega)$ to be the set of all n-walks ω' such that $V_I(\omega') = V_I(\omega)$, $V_F(\omega') = V_F(\omega)$, and for all $e \in E$, the number of times e is visited by ω' is equal to the number of times e is visited by ω. We define $R^+(\omega)$ as the set $\{\omega' \in R(\omega) | \mathrm{sgn}(\sigma(\omega')) = 1\}$, and we define $R^-(\omega)$ to be the set $R(\omega) - R^+(\omega)$.

Observation. Since $f(w)$ is the product of the edge probabilities $\rho(e)$ on all edges visited by w raised to the power of the number of times edge e was visited, we can rewrite $\hat{f}(\omega)$ as:

$$\begin{aligned} \hat{f}(\omega) &= \prod_{w \in \omega} f(w) \\ &= \prod_{e \in E} \rho(e)^{k_e} \end{aligned} \tag{6.1}$$

where k_e is the number of times edge e is visited by ω. Therefore, $\hat{f}(\omega') = \hat{f}(\omega)$ for all $\omega' \in R(\omega)$.

Observation. Since we know that $\hat{f}(\omega') = \hat{f}(\omega)$ for all $\omega' \in R(\omega)$, equation (5.2):

$$\det(\hat{B}(Q; S)) = \sum_{\omega \in \Omega(Q,S)} \hat{f}(\omega) \mathrm{sgn}(\sigma(\omega))$$

can be rewritten

$$\det(\hat{B}(Q; S)) = \sum_{R(\omega): \omega \in \Omega(Q,S)} \hat{f}(\omega)(|R^+(\omega)| - |R^-(\omega)|) \tag{6.2}$$

This allows us to consider various n-walks ω and look at the cancellations in (5.2) between the terms for elements in $R(\omega)$. For example, if none of the walks in ω share any vertices with each other, then an initial vertex $q \in V_I(\omega)$ is connected by edges used by the n-walk to the endpoint of the walk with initial vertex q, and to no other. Therefore any other n-walk in $R(\omega)$ must have the same final vertex for the walk with initial vertex q, and since q is arbitrary, we see that $\sigma(\omega') = \sigma(\omega)$ for all $\omega' \in R(\omega)$.

If two walks in ω each visit a vertex v once, and no other vertex is visited by more than one walk in ω, then to each n-walk $\omega' \in R(\omega)$ we can associate another n-walk $\tilde{\omega}' \in R(\omega)$ where all walks disjoint from v are the same, and the two walks passing through v are modified in the following way: let $w_1, w_2 \in \omega$ be the two walks which intersect v, and let $w_1^{(1)}$ and $w_1^{(2)}$ denote the portions of w_1 before (and including) and after v, respectively; declare $w_2^{(1)}$ and $w_2^{(2)}$ to be the before- and after-v portions of w_2 similarly; and replace w_1 and w_2 by two walks which are the concatenation of $w_1^{(1)}$ and $w_2^{(2)}$, and the concatenation of $w_2^{(1)}$ and $w_1^{(2)}$. Note that $\tilde{\tilde{\omega}}' = \omega'$, so the map $\omega' \mapsto \tilde{\omega}'$ is a bijection that reverses the sign of $\sigma(\omega')$. Therefore, $R(\omega)$ must contain the same number of elements ω' with $\operatorname{sgn}(\sigma(\omega'))$ positive as elements where it is negative.

Now we are ready to give a refined conjecture along the same lines as Card's Conjecture. This conjecture is meant to give sufficient conditions for a network Γ to be recoverable. [3]

Conjecture 6.1. (Refined Walk Network Recoverability Conjecture) *If Γ is a random walk with interior vertex q, where $Q = \{q_1, ..., q_n\}$ is the set of vertices adjacent to q, such that there is a set S of n boundary vertices so that for all n-walks $\omega = \{w_1, ..., w_n\} \in \Omega(Q, S)$ where the w_i are disjoint paths we have that $sgn(\sigma(\omega)) = 1$, and there exists an n-walk ν in $\Omega(Q, S)$ that is a set of disjoint paths, then the probabilities $\rho(q, q_i)$ are recoverable from G and B. If all interior vertices have that property, then Γ is recoverable.*

In order to give a justification for why I made this conjecture (6.1), I will make the following purely graph-theoretical conjecture (which makes no reference to edge weights) which will imply conjecture 6.1, and give an argument for why it should be true.

Conjecture 6.2. *If $Q \subset Int$ and $S \subset \partial V$ are two sets of n interior and boundary vertices, respectively, such that for all n-walks $\omega = \{w_1, ..., w_n\} \in \Omega(Q, S)$ where the w_i are disjoint paths we have that $sgn(\sigma(\omega)) = 1$, then for all $\omega \in \Omega(Q, S)$, $|R^+(\omega)| \geq |R^-(\omega)|$.*

An argument for why this conjecture might be true:

First we will consider a specific case where $n = 2$ for the sake of simplicity. Let the n-walk $\omega = \{w_1, w_2\}$ from the set of initial vertices $\{1, 2\}$ to the set of final vertices $\{3, 4\}$ be as in figure 6.1. Note that the only n-walk from $\{1, 2\}$ to $\{3, 4\}$ where the walks are disjoint paths has a path with initial vertex 1 and final vertex 3, and a path with initial vertex 2 and final vertex 4. Then in figure 6.2 where we see the set of 2-walks $R(\omega)$ there are four 2-walks with $1 \to 4$ and $2 \to 3$, and five 2-walks with $1 \to 3$ and $2 \to 4$, so the conjecture is true in this case.

In general, we see that in any case where a group of edges can be used in more than one way by the n-walk, those vertices form a loop. If the loop is used partially by one walk and partially by another walk, it's a crossing, and if it's used entirely by one walk, it is not a crossing. Then each crossing corresponds to a swap in the permutation, so we can write the permutation as a product of swaps that correspond to crossings like this, and we see that the permutation is of the same sign as the permutation of the n disjoint paths if and only if the number of crossings is even. The fewer the crossings, the more the loops that are not crossings, and such loops

[3]It is possible that this conjecture also gives necessary conditions, but I am hesitant to even conjecture that as of yet. A proof of necessity would require that it be shown that there is an appropriate selection of transition probabilities such that certain subdeterminants of $\hat{B}$ would be zero. This would be very similar to an unproven conjecture with regards to electrical networks, which have been better studied.

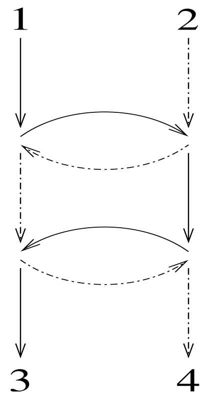

FIGURE 6.1. A sample 2-walk ω.

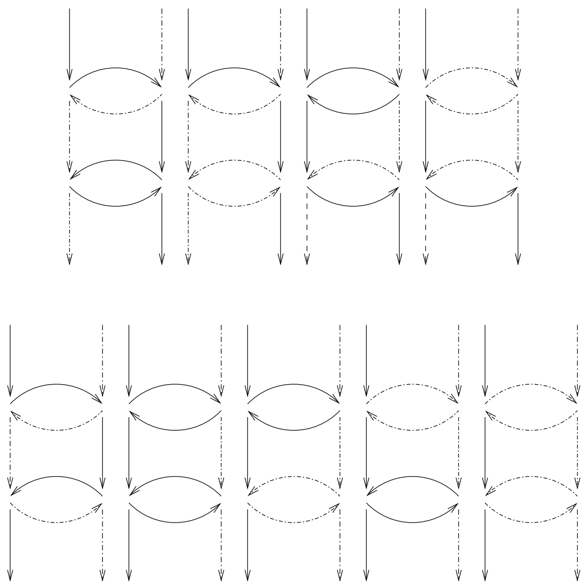

FIGURE 6.2. The set $R(\omega)$.

can be attached to any of the walks that pass through some part of the loop. In other words, there is some "stuff" that occurs where there are potential crossings. Where there is a crossing, that restricts how the "stuff" is used, but if there is no crossing, there is more freedom in how the n-walk uses the "stuff," so on average, there are more walks with fewer crossings than with more. Since there are walks with zero crossings and no walks with negative numbers of crossings, and there is a limit to how many crossings there can be, the numbers of possible crossings that are even are less, on average, than the numbers of possible crossings that are odd, so there are more n-walks with even numbers of crossings. However, there are many intricacies to this problem that will make difficult a rigorous proof.

Claim. *Conjecture 6.2 implies conjecture 6.1.*

Proof. To prove the claim, we will give a proof of conjecture 6.1, with the results of conjecture 6.2 assumed to be true.

If all n-walks ω from Q to S that are sets of disjoint paths have $\mathrm{sgn}(\sigma(\omega)) = 1$, then more things with a given value of $\hat{f}$ will have positive sign than negative, so equation (6.2)

$$\det(\hat{B}(Q;S)) = \sum_{R(\omega):\omega\in\Omega(Q,S)} \hat{f}(\omega)(|R^+(\omega)| - |R^-(\omega)|)$$

will be a sum of non-negative terms, since $\hat{f}$ is identically positive, and $|R^+(\omega)| - |R^-(\omega)|$ is non-negative. Since there exists a walk ν that is a set of disjoint paths, the term in the sum corresponding to $R(\nu)$ will be positive, so the sum will be positive, and we have $\det(\hat{B}(Q;S)) \neq 0$, for any transition probability ρ. Since $\det(\hat{B}(Q;S)) \neq 0$, the probabilities on the edges leaving q are recoverable from G and B. If all interior vertices have the property, then the probabilities on the edges leaving all interior vertices are recoverable, and since all edges are either leaving an interior vertex or are a loop where ρ must equal 1, all edge probabilities are recoverable, so Γ is recoverable from G and B. Since this is true for any transition probability ρ, the map φ_G is injective, so Γ is recoverable. ☐

7. Appendix: MATLAB code to recover a network from G and B

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% RWNrecover.m                                                             %
% When given the adjacency matrix and absorbtion matrix for a random walk %
% network, returns the transition matrix T if the network is recoverable  %
% from G and B, otherwise it prints a message stating that the pair (G,B) %
% is not recoverable.                                                      %
%                                                                          %
% Call as RWNrecover(G,B) where G is the graph adjacency matrix, and B is %
% the absorbtion matrix.                                                   %
%                                                                          %
% Author: David Diamondstone; diamondstone@brandeis.edu                    %
% August 11, 2004                                                          %
% Written as part of the 2004 Summer REU on inverse problems at the        %
% University of Washington.                                                %
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function T=RWNrecover(G,B)

% Recover |dV| and |Int| from the dimensions of B
v=size(B);
d=v(2); % |dV|
n=v(1); % |Int|
N=n+d; % |V|

% Create the extended absorption matrix B_hat
```

```
I=eye(d);
B_hat=zeros(N,d);
for i=1:d
    for j=1:d
        B_hat(i,j)=I(i,j);
    end
end
for i=1:n
    for j=1:d
        B_hat(d+i,j)=B(i,j);
    end
end

% Check that G is a legal adjacency matrix for a graph with d boundary
% vertices and n interior vertices

v=size(G);
if ~((v(1)==N)&&(v(2)==N))
    disp('G is not a valid adjacency matrix for B.')
    % G has wrong dimensions
else if ~(G(1:d,1:d)==eye(d))
    disp('G is not a valid adjacency matrix for B.')
    % either G has non-loop boundary-boundary edges, or at least one
    % boundary vertex has no loop
else if ~(G(1:d,1+d:n+d)==zeros(d,n))
    disp('G is not a valid adjacency matrix for B.')
    % G has boundary to interior edges
else if ~(G(1+d:n+d,1+d:n+d)==transpose(G(1+d:n+d,1+d:n+d)))
        % G has non-symmetric interior-interior adjacencies
    disp('G is not a valid adjacency matrix for B.')
else if ~(G(1+d:n+d,1+d:n+d).*eye(n,n)==zeros(n,n))
    disp('G is not a valid adjacency matrix for B.')
    % G hasloops on interior vertices
end; end; end; end; end;

% Check that row sums of B are 1

B_bad=0;
for i=1:n
    rowsum=0;
    for j=1:d
        rowsum=rowsum+B(i,j);
    end
    if ~((rowsum<1.0000001)&&(rowsum>.9999999))
        B_bad=1;
    end
end
if B_bad
```

```
    disp('B is not a valid absorption matrix')
end

% begin edge recovery
% q ranges over interior vertices
% for each value of q, either a system of linear equations is solved that
% uniquely determine the transition probabilities on edges leaving q, or it
% is determined that no such system exists (in which the pair G, B is not
% recoverable.)

T=G;

for q=1+d:n+d
    adj_n=0;
    for i=1:N
        adj_n=adj_n+G(q,i);
    end
    % adj_n now gives the number of vertices adjacent to q
    adj_v=zeros(1,adj_n);
    j=1;
    for i=1:adj_n
        while G(q,j)==0
            j=j+1;
        end
        adj_v(i)=j;
        j=j+1;
    end
    if rank(B_hat(adj_v,1:d))<adj_n
        disp('non recoverable pair G, B')
    else
        % find adj_n columns of B_hat(adj_v,1:d)) that are linearly
        % independent vectors
        boundary_v=(1:adj_n);
        for i=1:adj_n
            while rank(B_hat(adj_v,boundary_v(1:i)))<i
                boundary_v(i:adj_n)=boundary_v(i:adj_n)+ones(1,1+adj_n-i);
            end
        end
        % boundary_v is now a vector of boundary vertices such that
        % B_hat(adj_v,boundary_v) is invertible
    end
    % now that we have our set of boundary vertices, we can use them to
    % find the edge probabilities leaving q by
    % vec_rho_e=B_hat(adj_v,boundary_v)^(-1)*B_hat(q,boundary_v)
    vec_rho_e=transpose(B_hat(adj_v,boundary_v))^(-1)*transpose(B_hat(q,boundary_v));
    for i=1:adj_n
        T(q,adj_v(i))=vec_rho_e(i);
    end
```

```
    % The entries of the row of T corresponding to edges leaving q have now
    % been determined.
end
% by looping over all interior vertices q, T is completely determined.
% Check that the absorption matrix corresponding to T is B.
B_from_T=(eye(n)-T(1+d:n+d,1+d:n+d))^-1*T(1+d:n+d,1:d);
if ~(B==B_from_T)
    disp('error: B is not a valid absorption matrix for the graph G')
% note: this can only happen if there is no set of transition probabilities
% on G that gives rise to the absorption matrix B, because it means that
% some of the systems of equations which we solved a subset of to determine
% T had some equations outside of the subset we used that were inconsistent
% with the ones we used, which can't happen if B is actually an absorption
% matrix for the graph.
else
disp('The transition matrix for the pair G, B is:')
end
```

References

[1] DeVries, Tim, "Recoverability of Random Walk Networks," 2003.
[2] Lewis, Andrew, "Random Walk Networks," 2004.
[3] Menger, Karl, *Kurventheorie*, Teubner, Berlin, Germany, 1932.

Brandeis University
E-mail address: ded@brandeis.edu