2-9-2015

Efficiency and Betweenness Centrality of Graphs and some Applications

Bryan Ek

Efficiency and Betweenness Centrality of Graphs and some Applications

by

BRYAN EK

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Applied Mathematics
School of Mathematical Sciences, College of Science

Rochester Institute of Technology
Rochester, NY
February 9, 2015

Committee Approval:

R·I·T
College of SCIENCE

Dr. Darren Narayan Date
School of Mathematical Sciences
Thesis Advisor

Dr. Jobby Jacob Date
School of Mathematical Sciences
Committee Member

Dr. Paul Wenger Date
School of Mathematical Sciences
Committee Member

Dr. Rigoberto Florez Date
The Citadel
Committee Member

Nathan Cahill, D.Phil. Date
School of Mathematical Sciences
Director of Graduate Programs

Abstract

The distance $d_G(i,j)$ between any two vertices i and j in a graph G is the minimum number of edges in a path between i and j. If there is no path connecting i and j, then $d_G(i,j) = \infty$. In 2001, Latora and Marchiori introduced the measure of efficiency between vertices in a graph. The efficiency between two vertices i and j is defined to be $\in_{i,j} = \frac{1}{d_G(i,j)}$ for all $i \neq j$. The global efficiency of a graph is the average efficiency over all $i \neq j$. The power of a graph G^m is defined to be $V(G^m) = V(G)$ and $E(G^m) = \{(u,v) | d_G(u,v) \leq m\}$. In this paper we determine the global efficiency for path power graphs P_n^m, cycle power graphs C_n^m, complete multipartite graphs $K_{m,n}$, star and subdivided star graphs, and the Cartesian products $K_n \times P_m^t$, $K_n \times C_m^t$, $K_m \times K_n$, and $P_m \times P_n$.

The concept of global efficiency has been applied to optimization of transportation systems and brain connectivity. We show that star-like networks have a high level of efficiency. We apply these ideas to an analysis of the Metropolitan Atlanta Rapid Transit Authority (MARTA) Subway system, and show this network is 82% as efficient as a network where there is a direct line between every pair of stations. From BOLD fMRI scans we are able to partition the brain with consistency in terms of functionality and physical location. We also find that football players who suffer the largest number of high-energy impacts experience the largest drop in efficiency over a season.

Latora and Marchiori also presented two local properties. The local efficiency $E_{loc} = \frac{1}{n} \sum\limits_{i \in V(G)} E_{glob}(G_i)$ is the average of the global efficiencies over the subgraphs G_i, the subgraph induced by the neighbors of i. The clustering coefficient of a graph G is defined to be $CC(G) = \frac{1}{n} \sum\limits_{i} C_i$ where $C_i = |E(G_i)| / \binom{|V(G_i)|}{2}$ is a degree of completeness of G_i. In this paper, we compare and contrast the two quantities, local efficiency and clustering coefficient.

Betweenness centrality is a measure of the importance of a vertex to the optimal paths in a graph. Betweenness centrality of a vertex is defined as $bc(v) = \sum_{x,y} \frac{\sigma_{xy}(v)}{\sigma_{xy}}$ where σ_{xy} is the number of unique paths of shortest length between vertices x and y. $\sigma_{xy}(v)$ is the number of optimal paths that include the vertex v. In this paper, we examined betweenness centrality for vertices in C_n^m. We also include results for subdivided star graphs and C_3 star graphs.

A graph is said to have unique betweenness centrality if $bc(v_i) = bc(v_j)$ implies $i = j$: the betweenness centrality function is injective over the vertices of G. We describe the betweenness centrality for vertices in ladder graphs, $P_2 \times P_n$. An appended ladder graph U_n is $P_2 \times P_n$ with a pendant vertex attached to an "end". We conjecture that the infinite family of appended graphs has unique betweenness centrality.

Contents

I. Introduction

I.1 Efficiency

In this thesis, we are concerned with several measures of connectivity of graphs: global efficiency, local efficiency, clustering coefficient and betweenness centrality.

In 2001, Latora and Marchiori introduced the measure of efficiency between vertices in a graph [1]. The (unweighted) *efficiency* between two vertices v_i and v_j is defined to be $\in (v_i, v_j) = \frac{1}{d(v_i,v_j)}$ for all $i \neq j$. The *global efficiency* of a graph $E_{glob}(G) = \frac{1}{n(n-1)} \sum_{i \neq j} \in (v_i, v_j)$ which is simply the average of the efficiencies over all pairs of the distinct n vertices. Then note that $0 \leq E_{glob}(G) \leq 1$ with equality only when G has no edges and when G is a complete graph respectively.

The concept of reciprocal distance has been studied previously. In 1993, Plavšić, Nikolić, Trinajstić, and Mihalić introduced the *Harary index* of a simple graph [2]. For a simple graph G with vertices $v_1, v_2, ..., v_n$ the Harary index is denoted $H(G)$ and equals $\sum\limits_{1 \leq i < j \leq n} \frac{1}{d(v_i,v_i)}$. We note the close relationship between global efficiency and the Harary index, $E_{glob}(G) = \frac{2}{n(n-1)} H(G)$. There also have been other studies involving the Harary index and reciprocal distances [3, 4, 5, 6, 7].

In this thesis we determine the global efficiency for path power graphs P_n^m, cycle power graphs C_n^m, complete multipartite graphs $K_{m,n}$, star and subdivided star graphs, and the Cartesian Products $K_n \times P_m^t$, $K_n \times C_m^t$, $K_m \times K_n$, and $P_m \times P_n$. As a consequence, we determine new results involving the Harary index for these families of graphs.

Recently other papers have studied the concept of efficiency, [8, 9, 10, 11, 12]. A comprehensive analysis of all of these measures is given by Sporns [13].

The concept of global efficiency has been applied to optimization of transportation systems. In 2002, Latora and Marchiori explored the global efficiency of the Boston Subway (MBTA) and found that the MBTA network is 63% as efficient as a network where there is a direct line between any two stations[8]. Motivated by the design of the Metropolitan Atlanta Rapid Transportation Authority (MARTA) Subway network (see Figure II.9.1), we investigate the global efficiency of subdivided stars. We show that networks of this type have a high level of efficiency. We apply these ideas to an analysis of the MARTA Subway system and show that their network is 82% as efficient as a network where there is a direct line connecting each pair of stations.

Latora and Marchiori also presented two local properties[8]. The *local efficiency* $E_{loc} = \frac{1}{n} \sum_{i \in G} E_{glob}(G_i)$ is the average of the global efficiencies over the subgraphs G_i, the subgraph induced by the neighbors of i. The *clustering coefficient* of a graph G is defined to be $CC(G) = \frac{1}{n} \sum_i C_i$ where $C_i = |E(G_i)| / \binom{|V(G_i)|}{2}$ is a degree of completeness of G_i. In this thesis, we compare and contrast the two quantities, local efficiency and clustering coefficient. We include results of these local measurements for complete multipartite graphs $K_{n,m}$, cycle power graphs C_n^m, and Cartesian products $K_m \times K_n$ and $K_n \times C_m$.

RCBI scientists conducted functional MRI (fMRI) scans of 25 volunteers to find blood oxygen level-dependent (BOLD) correlations of various regions of the brain. We constructed graphs with edges based on correlation cutoffs and then partitioned the brain using efficiency. The partitions were found to be consistent with functionality and physical location within the brain. We also used these measurements to analyze the effects of a season of hard-contact football on University of Rochester athletes. Again, an outside source conducted BOLD pre and postseason fMRI scans of the players. We received matrices of the correlations in oxygen levels of various regions of the brain and modeled these as graphs. We were then able to measure the "efficiency" of each athlete. As was expected, the athletes who received the largest number of high-energy impacts during the season also experienced the largest drop in brain efficiency. For comparison, we calculated the measurements of a macaque brain using data (see Figure VIII.1.1) from Honey et al.[12]

It was stated by Latora and Marchiori [1] that "It can be shown that, when in a graph, most of its local subgraphs G_i are not sparse, then C [clustering coefficient] is a good approximation of E_{loc}. In summary, there are not two different types of analyses to be done for the global and local scales, just one with a very precise physical meaning: the efficiency in transporting information". Due to the vague wording of "not sparse" we provide an in-depth analysis of this statement, identifying graphs where the clustering coefficient and local efficiency are in fact non-negligibly different. We also identify certain graph families where the two quantities are the same.

I.2 Betweenness Centrality

Betweenness centrality is a measure of the importance of a vertex to the optimal paths in a graph. Betweenness centrality of a vertex is defined as $bc(v) = \sum_{x,y} \frac{\sigma_{xy}(v)}{\sigma_{xy}}$ where σ_{xy} is the number of unique paths of shortest length between vertices x and y. $\sigma_{xy}(v)$ is the number of optimal paths that include the vertex v. In this thesis, we examined betweenness centrality for vertices in C_n^m. By the symmetry of C_n^m, every vertex will have the same betweenness centrality. We also include results for subdivided star graphs and C_3 star graphs.

A graph is said to have unique betweenness centrality if $bc(v_i) = bc(v_j)$ implies $i = j$: the betweenness centrality function is injective over the vertices of G. We describe the betweenness centrality for vertices in ladder graphs, $P_2 \times P_n$. An appended ladder graph U_n is $P_2 \times P_n$ with a pendant vertex attached to an "end". We conjecture that the infinite family of appended graphs has unique betweenness centrality. We were able to construct a partial proof but were forced to leave the completion as future research.

I.3 Definitions

Definition I.3.1. A *graph*, G, is a collection of a set of vertices, $V(G)$, and a set of edges, $E(G)$. The graph can be denoted $G(V,E)$. An edge is an unordered pair of vertices. The *distance* $d_G(i,j)$ between any two vertices i and j in a graph G is the minimum number of edges in a path between i and j. The subscript notation is dropped if it is apparent with respect to which graph the distance is. If there is no path connecting i and j, G is disconnected, then $d(i,j) = \infty$.

Definition I.3.2. The *power* of a graph, G, denoted G^m, is defined to be $V(G^m) = V(G)$ and $E(G^m) = \{(u,v) | d_G(u,v) \leq m\}$. With this definition $G^1 = G$.

Definition I.3.3. Given a graph G and a vertex $v \in V(G)$, the *neighborhood subgraph induced by* v is the subgraph containing all vertices adjacent to v and all edges, if any, that may exist between the adjacent vertices.

Definition I.3.4. The *eccentricity* of a vertex v in a graph G is defined as $\epsilon(v) = \max\{d(v,u) | u \in V(G)\}$. The *diameter* of a graph, G, is defined as $diam(G) = \max\{\epsilon(v) | v \in V(G)\}$. Diameter is the largest distance between two vertices in the graph.

Remark I.3.5. Note that ϵ and $\in$ are separate symbols and $\in (x,y)$ denotes the efficiency between vertices x and y and $\in$ alone means contained in, as in "an element is contained in a set".

Remark I.3.6. When we mention a "step" in paths of graphs, we mean an intermediate vertex of the path.

Definition I.3.7. Due to the nature of the properties we examine, the Harmonic number $H_n = \sum_{i=1}^{n} \frac{1}{i}$ is very useful for simplifications. Note that for ease of use, we define $H_0 = 0$.

Definition I.3.8. An *automorphism* of a graph $G(V, E)$ is a bijective (one-to-one and onto) function on the vertices of G, $\phi : V \to V$, that preserves edges. i.e. ϕ is a permutation of V that preserves edges. Preserving edges means that $(v_1, v_2) \in E(G)$ if and only if $(\phi(v_1), \phi(v_2)) \in E(G(\phi(V), E))$. The set of automorphisms is denoted $Aut(G)$ and forms a group under composition[14].

Definition I.3.9. Let H be a group of permutations of a set S. For each $s \in S$, let $\text{orb}_H(s) = \{\phi(s) | \phi \in H\}$. $\text{orb}_H(s)$ is called the *orbit* of s under H. The orbits partition S into equivalence classes[15].

Definition I.3.10. Let $G(V, E)$ be a graph. G is said to be *vertex-transitive* if for all $u, v \in V$, we have that $u \in\text{orb}_{Aut(G)}(v)$ (or equivalently $v \in\text{orb}_{Aut(G)}(u)$). i.e. there exists some $\phi \in Aut(G)$ such that $\phi(v) = u$ (or there exists some $\phi \in Aut(G)$ such that $\phi(u) = v$)[16].

II. Global Efficiency

II.1 Definition and Example

Definition II.1.1. Consider a graph G of order n. The *global efficiency* is defined as

$$E_{glob}(G) = \frac{1}{n(n-1)} \sum_{i \neq j} \in (v_i, v_j), \tag{II.1.1}$$

where $\in (v_i, v_j) = \frac{1}{d(v_i, v_j)}$. The global efficiency is the average efficiency of all pairs of vertices.

Example II.1.2. Let $H = P_7$ with vertices A, B, C, D, E, F and G. See Figure II.1.1.

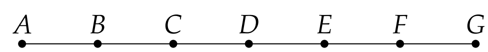


Figure II.1.1: P_7 for efficiency example.

The distances between each pair of vertices is given in the matrix shown below.

$DM =$

$L(H)$	A	B	C	D	E	F	G
A	0	1	2	3	4	5	6
B	1	0	1	2	3	4	5
C	2	1	0	1	2	3	4
D	3	2	1	0	1	2	3
E	4	3	2	1	0	1	2
F	5	4	3	2	1	0	1
G	6	5	4	3	2	1	0

Definition II.1.3. The average distance, $D(H)$, between vertices in a graph H shall be denoted:

$$D(H) = \frac{1}{n(n-1)} \sum_{i,j} d(v_i, v_j). \tag{II.1.2}$$

The inverse of $D(H)$ is a first approximation of the global efficiency.

In this case, $D(P_7) = \frac{2}{7 \cdot 6} [6(1) + 5(2) + 4(3) + 3(4) + 2(5) + 1(6)] = \frac{8}{3}$. The first approximation of the global efficiency is then $\frac{3}{8} = 0.375$.

The efficiency matrix is then as follows.

$EM =$

$E(H)$	A	B	C	D	E	F	G
A	0	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{6}$
B	1	0	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$
C	$\frac{1}{2}$	1	0	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$
D	$\frac{1}{3}$	$\frac{1}{2}$	1	0	1	$\frac{1}{2}$	$\frac{1}{3}$
E	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	0	1	$\frac{1}{2}$
F	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	0	1
G	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1	0

We note that the matrix is symmetric about the main diagonal. We can also sum the elements in the upper triangle of the matrix: $6(1) + 5(\frac{1}{2}) + 4\left(\frac{1}{3}\right) + 3\left(\frac{1}{4}\right) + 2\left(\frac{1}{5}\right) + 1\left(\frac{1}{6}\right)$. Finally we divide by the number of non-diagonal elements. Therefore $E_{glob}(P_7) = \frac{1}{7\cdot 6}\cdot 2\left(\sum_{i=1}^{7-1}\frac{7-i}{i}\right) = \frac{223}{420} \approx 0.531$. The first approximation in this case is off by nearly 30%: not very good.

II.2 Path Graphs: P_n

Let P_n denote the path on vertices $v_1, v_2, ..., v_n$ with edges $v_1v_2, v_2v_3, ..., v_{n-1}v_n$. The distance $d(v_i, v_j)$ between distinct vertices v_i and v_j is $|i - j|$. Hence the efficiency between v_i and v_j is $\in(v_i, v_j) = \frac{1}{d(v_i,v_j)} = \frac{1}{|i-j|}$.

Theorem II.2.1.

$$E_{glob}(P_n) = 2\left(\frac{H_{n-1}}{n-1} - \frac{1}{n}\right). \tag{II.2.1}$$

Proof. Consider the paths of various lengths in P_n. Without loss of generality we assume the "starting" vertex is located to the left of the ending vertex. Note that this will only account for half of the efficiencies. If we want to move i vertices to the right there are only $n - i$ starting vertices. Hence for the efficiency matrix of P_n, there are $n - i$ pairs of vertices whose efficiency is $\frac{1}{i}$. Then by doubling our efficiencies since the matrix is symmetric, and then normalizing, we have $E_{glob}(P_n) = \frac{2}{n\cdot(n-1)}\left(\sum_{i=1}^{n-1}\frac{n-i}{i}\right)$. Simple algebraic manipulation yields the theorem. ■

As expected, the global efficiency of a path will vary inversely to the number of vertices. We state this formally in our next theorem.

Theorem II.2.2.

$$\lim_{n\to\infty} E_{glob}(P_n) = 0$$

Proof. Using Lemma VIII.2.1,

$$\begin{aligned}
0 \le \lim_{n\to\infty} E_{glob}(P_n) &= \lim_{n\to\infty}\left[2\left(\frac{H_{n-1}}{n-1} - \frac{1}{n}\right)\right] \\
&= 2\lim_{n\to\infty}\frac{H_{n-1}}{n-1} - 2\lim_{n\to\infty}\frac{1}{n} \\
&\le 2\lim_{n\to\infty}\frac{\ln(n-1)+1}{n-1} - 2\lim_{n\to\infty}\frac{1}{n} \\
&= 2\lim_{n\to\infty}\frac{\ln(n-1)}{n-1} + 2\lim_{n\to\infty}\frac{1}{n-1} - 2\lim_{n\to\infty}\frac{1}{n} \\
&= 0 + 2\lim_{n\to\infty}\frac{1}{n} - 2\lim_{n\to\infty}\frac{1}{n} \\
&= 0.
\end{aligned}$$

■

II.3 Path Power Graphs: P_n^m

We next investigate the efficiency of powers of a path P_n. Su, Xiong, and Gutman obtained the Harary index of P_n^m, from which $E_{glob}(P_n^m)$ can easily be obtained. However, we include a computation of $E_{glob}(P_n^m)$, as it is useful for obtaining the global efficiency for the families $K_n \times P_n^m$ and $K_n \times C_n^m$.

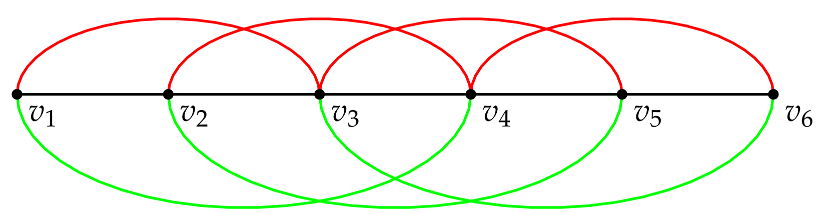


Figure II.3.1: A representation of the path power: P_6^3.

For the global efficiency of path power graphs: $E_{glob}(P_n^m)$, each element of the efficiency matrix is given by

$$\in_{ij} = \frac{1}{\left\lceil \frac{|i-j|}{m} \right\rceil}. \tag{II.3.1}$$

where i is the row and j is the column of the entry. This value corresponds to the efficiency between vertices i and j. The distance between the vertices in P_n is simply $|i-j|$. In P_n^m, each step can be up to m vertices away. Hence the distance between vertices equals $\left\lceil \frac{|i-j|}{m} \right\rceil$. Taking the inverse gives the formula in Eq. (II.3.1). Hence the matrix is:

$$\begin{array}{c|ccccccc}
 & v_1 & v_2 & v_3 & \cdots & v_{n-2} & v_{n-1} & v_n \\
\hline
v_1 & 0 & 1 & 1 & & \frac{1}{\lceil \frac{n-3}{m} \rceil} & \frac{1}{\lceil \frac{n-2}{m} \rceil} & \frac{1}{\lceil \frac{n-1}{m} \rceil} \\
v_2 & 1 & 0 & 1 & \cdots & \frac{1}{\lceil \frac{n-4}{m} \rceil} & \frac{1}{\lceil \frac{n-3}{m} \rceil} & \frac{1}{\lceil \frac{n-2}{m} \rceil} \\
v_3 & 1 & 1 & 0 & & \frac{1}{\lceil \frac{n-5}{m} \rceil} & \frac{1}{\lceil \frac{n-4}{m} \rceil} & \frac{1}{\lceil \frac{n-3}{m} \rceil} \\
\vdots & & \vdots & & \ddots & & \vdots & \\
v_{n-2} & \frac{1}{\lceil \frac{n-3}{m} \rceil} & \frac{1}{\lceil \frac{n-4}{m} \rceil} & \frac{1}{\lceil \frac{n-5}{m} \rceil} & & 0 & 1 & 1 \\
v_{n-1} & \frac{1}{\lceil \frac{n-2}{m} \rceil} & \frac{1}{\lceil \frac{n-3}{m} \rceil} & \frac{1}{\lceil \frac{n-4}{m} \rceil} & \cdots & 1 & 0 & 1 \\
v_n & \frac{1}{\lceil \frac{n-1}{m} \rceil} & \frac{1}{\lceil \frac{n-2}{m} \rceil} & \frac{1}{\lceil \frac{n-3}{m} \rceil} & & 1 & 1 & 0
\end{array}$$

Consider the first vertex of P_n^m. There are $(n-1)$ other vertices to compute the efficiency with. The sum of efficiencies from the first vertex is:

$$\sum_{j=2}^{n} \epsilon_{1,j} = \sum_{j=2}^{n} \frac{1}{\left\lceil \frac{|1-j|}{m} \right\rceil} = \sum_{i=1}^{n-1} \frac{1}{\left\lceil \frac{i}{m} \right\rceil}.$$

For the second vertex we have, $1 \leq |2-j| \leq n-2$ since $3 \leq j \leq n$, which yields:

$$\sum_{j=3}^{n} \epsilon_{2,j} = \sum_{j=3}^{n} \frac{1}{\left\lceil \frac{|2-j|}{m} \right\rceil} = \sum_{i=1}^{n-2} \frac{1}{\left\lceil \frac{i}{m} \right\rceil}.$$

Summing the terms for all vertices gives:

$$\sum_{i=1}^{n-1} \frac{1}{\left\lceil \frac{i}{m} \right\rceil} + \sum_{i=1}^{n-2} \frac{1}{\left\lceil \frac{i}{m} \right\rceil} + \cdots + \sum_{i=1}^{2} \frac{1}{\left\lceil \frac{i}{m} \right\rceil} + \sum_{i=1}^{1} \frac{1}{\left\lceil \frac{i}{m} \right\rceil} = \sum_{i=1}^{n-1} \frac{n-i}{\left\lceil \frac{i}{m} \right\rceil}.$$

Finally, we divide this term to get the result of Theorem II.3.1. We note as the matrix is symmetric, we can sum over the upper half of the matrix (ordered pairs) and then multiply our result by 2. The last step is normalization.

Theorem II.3.1.

$$E_{glob}(P_n^m) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \frac{n-i}{\left\lceil \frac{i}{m} \right\rceil}. \tag{II.3.2}$$

An alternate formula for faster computation can be found in Corollary VIII.1.5.

II.4 Cycle Power Graphs: C_n^m

Definition II.4.1. A *Cycle Graph* C_n has vertices $v_1, v_2, \ldots, v_n$ and edges (v_i, v_j) if $|i-j| = 1$ as well as the edge (v_1, v_n).

Definition II.4.2. A *Cycle Power Graph* C_n^m has vertices $v_1, v_2, \ldots, v_n$ and edge (v_i, v_j) if and only if $1 \leq \min\left(|i-j|, n-|i-j|\right) \leq m$. This condition is equivalent to $1 \leq \left\lfloor \frac{n}{2} \right\rfloor - \left\lfloor \left| |i-j| - \frac{n}{2} \right| \right\rfloor \leq m$.

Consider the Cycle C_8^3:

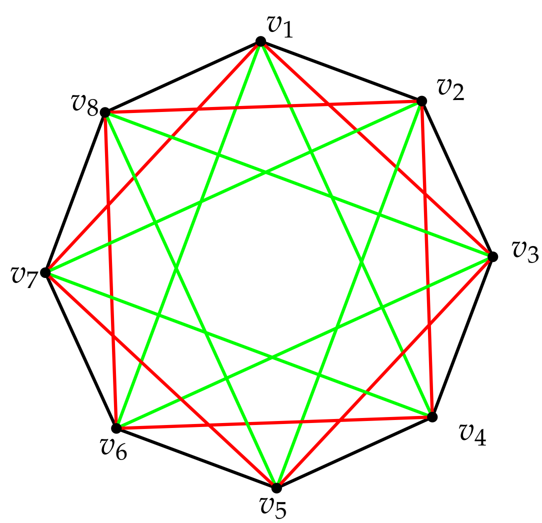


Figure II.4.1: A representation of the power cycle: C_8^3.

The efficiency matrix is given as:

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8
v_1	0	1	1	1	$\frac{1}{2}$	1	1	1
v_2	1	0	1	1	1	$\frac{1}{2}$	1	1
v_3	1	1	0	1	1	1	$\frac{1}{2}$	1
v_4	1	1	1	0	1	1	1	$\frac{1}{2}$
v_5	$\frac{1}{2}$	1	1	1	0	1	1	1
v_6	1	$\frac{1}{2}$	1	1	1	0	1	1
v_7	1	1	$\frac{1}{2}$	1	1	1	0	1
v_8	1	1	1	$\frac{1}{2}$	1	1	1	0

As previously stated for P_6^3, the global efficiency is found by summing all entries of the efficiency matrix and scaling it appropriately. Hence,

$$E_{glob}(C_8^3) = \frac{52}{8(8-1)} = \frac{13}{14} = 0.929.$$

Note that rows are identical; they are merely shifted representations of each other. This is the case due to the symmetry of the cycle. The following is the generic efficiency matrix for C_{2n}^m.

$$\begin{array}{c|cccccccccccc}
 & v_1 & v_2 & \cdots & v_i & \cdots & v_{\frac{n}{2}} & v_{\frac{n}{2}+1} & \cdots & v_j & \cdots & v_{n-1} & v_n \\
\hline
v_1 & 0 & 1 & \cdots & \frac{1}{\lceil \frac{i-1}{m} \rceil} & \cdots & \frac{1}{\lceil \frac{n/2-1}{m} \rceil} & \frac{1}{\lceil \frac{n/2}{m} \rceil} & \cdots & \frac{1}{\lceil \frac{n+1-j}{m} \rceil} & \cdots & 1 & 1 \\
\vdots & & & & & & & & & & & & \\
v_n & 1 & 1 & \cdots & \frac{1}{\lceil \frac{i}{m} \rceil} & \cdots & \frac{1}{\lceil \frac{n/2}{m} \rceil} & \frac{1}{\lceil \frac{n/2+1}{m} \rceil} & \cdots & \frac{1}{\lceil \frac{n+2-j}{m} \rceil} & \cdots & 1 & 0
\end{array}$$

Also, note that each row (by removing the zero: efficiency to itself) is symmetric to itself so the sum of the row is the same as twice the first half. For even cases, the center element is counted twice, thus it must be removed once. So considering the first portion of the last row, the i^{th} element is given as:

$$\in (v_i, v_n) = \frac{1}{\left\lceil \frac{i}{m} \right\rceil}.$$

So the sum of the row is almost given by:

$$2\sum_{i=1}^{\frac{n}{2}} \frac{1}{\left\lceil \frac{i}{m} \right\rceil}.$$

As indicated above, the center element is incorrectly doubled; however, the above sum does not take this into account. As a result, $\frac{1}{\lceil \frac{n}{2m} \rceil}$ must be subtracted off. The total efficiency is then the row sum multiplied by the number of rows: n. Therefore the global efficiency of any power cycle with an even number of vertices: n, is given by:

Lemma II.4.3. *For even n,*

$$\begin{aligned}
E_{glob}(C_n^m) &= \frac{1}{n(n-1)} \left[n \left(2\sum_{i=1}^{\frac{n}{2}} \frac{1}{\left\lceil \frac{i}{m} \right\rceil} - \frac{1}{\left\lceil \frac{n}{2m} \right\rceil} \right) \right] \\
&= \frac{1}{(n-1)} \left[2\sum_{i=1}^{\frac{n}{2}} \frac{1}{\left\lceil \frac{i}{m} \right\rceil} - \frac{1}{\left\lceil \frac{n}{2m} \right\rceil} \right]. \qquad \text{(II.4.1)}
\end{aligned}$$

Consider the Cycle C_9^2:

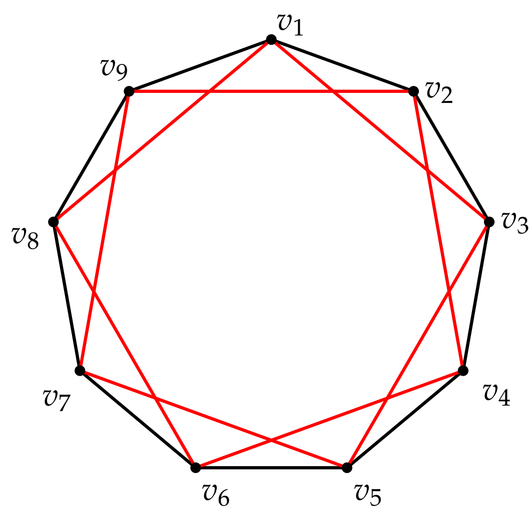


Figure II.4.2: A representation of the power cycle: C_9^2.

The efficiency matrix is given as:

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
v_1	0	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	1
v_2	1	0	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
v_3	1	1	0	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
v_4	$\frac{1}{2}$	1	1	0	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
v_5	$\frac{1}{2}$	$\frac{1}{2}$	1	1	0	1	1	$\frac{1}{2}$	$\frac{1}{2}$
v_6	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	1	0	1	1	$\frac{1}{2}$
v_7	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	1	0	1	1
v_8	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	1	0	1
v_9	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	1	0

From the efficiency matrix,

$$E_{glob}(C_9^2) = \frac{9(1+1+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+1+1)}{9(9-1)} = \frac{9\cdot 6}{9\cdot 8} = \frac{3}{4} = 0.75.$$

The generic efficiency matrix for C_n^m, where n is odd, is then given by:

	v_1	v_2	$\cdots$	v_i	$\cdots$	$v_{\frac{n+1}{2}}$	$v_{\frac{n+1}{2}+1}$	$\cdots$	v_j	$\cdots$	v_{n-1}	v_n
v_1	0	1	$\cdots$	$\frac{1}{\lceil\frac{i-1}{m}\rceil}$	$\cdots$	$\frac{1}{\lceil\frac{(n-1)/2}{m}\rceil}$	$\frac{1}{\lceil\frac{(n-1)/2}{m}\rceil}$	$\cdots$	$\frac{1}{\lceil\frac{n+1-j}{m}\rceil}$	$\cdots$	1	1
$\vdots$												
v_n	1	1	$\cdots$	$\frac{1}{\lceil\frac{i}{m}\rceil}$	$\cdots$	$\frac{1}{\lceil\frac{(n-1)/2}{m}\rceil}$	$\frac{1}{\lceil\frac{(n-3)/2}{m}\rceil}$	$\cdots$	$\frac{1}{\lceil\frac{n+2-j}{m}\rceil}$	$\cdots$	1	0

Considering the first portion of the first row, the i^{th} element is given as:

$$\in (v_i, v_1) = \frac{1}{\left\lceil \frac{i-1}{m} \right\rceil}.$$

So the sum of the row is given by:

$$2\sum_{i=2}^{\frac{n+1}{2}} \frac{1}{\left\lceil \frac{i-1}{m} \right\rceil} = 2\sum_{i=1}^{\frac{n-1}{2}} \frac{1}{\left\lceil \frac{i}{m} \right\rceil}.$$

Note the sums are identical; the index was merely shifted. And so the global efficiency for C_n^m, where n is odd, is found by multiplying this sum by the number of rows and normalizing:

Lemma II.4.4. *For odd n,*

$$E_{glob}(C_n^m) = \frac{2}{(n-1)} \sum_{i=1}^{\frac{n-1}{2}} \frac{1}{\left\lceil \frac{i}{m} \right\rceil}. \tag{II.4.2}$$

We can combine Lemmas II.4.3 and II.4.4 into Theorem II.4.5.

Theorem II.4.5.

$$E_{glob}(C_n^m) = \begin{cases} \frac{1}{2k-1}\left[\sum_{i=1}^{k} \frac{2}{\lceil \frac{i}{m} \rceil} - \frac{1}{\lceil \frac{k}{m} \rceil}\right] & \textit{if } n = 2k, \\ \frac{1}{k}\sum_{i=1}^{k} \frac{1}{\lceil \frac{i}{m} \rceil} & \textit{if } n = 2k+1. \end{cases} \tag{II.4.3}$$

An alternate formula for faster computation can be found in Corollary VIII.1.6.

II.5 Complete Multipartite Graphs

Definition II.5.1. A *complete multipartite graph* $G = K_{s_1,s_2,\ldots,s_t}$ is composed of t classes each with s_i vertices, $1 \le i \le t$, where each vertex in class i is adjacent to every vertex in class $j \neq i$, and is not adjacent to any vertex in class i.

We note that the distance between any pair of vertices in different classes is 1 and the distance between any pair of vertices in the same class is 2. This leads to our next theorem.

Theorem II.5.2. *Let $G = K_{s_1,s_2,\ldots,s_t}$, $t > 1$, and let $n = \sum_{i=1}^{t} s_i$. Then*

$$E_{glob}(G) = 1 - \frac{1}{2(n-1)} \left[\frac{1}{n} \sum_{i=1}^{t} s_i^2 - 1 \right]. \tag{II.5.1}$$

Proof. Let v be a vertex in a part with s_i vertices. Then the shortest path from v to any vertex in the same part is 2 and v is adjacent to all vertices in other parts. Hence for every vertex, v, in part i, the sum of efficiencies including v is $\left[\frac{(s_i-1)}{2} + (n - s_i)\right]$. Summing over all vertices in a given part and then over all parts gives $\sum_{i=1}^{t} \sum_{v \in \text{part } i} \left[\frac{(s_i-1)}{2} + (n-s_i)\right] = \sum_{i=1}^{t} s_i \left[\frac{(s_i-1)}{2} + (n-s_i)\right]$. Normalizing and some algebraic manipulation gives the desired result. ■

Remark II.5.3. If we increase the value of n, the efficiency doesn't necessarily tend toward 1. It all depends on the distribution of vertices within the classes. If one class is filled with nearly all the vertices, then the efficiency will tend towards $\frac{1}{2}$. Other ratios will tend toward different values in between $\frac{1}{2}$ and 1.

For the small case of $t = 2$, $s_1 = n$, $s_2 = m$: complete bipartite graphs, we have great simplifications.

Theorem II.5.4.

$$E_{glob}(K_{n,m}) = 1 - \frac{1}{2(n+m-1)} \left[\frac{n^2 + m^2}{n+m} - 1 \right]. \tag{II.5.2}$$

Another simplification due to symmetry is for $K_{r,\ldots,r}$: a complete multipartite graph with n classes, each with r vertices. We will denote the complete multipartite graph $K_{r,\ldots,r}$ as $K_{r;n}$. An example is shown below in Figure II.5.1

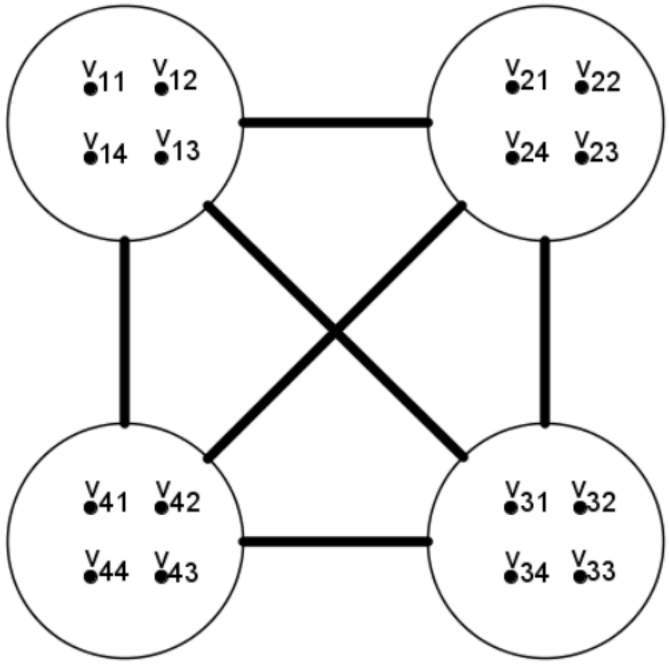


Figure II.5.1: A complete multipartite graph. $K_{4;4}$

Theorem II.5.5.

$$E_{glob}(K_{r;n}) = 1 - \frac{r-1}{2(nr-1)}. \tag{II.5.3}$$

Remark II.5.6. As the number of classes increases, the complete multipartite graph begins to approach the appearance of a complete graph. Thus we have a global efficiency approaching 1.

$$\lim_{n\to\infty} E_{glob}(K_{r;n}) = 1.$$

Increasing the number of vertices in each class tends to decrease the global efficiency since this increases the number of optimal paths of length 2 (the worst paths). However it also increases the number of optimal paths of length 1 depending on the number of classes. Thus we arrive at a lower bound for the global efficiency based on the number of classes.

$$\lim_{r\to\infty} E_{glob}(K_{r;n}) = 1 - \frac{1}{2n}.$$

II.6 Efficiency under the Euclidean Metric

When analyzing the efficiency of a transportation, it is natural to compare global efficiency under the graph metric $E_{glob}(G)$ versus a weighted metric $E^{w}_{glob}(G)$. We will refer to the former as unweighted efficiency and the latter as maximum weighted efficiency. In calculating the maximum weighted efficiency, we consider every pair of vertices to be adjacent with the weight of an edge as the Euclidean distance between the corresponding vertices. Note then that the weighted efficiency is highly dependent on the orientation of the graph as well as the plane in which it is embedded.

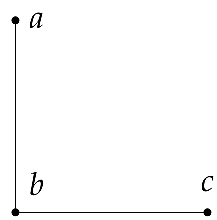


Figure II.6.1: Demonstration of the effect of considering Euclidean distance.

For the unweighted efficiency we have $\in (x, y) = 1$, $\in (x, z) = 1$, and $\in (y, z) = \frac{1}{2}$. Hence $E_{glob}(G) = \frac{1}{3 \cdot 2} \cdot 2(1 + 1 + \frac{1}{2}) = \frac{5}{6} \approx 0.83$. However for the maximum weighted efficiency we have $\in (x, y) = 1$, $\in (x, z) = 1$, and $\in (y, z) = \frac{1}{\sqrt{2}}$. Hence $E^{w}_{glob}(G) = \frac{1}{3 \cdot 2} \cdot 2(1 + 1 + \frac{1}{\sqrt{2}}) = \frac{1}{6}\sqrt{2} + \frac{2}{3} \approx 0.90$.

In Figure II.6.1, we compare the two types of efficiency of the graph G, drawn with a prescribed orientation. By examining the ratio of the unweighted efficiency to the maximum weighted efficiency, we can compare how efficient a graph network is compared to a Euclidean network. The ratio $E_{Ratio}(G) = E_{glob}(G)/E^{w}_{glob}(G) = \frac{5}{6}/(\frac{1}{6}\sqrt{2} + \frac{2}{3}) \approx 0.92$. Hence for the particular graph in Figure II.6.1, the graph is 92% as efficient as the completed graph under the Euclidean metric.

The case where $G = P_n$ is straightforward since the shortest distance between any points is a straight line. This assumes that the path is oriented in the "usual" fashion of a line. Hence

Theorem II.6.1. *$E_{glob}(P_n) = E^{w}_{glob}(P_n)$, and $E_{Ratio}(P_n) = 1$.*

II.7 Uniformly Subdivided Star Graphs: $S_{d,l}$

In this subsection we consider the efficiency of star-like networks. The graph $K_{1,r}$ is called a star and is a complete bipartite graph with a single vertex in one part and r vertices in the other. We next recall the graph operation known as an edge subdivision.

Definition II.7.1. An *edge subdivision* is an operation that is applied to an edge uv where a new vertex w is inserted, and the edge uv is replaced by edges uw and wv. A *subdivision* H of a graph G is a graph that can be obtained by performing a sequence of edge subdivisions.

Hence we can define a subdivided star.

Definition II.7.2. Let $S_{d,l}$ be the subdivision of the star $K_{1,r}$ where each edge is replaced by a path with l vertices. The vertex of degree d will be referred to as the center.

The subdivided star $S_{4,3}$ is shown in Figure II.7.1.

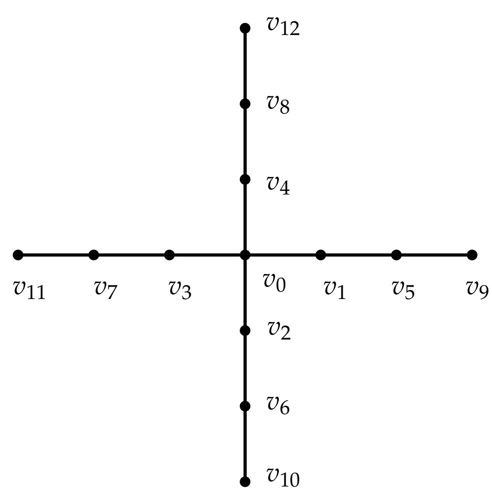


Figure II.7.1: An $S_{4,3}$ graph. See the accompanying efficiency matrix below: Table II.7.1.

Table II.7.1: Efficiency matrix for $S(4,3)$.

	v_0	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	v_{11}	v_{12}
v_0	0	1	1	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
v_1		0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
v_2			0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$
v_3				0	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$	1	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
v_4					0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
v_5						0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	1	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$
v_6							0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{5}$	1	$\frac{1}{5}$	$\frac{1}{5}$
v_7								0	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{5}$	1	$\frac{1}{5}$
v_8									0	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	1
v_9										0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
v_{10}											0	$\frac{1}{6}$	$\frac{1}{6}$
v_{11}												0	$\frac{1}{6}$
v_{12}													0

We first examine efficiencies between vertices on the same spoke including the center. Note that based on our labeling, there are three blocks of four identical entries across the top row and each continues in a "downward diagonal pattern". The total sum of these diagonals is: $\frac{4(3)}{1} + \frac{4(2)}{2} + \frac{4(1)}{3}$.

Next we examine efficiencies between vertices on different spokes. There are "patches" of $\binom{4}{2} = 6$ identical entries. There is one patch where the entries are equal to $\frac{1}{2}$, two patches where the entries equal $\frac{1}{3}$, three patches where the entries equal $\frac{1}{4}$, two patches where the entries equal $\frac{1}{5}$, and one patch where the entries equal $\frac{1}{6}$. This pattern is inherent from the labeling of our vertices. The vertices $v_{hi+1}, v_{hi+2}, v_{hi+3}, v_{hi+4}$, all have distance h from the center. We will consider paths between vertices on different spokes. Paths of length 2 must be between vertices where $h = 1$. Paths of length 3 must be between vertices where one vertex has $h = 1$ and another has $h = 2$. Paths of length 4 must be between vertices where both vertices have $h = 2$, or where one has $h = 1$ and the other has $h = 3$. Paths of length 5 must be between vertices where one vertex has $h = 2$ and another has $h = 3$. Paths of length 6 must be between vertices where $h = 3$. For each partition of a path length, there will be $\binom{4}{2}$ paths: picking the spokes to travel between.

The sum over all of the patches is $\frac{4(3)}{2} \cdot \frac{1}{2} + \frac{4(3)}{2} \cdot \frac{2}{3} + \frac{4(3)}{2} \cdot \frac{3}{4} + \frac{4(3)}{2} \cdot \frac{2}{5} + \frac{4(3)}{2} \cdot \frac{1}{6}$. Using symmetry about the main diagonal, the total sum over all efficiencies is

$$2 \cdot \left(\frac{4(3)}{1} + \frac{4(2)}{2} + \frac{4(1)}{3} + \frac{4(3)}{2} \cdot \frac{1}{2} + \frac{4(3)}{2} \cdot \frac{2}{3} + \frac{4(3)}{2} \cdot \frac{3}{4} + \frac{4(3)}{2} \cdot \frac{2}{5} + \frac{4(3)}{2} \cdot \frac{1}{6} \right) = \frac{967}{15}.$$

Dividing by the number of non-diagonal entries in our matrix gives: $\frac{1}{13 \cdot 12} \cdot \frac{967}{15} = \frac{967}{2340} = 0.413\,25$.

This example provides the structure for the proof of our next theorem.

Theorem II.7.3. *We have*

$$E_{glob}(S_{d,l}) = \frac{2}{l\,(dl+1)} \left((d-1) \left(l + \frac{1}{2} \right) H_{2l} - (d-2)(l+1)H_l - l \right). \tag{II.7.1}$$

Proof. First we consider the efficiencies between vertices on the same spoke including the center. Each spoke is isomorphic to P_{l+1}. By Theorem II.2.1 the sum of the efficiencies of this spoke is $(l+1)H_l - l$. Hence the total sum of the efficiencies over all d spokes is $d[(l+1)H_l - l]$.

Next we consider efficiencies between vertices on different spokes. In general the number of paths of length k in $S_{d,l}$ will equal the number of partitions of k into a and b where $a, b \leq l$. Each partition $k = a + b$ corresponds to a path in $S_{d,l}$ that travels through the center with a subpath of length a from the starting vertex to the center and a subpath of length b from the center to the end vertex. These correspond to the patches with entries equal to $\frac{1}{k}$.

Each of the patches will contain $\binom{d}{2}$ identical entries since this is the number of ways to choose the starting and ending spokes. Considering the various partitions of k there will be i patches where all of the entries are equal to $\frac{1}{i+1}$ for $1 \leq i \leq l$ and i patches where all of the entries are equal to $\frac{1}{2l+1-i}$ for $1 \leq i \leq l-1$.

The total sum is $\sum_{i=1}^{l} \frac{d(d-1)}{2} \cdot \frac{i}{i+1} + \sum_{i=1}^{l-1} \frac{d(d-1)}{2} \cdot \frac{i}{2l+1-i}$. The sum of the efficiencies for a subdivided star graph with d spokes, each of length l is then (doubling for the symmetry of the matrix):

$$\sum_{i,j} \epsilon_{ij} = 2 \left(d[(l+1)H_l - l] + \sum_{i=1}^{l} \frac{d(d-1)}{2} \cdot \frac{i}{i+1} + \sum_{i=1}^{l-1} \frac{d(d-1)}{2} \cdot \frac{i}{2l+1-i} \right),$$

which can simplify to

$$2d \left((d-1) \left(l + \frac{1}{2} \right) H_{2l} - (d-2)(l+1)H_l - l \right).$$

Normalizing with $n = dl + 1$ completes the proof. ■

II.7.1 Weighted Efficiencies

When applying these methods in a real-world situation, we consider edges weighted by the Euclidean distance between the corresponding vertices (See Figure II.7.2). For the weighted version we will consider the distance between any adjacent vertices to be 1. Furthermore, we consider all spokes to be linear and spaced at equal angles around the center vertex, v_0 in the plane. Weighted efficiency can effectively approximate real-world networks such as a subway system. This is found by dividing the unweighted global efficiency by the maximum weighted global efficiency.

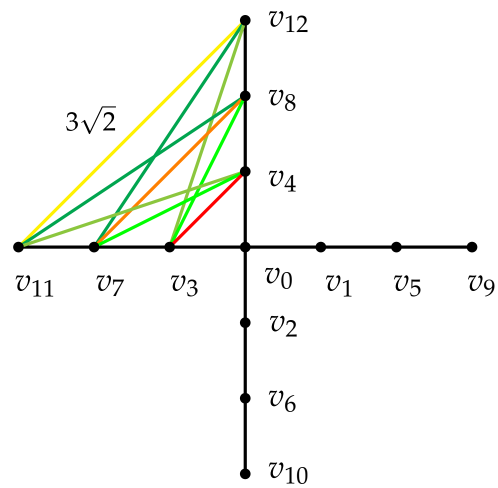


Figure II.7.2: An $S_{4,3}$ graph partially completed.

The following, Table II.7.2, is a matrix of the efficiency of a subdivided star graph as if each pair of vertices were connected with an edge weighted by the Euclidean distance between them, see Figure II.7.2. For example, v_8 and v_{11} would be connected by an edge of weight equal to the Euclidean distance between the points, $\sqrt{2^2+3^2} = \sqrt{13}$. Here the efficiency $\in (v_8, v_{11}) = \frac{1}{\sqrt{13}}$.

Table II.7.2: Euclidean efficiency matrix for $S(4,3)$.

			$j=1$				$j=2$				$j=3$			
		v_0	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}	v_{11}	v_{12}
	v_0	0	1	1	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
i=1	v_1	1	0	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	1	$\frac{1}{\sqrt{5}}$	$\frac{1}{3}$	$\frac{1}{\sqrt{5}}$	$\frac{1}{2}$	$\frac{1}{\sqrt{10}}$	$\frac{1}{4}$	$\frac{1}{\sqrt{10}}$
	v_2	1	$\frac{1}{\sqrt{2}}$	0	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	$\frac{1}{\sqrt{5}}$	1	$\frac{1}{\sqrt{5}}$	$\frac{1}{3}$	$\frac{1}{\sqrt{10}}$	$\frac{1}{2}$	$\frac{1}{\sqrt{10}}$	$\frac{1}{4}$
	v_3	1	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	0	$\frac{1}{\sqrt{2}}$	$\frac{1}{3}$	$\frac{1}{\sqrt{5}}$	1	$\frac{1}{\sqrt{5}}$	$\frac{1}{4}$	$\frac{1}{\sqrt{10}}$	$\frac{1}{2}$	$\frac{1}{\sqrt{10}}$
	v_4	1	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	0	$\frac{1}{\sqrt{5}}$	$\frac{1}{3}$	$\frac{1}{\sqrt{5}}$	1	$\frac{1}{\sqrt{10}}$	$\frac{1}{4}$	$\frac{1}{\sqrt{10}}$	$\frac{1}{2}$
i=2	v_5	$\frac{1}{2}$	1	$\frac{1}{\sqrt{5}}$	$\frac{1}{3}$	$\frac{1}{\sqrt{5}}$	0	$\frac{1}{\sqrt{8}}$	$\frac{1}{4}$	$\frac{1}{\sqrt{8}}$	1	$\frac{1}{\sqrt{13}}$	$\frac{1}{5}$	$\frac{1}{\sqrt{13}}$
	v_6	$\frac{1}{2}$	$\frac{1}{\sqrt{5}}$	1	$\frac{1}{\sqrt{5}}$	$\frac{1}{3}$	$\frac{1}{\sqrt{8}}$	0	$\frac{1}{\sqrt{8}}$	$\frac{1}{4}$	$\frac{1}{\sqrt{13}}$	1	$\frac{1}{\sqrt{13}}$	$\frac{1}{5}$
	v_7	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{\sqrt{5}}$	1	$\frac{1}{\sqrt{5}}$	$\frac{1}{4}$	$\frac{1}{\sqrt{8}}$	0	$\frac{1}{\sqrt{8}}$	$\frac{1}{5}$	$\frac{1}{\sqrt{13}}$	1	$\frac{1}{\sqrt{13}}$
	v_8	$\frac{1}{2}$	$\frac{1}{\sqrt{5}}$	$\frac{1}{3}$	$\frac{1}{\sqrt{5}}$	1	$\frac{1}{\sqrt{8}}$	$\frac{1}{4}$	$\frac{1}{\sqrt{8}}$	0	$\frac{1}{\sqrt{13}}$	$\frac{1}{5}$	$\frac{1}{\sqrt{13}}$	1
i=3	v_9	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{\sqrt{10}}$	$\frac{1}{4}$	$\frac{1}{\sqrt{10}}$	1	$\frac{1}{\sqrt{13}}$	$\frac{1}{5}$	$\frac{1}{\sqrt{13}}$	0	$\frac{1}{\sqrt{18}}$	$\frac{1}{6}$	$\frac{1}{\sqrt{18}}$
	v_{10}	$\frac{1}{3}$	$\frac{1}{\sqrt{10}}$	$\frac{1}{2}$	$\frac{1}{\sqrt{10}}$	$\frac{1}{4}$	$\frac{1}{\sqrt{13}}$	1	$\frac{1}{\sqrt{13}}$	$\frac{1}{5}$	$\frac{1}{\sqrt{18}}$	0	$\frac{1}{\sqrt{18}}$	$\frac{1}{6}$
	v_{11}	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{\sqrt{10}}$	$\frac{1}{2}$	$\frac{1}{\sqrt{10}}$	$\frac{1}{5}$	$\frac{1}{\sqrt{13}}$	1	$\frac{1}{\sqrt{13}}$	$\frac{1}{6}$	$\frac{1}{\sqrt{18}}$	0	$\frac{1}{\sqrt{18}}$
	v_{12}	$\frac{1}{3}$	$\frac{1}{\sqrt{10}}$	$\frac{1}{4}$	$\frac{1}{\sqrt{10}}$	$\frac{1}{2}$	$\frac{1}{\sqrt{13}}$	$\frac{1}{5}$	$\frac{1}{\sqrt{13}}$	1	$\frac{1}{\sqrt{18}}$	$\frac{1}{6}$	$\frac{1}{\sqrt{18}}$	0

Notice that the blocks of 4 identical terms with diagonals directed downward are identical to those appearing in the non-weighted case. These are the efficiencies between vertices on the same spoke or the center. For the pairs of vertices on different spokes, we focus on the squares which represent efficiencies between two vertices, where one is distance i from the center and the other is distance j from the center. For box, $i=1$ and $j=2$, the sum equals $8\cdot\frac{1}{\sqrt{5}}+4\cdot\frac{1}{3}=\frac{4}{\sqrt{5}}+\frac{4}{3}+\frac{4}{\sqrt{5}}$. In Figure 5, going from v_1 to v_8 requires a turn of an angle of $\frac{\pi}{2}$. The terms can be expressed using the law of cosines:

$$\frac{4}{\sqrt{1^2+2^2-2\cdot1\cdot2\cdot\cos\left(\frac{\pi}{2}\right)}}+\frac{4}{\sqrt{1^2+2^2-2\cdot1\cdot2\cdot\cos\left(\pi\right)}}+\frac{4}{\sqrt{1^2+2^2-2\cdot1\cdot2\cdot\cos\left(\frac{3\pi}{2}\right)}}.$$

Theorem II.7.4.

$$E^w_{glob}\left(S_{d,l}\right)=\frac{1}{l\left(dl+1\right)}\left[2(l+1)H_l-2l+\sum_{i=1}^{l}\sum_{j=1}^{l}\sum_{\theta=1}^{d-1}\frac{1}{\sqrt{i^2+j^2-2ij\cos\left(\frac{2\pi}{d}\theta\right)}}\right]. \tag{II.7.2}$$

Proof. The first step is to consider the orientation of the star graph. We assume that all spokes are straight lines in the Euclidean plane. We also assume that every spoke is spaced around the center vertex at equal angle intervals.

The sum of the efficiencies for vertices on the same spoke including the center is almost the same as in the proof of the previous theorem, $2d((l+1)H_l - l)$. We need to double this now as we are not doubling all terms later. Next we consider the pairs of vertices that are found on different spokes. In general the number of paths of length k in $S_{d,l}$ will equal the number of partitions of k into a and b where $a, b \leq l$. Each partition $k = a + b$ corresponds to a path in $S_{d,l}$ that travels through the center with a subpath of length a from the starting vertex to the center and a subpath of length b from the center to the end vertex. These form entries equal to $\frac{1}{\sqrt{a^2+b^2-2ab\cos\theta}}$ where θ is the angle between spokes. We focus on the $d \times d$ submatrices which represent efficiencies between two vertices, where one is distance i from the center and the other is distance j from the center.

The generic terms in a given $d \times d$ submatrix could then be written as $\frac{d}{\sqrt{i^2+j^2-2ij\cdot\cos\left(\frac{2\pi}{d}\theta\right)}}$ where θ varies from 1 to $d-1$. We then sum over all $d \times d$ submatrices and add the diagonal terms to yield the sum of the Euclidean efficiencies for $S_{d,l}$, $\sum\limits_{i \neq j} \in (v_i, v_j) = 2d((l+1)H_l - l) + \sum\limits_{i=1}^{l}\sum\limits_{j=1}^{l}\sum\limits_{\theta=1}^{d-1} \frac{d}{\sqrt{i^2+j^2-2ij\cos\left(\frac{2\pi}{d}\theta\right)}}$. Normalizing with $n = dl + 1$ gives the result. ∎

An alternate formula for faster computation can be found in Corollary VIII.1.7.

Instead of normalizing by the maximum number of edges, $n(n-1)$, we can normalize by the maximum weighted efficiency. The efficiency ratio, E_{Ratio} for a subdivided star graph $S_{d,l}$ is found by dividing the unweighted global efficiency by the maximum weighted global efficiency.

Theorem II.7.5.

$$E_{Ratio}\left(S_{d,l}\right) = \frac{(d-1)\left(2l+1\right)H_{2l} - (d-2)(2l+2)H_l - 2l}{(2l+2)H_l - 2l + \sum_{i=1}^{l}\sum_{j=1}^{l}\sum_{\theta=1}^{d-1}\frac{1}{\sqrt{i^2+j^2-2ij\cos\left(\frac{2\pi}{d}\theta\right)}}}. \tag{II.7.3}$$

Remark II.7.6. As expected, when d increases, the efficiency ratio decreases. In this case the spokes are getting closer but travel between spokes still requires traveling to the center vertex. However, an interesting aspect of this formula is that as l increases, the efficiency ratio increases. To see why this is true note that a straight line path has a weighted efficiency ratio of 1. We note that as the lengths of the spokes increases, the shape of a subdivided star bears a closer resemblance to a path.

II.7.2 New Unweighted Global Efficiency and Double Sum Reduction

The method used to find the weighted global efficiency also gives us a different way to calculate the unweighted global efficiency.

Corollary II.7.7.

$$E_{glob}\left(S_{d,l}\right) = \frac{1}{l\left(dl+1\right)}\left[2(l+1)H_l - 2l + \sum_{i=1}^{l}\sum_{j=1}^{l}\frac{d-1}{i+j}\right].$$

Proof. The law of cosines term in Theorem II.7.4 can be replaced with the graph path distance $(i+j)$. Then the triple sum reduces to the double sum above. ■

We can now equate the two formulas to find a reduction for the double sum.

Corollary II.7.8.

$$\left(2l+1\right)H_{2l} - (2l+2)H_l = \sum_{i=1}^{l}\sum_{j=1}^{l}\frac{1}{i+j}. \tag{II.7.4}$$

Proof. Equating the two formulas for the unweighted global efficiency of a star graph in Theorem II.7.3 and Corollary II.7.7 gives the following reduction:

$$\begin{aligned}
\frac{2}{l\left(dl+1\right)}\left((d-1)\left(l+\frac{1}{2}\right)H_{2l} - (d-2)(l+1)H_l - l\right) &= \frac{1}{l\left(dl+1\right)}\left[2(l+1)H_l - 2l + \sum_{i=1}^{l}\sum_{j=1}^{l}\frac{d-1}{i+j}\right],\\
\left(d-1\right)\left(2l+1\right)H_{2l} - (d-2)(2l+2)H_l - 2l &= (2l+2)H_l - 2l + \sum_{i=1}^{l}\sum_{j=1}^{l}\frac{d-1}{i+j},\\
\left(d-1\right)\left(2l+1\right)H_{2l} - (d-1)(2l+2)H_l &= \sum_{i=1}^{l}\sum_{j=1}^{l}\frac{d-1}{i+j},\\
\left(2l+1\right)H_{2l} - (2l+2)H_l &= \sum_{i=1}^{l}\sum_{j=1}^{l}\frac{1}{i+j}.
\end{aligned}$$

■

II.8 Cartesian Products

Definition II.8.1. The *Cartesian Product* of two graphs G and H is a graph $G \times H$, with the vertex set $V(G) \times V(H)$, where vertices $\{(i_1, i_2), (j_1, j_2)\}$ are adjacent if $\{i_1, j_1\} \in E(G)$ and $i_2 = j_2$, or $\{i_2, j_2\} \in E(H)$ and $i_1 = j_1$.

II.8.1 $K_r \times P_n^m$

In the figure below, we show the graph of the Cartesian product of K_4 and P_4^2.

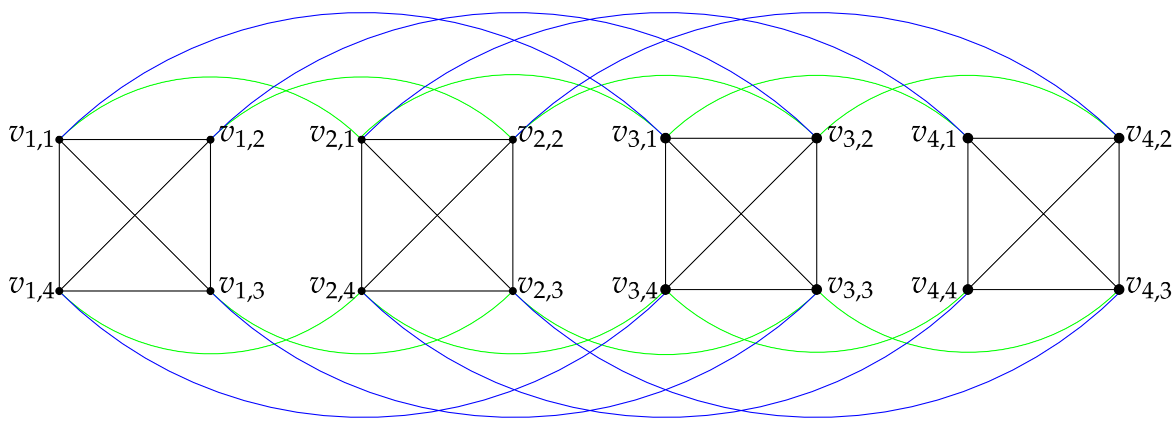


Figure II.8.1: The Cartesian product of a complete graph and a path power. The efficiency matrix is given in Table II.8.1 below.

Table II.8.1: The efficiency matrix for $K_4 \times P_4^2$.

	$v_{1,1}$	$v_{1,2}$	$v_{1,3}$	$v_{1,4}$	$v_{2,1}$	$v_{2,2}$	$v_{2,3}$	$v_{2,4}$	$v_{3,1}$	$v_{3,2}$	$v_{3,3}$	$v_{3,4}$	$v_{4,1}$	$v_{4,2}$	$v_{4,3}$	$v_{4,4}$
$v_{1,1}$	0	1	1	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$v_{1,2}$		0	1	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$
$v_{1,3}$			0	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$
$v_{1,4}$				0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{2}$
$v_{2,1}$					0	1	1	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
$v_{2,2}$						0	1	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$
$v_{2,3}$							0	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$
$v_{2,4}$								0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
$v_{3,1}$									0	1	1	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
$v_{3,2}$										0	1	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$
$v_{3,3}$											0	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$
$v_{3,4}$												0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
$v_{4,1}$													0	1	1	1
$v_{4,2}$														0	1	1
$v_{4,3}$															0	1
$v_{4,4}$																0

We next extend to the general case.

Theorem II.8.2.

$$E_{glob}(K_r \times P_n^m) = \frac{2}{n(nr-1)} \sum_{i=1}^{n-1} (n-i) \left(\frac{1}{\left\lceil \frac{i}{m} \right\rceil} + \frac{r-1}{\left\lceil \frac{i}{m} \right\rceil + 1} \right) + \frac{r-1}{nr-1}. \tag{II.8.1}$$

Proof. Notice that the matrix is very similar to that of a path power. Each i now corresponds to a block. Each block has r terms on the main diagonal of a block and these correspond to the pairs of vertices in the r adjacent path powers. All other terms correspond to the distance between vertices in different path powers and then within the copies of the complete subgraphs. For this, we have r vertices in the initial class to choose from and $r-1$ vertices in the final K_r. Since each class is complete, it will only take one additional step to reach the final vertex, and so the efficiency is only slightly smaller. There are also 'triangles' of 1s next to the main diagonal; these correspond to movements within a single K_r. The number of 1's is then n times the number of edges in K_r which equals $\frac{r(r-1)}{2}$. Averaging the efficiencies over all pairs yields Eq. (II.8.1). ■

An alternate formula for faster computation can be found in Corollary VIII.1.8.

II.8.2 $K_r \times C_n^m$

We next investigate the global efficiency of a Cartesian product of a complete graph and a cycle power. The graph of the Cartesian product of K_3 and C_6^2 is shown in Figure II.8.2.

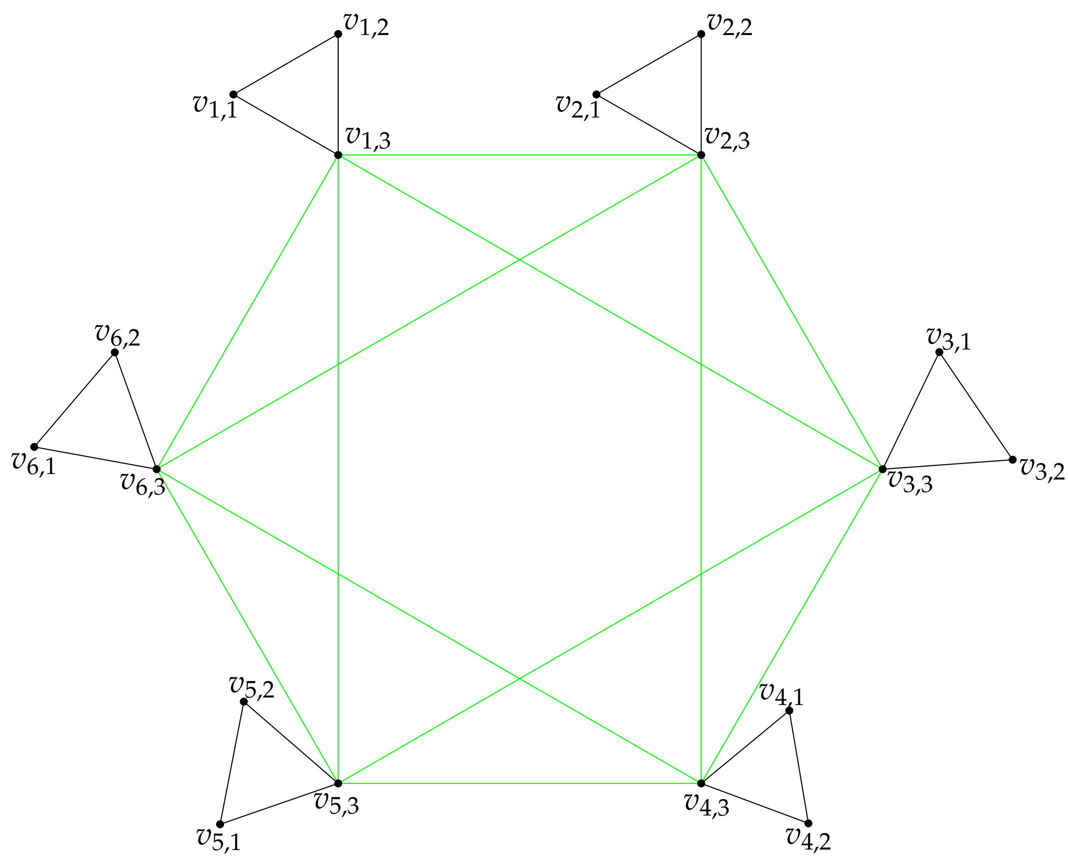


Figure II.8.2: The Cartesian product of a complete graph and a cycle power. Note that there is also a C_6^2 between the $v_{i,1}$ vertices and another C_6^2 connecting the $v_{j,2}$ vertices. See Table II.8.2 for efficiency matrix.

Table II.8.2: The efficiency matrix for $K_3 \times C_6^2$.

	$v_{1,1}$	$v_{1,2}$	$v_{1,3}$	$v_{2,1}$	$v_{2,2}$	$v_{2,3}$	$v_{3,1}$	$v_{3,2}$	$v_{3,3}$	$v_{4,1}$	$v_{4,2}$	$v_{4,3}$	$v_{5,1}$	$v_{5,2}$	$v_{5,3}$	$v_{6,1}$	$v_{6,2}$	$v_{6,3}$
$v_{1,1}$	0	1	1	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$
$v_{1,2}$	1	0	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$
$v_{1,3}$	1	1	0	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1
$v_{2,1}$	1	$\frac{1}{2}$	$\frac{1}{2}$	0	1	1	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$	1	$\frac{1}{2}$	$\frac{1}{2}$
$v_{2,2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	1	0	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	1	$\frac{1}{2}$
$v_{2,3}$	$\frac{1}{2}$	$\frac{1}{2}$	1	1	1	0	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1
$v_{3,1}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	0	1	1	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$
$v_{3,2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	1	0	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$
$v_{3,3}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	1	1	0	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{2}$
$v_{4,1}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	0	1	1	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$
$v_{4,2}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	1	0	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$
$v_{4,3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	1	1	0	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1
$v_{5,1}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	0	1	1	1	$\frac{1}{2}$	$\frac{1}{2}$
$v_{5,2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	1	0	1	$\frac{1}{2}$	1	$\frac{1}{2}$
$v_{5,3}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	1	1	0	$\frac{1}{2}$	$\frac{1}{2}$	1
$v_{6,1}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	0	1	1
$v_{6,2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	1	0	1
$v_{6,3}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	1	1	1	0

For a Cartesian product between a complete graph K_r and a cycle power C_n^m, we must divide the global efficiency into two cases where n is either odd or even.

Theorem II.8.3. *If $n = 2k + 1$ then*

$$E_{glob}(K_r \times C_n^m) = \frac{1}{r(2k+1)-1}\left[2\sum_{i=1}^{k}\left(\frac{1}{\left\lceil \frac{i}{m} \right\rceil} + \frac{r-1}{\left\lceil \frac{i}{m} \right\rceil + 1}\right) + r - 1\right]. \tag{II.8.2}$$

If $n = 2k$ then

$$E_{glob}(K_r \times C_n^m) = \frac{1}{2rk-1}\left[2\sum_{i=1}^{k}\left(\frac{1}{\left\lceil \frac{i}{m}\right\rceil} + \frac{r-1}{\left\lceil \frac{i}{m}\right\rceil + 1}\right) + r - 1 - \left(\frac{1}{\left\lceil \frac{k}{m}\right\rceil} + \frac{r-1}{\left\lceil \frac{k}{m}\right\rceil + 1}\right)\right]. \tag{II.8.3}$$

Proof. Each i corresponds to a single line of each block. Also, each i has one entry that falls on the main diagonal of an $r \times r$ block that corresponds to pairs of vertices within a cycle power. All other terms correspond to pairs of vertices that are in different cycle powers but in different copies of K_r. There are also 1's next to the main diagonal; these correspond to movements within a single complete graph. The number of 1's is then $r-1$: the number of vertices that are available for the final position. Averaging the efficiencies over all pairs of vertices yields Eqs. (II.8.2) and (II.8.3).

$$\begin{aligned} E_{glob}(K_r \times C_n^m) &= \frac{1}{nr(nr-1)} \cdot nr \cdot \left[2\sum_{i=1}^{\frac{n-1}{2}}\left(\frac{1}{\left\lceil \frac{i}{m}\right\rceil} + \frac{r-1}{1+\left\lceil \frac{i}{m}\right\rceil}\right) + r - 1\right] \\ &= \frac{1}{nr-1}\left[2\sum_{i=1}^{\frac{n-1}{2}}\left(\frac{1}{\left\lceil \frac{i}{m}\right\rceil} + \frac{r-1}{1+\left\lceil \frac{i}{m}\right\rceil}\right) + r - 1\right]. \end{aligned}$$

For the even case, note that the term corresponding to the efficiency of moving across the diameter is counted twice, so it must be subtracted to obtain Eq. (II.8.3). ■

An alternate formula for faster computation can be found in Corollary VIII.1.9.

II.8.3 $K_m \times K_n$

Theorem II.8.4.

$$E_{glob}(K_m \times K_n) = \frac{nm + m + n - 3}{2(nm-1)}. \tag{II.8.4}$$

Proof. We obtain the global efficiency for $K_m \times K_n$ using $E_{glob}\left(K_m \times P_n^{n-1}\right)$ and $E_{glob}\left(K_m \times C_n^{\lfloor \frac{n}{2} \rfloor}\right)$.

$$\begin{aligned}
E_{glob}(K_m \times K_n) &= E_{glob}(K_m \times P_n^{n-1}) \\
&= \frac{2}{nm(nm-1)}\left[\sum_{i=1}^{n-1}(n-i)\left(\frac{m}{\left\lceil \frac{i}{n-1} \right\rceil} + \frac{m(m-1)}{\left\lceil \frac{i}{n-1} \right\rceil + 1}\right) + \frac{nm(m-1)}{2}\right] \\
&= \frac{2}{n(nm-1)}\left[\sum_{i=1}^{n-1}(n-i)\left(\frac{1}{1} + \frac{m-1}{1+1}\right) + \frac{n(m-1)}{2}\right] \\
&= \frac{1}{n(nm-1)}\left[\sum_{i=1}^{n-1} i\,(m+1) + n(m-1)\right] \\
&= \frac{1}{n(nm-1)}\left[\left(\frac{(n-1)(n-1+1)}{2}\right)(m+1) + n(m-1)\right] \\
&= \frac{1}{nm-1}\left[\frac{n-1}{2}(m+1) + (m-1)\right] \\
&= \frac{1}{nm-1}\left[\frac{nm}{2} - \frac{m}{2} + \frac{n}{2} - \frac{1}{2} + m - 1\right] \\
&= \frac{nm+m+n-3}{2(nm-1)}.
\end{aligned}$$

We note that the ceiling functions were dropped since $1 \le i \le n-1$ implies $0 < \frac{1}{n-1} \le \frac{i}{n-1} \le 1$ which makes the ceiling terms always equal to 1. We can also use the Cartesian product of a complete graph and a cycle power graph. If we use the case of an odd cycle power graph, we have:

$$\begin{aligned}
E_{glob}(K_m \times K_n) &= E_{glob}\left(K_m \times C_n^{\frac{n-1}{2}}\right) \\
&= \frac{1}{nm-1}\left[2\sum_{i=1}^{\frac{n-1}{2}}\left(\frac{1}{\left\lceil \frac{i}{(n-1)/2} \right\rceil} + \frac{m-1}{\left\lceil \frac{i}{(n-1)/2} \right\rceil + 1}\right) + m - 1\right] \\
&= \frac{1}{nm-1}\left[2\sum_{i=1}^{\frac{n-1}{2}}\left(\frac{1}{1} + \frac{m-1}{1+1}\right) + m - 1\right] \\
&= \frac{1}{nm-1}\left[2\sum_{i=1}^{\frac{n-1}{2}}\frac{m+1}{2} + m - 1\right] \\
&= \frac{1}{nm-1}\left[\frac{n-1}{2}(m+1) + m - 1\right] \\
&= \frac{nm+m+n-3}{2(nm-1)}.
\end{aligned}$$

For the Cartesian product of a complete graph and an even cycle, we have:

$$
\begin{aligned}
E_{glob}(K_m \times K_n) &= E_{glob}\left(K_m \times C_n^{\frac{n}{2}}\right) \\
&= \frac{1}{nm-1}\left[2\sum_{i=1}^{\frac{n}{2}}\left(\frac{1}{\left\lceil\frac{i}{n/2}\right\rceil} + \frac{m-1}{\left\lceil\frac{i}{n/2}\right\rceil+1}\right) + m - 1 - \left(\frac{1}{\left\lceil\frac{n}{2(n/2)}\right\rceil} + \frac{m-1}{\left\lceil\frac{n}{2(n/2)}\right\rceil+1}\right)\right] \\
&= \frac{1}{nm-1}\left[2\sum_{i=1}^{\frac{n}{2}}\left(\frac{1}{1} + \frac{m-1}{1+1}\right) + m - 1 - \left(\frac{1}{1} + \frac{m-1}{1+1}\right)\right] \\
&= \frac{1}{nm-1}\left[2\sum_{i=1}^{\frac{n}{2}}\frac{m+1}{2} + m - 1 - \frac{1}{2}(m+1)\right] \\
&= \frac{1}{nm-1}\left[\frac{n}{2}(m+1) + m - 1 - \frac{1}{2}(m+1)\right] \\
&= \frac{1}{nm-1}\left[\frac{n-1}{2}(m+1) + m - 1\right] \\
&= \frac{nm+m+n-3}{2(nm-1)}.
\end{aligned}
$$

All three derivations agree. Thus $E_{glob}(K_m \times K_n)$ is given by Eq. (II.8.4). ∎

II.8.4 Grid Graphs: $P_m \times P_n$

Consider the grid graph $P_m \times P_n$ which is embedded in the cartesian plane. The vertex in the upper left corner is labeled $v_{1,1}$ and $v_{i,j}$ is used to label vertex that is obtained by starting at vertex $v_{1,1}$ and travelling $i-1$ positions to the right and then $j-1$ units downward.

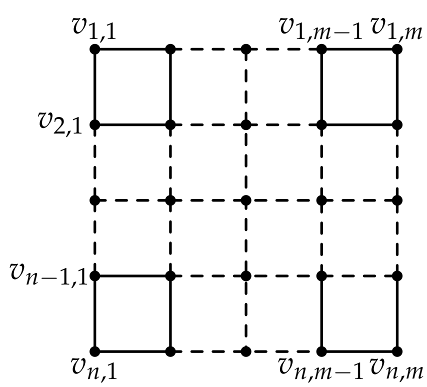


Figure II.8.3: A generic grid Ggaph composed of $P_n \times P_m$.

Now consider the graph $P_3 \times P_6$.

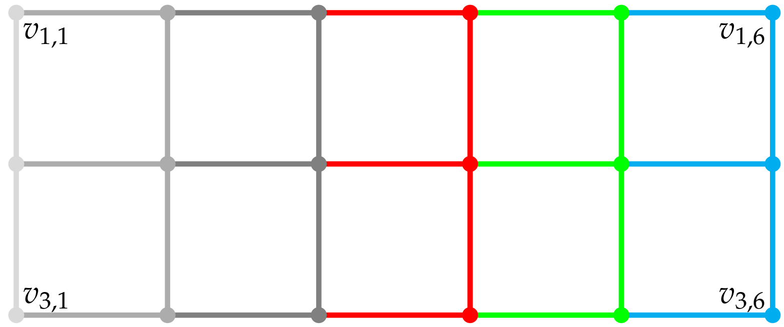


Figure II.8.4: A Grid Graph composed of $P_3 \times P_6$.

The initial block of 9 vertices from $v_{1,1}$ to $v_{3,3}$ creates the graph $P_3 \times P_3$. Adding sets of 3 additional vertices, $v_{1,4}$ to $v_{3,4}$ up to $v_{1,6}$ to $v_{3,6}$ we obtain the entire graph of $P_3 \times P_6$. This can be seen in Figure II.8.4. The efficiency matrix in Table II.8.3 is divided into subsections of $P_3 \times P_n$ where $n \leq 6$.

	$v_{1,1}$	$v_{2,1}$	$v_{3,1}$	$v_{1,2}$	$v_{2,2}$	$v_{3,2}$	$v_{1,3}$	$v_{2,3}$	$v_{3,3}$	$v_{1,4}$	$v_{2,4}$	$v_{3,4}$	$v_{1,5}$	$v_{2,5}$	$v_{3,5}$	$v_{1,6}$	$v_{2,6}$	$v_{3,6}$
$v_{1,1}$	0	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{7}$
$v_{2,1}$		0	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{6}$
$v_{3,1}$			0	$\frac{1}{3}$	$\frac{1}{2}$	1	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{5}$
$v_{1,2}$				0	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{6}$
$v_{2,2}$					0	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{5}$
$v_{3,2}$						0	$\frac{1}{3}$	$\frac{1}{2}$	1	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{4}$
$v_{1,3}$							0	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{5}$
$v_{2,3}$								0	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{4}$
$v_{3,3}$									0	$\frac{1}{3}$	$\frac{1}{2}$	1	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$
$v_{1,4}$										0	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$
$v_{2,4}$											0	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$
$v_{3,4}$												0	$\frac{1}{3}$	$\frac{1}{2}$	1	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$
$v_{1,5}$													0	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{3}$
$v_{2,5}$														0	1	$\frac{1}{2}$	1	$\frac{1}{2}$
$v_{3,5}$															0	$\frac{1}{3}$	$\frac{1}{2}$	1
$v_{1,6}$																0	1	$\frac{1}{2}$
$v_{2,6}$																	0	1
$v_{3,6}$																		0

Table II.8.3: The efficiency matrix for $P_3 \times P_6$.

Our first goal is to sum the efficiencies of $P_m \times P_n$. We shall consider the copies of P_m to be 'vertical' and the P_n copies to be 'horizontal'. To sum the efficiencies we begin by considering the n copies of P_m. The sum of efficiencies between a single P_m is simply $\sum_{k=1}^{m-1} \frac{m-k}{k}$. So our total for vertical connections is $n \sum_{k=1}^{m-1} \frac{m-k}{k}$. Similarly, our total for horizontal connections is $m \sum_{k=1}^{n-1} \frac{n-k}{k}$.

Next we determine the remaining efficiencies. Consider two copies of P_m. There are $n - i$ pairs of P_m that are separated by a horizontal distance of $i \leq n - 1$. There are $2(m - j)$ pairs of points in separate P_m that are separated by a vertical distance of $j \leq m - 1$. Thus the sum of efficiencies of the cross terms is $\sum_{i=1}^{n-1} \sum_{j=1}^{m-1} \frac{2(n-i)(m-j)}{i+j}$. Since the total number of vertices is nm, our global efficiency is:

$$\begin{aligned}
E_{glob}\left(P_m \times P_n\right) &= \frac{2}{mn(mn-1)}\left[n\sum_{k=1}^{m-1}\frac{m-k}{k} + m\sum_{k=1}^{n-1}\frac{n-k}{k} + \sum_{i=1}^{n-1}\sum_{j=1}^{m-1}\frac{2(n-i)(m-j)}{i+j}\right] \\
&= \frac{2}{mn(mn-1)}\left[\sum_{k=1}^{m-1}\frac{nm}{k} - n(m-1) + \sum_{k=1}^{n-1}\frac{mn}{k} - m(n-1) + 2\sum_{i=1}^{n-1}\sum_{j=1}^{m-1}\frac{(n-i)(m-j)}{i+j}\right] \\
&= \frac{2}{mn(mn-1)}\left[m+n-2nm + \sum_{k=1}^{m-1}\frac{nm}{k} + \sum_{k=1}^{n-1}\frac{nm}{k} + 2\sum_{i=1}^{n-1}\sum_{j=1}^{m-1}\frac{(n-i)(m-j)}{i+j}\right] \\
&= \frac{2}{mn(mn-1)}\left[\sum_{k=1}^{m}\frac{nm}{k} + \sum_{k=1}^{n}\frac{nm}{k} - 2nm + 2\sum_{i=1}^{n-1}\sum_{j=1}^{m-1}\frac{(n-i)(m-j)}{i+j}\right].
\end{aligned}$$

We shall restate this in a theorem.

Theorem II.8.5.

$$E_{glob}(P_n \times P_m) = \frac{2}{mn-1}\left[H_n + H_m - 2 + \frac{2}{nm}\sum_{i=1}^{n-1}\sum_{j=1}^{m-1}\frac{(n-i)(m-j)}{i+j}\right].$$

Conjecture II.8.6. *Without loss of generality, assume $m \geq n$.*

$$\begin{aligned}
E_{glob}(P_n \times P_m) = \frac{2}{3nm(nm-1)}\Bigg[&\frac{9mn^2 - 27mn + 11n^3 + 6n^2 - 5n + 6}{6} \\
&+ n\left(3mn + n^2 - 1\right)\sum_{k=n}^{m-1}\frac{1}{k} \\
&+ \sum_{k=0}^{n-2}\frac{(n-k)\left[(n-k)^2 - 1\right]}{k+m}\Bigg].
\end{aligned}$$

This formula was found to be consistent but is not called a theorem due to the inadequate description of derivation.

Using a weight corresponding to the Euclidean distance, we can obtain the global efficiency ratio which compares the efficiency using distances along the lines of the grid versus the ideal Euclidean distance.

Theorem II.8.7. *The global efficiency ratio is given by:*

$$E_{Ratio}(P_m \times P_n) = \frac{(H_n + H_m - 2)nm + 2\sum_{i=1}^{n-1}\sum_{j=1}^{m-1}\frac{(n-i)(m-j)}{i+j}}{(H_n + H_m - 2)nm + 2\sum_{i=1}^{n-1}\sum_{j=1}^{m-1}\frac{(n-i)(m-j)}{\sqrt{i^2+j^2}}}. \tag{II.8.5}$$

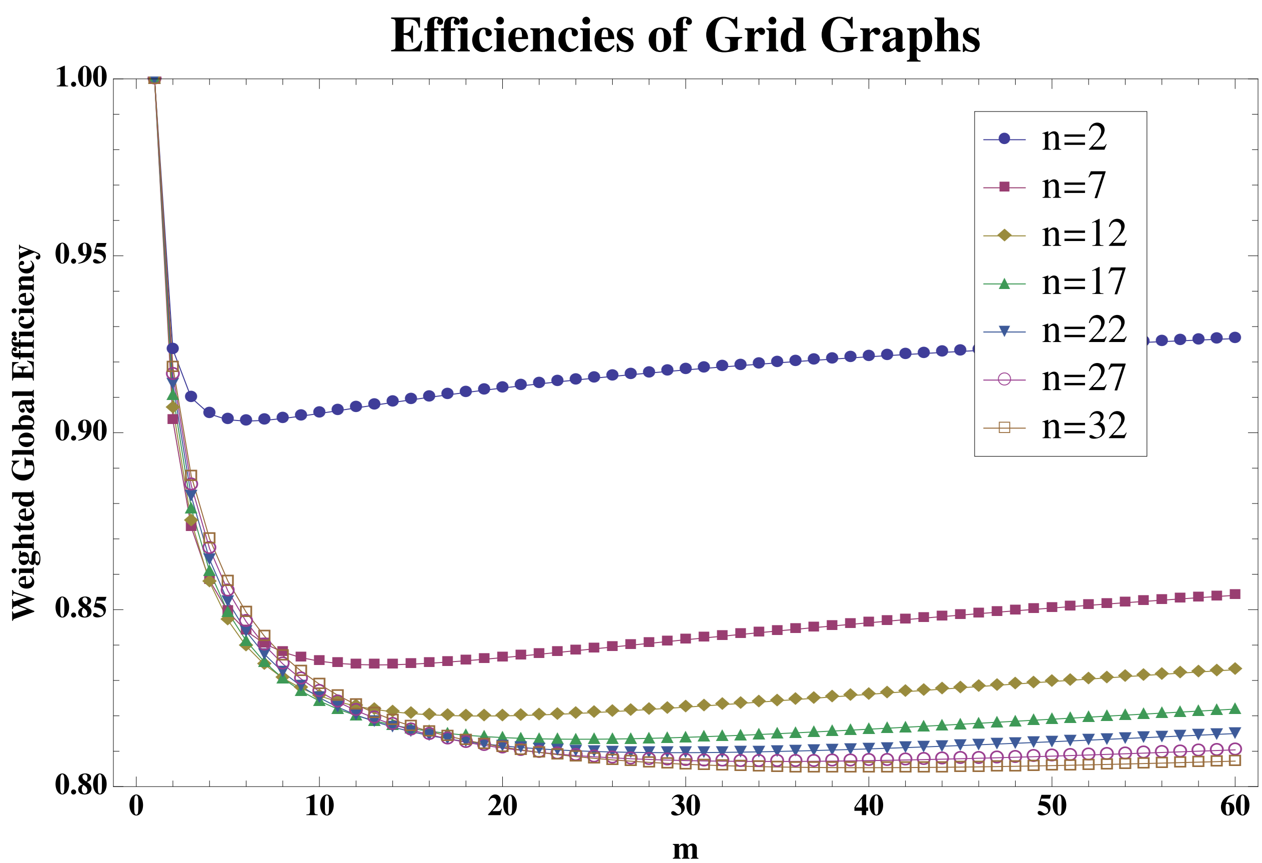


Figure II.8.5: A plot of $E_{Ratio}(P_m \times P_n)$ for fixed n as a function of m. Fixing n and increasing m tends to initially decrease E_{Ratio} until around $n = m$ and then increases asymptotically towards 1: resembling a path. The minimum value is due to the extreme non-path like nature of a square grid. Note though that the minimum occurs slightly past a square grid. Future research could be into why this is the case.

II.8.5 Harary Index

We can use the close relationship between global efficiency and the Harary index, $H(G) = \frac{n(n-1)}{2} E_{glob}(G)$, where n is the size of the vertex set, to obtain new results. Note that $H(G)$ denotes Harary index of a graph whereas H_n denotes the n^{th} harmonic number.

Corollary II.8.8. *Let $H(G)$ be the Harary index of a graph G. Then:*

1. $H\left(P_n^m\right) = \sum_{i=1}^{k} \frac{n-i}{\left\lceil \frac{i}{m} \right\rceil}$,
2. $n = 2k$: $H\left(C_n^m\right) = k\left[2\sum_{i=1}^{k} \frac{1}{\left\lceil \frac{i}{m} \right\rceil} - \frac{1}{\left\lceil \frac{k}{m} \right\rceil}\right]$,
3. $n = 2k+1$: $H\left(C_n^m\right) = (2k+1)\sum_{i=1}^{k} \frac{1}{\left\lceil \frac{i}{m} \right\rceil}$,

4. $H\left(K_{s_1,\ldots,s_t}\right) = \frac{n(n-1)}{2} - \frac{1}{4}\left[\sum_{i=1}^{t} s_i^2 - n\right]$,

5. $H(K_{n,m}) = nm + \frac{1}{4}\left[n^2 - n + m^2 - m\right]$,

6. $H(K_{r;n}) = \frac{nr(2nr-r-1)}{4}$,

7. $H(S_{d,l}) = d\left((d-1)\left(l+\frac{1}{2}\right)H_{2l} - (d-2)(l+1)H_l - l\right)$,

8. $H\left(K_r \times P_n^m\right) = r\sum_{i=1}^{k}(n-i)\left(\frac{1}{\lceil \frac{i}{m} \rceil} + \frac{r-1}{1+\lceil \frac{i}{m} \rceil}\right) + \frac{nr(r-1)}{2}$,

9. $n = 2k$: $H\left(K_r \times C_n^m\right) = rk\left[2\sum_{i=1}^{k}\left(\frac{1}{\lceil \frac{i}{m} \rceil} + \frac{r-1}{\lceil \frac{i}{m} \rceil + 1}\right) + r - 1 - \left(\frac{1}{\lceil \frac{k}{m} \rceil} + \frac{r-1}{\lceil \frac{k}{m} \rceil + 1}\right)\right]$,

10. $n = 2k+1$: $H\left(K_r \times C_n^m\right) = r(2k+1)\left[\sum_{i=1}^{k}\left(\frac{1}{\lceil \frac{i}{m} \rceil} + \frac{r-1}{\lceil \frac{i}{m} \rceil + 1}\right) + \frac{r-1}{2}\right]$,

11. $H\left(K_m \times K_n\right) = \frac{1}{4}nm(nm + m + n - 3)$,

12. $H\left(P_m \times P_n\right) = (H_n + H_m - 2)nm + 2\sum_{i=1}^{n-1}\sum_{j=1}^{m-1}\frac{(n-i)(m-j)}{i+j}$.

A listing of Harary index values using faster computation are available in Corollary VIII.1.10.

II.9 Applications

II.9.1 Metropolitan Atlanta Rapid Transit Authority Subway

The Metropolitan Atlanta Rapid Transit Authority Subway has 38 stations (see Figure II.9.1). We note that 33 of the 38 stations fall within a subdivided star formation. Approximating the network as the graph $S_{4,8}$, we have

$$E_{Ratio}\left(S_{4,8}\right) = \frac{(4-1)\left(2\cdot 8+1\right)H_{2\cdot 8} - (4-2)(2\cdot 8+2)H_8 - 2\cdot 8}{(2\cdot 8+2)H_8 - 2\cdot 8 + \sum_{i=1}^{8}\sum_{j=1}^{8}\sum_{\theta=1}^{4-1}\frac{1}{\sqrt{i^2+j^2-2ij\cos\left(\frac{2\pi}{4}\theta\right)}}} = 0.914\,27.$$

Next, we put our estimate to the test by considering the actual MARTA network.

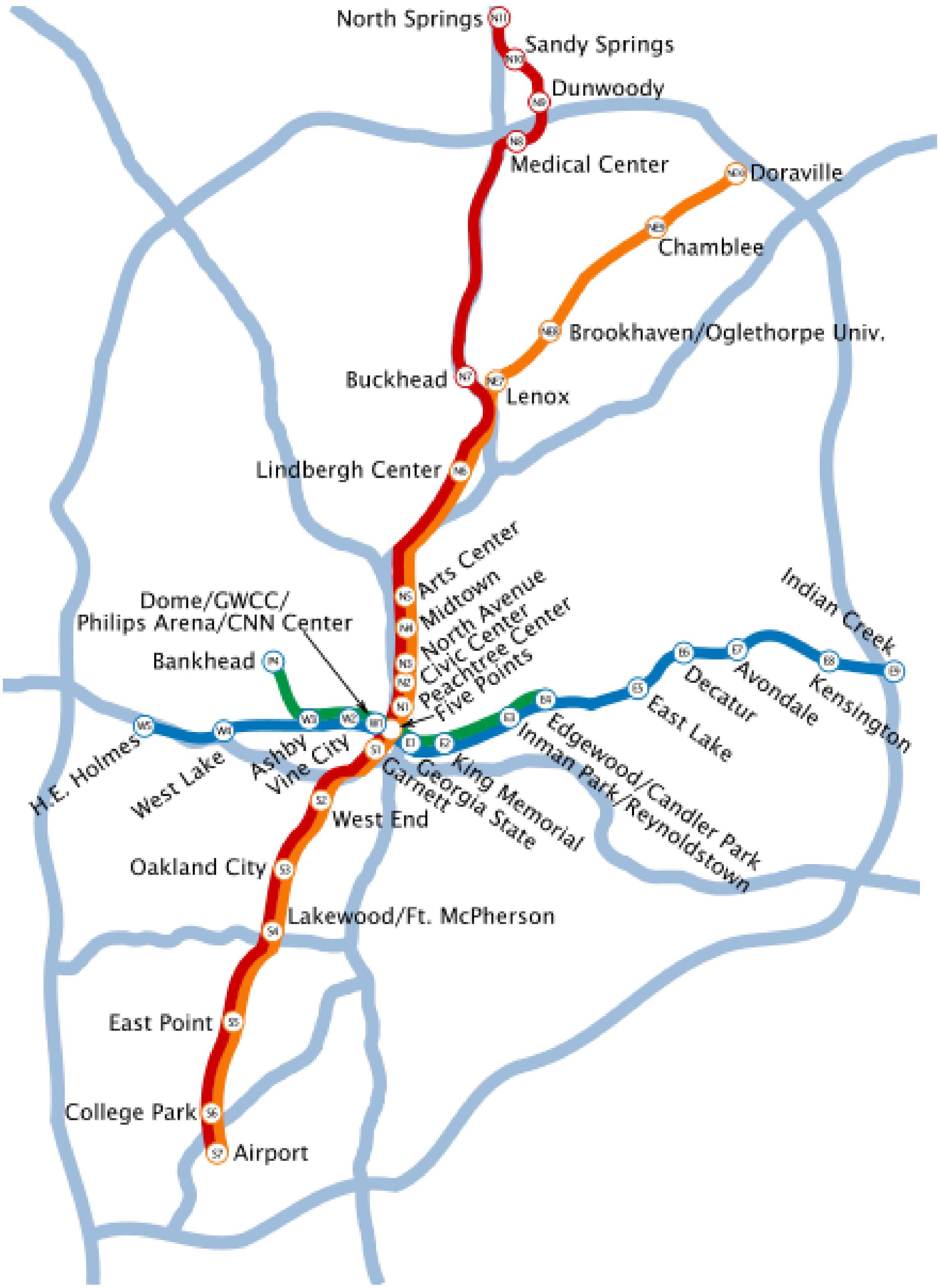


Figure II.9.1: A scale map of MARTA: the Atlanta metro. Note the star-like design. Five Points Station serves as the central vertex. The distance between North Springs station and Sandy Springs station is approximately 1 mile. Light blue lines denote the city's major interstate corridors. Used with permission from M. Casey.[17]

After obtaining rail distances along each of the lines directly from MARTA, we calculated the rail distance between every pair of stations. The distances are shown in Table VIII.1.1 (Appendix). Using Google Earth we determined the Euclidean distance between every pair of stations (see Table VIII.1.2 in Appendix).

In our analysis we only consider distances between stations, and not the length of a track in a particular station. Using Google Earth, we found the Euclidean distances (in miles) between every pair of rail stations. For a map of the MARTA Subway network where the scale is Euclidean distance, see Figure II.9.1. The sum of the maximum Euclidean efficiencies was then computed to be 379.8169. Using rail distances provided by MARTA we calculated the actual efficiencies with total sum of 311.7036. Hence, E_{Ratio}(MARTA)$= \frac{311.7036}{379.8169} = 0.8207$.

This means that the MARTA system is roughly 82% as efficient (in terms of distance) as a system that has every station connected to every other station by a direct rail line.

Our analysis shows that a main fingerprint of a subway network can be star-like in structure. We note that the graph $S_{d,l}$ is a star that is perfectly "balanced", meaning that all of pendant paths have the same length. As noted earlier if l is fixed and d is increased then the efficiency ratio decreases as the number of pairs of vertices on different spokes is increased. Also if d is fixed and l is increased then the efficiency ratio increases, as the network bears a closer resemblance to a path.

It would be a difficult problem indeed to generalize the result for balanced stars to stars where the pendant paths are of arbitrary length. However it is reasonable to derive some approximations. If the pendant paths are of similar lengths then the efficiency will be close to that of a balanced star. Given a star-like network where the pendant paths have different lengths, it is tempting to consider $S_{d,l}$ where l is the average of the pendant path lengths. However this will not work well in a case where there is significant variation in the path lengths. For example the efficiency of $S_{4,25}$ is much different than the efficiency of a star-like graph with one pendant path with 97 edges and three pendant edges. The latter will be much closer to 1. If there are a small number of pendant paths and one of the paths is much larger than the others (e.g.. there are four pendant paths with lengths 91, 1,1, and 1), this graph will resemble a path in structure and will thus have an efficiency close to 1.

There is also a further generalization where the pendant paths are replaced by pendant trees (as in the MARTA network). We noted previously that the MARTA network is similar to $S_{4,8}$, and the efficiencies are 82% and 91% respectively. There are two reasons for the discrepancy, the first being that distances from Five Points Station to the last station on each line are not all the same. The second is that the lines leaving Five Points

Station going north and west split into different routes. This split impacts the efficiency in a manner similar to increasing d in $S_{d,l}$ and decreases the efficiency ratio.

We conclude by posing a problem involving a broad generalization where the network is a tree and that incident edges are separated by angles that are equal. It would be interesting to investigate not only bounds but the complexity of this problem as well.

II.9.2 Brain Network

Anything that is relatable to our brains is an area of great importance and interest. Efficiency is another way of measuring the effect of an event on a brain's composition. RCBI scientists conducted functional MRI (fMRI) scans of 25 volunteers to find blood oxygen level-dependent (BOLD) correlations of various regions of the brain. Previous papers have looked at the BOLD correlation matrices of macaque brains[12]. They used binary connections with a given correlation threshold to find regions of high connectivity. We noticed definitive "hot spots" of communication in our matrices as well and divided the regions into cliques. We developed our matrix with threshold correlations, computed the efficiency matrix and used a partitioning algorithm[18] in order to find the cliques. The algorithm ranks the regions based on the values of the eigenvector corresponding to the second largest eigenvalue of the efficiency matrix.

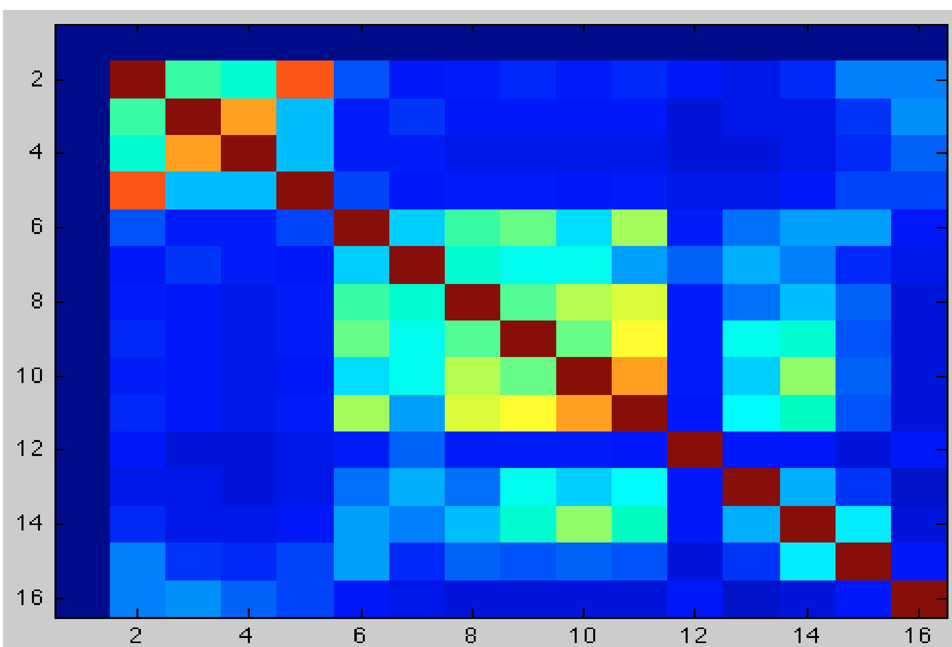


Figure II.9.2: The average correlations between regions of the brain ordered according to the efficiency partitioning algorithm. Warmer colors are more positively correlated. From left to right the regions are: LMTG, RMF, LMF, LMT, IPS, RFEF, LIP, RS, LM, LS, MPC, RM, LMLA, LPC, LLP.

The algorithm proved valid visually in the matrix as well as by partitioning motor system regions, sensory regions and left brain regions: expected high-connectivity cliques. Our preliminary findings demonstrated that while most brains varied in one or two region assignments, partitions were nearly constant across the 25 matrices (see Figure II.9.3).

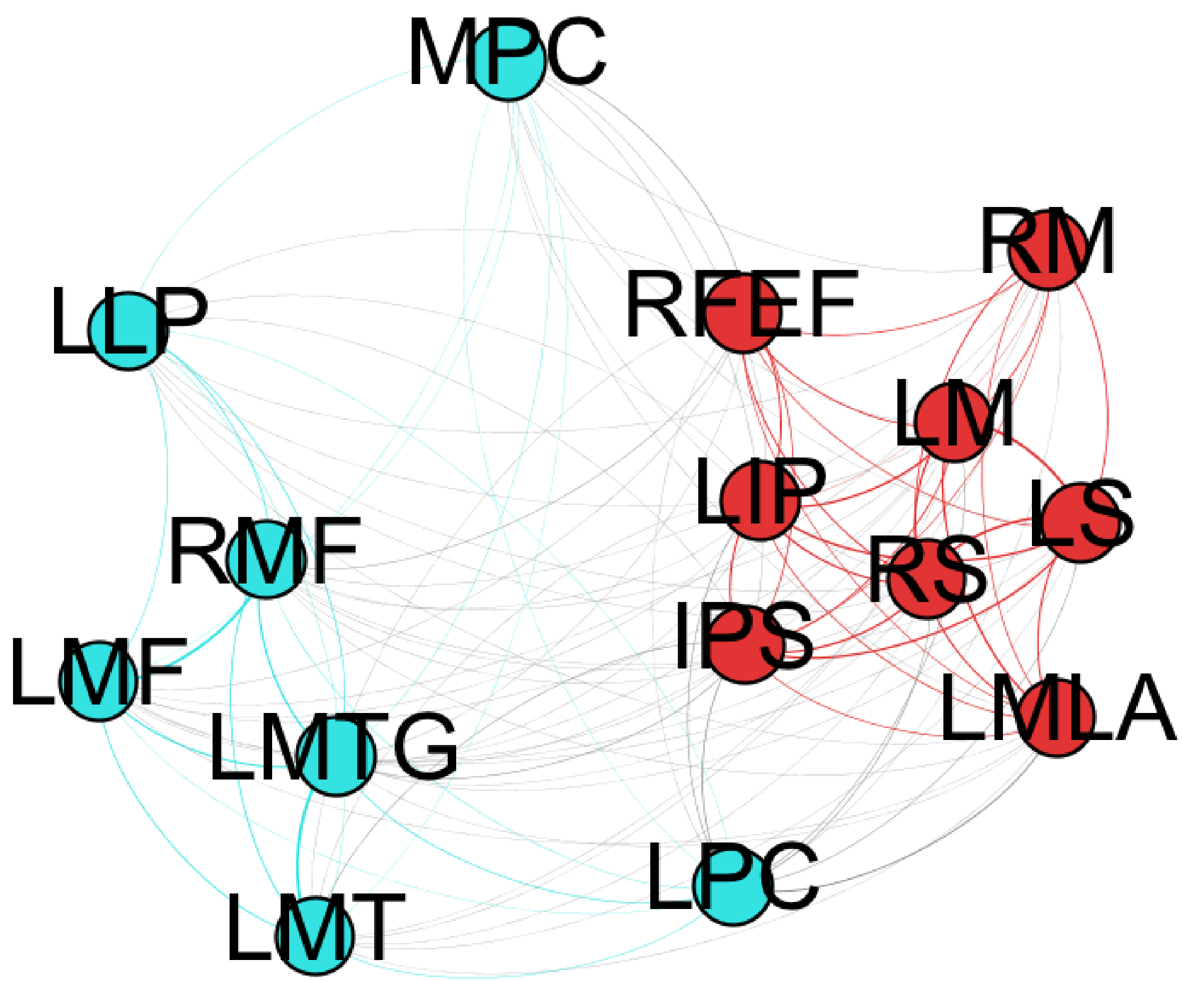


Figure II.9.3: Partition of the brain into two groups using the average of absolute values. The width of the lines correspond to the strength of the correlation. The difference is hard to see in most cases but the connection from LMF to RMF is clearly stronger than LLP to RFEF. See Table VIII.1.4 for legend.

This is great evidence for a "normal" brain that can be used as a template. Disregarding the physical distance between brain coordinates, the "average" brain graph in Figure II.9.4 was computed to have a global efficiency of 0.294. If one considered the physical distance as well, the efficiency would undoubtable increase.

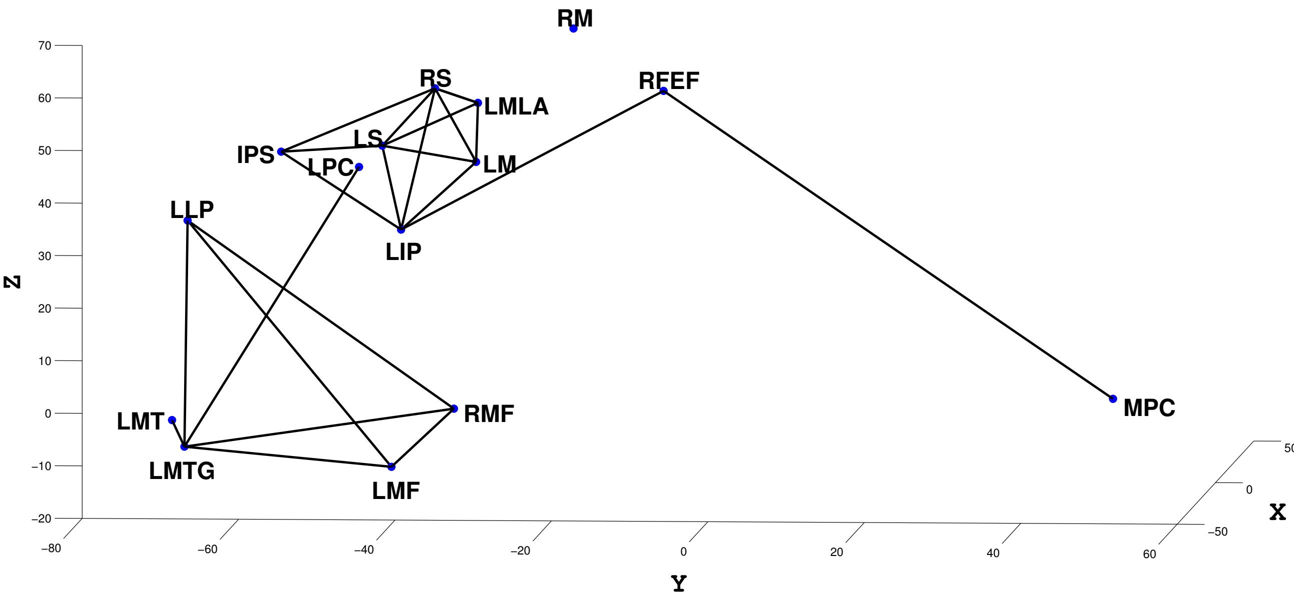


Figure II.9.4: Image of the graph created by the adjacency matrix for the average minimum matrix with a cut off of -0.2 and the average maximum matrix with a cut off of 0.4. The nodes are plotted according to talairach coordinates of the 15 measured regions in the brain. Notice that the regions of the brain tend to be connected to physically closer regions. Using the weights of the correlations instead of strict cutoffs, these regions were partitioned together as in Figure II.9.3.

We did not exhaustively try to recreate the results from Honey et. al.[12] but simply include their data as a comparison point. See Figure VIII.1.1 for their adjacency matrix of a macaque neocortex. Using their matrix we obtained a value of 0.571 for the global efficiency of their corresponding graph. If one was more relaxed in our cutoffs, the efficiency would likely approach that of the macaque's. In our further study we actually approached the same efficiency with a relaxed correlation cutoff of absolute value ≥ 0.325.

We continued our brain investigation in collaboration with the University of Rochester. Again, an outside source conducted BOLD pre and postseason fMRI scans of the players. We received matrices of the correlations in oxygen levels of various regions of the brain and modeled these as graphs. This time we received data from 52 regions of the brain (a superset of the above regions). We were then able to measure the "efficiency" of each athlete. As was expected, the athletes who received the largest number of high-energy impacts during the season also experienced the largest drop in brain efficiency. At a typical correlation cutoff of 0.325, one patient lost nearly half of the connections. This reduced the efficiency from 0.555 to 0.478. The average loss in efficiency (after removing the player who sat on the bench all season) was 0.047: a roughly 10% drop in measurement.

III. Local Efficiency and Clustering Coefficient

For simplicity of reading, we provide the definition of local efficiency and clustering coefficient again here.

Definition III.0.1. Local Efficiency Consider a graph G with vertex set $V(G)$. Let $n = |V(G)|$. Let G_i denote the neighborhood induced by vertex i.

$$E_{loc} = \frac{1}{n} \sum_{i \in V(G)} E_{glob}\left(G_i\right). \tag{III.0.1}$$

The local efficiency is the average global efficiency of all neighbor induced subgraphs.

Definition III.0.2. Clustering Coefficient Consider a graph G.

$$CC(G) = \frac{1}{n} \sum_{i} C_i, \tag{III.0.2}$$

where C_i is the number of edges in G_i divided by the maximum number of possible edges: $\binom{|V(G_i)|}{2}$.

Remark III.0.3. Note that because of these definitions, $E_{loc}(G) \geq CC(G)$.

Lemma III.0.4. *A graph G contains a K_3 subgraph if and only if $E_{loc}(G) \geq CC(G) > 0$.*

Proof. $\Longrightarrow$

If $E_{loc}(G) \geq CC(G) > 0$ then there must be some neighborhood subgraph such that $E_{glob}(G_i) > 0$. Thus G_i must contain an edge. Let this edge be between vertices u and v. Then $\{u, v, i\}$ form a K_3.

$\Longleftarrow$

If G contains a K_3 subgraph, then taking the neighborhood induced subgraph of one of these vertices will have an edge. Thus it will have nonzero C_i and nonzero E_{glob}. Therefore the local efficiency and clustering coefficient (averages) must also be nonzero.

■

III.1 Graphs where the Clustering Coefficient and Local Efficiency Differ

In this section we analyze the claim made by Latora and Marchiori and show cases where it does not hold. That is we present families of graphs where the local subgraphs are not sparse, but where CC and E_{loc} differ.

III.1.1 Clustering Coefficients vs. Local Efficiency for Complete Multipartite Graphs

A complete multipartite graph $G = K_{s_1, s_2, \ldots, s_t}$ is composed of t classes each with s_i vertices, $1 \leq i \leq t$, where each vertex in class i is adjacent to every vertex in class $j \neq i$, and is not adjacent to any vertex in class i. For bipartite graphs, the clustering coefficient and local efficiency are clearly 0.

Theorem III.1.1. *Let $G = K_{m,n}$ where $m, n \geq 1$. Then $CC(G) = 0$.*

Theorem III.1.2. *Let $G = K_{m,n}$ where $m, n \geq 1$. Then $E_{loc}(G) = 0$.*

However, for general multipartite graphs this is not the case. Let v be a vertex in $K_{s_1,s_2,\ldots,s_t}$. Let v be in part i. The subgraph G_v induced by the neighbors of v is isomorphic to $K_{s_1,\ldots,s_{i-1},s_{i+1},\ldots,s_t}$.

Theorem III.1.3. *Let $n = \sum_{i=1}^{t} s_i$.*

$$\begin{aligned} E_{loc}(K_{s_1,\ldots,s_t}) &= \frac{1}{n}\sum_{i=1}^{t} s_i E_{glob}(K_{s_1,\ldots,s_{i-1},s_{i+1},\ldots,s_t}) \\ &= \frac{1}{n}\sum_{i=1}^{t} \frac{s_i}{(n-s_i)(n-s_i-1)} \sum_{j=1, j\neq i}^{t} s_j \left[\frac{(s_j-1)}{2} + (n-s_i-s_j)\right]. \end{aligned} \tag{III.1.1}$$

Theorem III.1.4. *Let $n = \sum_{i=1}^{t} s_i$.*

$$CC(K_{s_1,\ldots,s_t}) = \frac{1}{n}\sum_{i=1}^{t} \frac{s_i}{(n-s_i)(n-s_i-1)} \sum_{j=1, j\neq i}^{t} s_j \left(n-s_i-s_j\right). \tag{III.1.2}$$

As before with the global efficiency, we shall look at a few special cases.

Lemma III.1.5. *Let $G = K_{r,r,r}$ with $r \geq 3$. Then $CC(K_{r,r,r}) = \frac{r}{2r-1}$ and $E_{loc}(K_{r,r,r}) = \frac{3r-1}{2(2r-1)}$.*

Proof. Let $V(G)$ be the union of the three classes of vertices $\{v_{1,1}, v_{1,2}, \ldots, v_{1,r}\}$, $\{v_{2,1}, v_{2,2}, \ldots, v_{2,r}\}$, and $\{v_{3,1}, v_{3,2}, \ldots, v_{3,r}\}$. Note that all vertices have isomorphic neighborhoods. Without loss of generality consider the vertex $v_{1,1}$. Then the neighbors of $v_{1,1}$ are $\{v_{2,1}, v_{2,2}, \ldots, v_{2,r}\}$ and $\{v_{3,1}, v_{3,2}, \ldots, v_{3,r}\}$. We first calculate the clustering coefficient. The only pairs of neighbors of $v_{1,1}$ that are adjacent are $(v_{2,i}, v_{3,j})$. The number of these edges is r^2. Dividing over the total number of possible edges gives $\frac{r^2}{\binom{2r}{2}} = \frac{r}{2r-1}$. This is the value of the clustering coefficient. Non-1 efficiencies create the difference between clustering coefficient and local efficiency. Therefore, we next examine efficiencies of each pair of these vertices. We have $\in (v_{2,i}, v_{2,j}) = \frac{1}{2}$ and $\in (v_{3,i}, v_{3,j}) = \frac{1}{2}$ for all $i \neq j$, $1 \leq j \leq r$, and $\in (v_{2,i}, v_{3,j}) = 1$ for all $1 \leq i, j \leq r$. Hence the sum of all efficiencies is $\binom{r}{2}\frac{1}{2} + \binom{r}{2}\frac{1}{2} + r^2 = r^2 + \binom{r}{2}$. Averaging over all efficiencies gives $\frac{r^2+\binom{r}{2}}{\binom{2r}{2}}$. ∎

We note the two small cases.

Corollary III.1.6.

$$\begin{aligned} \lim_{r\to\infty} CC(K_{r,r,r}) &= \lim_{r\to\infty} \frac{r^2}{r(2r-1)} = \frac{r}{2r-1} = \frac{1}{2}, \\ \lim_{r\to\infty} E_{loc}(K_{r,r,r}) &= \lim_{r\to\infty} \frac{r^2+\binom{r}{2}}{\binom{2r}{2}} = \lim_{r\to\infty} \frac{r^2+\frac{1}{2}(r-1)}{2r-1} = \frac{1+\frac{1}{2}}{2} = \frac{3}{4}, \\ \lim_{r\to\infty} \left(E_{loc}(K_{r,r,r}) - CC(K_{r,r,r})\right) &= \frac{1}{4}. \end{aligned}$$

We can then look at the case for general n.

Theorem III.1.7. *Let $K_{r;n}$ denote the complete multipartite graph consisting of n parts of order r. Then $E_{loc}(K_{r;n}) = 1 - \frac{r-1}{2[(n-1)r-1]}$ and $CC(K_{r;n}) = 1 - \frac{r-1}{(n-1)r-1}$.*

Proof. Let v be a vertex in $K_{r;n}$. The subgraph G_v induced by the neighbors of v is isomorphic to $K_{r;n-1}$. Then we can use Theorem II.5.5 to directly find E_{loc}.

Next, we consider $CC(K_{r;n})$. The number of edges in a neighborhood subgraph is $|E(K_{r;n-1})| = \binom{n-1}{2}r^2 = \frac{1}{2}(n-1)(n-2)r^2$. The maximum number of edges that can exist on this set of $r(n-1)$ vertices is $\binom{(n-1)r}{2} = \frac{1}{2}(n-1)r \cdot ((n-1)r-1)$. Therefore $CC(K_{r;n}) = \frac{\frac{1}{2}(n-1)(n-2)r^2}{\frac{1}{2}(n-1)r\cdot((n-1)r-1)} = \frac{(n-2)r}{(n-1)r-1} = 1 - \frac{r-1}{(n-1)r-1}$. ■

Corollary III.1.8. *As $n \to \infty$, $E_{loc}(K_{r;n}) \to 1$ and $CC(K_{r;n}) \to 1$. The sizes of the parts have become negligible with respect to the total composition of the graph: a complete graph.*

Corollary III.1.9.

$$\lim_{r\to\infty} CC(K_{r,r,r}) = 1 - \frac{1}{n-1},$$
$$\lim_{r\to\infty} E_{loc}(K_{r,r,r}) = 1 - \frac{1}{2(n-1)},$$
$$\lim_{r\to\infty} \left(E_{loc}(K_{r,r,r}) - CC(K_{r,r,r})\right) = \frac{1}{2(n-1)}.$$

III.1.2 Cycle Powers

A cycle power C_n^m is a graph with vertices $v_1, v_2, ..., v_m$ and edges v_iv_j where $|i-j| \leq m \mod n$. In the next two theorems we determine $E_{loc}(C_n^m)$ and $CC(C_n^m)$.

Theorem III.1.10.

$$E_{loc}(C_n^m) = \begin{cases} 1 & \textit{when} \quad m \leq n \leq 2m+1, \\ \frac{16m^2+n^2+4m+2-6mn-3n}{4m(2m-1)} & \textit{when} \quad 2m+2 \leq n \leq 3m, \\ \frac{21m^2-15m-2}{12m(2m-1)} & \textit{when} \quad n \geq 3m+1. \end{cases}$$

Proof. Because the cycle power graph is vertex-transitive, the average local efficiency is equal to the local efficiency of any particular neighborhood subgraph. Thus we consider a single vertex v_i with neighbors v_j where $i-m \leq j \leq i-1$ and $i+1 \leq j \leq i+m$. We consider three cases.

Case (i). Let $m \leq n \leq 2m+1$. Then all of the local subgraphs are complete and $E_{loc}(C_n^m) = 1$.

Case (ii). Let $2m+2 \leq n \leq 3m$. See Figure III.1.1.

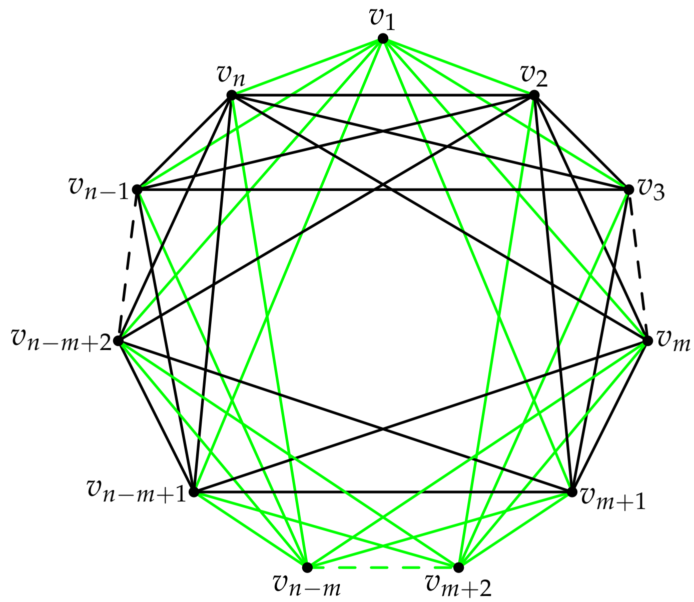


Figure III.1.1: A cycle power with the local subgraph shown in black.

Keeping m constant but increasing n increases the distance separating v_{i-m} and v_{i+m} until it becomes greater than m; when $n \geq 3m+1$. As a result, fewer edges are kept until a single edge remains. This prevents G_i from becoming a path. The single edge is between v_{i-m} and v_{i+m}. For each value of m, there are $(3m)-(2m+3)+1 = m-2$ different values of n that fall into this case. The greatest value of n in this interval: $3m$, will yield a path power plus a connection between v_{i-m} and v_{i+m}. Then,

$$\begin{aligned}
E_{loc}(C_n^m) &= \frac{2}{2m\,(2m-1)}\left[\sum_{k=1}^{2m-1}\frac{2m-k}{\left\lceil\frac{i}{m-1}\right\rceil} + \sum_{k=1}^{3m+1-n}(3m+2-n-k)\left(1-\frac{1}{\left\lceil\frac{n-m-2+k}{m-1}\right\rceil}\right)\right] \\
&= \frac{1}{m\,(2m-1)}\left[\sum_{k=1}^{2m-2}\frac{2m-k}{\left\lceil\frac{i}{m-1}\right\rceil} + 1 + \sum_{k=1}^{3m-n}(3m+2-n-k)\left(1-\frac{1}{\left\lceil\frac{n-m-2+k}{m-1}\right\rceil}\right)\right].
\end{aligned}$$

The second double sum corrects the efficiencies that were changed by having a connection between v_{i-m} and v_{i+m}. Once the final term is removed, we can use $n \geq 2m+2$, and $1 \leq k \leq 3m-n$ to see that $m+1 \leq n-m-2+k \leq 2m-2$. Therefore the fraction term will always have denominator 2. Then

$$\begin{aligned}
E_{loc}(C_n^m) &= \frac{1}{m\,(2m-1)}\left[\sum_{k=1}^{m-1}\frac{2m-k}{2} + \sum_{k=1}^{2m-2}\frac{2m-k}{2} + 1 + \sum_{k=1}^{3m-n}\frac{3m+2-n-k}{2}\right] \\
&= \frac{1}{m\,(2m-1)}\left[(m-1)\left(3m-\frac{m}{4}-\frac{(2m-1)}{2}\right) + 1 + (3m-n)\left(\frac{3m+2-n}{2}-\frac{3m-n+1}{4}\right)\right] \\
&= \frac{1}{4m\,(2m-1)}\left[16m^2+4m-6mn+n^2-3n+2\right].
\end{aligned}$$

Case (iii). Let $n \geq 3m+1$. Then the subgraph induced by the neighborhood of a vertex is a path power (see Figure III.1.2). Specifically, it is P_{2m}^{m-1}. The generic graphic is given by:

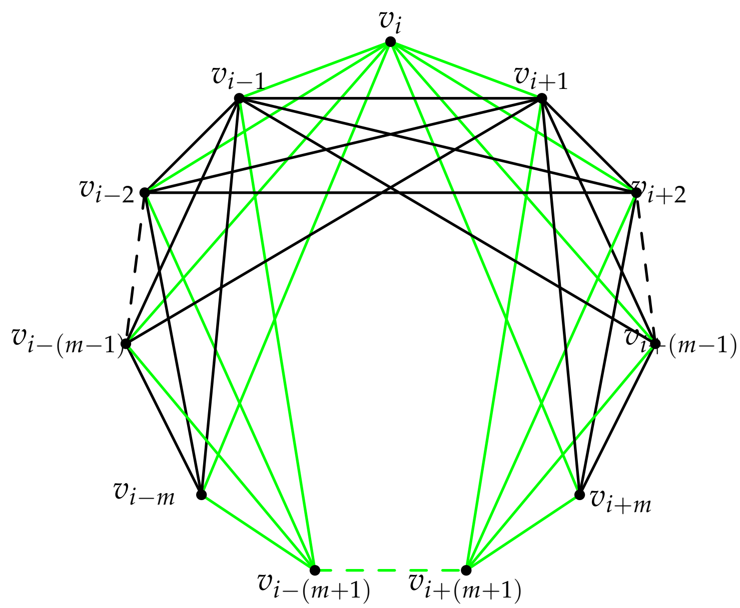


Figure III.1.2: A generic cycle power. Note the black edges indicate G_i of C_n^m where $n \geq 3m+1$

Then

$$\begin{aligned}
E_{loc}(C_n^m) &= \frac{2}{2m\,(2m-1)}\left[\sum_{k=1}^{2m-1}\sum_{i=1}^{k}\frac{1}{\left\lceil \frac{i}{m-1}\right\rceil}\right] \\
&= \frac{1}{m(2m-1)}\left[\sum_{k=1}^{m-1}\sum_{i=1}^{k}\frac{1}{1}+\sum_{k=m}^{2m-1}\left(k-1+(k-m+1)\frac{1}{2}\right)\right] \\
&= \frac{1}{2m\,(2m-1)}\left(\frac{7}{2}m^2-\frac{5}{2}m-\frac{1}{3}\right).
\end{aligned}$$

■

Next we investigate the clustering coefficient of a cycle power.

Theorem III.1.11.

$$CC(C_n^m) = \begin{cases} 1 & \textit{when} \quad n \leq 2m+1, \\ \frac{12m^2+6m-6mn+n^2-3n+2}{4m^2-2m} & \textit{when} \quad 2m+2 \leq n \leq 3m, \\ \frac{3m-3}{4m-2} & \textit{when} \quad n \geq 3m+1. \end{cases} \tag{III.1.3}$$

Proof. We consider three cases.

Case (i). $n \leq 2m+1$ Then we have a complete graph and $CC(C_n^m) = 1$.

Case (ii). $2m+2 \leq n \leq 3m$ $CC(C_n^m) = \frac{12m^2+6m-6mn+n^2-3n+2}{4m^2-2m}$.

The adjacency matrix of a subgraph is similar to the matrix of Case 1, however additional ones must be added since each subgraph is a path power with additional edges. Thus the clustering coefficient is found by adding the sum of Case 1, and the additional edges. In Case 2 for local efficiency, we changed entries equal to $\frac{1}{2}$ to 1. Now, we must add 1 to the same entries since they are 0 in this case. Thus we are adding twice as much to the adjacency matrix as in Case 2 of local efficiency, yielding $2\frac{(3m+2-n)(3m+1-n)}{2}$. The sum is $2 \cdot \frac{3}{2}m(m-1) + (3m+2-n)(3m+1-n) = 12m^2+6m-6mn+n^2-3n+2$. And normalizing yields Eq. (III.1.3).

Case (iii). $n \geq 3m+1$. Consider the subgraph created by C_n^m where $n \geq 3m+1$. Recall that this subgraph is a path power: P_{2m}^{m-1}. Hence, the adjacency matrix is simply the efficiency matrix of a path power but with every value less than 1 replaced by 0. Then the sum is $\frac{3}{2}m(m-1)$. However, this quantity is doubled since the summation only sums over the upper half of the symmetric matrix. Normalizing gives a clustering coefficient of $CC(C_n^m) = \frac{2 \cdot \frac{3}{2}m(m-1)}{2m(2m-1)} = \frac{3m-3}{4m-2}$. ∎

Remark III.1.12. We note that when $n \geq 3m+1$, $\lim_{m\to\infty} CC(C_n^m) = \frac{3}{4}$ and $\lim_{m\to\infty} E_{loc}(C_n^m) = \frac{7}{8}$.

III.1.3 Networks of the Brain

In 2007, Honey et al. [12] examined a large-scale anatomical data set known as a "macaque neocortex". This consisted of a binary connection matrix of brain regions connected by interregional pathways (see Figure VIII.1.1 in Appendix). The 47-node network was constructed by collating data from different macaques using tract-tracing studies. The network consists of 47 nodes and 505 unweighted directed edges. Here each vertex represents a particular region of the brain and each edge denotes the presence of a directed anatomical connection.

We used MATLAB to verify the following properties: $L = 2.0541$, $E_{glob} = 0.5714$, $CC = 0.6098$, and $E_{loc} = 0.7903$. We note the local subgraphs are not sparse as $CC > 0.5$ and yet there is a significant difference between CC and E_{loc}.

III.2 Graphs where C and E_{loc} are equal

In this section we consider graphs that are the Cartesian Product of a complete graph and another graph H where H is a cycle or complete graph.

III.2.1 Clustering Coefficients vs. Local Efficiency for Cartesian Products of Graphs

Theorem III.2.1. $E_{loc}(K_n \times C_m) = CC(K_n \times C_m) = \frac{n^2-3n+2}{n^2+n}$, $m \geq 4$.

Proof. We note that all vertices have isomorphic neighborhoods. Let v be a vertex in $K_n \times C_m$. Then $\deg(v) = n+1$. The vertex v has $n-1$ neighbors in the clique that contains v, and two outside of this clique. Among these neighbors in the clique there are $2\binom{n-1}{2}$ ordered pairs of adjacent vertices. The other two neighbors are non-adjacent. Hence $E_{loc}(K_n \times C_m) = \frac{2\binom{n-1}{2}}{(n+1)(n)} = \frac{n^2-3n+2}{n^2+n}$. Since all of the efficiencies between vertices are either 0 or 1, $CC(K_n \times C_m) = E_{loc}(K_n \times C_m)$. ∎

Theorem III.2.2. $E_{loc}(K_n \times K_m) = CC(K_n \times K_m) = \frac{(n-1)(n-2)+(m-1)(m-2)}{(n+m-2)(n+m-3)}$.

Proof. We note that all vertices have isomorphic neighborhoods. Let v be a vertex in $K_n \times K_m$. Then $\deg(v) = (n-1)+(m-1)$. The vertex v has $n-1$ neighbors in the clique that contains v. Among these neighbors in the clique there are $2\binom{n-1}{2}$ ordered pairs of adjacent vertices. The other neighbors form a clique of size $m-1$, with $2\binom{m-1}{2}$ ordered pairs of adjacent vertices. Hence $E_{loc}(K_n \times K_m) = \frac{2\binom{n-1}{2}+2\binom{m-1}{2}}{(n-1+m-1)(n-1+m-1-1)}$. Since all of the efficiencies between vertices are either 0 or 1, $CC(K_n \times K_m) = E_{loc}(K_n \times K_m) = \frac{2\binom{n-1}{2}+2\binom{m-1}{2}}{(n-1+m-1)(n-1+m-1-1)}$. ∎

III.2.2 Removing an edge from K_n.

Theorem III.2.3. *Let $G = K_n - e$. Then $CC(K_n - e) = 1 - \frac{2}{n(n-1)}$ and $E_{loc}(K_n - e) = 1 - \frac{1}{n(n-1)}$.*

Proof. Note that all but two of the entries in the C matrix are 1. In the E_{loc} matrix, all entries are 1 except one that is $\frac{1}{2}$. ∎

Remark III.2.4. Note that $\lim_{n\to\infty} CC(K_n - e) = 1$ and $\lim_{n\to\infty} E_{loc}(K_n - e) = 1$.

III.3 Summary of Differences in Local Efficiency and Clustering Coefficient

We identified graphs where the local efficiency and clustering coefficient were different. In Corollary III.1.6 we showed that these two quantities can asymptotically differ by $\frac{1}{4}$. It would be interesting to see how much these two quantities can differ. We formally state this problem in the future research section.

Latora and Marchiori mentioned that the clustering coefficient is a good approximation for the local efficiency of a graph when it is sparse. More accurately, the clustering coefficient is a good approximation when the vertices of neighborhoods of every vertex have low eccentricity.

IV. Betweenness Centrality

IV.1 Introduction

IV.1.1 Definition and Example

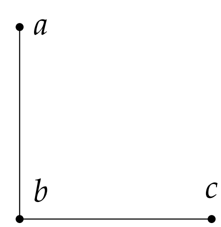
Let G be a connected graph. The betweenness centrality of a vertex $v \in G$, denoted $bc(v)$, measures the frequency at which v appears on a shortest path between two other distinct vertices x and y. Let σ_{xy} be the number of shortest paths between distinct vertices x and y, and let $\sigma_{xy}(v)$ be the number of shortest paths between x and y that contain v. Therefore

Definition IV.1.1.

$$bc(v) = \sum_{x,y} \frac{\sigma_{xy}(v)}{\sigma_{xy}}, \tag{IV.1.1}$$

for all distinct vertices x, and y.[19]

Note that $\sigma_{xy} \not\equiv \sigma_{yx}$. This definition can then be used for directed graphs. It can also be generalized to the components of disconnected graphs.



Example IV.1.2.

Figure IV.1.1: Here we have $bc(a) = bc(c) = 0$ and $bc(b) = 2$.

IV.1.2 Bounds on Betweenness Centrality

Lemma IV.1.3. *Consider a graph $G(V, E)$ with $|V(G)| = n$ and $|E(G)| = m$. Then*

$$\min\{bc(v)|v \in V(G)\} \geq 0, \tag{IV.1.2}$$

$$\max\{bc(v)|v \in V(G)\} \leq (n-1)(n-2). \tag{IV.1.3}$$

The lower bound occurs for vertices that do not lie on any optimal paths. This happens with vertices that are pendants or lie on cycles with chords that bypass it: in complete graphs, etc. The upper bound only occurs if the vertex to be considered is the central vertex of a star graph: $K_{1,r}$.

Remark IV.1.4. If the graph G is a tree, then all vertices will have integer betweenness centrality. Note that there are no cycles so an optimal path either always contains a given vertex, or never does: contribution of 0 or 1.

IV.2 Betweenness Centrality for Various Graphs

Because betweenness centrality is a measurement of a particular vertex, we choose to examine graphs with high vertex transitivity (few group orbits). This reduces the total number of derivations to completely analyze a given graph.

IV.2.1 Cycle Powers

Lemma IV.2.1. *Consider a cycle power graph C_n^m.*

$$diam\,(C_n^m) = \left\lceil \frac{n-1}{2m} \right\rceil . \tag{IV.2.1}$$

Example: with C_{19}^3, every vertex is reachable from any spot in at most 3 steps. For C_{20}^3, going across the circle requires a minimum of 4 steps.

Proof. This can be seen by picking one vertex to look at: v_i. The furthest vertex from v_i is the vertex halfway around the circle: once you move further clockwise or counterclockwise, you could simply choose to approach from the counterclockwise or clockwise direction respectively. Now we look at the odd or even cases. If n is odd, the distance along the circle would be $\frac{n-1}{2}$. The minimum number of steps is then $\left\lceil \frac{(n-1)/2}{m} \right\rceil$. If n is even, the distance along the circle would be $\frac{n}{2}$. The minimum number of steps is then $\left\lceil \frac{n/2}{m} \right\rceil = \left\lceil \frac{\lceil (n-1)/2 \rceil}{m} \right\rceil = \left\lceil \frac{(n-1)/2}{m} \right\rceil$. ■

Theorem IV.2.2. *Consider a cycle power graph C_n^m and a vertex $v \in C_n^m$. Let $d = diam\,(C_n^m) = \left\lceil \frac{n-1}{2m} \right\rceil$. Assume $n > 2m+1$ otherwise we have a complete graph. This also means $d \geq 2$. Then*

$$bc(v) = (d-1)\left(2\left\lceil \frac{n-1}{2} \right\rceil - dm\right). \tag{IV.2.2}$$

Proof. Consider C_n^m. Let $d = diam\,(C_n^m) = \left\lceil \frac{n-1}{2m} \right\rceil$ and pick $r \equiv -\left\lceil \frac{n-1}{2} \right\rceil \mod m$ such that $m > r \geq 0$. Then $r = dm - \left\lceil \frac{n-1}{2} \right\rceil$.

d is the maximum number of steps between vertices so $d-1$ is the maximum number of stepping stones used.

Let P_l be the set of unique optimal paths of length l: $m+1 \leq l \leq d$. The upper bound is the largest distance possible in C_n^m and the lower bound is so that we ignore pairs of adjacent vertices. The number of steps in each optimal path is $\left\lceil \frac{l}{m} \right\rceil$. Finding $|P_l| = p_l$ is equivalent to finding the number of partitions of an integer with restriction on number of integers and values of said integers[20]. It turns out that an explicit formula for p_l is not necessary in our search for $bc(v)$.

Consider a vertex $v \in C_n^m$ and paths that use s number of stepping stones: $1 \leq s \leq d$. For every unique path of length l with $s = \left\lceil \frac{l}{m} \right\rceil - 1$ mid-steps, there exist s pair(s) of vertices such that the unique path passes through v: s term(s) of $\frac{1}{p_l}$. Since we can reverse the order of vertices, this doubles the term. Counting all unique paths of length l, the sum of betweenness centrality for v will get a total contribution of $2p_l \frac{s}{p_l} = 2s$. Summing over all values of l gives:

$$\begin{aligned}
bc(v) &= \sum_{l=m+1}^{\lceil \frac{n-1}{2} \rceil} 2\left(\left\lceil \frac{l}{m} \right\rceil - 1 \right) \\
&= 2 \sum_{l=1}^{\lceil \frac{n-1}{2} \rceil - m} \left(\left\lceil \frac{l+m}{m} \right\rceil - 1 \right) \\
&= 2 \sum_{l=1}^{\lceil \frac{n-1}{2} \rceil - m} \left\lceil \frac{l}{m} \right\rceil \\
&= 2m \sum_{l=1}^{\lceil (\lceil \frac{n-1}{2} \rceil - m)/m \rceil} l - 2 \left\lceil \frac{\left\lceil \frac{n-1}{2} \right\rceil - m}{m} \right\rceil r \\
&= 2m \sum_{l=1}^{\lceil \frac{n-1}{2m} \rceil - 1} l - 2\left(\left\lceil \frac{n-1}{2m} \right\rceil - 1 \right) r \\
&= 2m \sum_{l=1}^{d-1} l - 2(d-1)\left(dm - \left\lceil \frac{n-1}{2} \right\rceil \right) \\
&= 2m \frac{d(d-1)}{2} + 2(d-1)\left(\left\lceil \frac{n-1}{2} \right\rceil - dm \right) \\
&= md(d-1) + (d-1)\left(2\left\lceil \frac{n-1}{2} \right\rceil - 2dm \right) \\
&= (d-1)\left(2\left\lceil \frac{n-1}{2} \right\rceil - dm \right).
\end{aligned}$$

There is one key point we glossed over. For the paths of the lengths with the largest number of steps, there are also possibilities of moving the opposite way around the cycle. The reason these are not separately considered is that we have already accounted for those paths. Moving l around one way is the same as $n-l$ around the other way. In the original counting method, we looked at s pairs of vertices for p_l paths each having a contribution of $\frac{s}{p_l}$: total term of s. Instead, we actually have $p_l + p_{n-l}$ paths and s pairs of vertices that contribute terms of $\frac{s}{p_l + p_{n-l}}$: still a total term of s. So we arrived at the same answer. ■

IV.2.2 Subdivided Star Graphs

Theorem IV.2.3. *Consider a subdivided star graph with arms of lengths $s_1, \ldots, s_n$ and vertices labeled $v_{l,k}$: k^{th} vertex from $v_{0,0}$ on l^{th} arm. $v_{0,0}$ is the center vertex. Then*

$$bc(v_{0,0}) = 2\sum_{j=2}^{n} s_j \sum_{i=1}^{j-1} s_i,$$
$$bc(v_{l,k}) = 2(s_l - k)\left(\sum_{i\neq l} s_i + k\right). \tag{IV.2.3}$$

Proof. The central vertex will lie on an optimal paths if and only if the path is between vertices on different spokes. Thus we sum the number of pairs of vertices between spokes. Vertices on an arm will lie on an optimal path if and only if the path is between a vertex further along the same arm ($s_l - k$ vertices) and a vertex closer to $v_{0,0}$ or on a different arm ($\sum_{i\neq l} s_i + k$ vertices). Simply using multiplication to find the number of ways to choose these pairs gives us our result. Doubling is to account for moving in either direction. ■

Note that $bc(v_{l,s_l}) = 0$. These are the pendant vertices.

IV.2.3 Subdivided Triangle Star Graphs

Theorem IV.2.4. *Consider a subdivided triangle star graph with arms of lengths $s_0 \leq s_1 \leq s_2$: a C_3 with arms appended to each vertex of the cycle. Label vertices $v_{l,k}$: k^{th} vertex from $v_{l,0}$ on l^{th} arm. $k = 0$ indicates a vertex on the cycle. By symmetries of C_3, the order in which we append the arms is irrelevant. Then*

$$bc(v_{l,k}) = 2(s_l - k)\left(k + 2 + \sum_{i\neq l} s_i\right). \tag{IV.2.4}$$

Proof. $v_{l,k}$ will lie on an optimal path if and only if the path is between a vertex on the l^{th} arm further from the C_3 ($s_l - k$ vertices) and a vertex closer to the C_3 or on a different spoke ($k + 2 + \sum_{i\neq l} s_i$ vertices). Simply using multiplication to find the number of ways to choose these pairs gives us our result. Doubling is to account for moving in either direction. ■

IV.2.4 Ladders

A very useful reduction from double sum to single sum is given in Lemma IV.2.5.

Lemma IV.2.5.

$$\sum_{j=1}^{k}\sum_{i=k}^{n} \frac{1}{i-j+1} = \sum_{i=n-k+1}^{n} \frac{n-k+1}{i} + \sum_{i=k}^{n} \frac{k}{i} - 1. \tag{IV.2.5}$$

Proof. We will proceed with proof by induction. Let $S(k;n)$, $1 \leq k \leq n$, be the statement

$$\sum_{j=1}^{k}\sum_{i=k}^{n}\frac{1}{i-j+1} = \sum_{i=n-k+1}^{n}\frac{n-k+1}{i} + \sum_{i=k}^{n}\frac{k}{i} - 1.$$

We note that the statement $S(1;1)$ is true by the following.

$$\begin{aligned}
\sum_{j=1}^{k}\sum_{i=k}^{n}\frac{1}{i-j+1} &= \sum_{j=1}^{1}\sum_{i=1}^{1}\frac{1}{i-j+1} \\
&= \frac{1}{1-1+1} \\
&= 1 \\
&= \frac{1-1+1}{1-1+1} + \frac{1}{1} - 1 \\
&= \sum_{i=1-1+1}^{1}\frac{1-1+1}{i} + \sum_{i=1}^{1}\frac{1}{i} - 1 \\
&= \sum_{i=n-k+1}^{n}\frac{n-k+1}{i} + \sum_{i=k}^{n}\frac{k}{i} - 1.
\end{aligned}$$

Thus the base case is proven.

Assume $S(k;n)$ is true. To prove: $S(k;n+1)$ is true. Let $1 \leq k \leq n+1$.

$$\begin{aligned}
\sum_{j=1}^{k}\sum_{i=k}^{n+1}\frac{1}{i-j+1} &= \sum_{j=1}^{k}\left[\sum_{i=k}^{n}\frac{1}{i-j+1} + \frac{1}{n+1-j+1}\right] \\
&= \sum_{j=1}^{k}\sum_{i=k}^{n}\frac{1}{i-j+1} + \sum_{j=1}^{k}\frac{1}{n-j+2} \\
&= \sum_{i=n-k+1}^{n}\frac{n-k+1}{i} + \sum_{i=k}^{n}\frac{k}{i} - 1 + \sum_{j=1}^{k}\frac{1}{n-j+2} \\
&= \sum_{i=n-k+1}^{n}\frac{n-k+1}{i} + \sum_{i=k}^{n+1}\frac{k}{i} - \frac{k}{n+1} - 1 + \sum_{j=n-k+2}^{n+1}\frac{1}{j} \\
&= \sum_{i=n-k+2}^{n+1}\frac{n-k+1}{i} + \frac{n-k+1}{n-k+1} - \frac{n-k+1}{n+1} + \sum_{i=k}^{n+1}\frac{k}{i} - \frac{k}{n+1} - 1 + \sum_{j=n-k+2}^{n+1}\frac{1}{j} \\
&= \sum_{i=n-k+2}^{n+1}\frac{n-k+2}{i} + \sum_{i=k}^{n+1}\frac{k}{i} - \sum_{i=n-k+2}^{n+1}\frac{1}{i} + \frac{n-k+1}{n-k+1} - \frac{n-k+1}{n+1} - \frac{k}{n+1} - 1 + \sum_{j=n-k+2}^{n+1}\frac{1}{j} \\
&= \sum_{i=n-k+2}^{n+1}\frac{n-k+2}{i} + \sum_{i=k}^{n+1}\frac{k}{i} + \sum_{j=n-k+2}^{n+1}\frac{1}{j} - \sum_{i=n-k+2}^{n+1}\frac{1}{i} + 1 - 1 - \frac{n+1}{n+1} \\
&= \sum_{i=(n+1)-k+1}^{n+1}\frac{(n+1)-k+1}{i} + \sum_{i=k}^{n+1}\frac{k}{i} - 1.
\end{aligned}$$

Thus $S(k;n+1)$ is true. By induction, the Lemma holds for all $n \in \mathbb{N}$. ■

Lemma IV.2.6. *Let L_n be a ladder graph on $2n$ vertices. Label the vertices so that the "top" vertices read $v_1, v_2, \ldots, v_n$ and the "bottom" vertices are $v_{1'}, v_{2'}, \ldots, v_{n'}$. See Figure IV.2.1. Then,*

$$bc(v_{k'}) = bc(v_k) = 2\left[2k(n-k+1) - 2(n+1) + \sum_{i=n-k+1}^{n} \frac{n-k+1}{i} + \sum_{i=k}^{n} \frac{k}{i}\right].$$

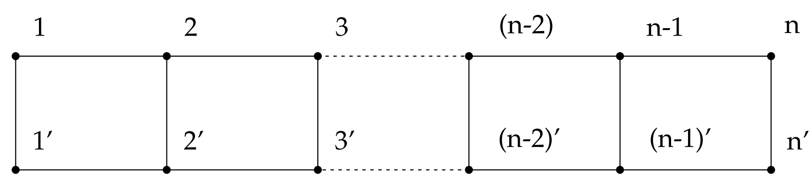


Figure IV.2.1: L_n. Note there are $2n$ vertices.

Proof. In summing the betweenness centrality, we will only consider paths moving from left to right and then double the final sum. Consider $v_k \in L_n$. First, the betweenness centrality from paths beginning on the top side of the ladder. There are $(k-1)$ vertices to the left of v_k and $(n-k)$ to the right. Picking one vertex from each side, the optimal path will be a straight path that passes through v_k: a contribution of 1. From picking any pair, we get the term: $(k-1)(n-k)$.

Next we look at optimal paths that begin on top and end on bottom. The path can start at any vertex v_j, $1 \leq j \leq k-1$. The ending vertex must be $v_{i'}$, $k \leq i \leq n$, otherwise the optimal path could not pass through v_k. The optimal path can drop from the top to bottom at any point of moving along the ladder. Thus there are $i-j+1$ distinct paths. However, we want the path to go through v_k. This means the drop must occur after passing v_k: $i-k+1$ total options. The proportion of optimal paths that go through v_k is then $\frac{i-k+1}{i-j+1}$. Summing over all possibilities gives $\sum_{j=1}^{k-1} \sum_{i=k}^{n} \frac{i-k+1}{i-j+1}$.

Similarly we can consider moving from bottom to top and obtain the term: $\sum_{j=1}^{k} \sum_{i=k+1}^{n} \frac{k-j+1}{i-j+1}$. Then summing the three terms and simplifying using Lemma IV.2.5 yields the final result of Lemma IV.2.6. ■

IV.2.5 Pendant Ladders

Lemma IV.2.7. *Let L_n be a ladder graph. Append another vertex, which will be labeled vertex v_0, to vertex v_1. Let U_n describe this graph. See Figure IV.2.2.*

$$
\begin{aligned}
bc(v_0) =& 0, \\
bc(v_k) =& 2\left[2k(n-k) - 1 + \sum_{i=n-k+1}^{n} \frac{n-k+1}{i} + \sum_{i=k}^{n} \frac{1}{i}\right], \\
bc(v_{k'}) =& 2\left[2(k-1)(n-k) - 3 + \sum_{i=n-k+1}^{n} \frac{n-k+1}{i} + 2\sum_{i=k}^{n} \frac{k}{i}\right].
\end{aligned}
$$

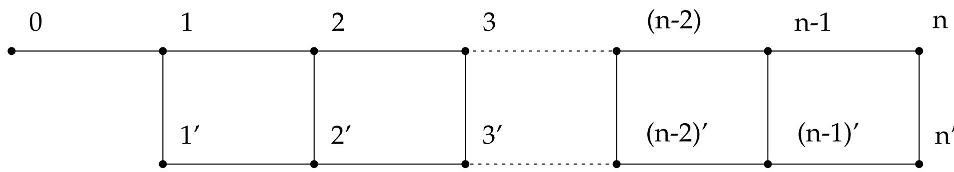


Figure IV.2.2: U_n. Note there are $2n+1$ vertices.

Proof. The expressions for $bc(v_k)$ and $bc(v_{k'})$ use Lemma IV.2.6 as the "base".
Note that by adding v_0, we have not changed any optimal paths between existing vertices: it is trivially seen that no optimal path would use v_0 as a midpoint. This means that $bc(v_0) = 0$. The only difference from Lemma IV.2.6 is that we have contributions from the optimal paths that include v_0 as an end point. For the betweenness centrality of a top vertex, v_k, every optimal path from v_0 to v_l, $k+1 \leq l \leq n$ will go through v_k. Hence we have an extra contribution of $(n-k)$ for top vertices. There is also the contribution for optimal paths to a bottom vertex. The fraction is explained in Lemma IV.2.6, though there are slightly restricted options since the first step from v_0 MUST be to v_1. The contribution for the betweenness centrality of bottom vertices is also very similar and only differs by some slight restrictions.
We also make use of Lemma IV.2.5 to simplify our expressions. ■

IV.3 Unique Betweenness Centrality

Definition IV.3.1. Let a graph $G(V,E)$ be said to have *unique betweenness centrality* if for all $v_i, v_j \in V(G)$, we have $bc(v_i) = bc(v_j)$ implies $i = j$. i.e. the betweenness centrality function is injective.

IV.3.1 Necessary Conditions

Theorem IV.3.2. *If a graph, G, has unique betweenness centrality, then $Aut(G) = \{id\}$.*

Proof. Suppose a graph, G, has unique betweenness centrality. Automorphisms preserve edge connectivity and thus preserve path connections. Therefore the betweenness centrality of a vertex is preserved across orbits of automorphisms. Thus two similar vertices have the same betweenness centrality. Then each vertex must be in its own orbit: $|Aut(G)| = 1^n = 1$. ■

Theorem IV.3.3. *If a graph, G, has unique betweenness centrality, then there is at most 1 pendant vertex.*

Proof. Suppose a graph, G, has 2 distinct pendant vertices u, v. Then it is clear that each pendant vertex does not lie on any optimal paths between other vertices. Thus $bc(u) = 0 = bc(v)$. Therefore G does not have unique betweenness centrality. ■

Corollary IV.3.4. *If a graph, G, of order $n \geq 2$, has unique betweenness centrality, then it is not a tree.*

IV.3.2 Infinite Family of Graphs with Unique Betweenness Centrality

Conjecture IV.3.5. *U_n has unique betweenness centrality for $n > 2$.*

Proof. Our goal is to prove that $bc(v_k) = bc(v_l)$ implies $k = l$. We also need to prove $bc(v_{k'}) = bc(v_{l'})$ implies $k = l$ and $bc(v_k)$ can never equal $bc(v_{l'})$. Clearly every vertex has nonzero betweenness centrality except for v_0. The cases $n = 3, 4, 5, 6, 7, 8$ are enumerated in Table VIII.2.1. Thus we may consider now that $n \geq 9$.

For $1 \leq l \leq \lceil \frac{n}{2} \rceil - 1$, $bc(v_{n-l+1}) < bc(v_l) < bc(v_{n-l})$.

$$\begin{aligned}
bc(v_{n-l}) &= 2\left[2(n-l)(n-(n-l)) - 1 + \sum_{i=n-(n-l)+1}^{n} \frac{n-(n-l)+1}{i} + \sum_{i=(n-l)}^{n} \frac{1}{i}\right] \\
&= 2\left[2l(n-l) - 1 + \sum_{i=l+1}^{n} \frac{l+1}{i} + \sum_{i=n-l}^{n} \frac{1}{i}\right] \\
&= 2\left[2l(n-l) - 1 + \sum_{i=l}^{n} \frac{l+1}{i} - \frac{l+1}{l} + \sum_{i=n-l+1}^{n} \frac{1}{i} + \frac{1}{n-l}\right] \\
&= 2\left[2l(n-l) - 1 + \sum_{i=l}^{n} \frac{1}{i} + \sum_{i=l}^{n} \frac{l}{i} + \sum_{i=n-l+1}^{n} \frac{n-l+1}{i} - \sum_{i=n-l+1}^{n} \frac{n-l}{i} + \frac{1}{n-l} - \frac{l+1}{l}\right] \\
&= 2\left[2l(n-l) - 1 + \sum_{i=n-l+1}^{n} \frac{n-l+1}{i} + \sum_{i=l}^{n} \frac{1}{i} + \sum_{i=l}^{n} \frac{l}{i} - \sum_{i=n-l+1}^{n} \frac{n-l}{i} + \frac{1}{n-l} - \frac{l+1}{l}\right] \\
&= bc(v_l) + 2\left[\sum_{i=l}^{n} \frac{l}{i} - \sum_{i=n-l+1}^{n} \frac{n-l}{i} + \frac{1}{n-l} - \frac{l+1}{l}\right] \\
&> bc(v_l) \\
&= 2\left[2l(n-l) - 1 + \sum_{i=n-l+1}^{n} \frac{n-l+1}{i} + \sum_{i=l}^{n} \frac{1}{i}\right] \\
&= 2\left[2l(n-l) - 1 + \sum_{i=n-l+1}^{n} \frac{n-l+1}{i} + \sum_{i=l}^{n} \frac{1}{i}\right] - bc(v_{n-l+1}) + bc(v_{n-l+1}) \\
&= bc(v_{n-l+1}) + 2\Bigg[2l(n-l) - 1 + \sum_{i=n-l+1}^{n} \frac{n-l+1}{i} + \sum_{i=l}^{n} \frac{1}{i} \\
&\qquad - \left(2(n-l+1)(n-(n-l+1)) - 1 + \sum_{i=n-(n-l+1)+1}^{n} \frac{n-(n-l+1)+1}{i} + \sum_{i=n-l+1}^{n} \frac{1}{i}\right)\Bigg] \\
&= bc(v_{n-l+1}) + 2\left[2l(n-l) + \sum_{i=n-l+1}^{n} \frac{n-l+1}{i} + \sum_{i=l}^{n} \frac{1}{i} - 2(n-l+1)(l-1) - \sum_{i=l}^{n} \frac{l}{i} - \sum_{i=n-l+1}^{n} \frac{1}{i}\right] \\
&= bc(v_{n-l+1}) + 2\left[2ln - 2l^2 + \sum_{i=n-l+1}^{n} \frac{n-l}{i} + \sum_{i=l}^{n} \frac{1-l}{i} - 2nl + 2l^2 - 2l + 2n - 2l + 2\right] \\
&= bc(v_{n-l+1}) + 2\left[\sum_{i=n-l+1}^{n} \frac{n-l}{i} - \sum_{i=l}^{n} \frac{l-1}{i} + 2n - 4l + 2\right] \\
&> bc(v_{n-l+1}).
\end{aligned}$$

This uses Lemmas VIII.2.2 and VIII.2.5 in the bounding lines. Also note that if n is even

$$\begin{aligned}
bc(v_{\frac{n}{2}}) - bc(v_{\frac{n}{2}+1}) &= 2\left[2\frac{n}{2}\left(n-\frac{n}{2}\right) - 1 + \sum_{i=n-\frac{n}{2}+1}^{n} \frac{n-\frac{n}{2}+1}{i} + \sum_{i=\frac{n}{2}}^{n} \frac{1}{i}\right] \\
&\quad - 2\left[2\left(\frac{n}{2}+1\right)\left(n-\frac{n}{2}-1\right) - 1 + \sum_{i=n-\frac{n}{2}-1+1}^{n} \frac{n-\frac{n}{2}-1+1}{i} + \sum_{i=\frac{n}{2}+1}^{n} \frac{1}{i}\right] \\
&= 2\Bigg[2\frac{n}{2}\left(n-\frac{n}{2}\right) - 1 + \sum_{i=n-\frac{n}{2}+1}^{n} \frac{n-\frac{n}{2}+1}{i} + \sum_{i=\frac{n}{2}}^{n} \frac{1}{i} \\
&\quad - 2\left(\frac{n}{2}+1\right)\left(n-\frac{n}{2}-1\right) + 1 - \sum_{i=n-\frac{n}{2}-1+1}^{n} \frac{n-\frac{n}{2}-1+1}{i} - \sum_{i=\frac{n}{2}+1}^{n} \frac{1}{i}\Bigg] \\
&= 2\left[n\left(\frac{n}{2}\right) + \sum_{i=\frac{n}{2}+1}^{n} \frac{\frac{n}{2}+1}{i} + \sum_{i=\frac{n}{2}}^{n} \frac{1}{i} - (n+2)\left(\frac{n}{2}-1\right) - \sum_{i=\frac{n}{2}}^{n} \frac{\frac{n}{2}}{i} - \sum_{i=\frac{n}{2}+1}^{n} \frac{1}{i}\right] \\
&= n^2 + \sum_{i=\frac{n}{2}+1}^{n} \frac{n+2}{i} + \sum_{i=\frac{n}{2}}^{n} \frac{2}{i} - (n+2)(n-2) - \sum_{i=\frac{n}{2}}^{n} \frac{n}{i} - \sum_{i=\frac{n}{2}+1}^{n} \frac{2}{i} \\
&= n^2 - n^2 + 4 + \sum_{i=\frac{n}{2}+1}^{n} \frac{n+2}{i} - \sum_{i=\frac{n}{2}}^{n} \frac{n}{i} + \sum_{i=\frac{n}{2}}^{n} \frac{2}{i} - \sum_{i=\frac{n}{2}+1}^{n} \frac{2}{i} \\
&= 4 + \sum_{i=\frac{n}{2}+1}^{n} \frac{2}{i} + \sum_{i=\frac{n}{2}}^{n} \frac{n}{i} - \frac{n}{\frac{n}{2}} - \sum_{i=\frac{n}{2}}^{n} \frac{n}{i} + \sum_{i=\frac{n}{2}}^{n} \frac{2}{i} - \sum_{i=\frac{n}{2}}^{n} \frac{2}{i} + \frac{2}{\frac{n}{2}} \\
&= 4 + \sum_{i=\frac{n}{2}+1}^{n} \frac{2}{i} - 2 + \frac{4}{n} \\
&= 2 + \frac{4}{n} + \sum_{i=\frac{n}{2}+1}^{n} \frac{2}{i} \\
&> 0.
\end{aligned}$$

Thus we can state that $0 = bc(v_0) < bc(v_n) < bc(v_1) < bc(v_{n-1}) < bc(v_2) < bc(v_{n-2}) < \cdots < bc(v_{\lceil \frac{n}{2} \rceil})$. Thus if $bc(v_k) = bc(v_l)$, then $k = l$ since the vertices are completed ordered by betweenness centrality.

For $1 \leq l \leq \lceil \frac{n}{2} \rceil - 1$, $bc(v_{(n-l+1)'}) < bc(v'_l) < bc(v_{(n-l)'})$.

$$
\begin{aligned}
bc(v_{(n-l)'}) &= 2\left[2((n-l)-1)(n-(n-l))-3+\sum_{i=n-(n-l)+1}^{n}\frac{n-(n-l)+1}{i}+2\sum_{i=n-l}^{n}\frac{n-l}{i}\right]\\
&= 2\left[2(n-l-1)l-3+\sum_{i=l+1}^{n}\frac{l+1}{i}+2\sum_{i=n-l}^{n}\frac{n-l}{i}\right]-bc(v_{l'})+bc(v_{l'})\\
&= bc(v_{l'})+2\Bigg[2(n-l-1)l-3+\sum_{i=l+1}^{n}\frac{l+1}{i}+2\sum_{i=n-l}^{n}\frac{n-l}{i}\\
&\qquad -\left(2(l-1)(n-l)-3+\sum_{i=n-l+1}^{n}\frac{n-l+1}{i}+2\sum_{i=l}^{n}\frac{l}{i}\right)\Bigg]\\
&= bc(v_{l'})+2\Bigg[2(n-l-1)l+\sum_{i=l+1}^{n}\frac{l+1}{i}+2\sum_{i=n-l}^{n}\frac{n-l}{i}\\
&\qquad -2ln+2l^2+2n-2l-\sum_{i=n-l+1}^{n}\frac{n-l+1}{i}-2\sum_{i=l}^{n}\frac{l}{i}\Bigg]\\
&= bc(v_{l'})+2\left[2n-4l+\sum_{i=l+1}^{n}\frac{l+1}{i}-2\sum_{i=l}^{n}\frac{l}{i}+2\sum_{i=n-l}^{n}\frac{n-l}{i}-\sum_{i=n-l+1}^{n}\frac{n-l+1}{i}\right]\\
&= bc(v_{l'})+2\Bigg[2n-4l+\sum_{i=l}^{n}\frac{l}{i}+\sum_{i=l}^{n}\frac{1}{i}-\frac{l+1}{l}-2\sum_{i=l}^{n}\frac{l}{i}\\
&\qquad +2\sum_{i=n-l}^{n}\frac{n-l}{i}-\sum_{i=n-l}^{n}\frac{n-l}{i}-\sum_{i=n-l}^{n}\frac{1}{i}+\frac{n-l+1}{n-l}\Bigg]\\
&= bc(v_{l'})+2\left[2n-4l-\sum_{i=l}^{n}\frac{l-1}{i}-1-\frac{1}{l}+\sum_{i=n-l}^{n}\frac{n-l-1}{i}+1+\frac{1}{n-l}\right]\\
&= bc(v_{l'})+2\left[2n-4l-\sum_{i=l}^{n}\frac{l-1}{i}+\sum_{i=n-l}^{n}\frac{n-l-1}{i}+\frac{1}{n-l}-\frac{1}{l}\right]\\
&> bc(v_{l'})\\
&= 2\left[2(l-1)(n-l)-3+\sum_{i=n-l+1}^{n}\frac{n-l+1}{i}+2\sum_{i=l}^{n}\frac{l}{i}\right]-bc(v_{(n-l+1)'})+bc(v_{(n-l+1)'})\\
&= bc(v_{(n-l+1)'})+2\Bigg[2(l-1)(n-l)-3+\sum_{i=n-l+1}^{n}\frac{n-l+1}{i}+2\sum_{i=l}^{n}\frac{l}{i}\\
&\qquad -\Bigg(2(n-l+1-1)(n-(n-l+1))-3\\
&\qquad\qquad +\sum_{i=n-(n-l+1)+1}^{n}\frac{n-(n-l+1)+1}{i}+2\sum_{i=n-l+1}^{n}\frac{n-l+1}{i}\Bigg)\Bigg]\\
&= bc(v_{(n-l+1)'})+2\left[\sum_{i=l}^{n}\frac{l}{i}-\sum_{i=n-l+1}^{n}\frac{n-l+1}{i}\right]\\
&> bc(v_{(n-l+1)'}).
\end{aligned}
$$

This uses Lemmas VIII.2.4 and VIII.2.3 in the bounding lines. Again, if n is even,

$$\begin{aligned}
bc(v_{(\frac{n}{2})'}) - bc(v_{(\frac{n}{2}+1)'}) &= 2\left[2\left(\frac{n}{2}-1\right)\left(n-\frac{n}{2}\right) - 3 + \sum_{i=n-\frac{n}{2}+1}^{n} \frac{n-\frac{n}{2}+1}{i} + 2\sum_{i=\frac{n}{2}}^{n} \frac{\frac{n}{2}}{i}\right] \\
&\quad - 2\left[2\left(\frac{n}{2}+1-1\right)\left(n-\frac{n}{2}-1\right) - 3 + \sum_{i=n-\frac{n}{2}-1+1}^{n} \frac{n-\frac{n}{2}-1+1}{i} + 2\sum_{i=\frac{n}{2}+1}^{n} \frac{\frac{n}{2}+1}{i}\right] \\
&= 2\left[\left(\frac{n}{2}-1\right)n - 3 + \sum_{i=\frac{n}{2}+1}^{n} \frac{\frac{n}{2}+1}{i} + \sum_{i=\frac{n}{2}}^{n} \frac{n}{i}\right] \\
&\quad - 2\left[n\left(\frac{n}{2}-1\right) - 3 + \sum_{i=\frac{n}{2}}^{n} \frac{\frac{n}{2}}{i} + \sum_{i=\frac{n}{2}+1}^{n} \frac{n+2}{i}\right] \\
&= \sum_{i=\frac{n}{2}+1}^{n} \frac{n+2}{i} + \sum_{i=\frac{n}{2}}^{n} \frac{2n}{i} - \sum_{i=\frac{n}{2}}^{n} \frac{n}{i} - \sum_{i=\frac{n}{2}+1}^{n} \frac{2n+4}{i} \\
&= \sum_{i=\frac{n}{2}}^{n} \frac{n}{i} - \sum_{i=\frac{n}{2}+1}^{n} \frac{n+2}{i} \\
&= \sum_{i=\frac{n}{2}}^{n} \frac{n}{i} - \sum_{i=\frac{n}{2}+1}^{n} \frac{n}{i} - \sum_{i=\frac{n}{2}+1}^{n} \frac{2}{i} \\
&= \frac{n}{\frac{n}{2}} - \sum_{i=\frac{n}{2}+1}^{n} \frac{2}{i} \\
&= 2\left(1 - \sum_{i=\frac{n}{2}+1}^{n} \frac{1}{i}\right) \\
&= 2\left(1 + \frac{2}{n} - \sum_{i=\frac{n}{2}}^{n} \frac{1}{i}\right) \\
&> 2\left(\ln(2) + \frac{3n+15}{2n(n+3)} - \sum_{i=\frac{n}{2}}^{n} \frac{1}{i}\right) \\
&> 0.
\end{aligned}$$

by Lemmas VIII.2.6 and VIII.2.7. Thus we can state that $bc(v_0) = 0 < bc(v'_n) < bc(v'_1) < bc(v_{(n-1)'}) < bc(v'_2) < bc(v_{(n-2)'}) < \cdots < bc(v_{(\frac{n}{2})'})$. Thus if $bc(v'_k) = bc(v'_l)$, then $k = l$ since the vertices are completely ordered by betweenness centrality.

$$
\begin{aligned}
bc(v_{\lceil \frac{n}{2} \rceil}) &= 2\left[2\left\lceil \frac{n}{2} \right\rceil \left(n - \left\lceil \frac{n}{2} \right\rceil\right) - 1 + \sum_{i=n-\lceil \frac{n}{2} \rceil+1}^{n} \frac{n - \lceil \frac{n}{2} \rceil + 1}{i} + \sum_{i=\lceil \frac{n}{2} \rceil}^{n} \frac{1}{i}\right] \\
&= 2\left[2\left(\left\lceil \frac{n}{2} \right\rceil - 1\right)\left(n - \left\lceil \frac{n}{2} \right\rceil\right) + 2n - 2\left\lceil \frac{n}{2} \right\rceil - 3 + 2 + \sum_{i=n-\lceil \frac{n}{2} \rceil+1}^{n} \frac{n - \lceil \frac{n}{2} \rceil + 1}{i} + \sum_{i=\lceil \frac{n}{2} \rceil}^{n} \frac{1}{i}\right] \\
&= 2\Bigg[2\left(\left\lceil \frac{n}{2} \right\rceil - 1\right)\left(n - \left\lceil \frac{n}{2} \right\rceil\right) - 3 + \sum_{i=n-\lceil \frac{n}{2} \rceil+1}^{n} \frac{n - \lceil \frac{n}{2} \rceil + 1}{i} \\
&\qquad + 2\sum_{i=\lceil \frac{n}{2} \rceil}^{n} \frac{\lceil \frac{n}{2} \rceil}{i} - 2\sum_{i=\lceil \frac{n}{2} \rceil}^{n} \frac{\lceil \frac{n}{2} \rceil}{i} + \sum_{i=\lceil \frac{n}{2} \rceil}^{n} \frac{1}{i} + 2n - 2\left\lceil \frac{n}{2} \right\rceil + 2\Bigg] \\
&= bc\left(v_{\lceil \frac{n}{2} \rceil'}\right) + 2\left[-2\sum_{i=\lceil \frac{n}{2} \rceil}^{n} \frac{\lceil \frac{n}{2} \rceil}{i} + \sum_{i=\lceil \frac{n}{2} \rceil}^{n} \frac{1}{i} + 2n - 2\left\lceil \frac{n}{2} \right\rceil + 2\right] \\
&> bc\left(v_{\lceil \frac{n}{2} \rceil'}\right).
\end{aligned}
$$

The last step uses Lemma VIII.2.9. Thus we have $bc(v_{k'}) < bc(v_{\lceil \frac{n}{2} \rceil'}) < bc(v_{\lceil \frac{n}{2} \rceil})$ for all k.

■

We have shown that betweenness centrality is injective when considering only the top or bottom vertices separately. Only a few cases of distinct values have been shown for inter-row consideration. Initially we though of attempting to order every vertex like the top and bottom rows had been. Problems arose in much higher cases of n when suspected orderings would change, seemingly randomly.

V. Future Research

Though I have tried hard to complete as much as I could, there is always more that could be done. This section is a list of possible avenues for research or fiddling around.

Problem V.0.6. *Let T be a tree. Determine $E_{glob}(T)$, $E^{w}_{glob}(T)$, and $E_{Ratio}(T)$. One might have to begin by restricting to specific trees.*

Problem V.0.7. *Find global efficiency for some families of directed graphs: tournaments, directed cycles, etc.*

Problem V.0.8. *Find the asymptotic nature of $E_{glob}(P_n^m)$ for various values of $\frac{m}{n}$. Or other values of m such as $\sqrt{n}$.*

Problem V.0.9. *The minimum value for the global efficiency of $P_n \times P_m$ does not occur exactly at a square grid. Why is this the case and what $n - m$ ratio produces the minimum value? Also, what is the asymptotic value of this worst grid as $n \to \infty$?*

Problem V.0.10. *Consider the efficiency of a graph PER EDGE. This would be useful as usually roads have a cost attributed to them. What size/graph maximizes this value?*

Problem V.0.11. *Determine the maximum value of $E_{loc}(G) - CC(G)$ over all graphs G. I suspect that the difference has to be less than or equal to $\frac{1}{2}$.*

Problem V.0.12. *Determine the betweenness centrality for C_n star graphs. Trouble in defining the graph begins to appear as the order of appending arms matters: less automorphisms.*

Problem V.0.13. *Complete the proof of unique betweenness centrality for U_n: hooked ladder graphs. Or find other families of graphs with this property.*

Problem V.0.14. *One could find the average difference for the graphs discussed in this thesis by going through the same processes and taking the reciprocal of each vertex-to-vertex efficiency term.*

VI. Conclusion

VI.1 Efficiency

In this thesis, we are concerned with several measures of connectivity of graphs: global efficiency, local efficiency, clustering coefficient and betweenness centrality.

We determined the global efficiency for path power graphs P_n^m, cycle power graphs C_n^m, complete multipartite graphs $K_{m,n}$, star and subdivided star graphs, and the Cartesian Products $K_n \times P_m^t$, $K_n \times C_m^t$, $K_m \times K_n$, and $P_m \times P_n$. As a consequence, we also determined new results involving the Harary index for these families of graphs.

Just as Latora and Marchiori explored the global efficiency of the Boston Subway (MBTA)[8], we investigated the global efficiency of the Metropolitan Atlanta Rapid Transportation Authority (MARTA) Subway network. Motivated by the design of MARTA (see Figure II.9.1), we investigated the global efficiency of subdivided stars. We showed that networks of this type have a high level of efficiency. We applied these ideas to an analysis of the MARTA Subway system and show that their network is 82% as efficient as a network where there is a direct line connecting each pair of stations.

RCBI scientists conducted functional magnetic resonance imaging (fMRI) scans of 25 volunteers to find blood oxygen level-dependent (BOLD) correlations of various regions of the brain. We constructed graphs with edges based on correlation cutoffs and then partitioned the brain using efficiency. The partitions were found to be consistent with functionality and physical location within the brain. We also used these measurements to analyze the effects of a season of hard-contact football on University of Rochester athletes. Again, an outside source conducted BOLD pre and postseason fMRI scans of the players. We received matrices of the correlations in oxygen levels of various regions of the brain and modeled these as graphs. We were then able to measure the “efficiency” of each athlete. As was expected, the athletes who received the largest number of high-energy impacts during the season also experienced the largest drop in brain efficiency. For comparison, we calculated the measurements of a macaque brain using data from Honey et al[12].

It was stated by Latora and Marchiori [1] that “It can be shown that, when in a graph, most of its local subgraphs G_i are not sparse, then C [clustering coefficient] is a good approximation of E_{loc}. In summary, there are not two different types of analyses to be done for the global and local scales, just one with a very precise physical meaning: the efficiency in transporting information". However we provided an in-depth analysis of this statement, identifying graphs where the clustering coefficient and local efficiency are in fact non-negligibly different. We also identified certain graph families where the two quantities are the same. In this thesis, we compared and contrasted the two quantities, local efficiency and clustering coefficient. We included results of these local measurements for complete multipartite graphs $K_{n,m}$, cycle power graphs C_n^m, and Cartesian products $K_m \times K_n$ and $K_n \times C_m$.

VI.2 Betweenness Centrality

In this thesis, we examined betweenness centrality for vertices in C_n^m. By the symmetry of C_n^m, every vertex will have the same betweenness centrality. We also include results for subdivided star graphs and C_3 star graphs. We also describe the betweenness centrality for vertices in ladder graphs, $P_2 \times P_n$, and appended ladder graphs U_n: a $P_2 \times P_n$ with a pendant vertex attached to an "end". We conjectured that the infinite family of appended graphs has unique betweenness centrality. We were able to construct a partial proof but were forced to leave the completion as future research.

VII. Acknowledgments

We are grateful to two anonymous referees whose comments greatly improved this thesis. Research was partially supported by the National Science Foundation and the Department of Defense (NSF-REU #1062128), and the College of Science at RIT. The authors also thank the staff at MARTA for providing data pertaining to the subway system.

We are thankful to an anonymous referee, for their valuable comments and alerting us that the global efficiency measure is closely related to the Harary index. Research was partially supported by an National Science Foundation Research Experience for Undergraduates Site Award #1062128.

I am very grateful to Caitlin VerSchneider for being my partner during my first REU which contributed greatly to this thesis. And thank you to everyone in the 2012 RIT REU for showing me that math research is what I most enjoy doing. Thank you as well to the 2013 REU at RIT for their initial work on betweenness centrality.

Thank you to Dr. Narayan for being willing to work with and help me on my first undergraduate research and beyond. You have done so much to help me get published and find new projects to tackle. Thank you.

Thank you to my committee who have read so many versions and worked hard with me to "perfect" this thesis.

Many thanks to my fiancée, now wife, Kaitlin for being so supportive, specifically for taking care of me when I was too busy focusing on writing or deriving. I love you. Thank you very much to my family for always viewing me in a wonderful light; your pride is my pride. Thank you to all my friends who have been a wonderful distraction when my mind verged on the edge of exploding. I love you all.

VIII. Appendix

VIII.1 Efficiency

VIII.1.1 Sum Simplifications

Lemma VIII.1.1. *Let $n, m \in \mathbb{N}$ and pick $d = \left\lceil \frac{n}{m} \right\rceil$. Then*

$$\sum_{i=1}^{n} \frac{1}{\left\lceil \frac{i}{m} \right\rceil} = m(H_d - 1) + \frac{n}{d}. \tag{VIII.1.1}$$

Proof. Let $n, m \in \mathbb{N}$ and pick $d = \left\lceil \frac{n}{m} \right\rceil$ and $r \equiv -n \mod m$ such that $m > r \geq 0$. We can write r as $-n + \left\lceil \frac{n}{m} \right\rceil m$, or $r = dm - n$.

Consider the sum $\sum_{i=1}^{n} \frac{1}{\left\lceil \frac{i}{m} \right\rceil}$. The first m terms of the sum, $i = 1, \ldots, m$, will have a value of $\frac{1}{1}$. The second m terms will have a value of $\frac{1}{2}$ and so on up to terms of $\frac{1}{d}$. Thus the total is almost the same as mH_d. We did however over count the number of terms of $\frac{1}{d}$. The number of terms of $\frac{1}{d}$ should be $\equiv n \mod m$. Thus we need to subtract $\frac{r}{d} = m - \frac{n}{d}$. ■

Lemma VIII.1.2. *Let $n, m \in \mathbb{N}$ and pick $d = \left\lceil \frac{n}{m} \right\rceil$. Then*

$$\sum_{i=1}^{n} \frac{i}{\left\lceil \frac{i}{m} \right\rceil} = \frac{1}{2}\left[dm^2 - m(m-1)H_d - m + \frac{n(n+1)}{d}\right]. \tag{VIII.1.2}$$

Proof. Let $n, m \in \mathbb{N}$ and pick $d = \left\lceil \frac{n}{m} \right\rceil$ and $r \equiv -n \mod m$ such that $m > r \geq 0$. We can write r as $-n + \left\lceil \frac{n}{m} \right\rceil m$, or $r = dm - n$. This proof is similar in thought to the proof of Lemma VIII.1.1.

Consider the sum $\sum_{i=1}^{n} \frac{1}{\left\lceil \frac{i}{m} \right\rceil}$. The first m terms of the sum will have a denominator of 1 and the numerators will be $1, \ldots, m$. The second m terms will have a denominator of 2 and the numerators will be $m+1, \ldots, 2m$. In general the i^{th} group, $1 \leq i \leq d$, of m terms will have denominator i and numerators $m(i-1)+1, \ldots, mi$. Thus the total numerator for the i^{th} group is

$$\begin{aligned}
(m(i-1)+1) + (m(i-1)+2) + \ldots + (mi-1) + (mi) &= \sum_{k=1}^{mi} k - \sum_{k=1}^{m(i-1)} k \\
&= \frac{mi(mi+1)}{2} - \frac{m(i-1)(m(i-1)+1)}{2} \\
&= \frac{m^2i^2 + mi - m^2i^2 + m^2i - mi + m^2i - m^2 + m}{2} \\
&= m^2i - \frac{m(m-1)}{2}.
\end{aligned}$$

Our sum now is almost equal to

$$\sum_{i=1}^{d} \frac{m^2i - \frac{m(m-1)}{2}}{i} = dm^2 - \frac{m(m-1)}{2}H_d.$$

However, again we over counted the final terms with denominator d. Thus we need to subtract off the last r terms:

$$\begin{aligned}
\frac{1}{d}\sum_{k=md-r+1}^{md} k &= \frac{1}{d}\left[\sum_{k=1}^{md} k - \sum_{k=1}^{md-r} k\right] \\
&= \frac{md(md+1)-(md-r)(md-r+1)}{2d} \\
&= \frac{m^2d^2+md+mdr-r^2+r-m^2d^2+mdr-md}{2d} \\
&= r\left(m-\frac{r-1}{2d}\right) \\
&= (dm-n)\left(m-\frac{dm-n-1}{2d}\right) \\
&= dm^2-\frac{dm^2-mn-m}{2}-nm+\frac{dmn-n^2-n}{2d} \\
&= \frac{dm^2}{2}+\frac{m}{2}-\frac{n(n+1)}{2d}.
\end{aligned}$$

Subtracting this yields the final formula. ■

Because of the number of times the following combination of Lemmas VIII.1.1 and VIII.1.2 is used, we should write it as its own separate corollary.

Corollary VIII.1.3. *Let $n, m \in \mathbb{N}$ and pick $d = \left\lceil \frac{n-1}{m} \right\rceil$. Then*

$$\sum_{i=1}^{n-1} \frac{n-i}{\left\lceil \frac{i}{m} \right\rceil} = m\left[\left(n+\frac{m-1}{2}\right)H_d - n - \frac{dm}{2} + \frac{1}{2}\right] + \frac{n(n-1)}{2d}. \tag{VIII.1.3}$$

Lemma VIII.1.4. *Let $f : \mathbb{N}\times\mathbb{N} \to \mathbb{R}$ with the property that $f(i,j) = f(j,i)$ for all $i, j \in \mathbb{N}$. Let $l \in \mathbb{N}$. Then*

$$\sum_{i=1}^{l}\sum_{j=1}^{l} f(i,j) = 2\sum_{i=2}^{l}\sum_{j=1}^{i-1} f(i,j) + \sum_{i=1}^{l} f(i,i) = 2\sum_{i=1}^{l-1}\sum_{j=i+1}^{l} f(i,j) + \sum_{i=1}^{l} f(i,i). \tag{VIII.1.4}$$

Proof. The left hand sum can be thought of as a bunch of ordered pairs (i, j). Since $f(i, j) = f(j,i)$ we are able to count only those pairs which have $i > j$ (middle sums) or $j < i$ (right hand sums) and then double them. We also must then add in the sum for $i = j$. ■

VIII.1.2 Faster Efficiency Formulae

We begin by simplifying the formula for global efficiency of a path power in Theorem II.3.1.

Corollary VIII.1.5. *Consider a path power graph P_n^m. Pick $d = \left\lceil \frac{n-1}{m} \right\rceil$. Then*

$$E_{glob}(P_n^m) = \frac{m\left[(2n+m-1)\,H_d - 2n - dm + 1\right]}{n(n-1)} + \frac{1}{d}. \tag{VIII.1.5}$$

Proof. This formula follows immediately from Corollary VIII.1.3. ■

Just as for the path power, we can somewhat simplify the expressions in Theorem II.4.5.

Corollary VIII.1.6. *Consider a cycle power graph C_n^m. We have $n = 2k$ or $n = 2k+1$. Pick $d = \left\lceil \frac{k}{m} \right\rceil$. Then*

$$E_{glob}(C_n^m) = \begin{cases} \frac{2m}{2k-1}(2H_d - 1) + \frac{1}{d} & \text{if } n = 2k, \\ \frac{m}{k}(2H_d - 1) + \frac{1}{d} & \text{if } n = 2k+1. \end{cases} \tag{VIII.1.6}$$

Proof. A consequence of using Lemma VIII.1.1. ■

Corollary VIII.1.7. *Consider a star graph $S_{d,l}$. If $d = 2k+1$,*

$$E_{glob}^{w}\left(S_{d,l}\right) = \frac{2}{l\,(dl+1)}\left[(l+1)H_l - l + 2\sum_{\theta=1}^{k}\left(\frac{H_l}{\sqrt{2\left(1-\cos\left(\frac{2\pi}{d}\theta\right)\right)}} + \sum_{i=2}^{l}\sum_{j=1}^{i-1}\frac{2}{\sqrt{i^2+j^2-2ij\cos\left(\frac{2\pi}{d}\theta\right)}}\right)\right]. \tag{VIII.1.7}$$

If $d = 2k$,

$$E_{glob}^{w}\left(S_{d,l}\right) = \frac{2}{l\,(dl+1)}\Bigg[(2l+1)H_{2l} - (l+1)H_l - l + 2\sum_{\theta=1}^{k-1}\left(\frac{H_l}{\sqrt{2\left(1-\cos\left(\frac{2\pi}{d}\theta\right)\right)}} + \sum_{i=2}^{l}\sum_{j=1}^{i-1}\frac{2}{\sqrt{i^2+j^2-2ij\cos\left(\frac{2\pi}{d}\theta\right)}}\right)\Bigg]. \tag{VIII.1.8}$$

Proof. The triple sum of Theorem II.7.4 can be slightly reduced by noting that the denominator is symmetric with respect to i and j. Then we can use Lemma VIII.1.4. We can also use the symmetry of cos and break the formula into even (an angle of π is counted) and odd (an angle of π is not counted) cases. In reducing, Corollary II.7.8 proved to be useful. ■

Corollary VIII.1.8. *Consider a complete graph crossed with a path power, $K_r \times P_n^m$. Pick $d = \left\lceil \frac{n-1}{m} \right\rceil$. Then*

$$E_{glob}(K_r \times P_n^m) = \frac{2}{n(nr-1)}\Bigg[m\left(\left[r\left(n + \frac{3m-1}{2}\right) - m\right]H_d - r\left[2n + 2m + \frac{dm}{2} - 1\right] + n + 2m - \frac{1}{2}\right) + \frac{n(n-1)}{2d} + (r-1)\frac{n^2 - n + 4nm + 4m^2 - 2m}{2(d+1)}\Bigg] + \frac{r-1}{nr-1}.$$

Proof. We can simplify the expression in Theorem II.8.2.

$$\begin{aligned}
E_{glob}(K_r \times P_n^m) &= \frac{2}{n(nr-1)} \sum_{i=1}^{n-1} (n-i) \left(\frac{1}{\left\lceil \frac{i}{m} \right\rceil} + \frac{r-1}{\left\lceil \frac{i}{m} \right\rceil + 1} \right) + \frac{r-1}{nr-1} \\
&= \frac{2}{n(nr-1)} \left[\sum_{i=1}^{n-1} \frac{n-i}{\left\lceil \frac{i}{m} \right\rceil} + (r-1) \sum_{i=1}^{n-1} \frac{n-i}{\left\lceil \frac{i}{m} \right\rceil + 1} \right] + \frac{r-1}{nr-1} \\
&= \frac{2}{n(nr-1)} \left[\sum_{i=1}^{n-1} \frac{n-i}{\left\lceil \frac{i}{m} \right\rceil} + (r-1) \sum_{i=m+1}^{n+m-1} \frac{n-(i-m)}{\left\lceil \frac{i}{m} \right\rceil} \right] + \frac{r-1}{nr-1} \\
&= \frac{2}{n(nr-1)} \left[\sum_{i=1}^{n-1} \frac{n-i}{\left\lceil \frac{i}{m} \right\rceil} + (r-1) \sum_{i=1}^{n+m-1} \frac{n+m-i}{\left\lceil \frac{i}{m} \right\rceil} - (r-1) \sum_{i=1}^{m} \frac{n+m-i}{\left\lceil \frac{i}{m} \right\rceil} \right] + \frac{r-1}{nr-1}.
\end{aligned}$$

Pick $d = \left\lceil \frac{n-1}{m} \right\rceil$. Then $\left\lceil \frac{n+m-1}{m} \right\rceil = d+1$. Then implementing Corollary VIII.1.3 twice and noting the final sum always has denominator 1,

$$\begin{aligned}
E_{glob}(K_r \times P_n^m) &= \frac{2}{n(nr-1)} \left[\sum_{i=1}^{n-1} \frac{n-i}{\left\lceil \frac{i}{m} \right\rceil} + (r-1) \sum_{i=1}^{n+m-1} \frac{n+m-i}{\left\lceil \frac{i}{m} \right\rceil} - (r-1) \sum_{i=1}^{m} (n+m-i) \right] + \frac{r-1}{nr-1} \\
&= \frac{2}{n(nr-1)} \Bigg[m \left[\left(n + \frac{m-1}{2} \right) H_d - n - \frac{dm}{2} + \frac{1}{2} \right] + \frac{n(n-1)}{2d} \\
&\qquad + (r-1) \left(m \left[\left(n + \frac{3m-1}{2} \right) H_{d+1} - n - \frac{(d+3)m}{2} + \frac{1}{2} \right] + \frac{(n+m)(n+m-1)}{2(d+1)} \right) \\
&\qquad - (r-1) \left[(n+m)m - \frac{m(m+1)}{2} \right] \Bigg] + \frac{r-1}{nr-1} \\
&= \frac{2}{n(nr-1)} \Bigg[m \left(n + \frac{m-1}{2} \right) H_d - mn - \frac{dm^2}{2} + \frac{m}{2} + \frac{n(n-1)}{2d} - (r-1) \left[nm + \frac{m(m-1)}{2} \right] \\
&\qquad + (r-1) \Bigg(m \left[n + \frac{m-1}{2} \right] H_d + m^2 H_d + \frac{m}{d+1} \left[n + \frac{3m-1}{2} \right] - nm \\
&\qquad - \frac{dm^2}{2} - \frac{3m^2}{2} + \frac{m}{2} + \frac{n^2+2nm+m^2-n-m}{2(d+1)} \Bigg) \Bigg] + \frac{r-1}{nr-1} \\
&= \frac{2}{n(nr-1)} \Bigg[rm \left(n + \frac{m-1}{2} \right) H_d - (2r-1)mn - r\frac{dm^2}{2} + \frac{m}{2} + \frac{n(n-1)}{2d} + (r-1) \left[m - 2m^2 \right] \\
&\qquad + (r-1) \left(m^2 H_d + \frac{n^2-n+4nm+4m^2-2m}{2(d+1)} \right) \Bigg] + \frac{r-1}{nr-1} \\
&= \frac{2}{n(nr-1)} \Bigg[m \left(\left[r \left(n + \frac{3m-1}{2} \right) - m \right] H_d - r \left[2n + 2m + \frac{dm}{2} - 1 \right] + n + 2m - \frac{1}{2} \right) \\
&\qquad + \frac{n(n-1)}{2d} + (r-1) \frac{n^2-n+4nm+4m^2-2m}{2(d+1)} \Bigg] + \frac{r-1}{nr-1}.
\end{aligned}$$

■

Corollary VIII.1.9. *Consider a complete graph crossed with a cycle power, $K_r \times C_n^m$. If $n = 2k+1$ then let $d = \left\lceil \frac{k}{m} \right\rceil$ and*

$$E_{glob}(K_r \times C_n^m) = \frac{1}{r(2k+1)-1}\left[2rm(H_d - 2) + r - 1 + 2m + 2\frac{k}{d} + 2\frac{rm+k}{d+1}\right].$$

If $n = 2k$ then let $d = \left\lceil \frac{k}{m} \right\rceil$ and

$$E_{glob}(K_r \times C_n^m) = \frac{1}{2rk-1}\left[2rm(H_d - 2) + r - 1 + 2m + \frac{2k-1}{d} + \frac{r(2m-1)+2k+1}{d+1}\right].$$

Proof. Begin with Theorem II.8.3. Using Lemma VIII.1.1, if $n = 2k+1$ then let $d = \left\lceil \frac{k}{m} \right\rceil$ and thus

$$\begin{aligned}
E_{glob}(K_r \times C_n^m) &= \frac{1}{r(2k+1)-1}\left[2\sum_{i=1}^{k}\left(\frac{1}{\left\lceil \frac{i}{m} \right\rceil} + \frac{r-1}{\left\lceil \frac{i}{m} \right\rceil + 1}\right) + r - 1\right] \\
&= \frac{1}{r(2k+1)-1}\left[2\left(m(H_d - 1) + \frac{k}{d} + (r-1)m(H_{d+1} - 1) + \frac{k+m}{d+1} - (r-1)m\right) + r - 1\right] \\
&= \frac{1}{r(2k+1)-1}\left[2rm\left(H_d - 2\right) + r + 2m - 1 + 2\frac{k}{d} + 2\frac{rm+k}{d+1}\right].
\end{aligned}$$

And if $n = 2k$ then let $d = \left\lceil \frac{k}{m} \right\rceil$ and therefore

$$\begin{aligned}
E_{glob}(K_r \times C_n^m) &= \frac{1}{2rk-1}\left[2\sum_{i=1}^{k}\left(\frac{1}{\left\lceil \frac{i}{m} \right\rceil} + \frac{r-1}{\left\lceil \frac{i}{m} \right\rceil + 1}\right) + r - 1 - \left(\frac{1}{\left\lceil \frac{k}{m} \right\rceil} + \frac{r-1}{\left\lceil \frac{k}{m} \right\rceil + 1}\right)\right] \\
&= \frac{1}{2rk-1}\left[2rm(H_d - 2) + r + 2m - 1 + 2\frac{k}{d} + 2\frac{rm+k}{d+1} - \frac{1}{d} - \frac{r-1}{d+1}\right] \\
&= \frac{1}{2rk-1}\left[2rm(H_d - 2) + r + 2m - 1 + \frac{2k-1}{d} + \frac{r(2m-1)+2k+1}{d+1}\right]
\end{aligned}$$

■

Using the formulae for faster computation that we just discovered, we repeat the list of Harary indices in Corollary II.8.8 with updated values.

Corollary VIII.1.10. *Let $H(G)$ be the Harary index of a graph G. Then:*

1. $H\left(P_n^m\right) = \frac{m}{2}\left[(2n + m - 1)\, H_d - 2n - dm + 1\right] + \frac{n(n-1)}{2d}$,
2. $n = 2k$: $H\left(C_n^m\right) = 2mk(2H_d - 1) + \frac{k(2k-1)}{d}$,
3. $n = 2k+1$: $H\left(C_n^m\right) = m(2k+1)(2H_d - 1) + \frac{k(2k-1)}{d}$,
4. $H\left(K_{s_1,\ldots,s_t}\right) = \frac{n(n-1)}{2} - \frac{1}{4}\left[\sum_{i=1}^{t} s_i^2 - n\right]$,

5. $H(K_{n,m}) = nm + \frac{1}{4}\left[n^2 - n + m^2 - m\right]$,

6. $H(K_{r;n}) = \frac{nr(2nr-r-1)}{4}$,

7. $H(S_{d,l}) = d\left((d-1)\left(l+\frac{1}{2}\right)H_{2l} - (d-2)(l+1)H_l - l\right)$,

8.

$$H\left(K_r \times P_n^m\right) = r\left[m\left(\left[r\left(n + \frac{3m-1}{2}\right) - m\right]H_d - r\left[2n + 2m + \frac{dm}{2} - 1\right] + n + 2m - \frac{1}{2}\right)\right.$$
$$\left. + \frac{n(n-1)}{2d} + (r-1)\frac{n^2 - n + 4nm + 4m^2 - 2m}{2(d+1)}\right] + \frac{nr(r-1)}{2},$$

9. $n = 2k+1$: $H\left(K_r \times C_n^m\right) = r(2k+1)\left[rm(H_d - 2) + \frac{r-1}{2} + m + \frac{k}{d} + \frac{rm+k}{d+1}\right]$,

10. $n = 2k$: $H\left(K_r \times C_n^m\right) = rk\left[2rm(H_d - 2) + r + 2m - 1 + \frac{2k-1}{d} + \frac{r(2m-1)+2k+1}{d+1}\right]$,

11. $H\left(K_m \times K_n\right) = \frac{1}{4}nm(nm + m + n - 3)$,

12. $H\left(P_m \times P_n\right) = (H_n + H_m - 2)nm + 2\sum_{i=1}^{n-1}\sum_{j=1}^{m-1}\frac{(n-i)(m-j)}{i+j}$.

VIII.1.3 MARTA

NS	SS	D	MC	BH	LC	LNX	BHOG	CH	DRA	Arts	MT	NA	CVC	PCH	5PT	GS	K	I	EW	EL	DCT	AVD	KNS	IND	GNT	WE	OAK	LW	EP	COL	APT	DOME	VC	ASH	WL	HAM	BNK	
0	1.1	2.0	3.0	7.7	9.9	11.8	13.3	16.0	18.0	13.1	13.1	13.7	14.1	14.6	15.1	15.5	16.2	17.6	18.4	20.1	21.4	22.2	24.1	25.4	15.5	17.0	18.5	19.6	21.5	23.3	24.1	15.5	15.9	16.6	18.3	19.8	18.0	NS
	0	0.9	1.9	6.6	8.8	10.7	12.2	14.9	16.9	12.0	12.0	12.6	13.0	13.5	14.0	14.4	15.1	16.5	17.3	19.0	20.3	21.1	23.0	24.3	14.4	15.9	17.4	18.5	20.4	22.2	23.0	14.4	14.8	15.5	17.2	18.7	16.9	SS
		0	1.0	5.7	7.9	9.8	11.3	14.0	16.0	11.1	11.1	11.7	12.1	12.6	13.1	13.5	14.2	15.6	16.4	18.1	19.4	20.2	22.1	23.4	13.5	15.0	16.5	17.6	19.5	21.3	22.1	13.5	13.9	14.6	16.3	17.8	16.0	D
			0	4.7	6.9	8.8	10.3	13.0	15.0	10.1	10.1	10.7	11.1	11.6	12.1	12.5	13.2	14.6	15.4	17.1	18.4	19.2	21.1	22.4	12.5	14.0	15.5	16.6	18.5	20.3	21.1	12.5	12.9	13.6	15.3	16.8	15.0	MC
				0	2.2	4.1	5.6	8.3	10.3	5.4	5.4	6.0	6.4	6.9	7.4	7.8	8.5	9.9	10.7	12.4	13.7	14.5	16.4	17.7	7.8	9.3	10.8	11.9	13.8	15.6	16.4	7.8	8.2	8.9	10.6	12.1	10.3	BH
					0	1.9	3.4	6.1	8.1	3.2	3.7	4.3	4.7	5.2	5.7	6.1	6.8	8.2	9.0	10.7	12.0	12.8	14.7	16.0	6.1	7.6	9.1	10.2	12.1	13.9	14.7	6.1	6.5	7.2	8.9	10.4	8.6	LC
						0	1.5	4.2	6.2	5.1	5.1	5.7	6.1	6.6	7.1	7.5	8.2	9.6	10.4	12.1	13.4	14.2	16.1	17.4	7.5	9.0	10.5	11.6	13.5	15.3	16.1	7.5	7.9	8.6	10.3	11.8	10.0	LNX
							0	2.7	4.7	6.6	6.6	7.2	7.6	8.1	8.6	9.0	9.7	11.1	11.9	13.6	14.9	15.7	17.6	18.9	9.0	10.5	12.0	13.1	15.0	16.8	17.6	9.0	9.4	10.1	11.8	13.3	11.5	BHOG
								0	2.0	9.3	9.3	9.9	10.3	10.8	11.3	11.7	12.4	13.8	14.6	16.3	17.6	18.4	20.3	21.6	11.7	13.2	14.7	15.8	17.7	19.5	20.3	11.7	12.1	12.8	14.5	16.0	14.2	CH
									0	11.3	11.3	11.9	12.3	12.8	13.3	13.7	14.4	15.8	16.6	18.3	19.6	20.4	22.3	23.6	13.7	15.2	16.7	17.8	19.7	21.5	22.3	13.7	14.1	14.8	16.5	18.0	16.2	DRA
										0	0.5	1.1	1.5	2.0	2.5	2.9	3.6	5.0	5.8	7.5	8.8	9.6	11.5	12.8	2.9	4.4	5.9	7.0	8.9	10.7	11.5	2.9	3.3	4.0	5.7	7.2	5.4	ARTS
											0	0.6	1.0	1.5	2.0	2.4	3.1	4.5	5.3	7.0	8.3	9.1	11.0	12.3	2.4	3.9	5.4	6.5	8.4	10.2	11.0	2.4	2.8	3.5	5.2	6.7	4.9	MT
												0	0.6	0.9	1.4	1.8	2.5	3.9	4.7	6.4	7.7	8.5	10.4	11.7	1.8	3.3	4.8	5.9	7.8	9.6	10.4	1.8	2.2	2.9	4.6	6.1	4.3	NA
													0	0.5	1.0	1.4	2.1	3.5	4.3	6.0	7.3	8.1	10.0	11.3	1.4	2.9	4.4	5.5	7.4	9.2	10.0	1.4	1.8	2.5	4.2	5.7	3.9	CVC
														0	0.5	0.9	1.6	3.0	3.8	5.5	6.8	7.6	9.5	10.8	0.9	2.4	3.9	5.0	6.9	8.7	9.5	0.9	1.3	2.0	3.7	5.2	3.4	PCH
															0	0.4	1.1	2.5	3.3	5.0	6.3	7.1	9.0	10.3	0.4	1.9	3.4	4.5	6.4	8.2	9.0	0.4	0.8	1.5	3.2	4.7	2.9	5PT
																0	0.7	2.1	2.9	4.6	5.9	6.7	8.6	9.9	0.8	2.3	3.8	4.9	6.8	8.6	9.4	0.8	1.2	1.9	3.6	5.1	3.3	GS
																	0	1.4	2.2	3.9	5.2	6.0	7.9	9.2	1.5	3.0	4.5	5.6	7.5	9.3	10.1	1.5	1.9	2.6	4.3	5.8	4.0	K
																		0	0.8	2.5	3.8	4.6	6.5	7.8	2.9	4.4	5.9	7.0	8.9	10.7	11.5	2.9	3.3	4.0	5.7	7.2	5.4	I
																			0	1.7	3.0	3.8	5.7	7.0	3.7	5.2	6.7	7.8	9.7	11.5	12.3	3.7	4.1	4.8	6.5	8.0	6.2	EW
																				0	1.3	2.1	4.0	5.3	5.4	6.9	8.4	9.5	11.4	13.2	14.0	5.4	5.8	6.5	8.2	9.7	7.9	EL
																					0	0.8	2.7	4.0	6.7	8.2	9.7	10.8	12.7	14.5	15.3	6.7	7.1	7.8	9.5	11.0	9.2	DCT
																						0	1.9	3.2	7.5	9.0	10.5	11.6	13.5	15.3	16.1	7.5	7.9	8.6	10.3	11.8	10.0	AVD
																							0	1.3	9.4	10.9	12.4	13.5	15.4	17.2	18.0	9.4	9.8	10.5	12.2	13.7	11.9	KNS
																								0	10.7	12.2	13.7	14.8	16.7	18.5	19.3	10.7	11.1	11.8	13.5	15.0	13.2	IC
																									0	1.5	3.0	4.1	6.0	7.8	8.6	0.8	1.2	1.9	3.6	5.1	3.3	GNT
																										0	1.5	2.6	4.5	6.3	7.1	2.3	2.7	3.4	5.1	6.6	4.8	WE
																											0	1.1	3.0	4.8	5.6	3.8	4.2	4.9	6.6	8.1	6.3	OAK
																												0	1.9	3.7	4.5	4.9	5.3	6.0	7.7	9.2	7.4	LW
																													0	1.8	2.6	6.8	7.2	7.9	9.6	11.1	9.3	EP
																														0	0.8	8.6	9.0	9.7	11.4	12.9	11.1	COL
																															0	9.4	9.8	10.5	12.2	13.7	11.9	APT
																																0	0.4	1.1	2.8	4.3	2.5	DOME
																																	0	0.7	2.4	3.9	2.1	VC
																																		0	1.7	3.2	1.4	ASH
																																			0	1.5	3.1	WL
																																				0	4.6	HAM
																																					0	BNK

Table VIII.1.1: A table of the rail distances between all 38 stations of the MARTA network[21]. Distances are in miles. See Table VIII.1.3 for acronym key.

NS	SS	D	MC	BH	LC	LNX	BHOG	CH	DRA	Arts	MT	NA	CVC	PCH	5PT	GS	K	I	EW	EL	DCT	AVD	KNS	IND	GNT	WE	OAK	LW	EP	COL	APT	DOME	VC	ASH	WL	HAM	BNK	
0	1.04	1.86	2.31	6.76	8.42	6.93	5.98	4.97	5.31	10.90	11.44	12.07	12.45	12.91	13.34	13.58	13.55	12.98	12.69	12.70	12.30	12.52	13.39	14.21	13.72	14.79	16.19	17.38	19.14	20.97	21.68	13.24	13.30	13.24	14.21	14.61	12.64	NS
	0	0.83	1.30	5.84	7.50	5.97	4.92	3.94	4.49	10.02	10.59	11.21	11.57	12.04	12.45	12.66	12.62	12.00	11.71	11.67	11.25	11.49	12.36	13.19	12.84	13.94	15.25	16.54	18.31	20.13	20.84	12.36	12.43	12.66	13.46	13.92	11.86	SS
		0	0.68	5.18	6.84	5.26	4.19	3.20	3.95	9.39	9.95	10.57	10.94	11.37	11.79	12.00	11.93	11.27	10.94	10.88	10.44	10.66	11.54	12.37	12.19	13.33	14.75	15.93	17.70	19.52	20.34	11.73	11.82	12.08	12.94	13.46	11.33	D
			0	4.55	6.20	4.67	3.67	3.15	4.16	8.73	9.30	9.91	10.28	10.73	11.15	11.36	11.31	10.69	10.39	10.39	10.01	10.28	11.21	12.11	11.55	12.66	14.09	15.27	17.04	18.86	19.57	11.08	11.16	11.41	12.26	12.77	10.65	MC
				0	1.67	0.58	1.86	4.48	6.26	4.20	4.76	5.37	5.75	6.20	6.63	6.82	6.79	6.30	6.13	6.52	6.49	7.05	8.44	9.63	7.01	8.15	9.59	10.76	12.53	14.35	15.05	6.55	6.63	6.94	7.96	8.64	6.31	BH
					0	1.63	3.09	5.71	7.45	2.57	3.12	3.73	4.10	4.55	4.97	5.16	5.12	4.67	4.57	5.18	5.37	6.09	7.61	8.88	5.36	6.55	8.00	9.16	10.94	12.74	13.43	4.92	5.05	5.41	6.58	7.41	4.93	LC
						0	1.51	4.18	5.94	4.20	4.75	5.35	5.71	6.16	6.57	6.75	6.67	6.06	5.83	6.10	6.00	6.53	7.89	9.06	6.98	8.17	9.62	10.78	12.56	14.37	15.05	6.54	6.66	7.03	8.14	8.89	6.48	LNX
							0	2.67	4.44	7.52	6.16	6.73	7.34	7.58	7.95	8.08	7.95	7.16	6.80	6.75	6.41	6.75	7.87	8.91	8.36	9.60	11.05	12.20	13.47	15.76	16.44	7.95	8.09	8.50	9.66	10.42	8.00	BHOG
								0	1.78	8.24	8.72	9.30	9.62	10.04	10.46	10.55	10.36	9.40	8.90	8.47	7.84	7.89	8.55	9.30	10.88	12.15	13.61	14.73	16.50	18.27	18.92	10.50	10.67	11.11	12.32	13.09	10.66	CH
									0	9.95	10.41	10.88	11.21	11.59	12.08	12.14	11.91	10.86	10.31	9.68	8.91	8.78	9.17	9.68	12.50	13.79	15.25	16.36	18.11	19.86	20.49	12.14	12.34	12.81	14.06	14.84	12.40	DRA
										0	0.59	1.24	1.61	2.09	2.45	2.71	2.81	2.97	3.30	4.60	5.30	6.19	7.85	9.19	2.82	3.97	5.42	6.58	8.36	10.17	10.86	2.35	2.46	2.87	4.23	5.23	2.67	ARTS
											0	0.62	1.01	1.47	1.87	2.12	2.24	2.54	2.98	4.39	5.19	6.10	7.76	9.10	2.25	3.44	4.90	6.05	7.83	9.63	10.31	1.80	1.95	2.45	3.93	5.01	2.49	MT
												0	0.32	0.84	1.25	1.50	1.66	2.23	2.81	4.32	5.22	6.13	7.78	9.11	1.63	2.87	4.33	5.46	7.24	9.03	9.71	1.22	1.43	2.04	3.64	4.78	2.37	NA
													0	0.52	0.93	1.13	1.33	2.10	2.75	4.31	5.25	6.18	7.80	9.11	1.28	2.56	4.01	5.13	6.90	8.69	9.36	0.90	1.18	1.86	3.52	4.70	2.40	CVC
														0	0.45	0.67	0.96	2.01	2.73	4.32	5.32	6.25	7.83	9.14	0.86	2.19	3.63	4.72	6.49	8.26	8.93	0.63	0.99	1.74	3.43	4.64	2.51	PCH
															0	0.45	0.96	2.25	3.01	4.60	5.63	6.55	8.11	9.40	0.41	1.75	3.18	4.26	6.03	7.81	8.48	0.39	0.76	1.50	3.17	4.39	2.46	5PT
																0	0.53	1.94	2.72	4.31	5.36	6.27	7.00	9.08	0.59	1.86	3.21	4.32	5.96	7.72	8.35	0.84	1.18	1.40	3.52	4.76	2.91	GS
																	0	1.44	2.23	3.79	4.86	5.76	7.29	8.56	1.14	2.34	3.61	4.56	6.26	7.97	8.58	1.34	1.70	2.43	4.07	5.30	3.40	K
																		0	0.78	2.35	3.42	4.32	5.87	7.15	2.56	3.79	5.02	5.91	7.53	9.17	9.73	2.60	2.98	3.74	5.42	6.64	4.49	I
																			0	1.59	2.63	3.54	5.11	6.40	3.32	4.57	5.79	6.65	8.24	9.85	10.40	3.34	3.71	4.48	6.16	7.38	5.15	EW
																				0	2.32	3.22	4.77	6.07	4.90	6.11	7.25	8.05	9.57	11.08	11.57	4.95	5.32	6.07	7.76	8.98	6.69	EL
																					0	0.93	2.56	3.91	5.94	7.20	8.36	9.11	10.67	12.18	12.66	5.93	6.31	7.06	8.74	9.94	7.58	DCT
																						0	1.66	2.99	6.85	8.09	9.23	9.98	11.44	12.90	13.35	6.86	7.23	7.99	9.67	10.87	8.52	AVD
																							0	1.34	8.40	9.59	10.66	11.32	12.69	14.04	14.42	8.44	8.82	4.58	11.26	12.47	10.15	KNS
																								0	9.67	10.83	11.85	12.45	13.75	15.01	15.36	9.74	10.11	10.88	12.56	13.78	11.48	IC
																									0	1.35	2.77	3.86	5.62	7.41	8.08	0.53	0.73	1.37	2.95	4.19	2.48	GNT
																										0	1.45	2.60	4.38	6.20	6.90	1.66	1.52	1.43	2.27	3.42	2.65	WE
																											0	1.18	2.94	4.77	5.48	3.11	2.96	2.71	2.75	3.55	3.76	OAK
																												0	1.78	3.60	4.31	4.26	4.12	3.90	3.76	4.34	4.92	LW
																													0	1.83	2.56	6.02	5.90	5.66	5.29	5.60	6.61	EP
																														0	0.77	7.83	7.72	7.48	7.07	7.25	8.43	COL
																															0	8.50	8.40	8.19	7.80	8.00	9.16	APT
																																0	0.38	1.14	2.83	4.05	2.08	DOME
																																	0	0.75	2.45	3.66	1.75	VC
																																		0	1.68	2.91	1.25	ASH
																																			0	1.24	1.66	WL
																																				0	2.57	HAM
																																					0	BNK

Table VIII.1.2: A table of the Euclidean (Earth) distances between all 38 stations of the MARTA network[22]. Distances are in miles. See Table VIII.1.3 for acronym key. There is error in these measurements but our analysis was not concerned with this.

Table VIII.1.3: Acronym key for MARTA stations of distance matrices.

Acronym	**MARTA Station**	**Acronym**	**MARTA Station**
NS	North Springs	SS	Sandy Springs
D	Dunwoody	MC	Medical Center
BH	Buckhead	LC	Lindbergh Center
LNX	Lenox	BHOG	Brookhaven/Oglethorpe Univ.
CH	Chamblee	DRA	Doraville
Arts	Arts Center	MT	Midtown
NA	North Avenue	CVC	Civic Center
PCH	Peachtree Center	5PT	Five Points
GS	Georgia State	K	King Memorial
I	Inman Park/Reynoldstown	EW	Edgewood/Candler Park
EL	East Lake	DCT	Decatur
AVD	Avondale	KNS	Kensington
IND	Indian Creek	GNT	Garnett
WE	West End	OAK	Oakland City
LW	Lakewood/Ft. McPherson	EP	East Point
COL	College Park	APT	Airport
DOME	Dome/GWCC/Philips Arena/CNN Center	VC	Vine City
ASH	Ashby	WL	West Lake
HAM	H. E. Holmes	BNK	Bankhead

VIII.1.4 Brain Network

Table VIII.1.4: Acronym key for the Regions of the Brain in Global Efficiency Analysis. See Figure II.9.3.

Acronym	Part of Brain
LMF	Left Medial Fusiform Gyrus
RMF	Right Medial Fusiform Gyrus
LMTG	Left Middle Temporal Gyrus
LMT	Left Motion Selective Cortex
LIP	Left Inferior Parietal Cortex
IPS	Intra Parietal Sulcus
RFEF	Right Frontal Eye Fields
MPC	Medial Prefrontal Cortex
LPC	Left Posterior Cingulate
LLP	Left Lateral Parietal
RS	Right Somato
LS	Left Somato
LM	Left Motor
RM	Right Motor
LMLA	Left Medial Leg Area

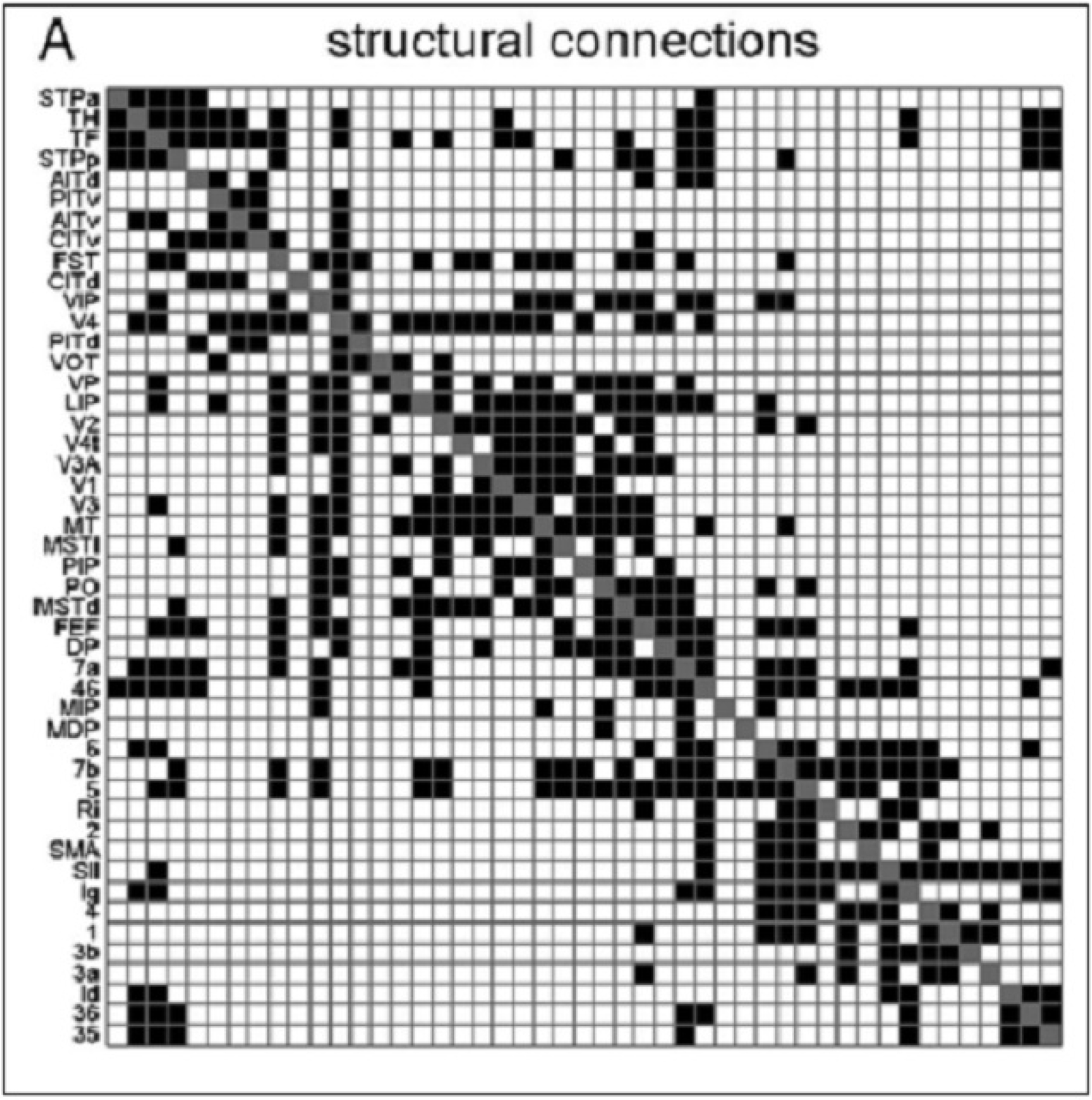


Figure VIII.1.1: This is the adjacency matrix of a graph corresponding to regions of a macaque neocortex. Black cells indicate a connection from row to column. Note that the matrix is not symmetric; the graph is directed. The corresponding graph has global efficiency 0.571. Used with permission from Honey et al.[12]

VIII.2 Betweenness Centrality

VIII.2.1 Unique Betweenness Centrality Lemmas and Table

Table VIII.2.1: Demonstration of unique betweenness centrality for small U_n graphs.

n	3	4	5	6	7	8
$bc(v_0)$	0	0	0	0	0	0
$bc(v_1)$	$\frac{35}{3}$	$\frac{97}{6}$	$\frac{617}{30}$	$\frac{249}{10}$	$\frac{2043}{70}$	$\frac{4681}{140}$
$bc(v_2)$	11	$\frac{59}{3}$	$\frac{169}{6}$	$\frac{1097}{30}$	$\frac{449}{10}$	$\frac{3723}{70}$
$bc(v_3)$	$\frac{7}{3}$	$\frac{31}{2}$	$\frac{424}{15}$	$\frac{245}{6}$	$\frac{11189}{210}$	$\frac{1313}{20}$
$bc(v_4)$	–	$\frac{8}{3}$	$\frac{601}{30}$	$\frac{554}{15}$	$\frac{2251}{42}$	$\frac{2454}{35}$
$bc(v_5)$	–	–	$\frac{89}{30}$	$\frac{368}{15}$	$\frac{9571}{210}$	$\frac{5573}{84}$
$bc(v_6)$	–	–	–	$\frac{97}{30}$	$\frac{3044}{105}$	$\frac{11377}{210}$
$bc(v_7)$	–	–	–	–	$\frac{243}{70}$	$\frac{4677}{140}$
$bc(v_8)$	–	–	–	–	–	$\frac{129}{35}$
$bc(v'_1)$	$\frac{10}{3}$	$\frac{13}{3}$	$\frac{77}{15}$	$\frac{29}{5}$	$\frac{223}{35}$	$\frac{481}{70}$
$bc(v'_2)$	8	$\frac{85}{6}$	$\frac{298}{15}$	$\frac{379}{15}$	$\frac{1066}{35}$	$\frac{4969}{140}$
$bc(v'_3)$	$\frac{5}{3}$	$\frac{40}{3}$	$\frac{241}{10}$	$\frac{103}{3}$	$\frac{4642}{105}$	$\frac{1884}{35}$
$bc(v'_4)$	–	$\frac{13}{6}$	$\frac{55}{3}$	$\frac{1007}{30}$	$\frac{1688}{35}$	$\frac{26249}{420}$
$bc(v'_5)$	–	–	$\frac{77}{30}$	$\frac{347}{15}$	$\frac{8977}{210}$	$\frac{1853}{30}$
$bc(v'_6)$	–	–	–	$\frac{29}{10}$	$\frac{139}{5}$	$\frac{7243}{140}$
$bc(v'_7)$	–	–	–	–	$\frac{223}{70}$	$\frac{1133}{35}$
$bc(v'_8)$	–	–	–	–	–	$\frac{481}{140}$

We use some bounding properties of the Harmonic numbers in a few places of this thesis:

Theorem VIII.2.1. *For $n \in \mathbb{N}$,*

$$\frac{1}{2(n+1)} < H_n - \ln(n) - \gamma < \frac{1}{2n},$$

where $\gamma = 0.5772\ldots$ is the Euler-Mascheroni constant[23].

Lemma VIII.2.2. *For $n \geq 9$ and $1 \leq l \leq \frac{n}{2} - 1$,*

$$\sum_{i=l}^{n} \frac{l}{i} - \sum_{i=n-l+1}^{n} \frac{n-l}{i} + \frac{1}{n-l} - \frac{1}{l} - 1 > 0. \tag{VIII.2.1}$$

Proof. This relies initially on Theorem VIII.2.1.

First we shall remove the cases where $l = 1$ and $l = \frac{n}{2} - 1$.

$$\begin{aligned} \sum_{i=1}^{n} \frac{1}{i} - \sum_{i=n-1+1}^{n} \frac{n-1}{i} + \frac{1}{n-1} - \frac{1}{1} - 1 &= H_n - \frac{n-1}{n} + \frac{1}{n-1} - 2 \\ &> \ln(n) + \gamma + \frac{1}{2(n+1)} + \frac{1}{n} + \frac{1}{n-1} - 3. \end{aligned}$$

Let $h(n) = \ln(n) + \gamma + \frac{1}{2(n+1)} + \frac{1}{n} + \frac{1}{n-1} - 3$. When $n = 9, 10, 11$, $h(n) \approx 0.06, 0.13, 0.20$ respectively. And when $n \geq 12$, $h(n) \geq \ln(12) + \gamma - 3 > 0$. Thus the case of $l = 1$ has been shown.

For $n = 2k$ and $l = \frac{n}{2} - 1$,

$$\begin{aligned} \sum_{i=\frac{n}{2}-1}^{n} \frac{\frac{n}{2}-1}{i} - \sum_{i=\frac{n}{2}+2}^{n} \frac{\frac{n}{2}+1}{i} + \frac{2}{n-4} - \frac{2}{n-2} - 1 &= \left(\frac{n}{2} - 1\right) \sum_{i=\frac{n}{2}}^{n} \frac{1}{i} + 1 - \left(\frac{n}{2} + 1\right) \sum_{i=\frac{n}{2}}^{n} \frac{1}{i} + 1 \\ &\quad + \frac{\frac{n}{2}+1}{n/2} + \frac{2}{n-4} - \frac{2}{n-2} - 1 \\ &= 2 - 2\sum_{i=\frac{n}{2}}^{n} \frac{1}{i} + \frac{2}{n} + \frac{2}{n-4} - \frac{2}{n-2} \\ &> 2 - 2(1) + \frac{2}{n} + \frac{2}{n-4} - \frac{2}{n-2} \\ &> 0. \end{aligned}$$

The last step made use of Theorem VIII.2.6.

Now it is important to remember throughout this proof that $n \geq 9$ and $2 \leq l \leq \frac{n-3}{2}$. In the case $n = 2k$, $l = \frac{n}{2} - 1$ is taken care of above. If $n = 2k + 1$, these are the effective bounds anyway. Now considering the other cases,

$$\begin{aligned}
\sum_{i=l}^{n}\frac{l}{i}-\sum_{i=n-l+1}^{n}\frac{n-l}{i}+\frac{1}{n-l}-\frac{1}{l}-1 &= l\left(H_n-H_l+\frac{1}{l}\right)-(n-l+1)(H_n-H_{n-l})+\frac{1}{n-l}-\frac{1}{l}-1\\
&= 1+2lH_n+(n-l+1)H_{n-l}-[lH_l+(n+1)H_n]+\frac{1}{n-l}-\frac{1}{l}-1\\
&> 2l(\ln(n)+\gamma+\frac{1}{2(n+1)})+(n-l+1)\left(\ln(n-l)+\gamma+\frac{1}{2(n-l+1)}\right)\\
&\quad -l\left(\ln(l)+\gamma+\frac{1}{2l}\right)-(n+1)(\ln(n)+\gamma+\frac{1}{2n})+\frac{1}{n-l}-\frac{1}{l}\\
&= (2l-(n+1))\ln(n)+\frac{l}{n+1}+(n-l+1)\ln(n-l)\\
&\quad -l\ln(l)-\frac{n+1}{2n}+\frac{1}{n-l}-\frac{1}{l}\\
&= \frac{l}{n+1}-\frac{1}{2}-\frac{1}{2n}+(2l-(n+1))\ln(n)+(n-l+1)\ln(n-l)\\
&\quad -l\ln(l)+\frac{1}{n-l}-\frac{1}{l}.
\end{aligned}$$

Thus we have created a continuous expression of a lower bound for our discrete expression. Consider a function of l and n: f. Let $f(l,n)=\frac{l}{n+1}-\frac{1}{2}-\frac{1}{2n}+(2l-(n+1))\ln(n)+(n-l+1)\ln(n-l)-l\ln(l)+\frac{1}{n-l}-\frac{1}{l}$. For notation, let $\frac{\partial}{\partial l}f=f'$ and $\frac{\partial}{\partial n}f=f^*$. Then

$$\begin{aligned}
f''(l,n) &= \frac{1}{n-l}+\frac{1}{(n-l)^2}-\frac{1}{l}+\frac{2}{(n-l)^3}-\frac{2}{l^3}\\
&= \frac{2+n+n^2-l-2nl+l^2-(n-l)^3(l^2+2)}{l^3(n-l)^3}\\
&< \frac{2+n+n^2-1-2n+\frac{n^2}{4}-\left(\frac{n}{2}\right)^3(1^2+2)}{l^3(n-l)^3}\\
&= \frac{1-n+\left(\frac{5}{4}-\frac{3}{8}n\right)n^2}{l^3(n-l)^3}\\
&< 0.
\end{aligned}$$

Therefore $f''(l,n)<0$: concave down, over the domain of l. And so we can say that for a given n, for all $x,y,z\in[1,\frac{n}{2}-1]$, if $x\le y\le z$, then $f(x,n)\le f(y,n)$ or $f(z,n)\le f(y,n)$.

Note that

$$\begin{aligned}
f(2,n) &= \frac{2}{n+1}-\frac{1}{2}-\frac{1}{2n}+(4-(n+1))\ln(n)+(n-2+1)\ln(n-2)-2\ln(2)+\frac{1}{n-2}-\frac{1}{2}\\
&= \frac{2}{n+1}-\frac{1}{2n}+\frac{1}{n-2}-(n-3)\ln(n)+(n-1)\ln(n-2)-1-2\ln(2).
\end{aligned}$$

And taking the derivative with respect to n this time we can see

$$\begin{aligned}
f^*(2,n) &= -\frac{2}{(n+1)^2} + \frac{1}{2n^2} - \frac{1}{(n-2)^2} - \ln(n) - \frac{n-3}{n} + \ln(n-2) + \frac{n-1}{n-2} \\
&= \frac{3}{n} - \frac{2}{(n+1)^2} + \frac{1}{2n^2} - \frac{1}{(n-2)^2} + \ln\left(1 - \frac{2}{n}\right) + \frac{n-1}{n-2} - 1 \\
&> \frac{1}{2n^2} + \frac{3}{n} - \frac{2}{(n+1)^2} - \frac{1}{(n-2)^2} + 1 - \frac{1}{1-\frac{2}{n}} + \frac{n-1}{n-2} - 1 \\
&= \frac{1}{2n^2} + \frac{3}{n} - \frac{2}{(n+1)^2} - \frac{1}{(n-2)^2} - \frac{n}{n-2} + \frac{n-1}{n-2} \\
&= \frac{4n^5 - 17n^4 - 2n^3 + 7n^2 + 28n + 4}{2(n-2)^2 n^2 (n+1)^2} \\
&> \frac{n^4(2n-17) + 2n^3(n^2-1) + 7n^2 + 28n + 4}{2(n-2)^2 n^2 (n+1)^2} \\
&> 0.
\end{aligned}$$

since $n \geq 9$. The logarithmic bound came from F. Topsøe, [24]. Thus $f(2,n)$ is a monotonically increasing function of n. And $f(2,9) = \frac{2}{10} - \frac{1}{18} + \frac{1}{7} - (6)\ln(9) + (8)\ln(7) - 1 - 2\ln(2) \approx 0.285 > 0$. Thus $f(2,n) > 0$ for all $n \geq 9$. The goal is to show that $f\left(\frac{n-3}{2}, n\right) > 0$ as well.

Now if we remember that $f(l,n)$ as a function of l is concave down over the domain, we now know that if $2 \leq l \leq \frac{n-3}{2}$, $f(l,n) > f(2,n)$ or $f(l,n) > f\left(\frac{n-3}{2}, n\right)$. Either way $f(l,n) > 0$. Combine this with the boundary cases for l and we have proven the lemma. ■

Lemma VIII.2.3. *For $n \geq 9$ and $1 \leq l \leq \frac{n}{2} - 1$.*

$$\sum_{i=l}^{n} \frac{l}{i} - \sum_{i=n-l+1}^{n} \frac{n-l+1}{i} > 0. \tag{VIII.2.2}$$

Proof. This uses Lemmas VIII.2.2,VIII.2.6 and VIII.2.7

$$\begin{aligned}
\sum_{i=l}^{n}\frac{l}{i}-\sum_{i=n-l+1}^{n}\frac{n-l+1}{i} &= \sum_{i=l}^{n}\frac{l}{i}-\sum_{i=n-l+1}^{n}\frac{n-l}{i}-\sum_{i=n-l+1}^{n}\frac{1}{i}\\
&\geq \sum_{i=l}^{n}\frac{l}{i}-\sum_{i=n-l+1}^{n}\frac{n-l}{i}-\sum_{i=n-(\frac{n}{2}-1)+1}^{n}\frac{1}{i}\\
&= \sum_{i=l}^{n}\frac{l}{i}-\sum_{i=n-l+1}^{n}\frac{n-l}{i}-\sum_{i=\frac{n}{2}+2}^{n}\frac{1}{i}\\
&> \sum_{i=l}^{n}\frac{l}{i}-\sum_{i=n-l+1}^{n}\frac{n-l}{i}-\sum_{i=\lceil\frac{n}{2}\rceil}^{n}\frac{1}{i}\\
&> \sum_{i=l}^{n}\frac{l}{i}-\sum_{i=n-l+1}^{n}\frac{n-l}{i}-1\\
&> \sum_{i=l}^{n}\frac{l}{i}-\sum_{i=n-l+1}^{n}\frac{n-l}{i}+\frac{1}{n-l}-\frac{1}{l}-1\\
&> 0.
\end{aligned}$$

■

A better bound of 0.898 was found for this expression using the same method as the proof of Lemma VIII.2.2. It was not included as it was not needed.

Lemma VIII.2.4. *For $n \geq 9$ and $1 \leq l \leq \frac{n}{2}-1$.*

$$2n-4l-\sum_{i=l}^{n}\frac{l-1}{i}+\sum_{i=n-l}^{n}\frac{n-l-1}{i}+\frac{1}{n-l}-\frac{1}{l}>0. \tag{VIII.2.3}$$

Proof. By the same method as Lemma VIII.2.2 except for the opposite bound. ■

Lemma VIII.2.5. *For $n \geq 9$ and $1 \leq l \leq \frac{n}{2}-1$.*

$$2n-4l+2+\sum_{i=n-l+1}^{n}\frac{n-l}{i}-\sum_{i=l}^{n}\frac{l-1}{i}>0. \tag{VIII.2.4}$$

Proof.

$$\begin{aligned}
2n-4l+2+\sum_{i=n-l+1}^{n}\frac{n-l}{i}-\sum_{i=l}^{n}\frac{l-1}{i} &= 2n-4l+2-1+\sum_{i=n-l}^{n}\frac{n-l-1}{i}+\sum_{i=n-l}^{n}\frac{1}{i}-\sum_{i=l}^{n}\frac{l-1}{i}\\
&> 2n-4l+\sum_{i=n-l}^{n}\frac{n-l-1}{i}-\sum_{i=l}^{n}\frac{l-1}{i}+\frac{1}{n-l}-\frac{1}{l}\\
&> 0.
\end{aligned}$$

■

Lemma VIII.2.6. *For $n \geq 1$,*

$$\sum_{i=\lceil \frac{n}{2} \rceil}^{n} \frac{1}{i} < \ln(2) + \frac{3n+15}{2n(n+3)}. \tag{VIII.2.5}$$

Proof. This uses Theorem VIII.2.1.

$$\begin{aligned}
\sum_{i=\lceil \frac{n}{2} \rceil}^{n} \frac{1}{i} &= H_n - H_{\lceil \frac{n}{2} \rceil} + \frac{1}{\lceil \frac{n}{2} \rceil} \\
&< \ln(n) + \gamma + \frac{1}{2n} - \ln\left(\left\lceil \frac{n}{2} \right\rceil\right) - \gamma - \frac{1}{2(\lceil \frac{n}{2} \rceil + 1)} + \frac{1}{\lceil \frac{n}{2} \rceil} \\
&\leq \ln(n) + \frac{1}{2n} - \ln\left(\frac{n}{2}\right) - \frac{1}{2(\lceil \frac{n}{2} \rceil + 1)} + \frac{1}{\lceil \frac{n}{2} \rceil} \\
&\leq \ln(2) + \frac{1}{2n} - \frac{1}{2(\frac{n+1}{2} + 1)} + \frac{1}{\frac{n}{2}} \\
&= \ln(2) + \frac{3n+15}{2n(n+3)}.
\end{aligned}$$

■

Lemma VIII.2.7. *For $n \geq 1$,*

$$\ln(2) + \frac{3n+15}{2n(n+3)} < 1 + \frac{2}{n}.$$

Proof.

$$\begin{aligned}
\frac{d}{dn}\left[\ln(2) + \frac{3n+15}{2n(n+3)}\right] &= \frac{3(2n(n+3)) - (4n+6)(3n+15)}{4n^2(n+3)^2} \\
&= \frac{6n(n+3) - 6(n+3)(n+5)}{4n^2(n+3)^2} \\
&= \frac{3n - 3(n+5)}{2n^2(n+3)} \\
&= \frac{-15}{2n^2(n+3)} \\
&< 0.
\end{aligned}$$

Thus $\ln(2) + \frac{3n+15}{2n(n+3)}$ is monotonically decreasing for $n \geq -3$. We can then use this to say that

n	Lemma VIII.2.6 Bound	$1 + \frac{2}{n}$
1	2.943	3
2	1.743	2
3	1.360	1.666
4	1.175	1.5
5	1.068	1.4
6	0.999	1.333
≥ 7	< 1	> 1

And therefore $\ln(2) + \frac{3n+15}{2n(n+3)} < 1 + \frac{2}{n}$, for all $n \in \mathbb{N}$. ■

Corollary VIII.2.8. *For $n \geq 7$,*

$$\sum_{i=\lceil \frac{n}{2} \rceil}^{n} \frac{1}{i} < 1.$$

Lemma VIII.2.9. *For $n \geq 9$*

$$2n - 2\left\lceil \frac{n}{2} \right\rceil + 2 - 2\sum_{i=\lceil \frac{n}{2} \rceil}^{n} \frac{\lceil \frac{n}{2} \rceil}{i} + \sum_{i=\lceil \frac{n}{2} \rceil}^{n} \frac{1}{i} > 0. \tag{VIII.2.6}$$

Proof. The second line uses Lemma VIII.2.6 and the third line uses Lemma VIII.2.7 and that $n \geq 9$.

$$\begin{aligned} 2n - 2\left\lceil \frac{n}{2} \right\rceil + 2 - 2\sum_{i=\lceil \frac{n}{2} \rceil}^{n} \frac{\lceil \frac{n}{2} \rceil}{i} + \sum_{i=\lceil \frac{n}{2} \rceil}^{n} \frac{1}{i} &\geq 2n - 2\frac{n+1}{2} + 2 - \left(2\left\lceil \frac{n}{2} \right\rceil - 1\right)\sum_{i=\lceil \frac{n}{2} \rceil}^{n} \frac{1}{i} \\ &> n + 1 - \left(2\frac{n+1}{2} - 1\right)\left(\ln(2) + \frac{3n+15}{2n(n+3)}\right) \\ &> n + 1 - n(1) \\ &> 0. \end{aligned}$$

■

IX. Bibliography

References

[1] V. Latora and M. Marchiori (2001). Efficient Behavior of Small-World Networks. *Physical Review Letters*, Vol. **87**, No 19, (2001).

[2] D. Plavšić, S. Nikolić, N. Trinajstić, and Z. Mihalić, On the Harary index for the characterization for chemical graphs, *J. Math. Chem.*, **12** (1993), 235-250.

[3] O. Ivanciuc, T.S. Balaban, and A.T. Balaban, J. Math. Chem, 1993, **12**, 309-318.

[4] G. Su, L. Xiong, and I. Gutman, Harary index of the k-th power of a graph, *Appl. Anal. Discrete Math.* **7** (2013), 94–105.

[5] H. Wang and L.Kang, More on the Harary index of cacti. *J. Appl. Math. Comput.* **43** (2013), no. 1-2, 369–386.

[6] K. C. Das, K. Xu, I.N. Cangul, A. S Cevik, and A. Graovac, On the Harary index of graph operations, Journal of Inequalities and Applications (2013), 339.

[7] A. Ilić, G. Yu, and L. Feng, The Harary index of trees, arXiv:1104.0920v3 [math.CO].

[8] V. Latora and M. Marchiori, Is the Boston subway a small-world network?, *Physica A* **314**, (2002), 109-113.

[9] B. Ek, C. VerSchneider, J. Lind and D. A. Narayan, Real world graph efficiency, to appear in the *Bulletin of the Institute of Combinatorics and Its Applications*

[10] B. Ek, C. VerSchneider, and D. A. Narayan, Efficiency of star-like networks and the Atlanta subway network, *Physica A*, **392** (2013) 5481-5489.

[11] B. Ek, C. VerSchneider, and D. A. Narayan, Local Efficiency and Clustering Coefficients of Graphs, submitted.

[12] C. Honey, R. Kötter, M. Breakspear, and O. Sporns. Network structure of cerebral cortex shapes functional connectivity on multiple time scales, *PNAS*, Vol. 104, No. 24, (2007), 10240-10245.

[13] O. Sporns, Networks of the Brain, *MIT Press*, 2010.

[14] Weisstein, Eric W. "Graph Automorphism". From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/GraphAutomorphism.html

[15] Gallian, Joseph A. *Contemporary Abstract Algebra*, 7th edition. Belmont, CA: Brooks/Cole. (2010). ISBN 978-0-547-16509-7.

[16] Weisstein, Eric W. "Vertex-Transitive Graph". From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/Vertex-TransitiveGraph.html

[17] Casey, Michael J. "Atlanta Transit > MARTA Midtown". Web blog post. WordPress, 8 Aug. 2010. Web. June 2012.

[18] M. Miller, B. Shah, A. Harkin, "Community Detection in Networks by Maximizing Partition Efficiency", unpublished

[19] Dr. Narayan, et. al. RIT REU 2013

[20] Weisstein, Eric W. "Partition Function P". From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/PartitionFunctionP.html

[21] Metropolitan Atlanta Rapid Transit Authority, personal communication and www.itsmarta.com

[22] "Atlanta, Georgia". 33°44′53.17″N and 84°23′16.38″W. **Google Earth**. October 16, 2011. June 2012.

[23] Sondow, Jonathan and Weisstein, Eric W. "Harmonic Number". From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/HarmonicNumber.html

[24] F. Topsøe, "Some bounds for the logarithmic function". Copenhagen University (2005).