

BRAESS'S PARADOX IN EXPANDERS

FAN CHUNG, STEPHEN J. YOUNG, AND WENBO ZHAO

ABSTRACT. Expander graphs are known to facilitate effective routing and most real-world networks have expansion properties. At the other extreme, it has been shown that in some special graphs, *removing* certain edges can lead to *more* efficient routing. This phenomenon is known as Braess's paradox and is usually regarded as a rare event. In contrast to what one might expect, we show that Braess's paradox is ubiquitous in expander graphs. Specifically, we prove that Braess's paradox occurs in a large class of expander graphs with continuous convex latency functions. Our results extend previous work which held only when the graph was both denser and random and for random linear latency functions. We identify deterministic sufficient conditions for a graph with as few as a linear number of edges, such that Braess's Paradox almost always occurs, with respect to a general family of random latency functions.

1. INTRODUCTION

Over the last few decades the study of graph expanders has grown from a mere whisper to a widespread phenomenon with diverse consequences, contributing to deep results in number theory, group theory, and geometry (see [24]). However, at its heart the study of expanders continues to be focused on the applications to computer science, including the construction of super concentrators, good error correcting codes, and randomness amplifiers (see [16]).

In the effort to understand the “small world” phenomenon in real world networks, there has been extensive research attempting to capture the universal characteristics of numerous social, technological, and biological networks. The expander property is one of the primary behaviors that has been extensively studied in this context. Some of the many real world networks found to exhibit expansion include the internet, collaboration graphs, social networks, and the power grid, among others (see [27]). In many cases the good behavior of these networks has been tied to expansion properties of the network which serves as the basis for the efficient approximation of the maximum multi-commodity flow problem [21, 23, 34], even with online distributed algorithms [10, 18].

Together these results have formed a core of folklore to the effect that any network which is an expander (or close enough to an expander, in the case of the internet [11, 25]) will automatically behave well under various routing situations. However, this folklore has a potentially significant shortcoming in that it fails to account for the “game theoretic” aspects of routing. More specifically, these results fail to account for the selfish routing of individual units of flow, which has long been known to have an adverse affect on the performance of a network [3, 4]. In fact, our results show that this effect is present even in the supposedly efficient routing on expanders and thus, Braess's paradox must be accounted for in real world routing situations.

To be precise, in 1968 Dietrich Braess made the counter-intuitive observation that there exists a road network such that the selfish routing behavior can be improved by closing a road, even though (or perhaps because) the removed road can be transited quickly [3, 4]. A simplified version of the Braess's original network due to Leonard Schulman appears in the work of Roughgarden [31]. This simplified version of the network considers selfishly routing one unit of flow from s to t in the network in Figure 1. In Figure 1(a), the presence of the edge with zero transit cross allows every “piece” of flow to avoid the constant travel time edges, resulting in an overall latency of 2.

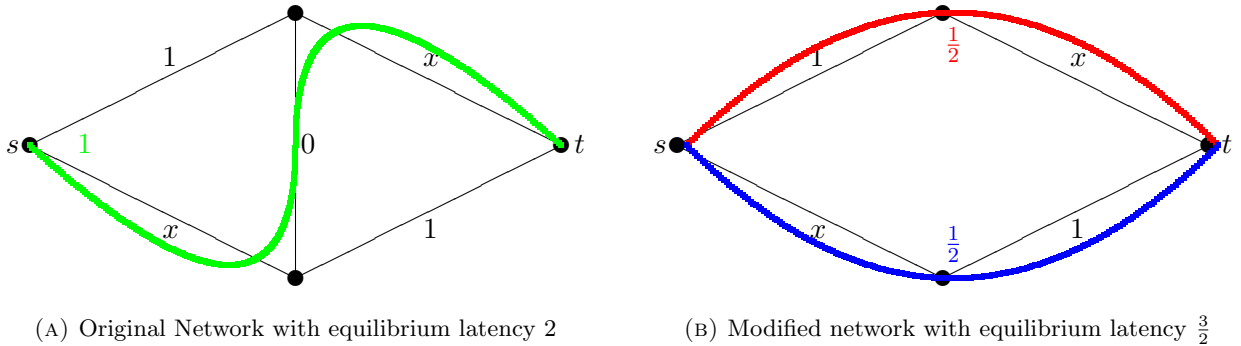


FIGURE 1. Braess's Paradox

In contrast, when the zero transit cost edge is removed, as in Figure 1(b), every “piece” of the flow is forced to use one of the constant travel time edges, yielding an overall latency of $3/2$.

This observation has come to be called *Braess's paradox* and in many ways can be viewed as the start of the modern study of the “Price of Anarchy” [20, 28]. In this context, our results can be informally stated as follows.

Theorem 1 (nontechnical). *If G is a sufficiently good expander with source-sink pair (s, t) with $\deg(s) \approx \deg(t)$, and the latency functions are randomly chosen from a reasonable class of continuous, convex, latency functions, then with high probability there is a subgraph G' and a traffic rate R such that the selfish routing on G' incurs less latency than the selfish routing on G .*

Thus, the exceedingly counter-intuitive, and hence presumably rare, phenomenon of Braess's paradox is anything but rare. In fact because of the widespread occurrence of expanders in numerous real world networks, not only is Braess's paradox not a rare phenomenon, it can occur in the networks we interact with on a daily basis.

1.1. Previous Work. Given the potential real world applicability of Braess's paradox, it is unsurprising that since its discovery a significant amount of work has gone into attempting to understand the paradox, both anecdotally [7, 19] and theoretically [9, 17, 29, 30]. Unfortunately, many of the theoretical arguments regarding Braess's paradox are based on a “worst case” analysis of designer instances, and fail to give insight into the real world occurrences of Braess's paradox. In many ways, these designer instances are captured by the result of Milchtaich which says that for an undirected non-series parallel network G there is a choice of latency functions and flow so that Braess's paradox occurs [26]. Furthermore, if G is series parallel there is no such choice of latency functions and flow. Alternatively, this may be viewed as saying that Braess's paradox occurs if and only if Figure 1 is a topological minor of the network with the given source-sink pair. However, since this result requires the freedom to determine the latency functions on each edge, it is not useful in practically recognizing Braess's paradox. In fact, Roughgarden has shown that (unless $P = NP$) there is not a polynomial time algorithm to approximate the best selfish routing subnetwork to within a factor of $n/2$ [31]. Further, Roughgarden and Tardos showed that even if the latency functions are restricted to be affine, there is not a $4/3$ -approximation algorithm [32]. Although there have been some indications that Braess's paradox can be recognized by limiting to latency functions near the class of strictly affine latency functions [8], the results of Roughgarden and Tardos imply that, in general, it is impossible to recognize when Braess's paradox will occur. Thus, from a real world point of view, the natural question is: how prevalent is Braess's paradox?

Valiant and Roughgarden began to answer this question by showing that Braess's paradox occurs in dense random graphs for reasonably chosen latency functions, that is, in $\mathcal{G}(n, p)$ with $p \in \Omega(n^{-1/2+\zeta})$ [33]. For a similar class of latency functions, Chung and Young, were able to extend this

to $p > c \log(n)/n$ [5]. As Erdős-Rényi graphs have nearly uniform degrees, and in particular are very far from observed real world networks, these previous results have limited practical applications. The goal of this work is to extend these results to a much broader class of latency functions and graphs, and in doing so, capture the behavior of Braess’s paradox on real world networks.

1.2. Comparison with Previous Work. For any $\alpha > 0$ and $\beta \in (0, 1)$ we will say that a graph on n vertices is an (α, β) -expander if for every set of vertices U , $|\Gamma(U) \cup U| \geq \min \left\{ (1 + \alpha) |U|, (1 + \beta) \frac{n}{2} \right\}$, where $\Gamma(U)$ is the collection of vertices adjacent to some vertex in U . That is, we define $\Gamma(U) = \{v \in V \mid \exists u \in U, v \sim u\}$. It can be readily seen that with high probability the standard Erdős-Rényi graph with edge density $p > c \log(n)/n$ is a $(3/5np, 1/4)$ -expander. We will show that Braess’s paradox occurs in the class of (α, β) -expanders which includes graphs with highly skewed degree distributions, such as power-laws. As a consequence, (α, β) -expanders encompass such real world networks as the internet [6], collaboration networks [12], the power grid [35], etc., which is in contrast with previous results which focused on nearly regular graphs [5, 33]. Additionally, our result does not require a global minimum degree condition, only a local degree condition around s and t . In fact, unlike [33] and [5] which require average degree $\Omega\left(n^{\frac{1}{2} + \zeta}\right)$ and $\Omega(\log(n))$, respectively, our results hold for graphs with constant average degree.

Furthermore, by understanding the commutation relationships between convex functions and their inverses, we extend the class of latency functions from the affine functions considered in [5, 33] to a general class of convex latency functions, reflecting the greater diversity of real world latency functions. In this process, we reveal a fundamental commutation relationship which fails to hold for a general concave function, indicating that in a network where the latency per unit of flow is decreasing, Braess’s paradox is unlikely to occur.

Finally, by focusing on specific deterministic properties of the underlying network and latency functions, we are able to decouple the total flow routed from the expansion properties of the network. Specifically, we focus properties such as the expansion, the degrees of s and t , and the parameter κ defined by the neighborhoods of s and t . In contrast to the results of [5, 33] for $\mathcal{G}(n, p)$, where $\alpha, \kappa, \deg(s), \deg(t) \in \Theta(np)$, our generalizations allow these parameters to vary from each other significantly, needing only to satisfy constraints similar to those in [5, 33].

Before formally stating our results we will need to be more explicit about the type of instances we are considering. Specifically, we will impose specific structural properties on the underlying network parametrized by κ , certain “smoothness” conditions on the distribution of latency functions over that graph, and that these properties are sufficiently strong relative to each other. We will refer to instances which satisfy the necessary technical conditions as δ -reasonable instances.

Theorem 1. *Let G be an (α, β) -expander, let s and t be a source sink pair such that $\deg(t) = (1 + o(1)) \deg(s)$, let $\ell = \{\ell_e\}_{e \in E(G)}$ be the latency functions for G . If for any sufficiently small fixed $\delta > 0$, (G, ℓ) is a δ -reasonable instance, then with probability at least $1 - e^{-\Omega(\min\{\kappa, \deg(s)\})}$ there is a traffic rate R such that Braess’s paradox occurs.*

We will defer the full definition of a δ -reasonable instance until Section 2, and instead provide here a large class of δ -reasonable instances in order to illustrate to the broad scope of our results. We first note that κ is a locally defined functions of the degrees which depends on G , s and t , and that $\kappa + 1$ is at least the minimum degree of G . Thus, if G is an (α, β) -expander with minimum degree going to infinity and, for some fixed k , the latency functions are of the form $A_k x^k + A_{k-1} x^{k-1} + \dots + A_1 x + B$ where each A_i is a strictly positive, bounded, continuous random variable and B is a positive, continuous (except at 0) random variable with $\mathbb{P}(B = 0) > 0$, then (G, ℓ) is a δ -reasonable instance for all $\delta > 0$ and Braess’s paradox occurs for some traffic rate on the order of $\deg(s)$. Although this class of instances captures the essential features of δ -reasonable

instances, as we will see in Section 2, the class of δ -reasonable instances is larger than this class, and in particular includes graphs with constant average degree.

Broadly speaking our proof follows the techniques previously employed by Valiant and Roughgarden in [33] and the first two authors in [5]. Specifically, in Section 3, we use the expansion properties of the network to show that latency cost of traversing the central portion of the network is essentially zero, much like the removed edge in the example depicted in Figure 1. The structural properties necessary for this proof are similar to those in [33] in that the β expansion property assures that there are $\Theta(n)$ edge disjoint paths of length at most two between sufficiently large sets, much as there are many edge disjoint paths of length 2 in $\mathcal{G}(n, n^{-1/2+\zeta})$ when $\zeta > 0$. The α expansion property captures the expansion properties necessary to iteratively grow these large sets in a similar manner as in [5]. The observation that the latency cost of traversing the central part of the network may alternatively be thought of as the observation that the latency distances from s are concentrated about given value. In Section 4 we give an implicit formula for this value. Combining the observation that the latency distance is essentially concentrated with some additional balance considerations (in particular that $\deg s = (1 + o(1)) \deg t$), allows the whole network to be treated similarly to a giant “fuzzy” instance of the example depicted in Figure 1. From this viewpoint, in Section 5 we analyze a specific class of subgraphs to prove Theorem 1. We also provide an infinite class of examples to illustrate the importance of the condition $\deg(t) = (1 + o(1)) \deg(s)$.

2. PRELIMINARIES

2.1. Nash Equilibrium Flows. In order to more formally define “selfish routing”, we consider a single commodity flow on an undirected network G with a specified source-sink pair (s, t) and latency functions $\ell = \{\ell_e\}_{e \in E(G)}$. Let \mathcal{P} be the set of simple paths from s to t . A flow f on G from s to t is defined as a function $f: \mathcal{P} \rightarrow \mathbb{R}^{\geq 0}$. The flow on an edge, denoted f_e , is $\sum_{\{P \in \mathcal{P} | e \in E(P)\}} f(P)$. We will say an edge e is flow carrying if $f_e > 0$ and a vertex is flow carrying if it is adjacent to a flow carrying edge. The latency of a path P with respect to a flow f is then $\ell_f(P) = \sum_{e \in E(P)} \ell_e(f_e)$. It is important to note that the latency of a path depends on the entire flow function, not just the value of the flow along a given path.

We wish to consider specific classes of flows that capture the idea of “selfish routing.” One of the ways of quantifying this is to insist that the flow is at Nash equilibrium, that is, the flow f is such that for any two $P_1, P_2 \in \mathcal{P}$, $\ell_f(P_1) \leq \ell_f(P_2)$ whenever $f(P_1) > 0$. Notice that this implies that if P_1 and P_2 are flow carrying paths, then $\ell_f(P_1) = \ell_f(P_2)$ and thus no “unit” of flow has any incentive to change its routing. And so, given a traffic rate R and a Nash equilibrium flow f , we can define $\mathcal{L}(G, \ell, R, f)$ as the common latency of all flow carrying paths at Nash equilibrium. We note that this value has been shown to exist and be unique [2] and so the dependence on f can be dropped. With this notation, Braess’s paradox is the observation that there exists a selfish routing instance (G, ℓ, R) and a subgraph G' of G such that $\mathcal{L}(G', \ell, R) < \mathcal{L}(G, \ell, R)$. For a given instance,

$$\rho = \max_{G' \subset G} \frac{\mathcal{L}(G, \ell, R)}{\mathcal{L}(G', \ell, R)}$$

is referred to as the Braess ratio of the instance.

In order to prove our results we will exploit some known structural properties of Nash equilibrium flows, and so to that end, we collect a few such results which will be important to our purposes in the following lemma.

Lemma 2. *Given an instance (G, ℓ, R) and a Nash equilibrium flow f , define for each vertex v , $d_s(v)$ as the shortest path from v to s with respect to the latencies $\ell(f)$. Define $d_t(v)$ analogously. The following properties then hold for all Nash equilibrium flows f .*

- (1) If f is a Nash equilibrium flow for traffic rate R on the network G with latencies ℓ , then for every vertex v we have $d_s(v) + d_t(v) \geq \mathcal{L}(G, \ell, R)$ with equality if v is a flow carrying vertex.
- (2) [31] If f is a Nash equilibrium flow achieving traffic rate R for the instance (G, ℓ) , then for all edges $e = \{u, v\}$, $|d_s(v) - d_s(u)| \geq \ell_e(f_e)$ with equality whenever $f_e > 0$.
- (3) [13, 22] For every network G and continuous, strictly increasing latency functions ℓ , $\mathcal{L}(G, \ell, R)$ is continuous and strictly increasing function of R .
- (4) [31] There is a Nash equilibrium flow f so that the set of edges with $f_e > 0$ is acyclic when considered as a directed graph.

It will be useful to note that by part (3) of this lemma, for any sufficiently large \mathcal{L} and fixed G and ℓ , there is a solution to $\mathcal{L} = \mathcal{L}(G, \ell, R)$, since the overall latency is continuous and strictly increasing. Thus we define the function $R_G(\mathcal{L})$, which is the unique traffic rate R such that the instance (G, ℓ, R) has latency \mathcal{L} . If the underlying network is clear from context, we will denote this as $R(\mathcal{L})$.

Additionally, at various times we will need to use Chernoff and Hoeffding bounds in order to concentrate various random variables around their expectation. For the sake of concreteness we provide the two specific forms in the following lemma.

Lemma 3 (Chernoff-Hoeffding Bounds [1, 15]). *Let X_i be a collection of independent random variables such that $a_i \leq X_i \leq b_i$. Define $X = \sum_i X_i$, then for all $\epsilon \in (0, 1)$*

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq \epsilon \mathbb{E}[X]) \leq 2e^{-\frac{2\epsilon^2 \mathbb{E}[X]^2}{\sum_i (b_i - a_i)^2}}.$$

If in addition the X_i are identically distributed with $X_i \in \{0, 1\}$ and, then

$$\mathbb{P}(X - \mathbb{E}[X] \geq \epsilon \mathbb{E}[X]) \leq e^{-\frac{\epsilon^2 \mathbb{E}[X]}{3}} \quad \text{and} \quad \mathbb{P}(X - \mathbb{E}[X] \leq -\epsilon \mathbb{E}[X]) \leq e^{-\frac{\epsilon^2 \mathbb{E}[X]}{2}}.$$

2.2. δ -Reasonable Instances. Our characterization of an instance as a δ -reasonable instance can be thought of as controls on three properties, the structural and expansion properties of the underlying network, the “smoothness” and “compactness” properties of the distribution of latency functions, and a restriction on the quantitative relationships between these properties.

Often we will abuse notation and use a distribution as a placeholder for a random variable distributed according to that distribution. Now, inspired by the work of Valiant and Roughgarden [33] and building on our previous work [5], we will say a pair of independent distributions $(\mathcal{H}, \mathcal{B})$ form a reasonable pair of distributions if

- $\mathbb{P}(\mathcal{H} \in \mathcal{H}^*) = 1$, where \mathcal{H}^* is the set of convex, strictly increasing functions h such that $h(0) = 0$,
- there exist $H_{\min}, H_{\max} \in \mathcal{H}^*$ such that $H_{\max}(0) = H_{\min}(0) = 0$ and $\mathbb{P}(H_{\min} \leq \mathcal{H} \leq H_{\max}) = 1$,
- there exist $H_1, H_2 \in \mathcal{H}^*$ where $H_1(0) = H_2(0) = 0$ and $H_1 < H_2$ otherwise, such that for all $\epsilon > 0$, $\mathbb{P}((1 - \epsilon)H_1 \leq \mathcal{H} \leq H_1) > 0$ and $\mathbb{P}((1 - \epsilon)H_2 \leq \mathcal{H} \leq H_2) > 0$, and
- there is some interval $I_B = [0, \nu]$ with $\nu > 0$ such that for every non-trivial subinterval $J \subseteq I_B$, $\mathbb{P}(\mathcal{B} \in J) > 0$.

Although the conditions on \mathcal{H} may seem unnatural at first glance, it includes, as we have noted previously, a large natural class of random convex functions. Specifically, if we have continuously distributed positive random variables A_1, \dots, A_k , then the distribution over polynomials $A_1x + A_2x^2 + \dots + A_kx^k$ satisfies the conditions for \mathcal{H} . As a consequence, these reasonable distributions can be viewed as a strict generalization of the reasonable distributions of Valiant and Roughgarden [33].

Given a graph G we will say that it has reasonable latency functions if the latency on each edge e is $h_e(x) + b_e$ where (h_e, b_e) is independently generated according to a reasonable pair of distributions $(\mathcal{H}, \mathcal{B})$. For later convenience of notation, we will assume without loss of generality that for all $h \in \mathcal{H}$, $h(x) = 0$ for any $x \leq 0$. If in addition, the graph G is an (α, β) -expander, we will say the instance (G, ℓ) is a δ -reasonable instance if

- For some $\nu > 0$, \mathcal{B} has a continuous distribution on $[0, \nu]$ except potentially at 0,
- $\alpha \mathbb{P}(\mathcal{B} \leq \delta/\log(n)) > 4$,
- $\kappa = \min_{v \in \Gamma(s) \cup \Gamma(t)} \deg(v) - 1$ is such that κ is $\omega(1)$,
- $\min\{\deg(s), \deg(t)\}$ is $o(\alpha \kappa \mathbb{P}(\mathcal{B} \leq \delta) \mathbb{P}(\mathcal{B} \leq \delta/\log(n)))$, and

Although not all of the proofs that will follow need the full power of δ -reasonable instances, for simplicity we will assume a δ -reasonable instance for all subsequent results since the entirety of the proof of the occurrence of Braess's paradox requires δ -reasonable instances. We note as well that the additional constraints imposed on $(\mathcal{H}, \mathcal{B})$ by being a δ -reasonable instance can be satisfied by assuming that \mathcal{B} has some positive probability of being 0, and otherwise is a continuous distribution over some range of positive values.

2.3. Braess Subgraphs. Fundamental to recognizing Braess's paradox is the construction of the underlying subgraph with better selfish routing behavior. In order to simplify our search for an improving subgraph we restrict ourselves to the class of subgraphs detailed below.

If (G, ℓ) is an instance with reasonable latency functions, for constants $c, \epsilon > 0$ we define by $G'(c, \epsilon)$ the random subgraph created by the following procedure which is similar to the one originally appearing in Valiant and Roughgarden [33]. For any subset $S \subseteq V(G)$, we will denote by $G[S]$ the subgraph of G induced by S . We will partition $V(G) - \{s, t\}$ into three sets of vertices V_{X1} , V_{1X} and V_U and then $G'(c, \epsilon)$ will be $G[V_{X1} \cup \{s, t\}] \cup G[V_{1X} \cup \{s, t\}] \cup G[V_U \cup \{s, t\}]$. That is, $G'(c, \epsilon)$ will have the same vertex set as G and the subset of edges that are either adjacent to s or t , together with edges that are entirely contained in one of V_{1X} , V_{X1} , or V_U . It will be convenient to be able to discuss the graphs generated by these sets as well, thus we define $G'_S(c, \epsilon)$ as the graph $G[\{s, t\} \cup V_S]$. Now define an edge e adjacent to s or t as a 1-type edge if its latency function $h_e(x) + b_e$ is such that $(1 - \epsilon)H_1 \leq h_e \leq H_1$ and $c \leq b_e \leq (1 + \epsilon)c$ and let p_1 be the probability that an edge is 1-type. Similarly we say that an edge is X -type if it is adjacent to s or t and its latency function is such that $(1 - \epsilon)H_2 \leq h_e \leq H_2$ and $0 \leq b_e \leq \epsilon c$ and let the probability of an edge being X -type be p_X .

For any vertex in $\Gamma(s) \cup \Gamma(t) - \Gamma(s) \cap \Gamma(t)$, if it is adjacent to s via a 1-type edge assign it to V_{1X} with probability p_X and to V_U otherwise. If it is adjacent to s via a X -type edge assign it to V_{X1} with probability p_1 and V_U otherwise. Conversely, if it is adjacent to t , assign it to V_{1X} if it is via an X -type edge with probability p_1 and if it is via a 1-type edge with probability p_X . Otherwise, assign it to V_U .

Now consider the vertices in $\Gamma(s) \cap \Gamma(t)$. Those vertices which are adjacent to s via a 1-type edge and t via an X -type edge are assigned to V_{1X} , while those edges that are adjacent to s via an X -type edge and t via a 1-type edge are assigned to V_{X1} . The remaining vertices in $\Gamma(s) \cap \Gamma(t)$ are assigned to V_U .

The remaining vertices are assigned with probability $p_1 p_X$ to V_{1X} , with probability $p_1 p_X$ to V_{X1} and with probability $1 - 2p_1 p_X$ to V_U . For an illustration of this distribution see Figure 2. Note that if G has $n + 2$ vertices the expected sizes of V_{1X} and V_{X1} are $p_1 p_X n$ and the expected size of V_U is $(1 - 2p_1 p_X)n$.

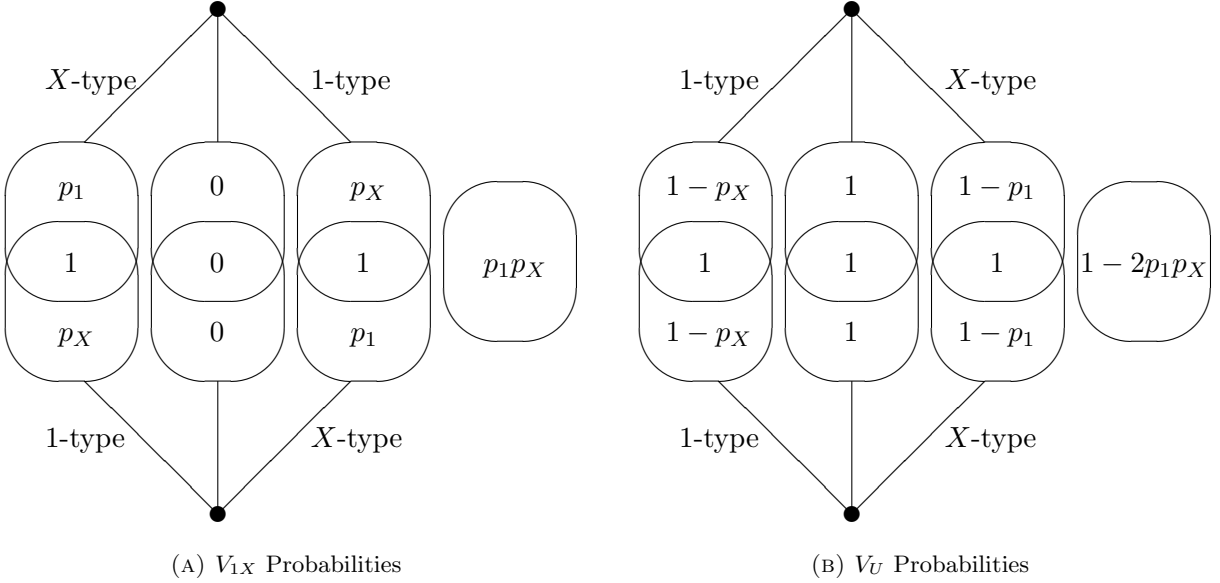


FIGURE 2. Illustration of probability of belonging to a given set

Lemma 4. *If $(\mathcal{H}, \mathcal{B})$ is a reasonable distribution, then for any $c \leq \nu$ and sufficiently small $\epsilon, \epsilon' > 0$,*

$$\begin{aligned}
(1 - \epsilon')p_1 p_X n &\leq |V_{1X}|, |V_{X1}| &&\leq (1 + \epsilon')p_1 p_X n, \\
(1 - \epsilon')p_1 p_X \deg(s) &\leq |V_{1X} \cap \Gamma(s)|, |V_{X1} \cap \Gamma(s)| &&\leq (1 + \epsilon')p_1 p_X \deg(s), \\
(1 - \epsilon')p_1 p_X \deg(t) &\leq |V_{1X} \cap \Gamma(t)|, |V_{X1} \cap \Gamma(t)| &&\leq (1 + \epsilon')p_1 p_X \deg(t), \\
(1 - \epsilon')(1 - 2p_1 p_X)n &\leq |V_U| &&\leq (1 + \epsilon')(1 - 2p_1 p_X)n, \\
(1 - \epsilon')(1 - 2p_1 p_X) \deg(s) &\leq |V_U \cap \Gamma(s)| &&\leq (1 + \epsilon')(1 - 2p_1 p_X) \deg(s), \quad \text{and} \\
(1 - \epsilon')(1 - 2p_1 p_X) \deg(t) &\leq |V_U \cap \Gamma(t)| &&\leq (1 + \epsilon')(1 - 2p_1 p_X) \deg(t),
\end{aligned}$$

with probability at least $1 - \mathcal{O}(e^{-\min\{\deg(s), \deg(t)\}})$.

Proof. We will only prove the results for V_{1X} as all the other arguments are analogous. We consider the vertices in four classes, $\Gamma(s) - \Gamma(t)$, $\Gamma(t) - \Gamma(s)$, $\Gamma(s) \cap \Gamma(t)$, and the remaining vertices. It suffices to show that for each of these four classes the probability of belonging to V_{1X} is independent and equal to $p_1 p_X$, and the result follows by Chernoff bounds.

Now we note that the probability of a vertex being in V_{1X} depends only on its adjacencies to s and t and thus no vertex depends on the same collection of edges, and thus the assignments to V_{1X} are independent. Consider then a vertex in $\Gamma(s) - \Gamma(t)$. Such a vertex is in V_{1X} only if the edge adjacent to s is a 1-type edge, which occurs with probability p_1 . Then with probability p_X this vertex is in V_{1X} , and so the total probability of a vertex in $\Gamma(s) - \Gamma(t)$ being in V_{1X} is $p_1 p_X$. A similar argument holds for a vertex in $\Gamma(t) - \Gamma(s)$. Now if a vertex is in $\Gamma(s) \cap \Gamma(t)$, it is V_{1X} if and only if the edge adjacent to s is 1-type and the edge adjacent to t is X -type, which occurs with probability $p_1 p_X$. Finally, if the vertex is not adjacent to s or t , then the probability it is in V_{1X} is $p_1 p_X$ by construction. \square

3. CONCENTRATION OF LATENCY

If we were to fix a graph G and a traffic rate R (or overall latency \mathcal{L}), then for any vertex v , $d_s(v)$ can be thought of as a random variable depending on the choices of latency function. In this context, the following lemma can be thought of as saying that for flow carrying vertices $d_s(v)$ concentrates around a particular (unknown) value.

Lemma 5. *Let G be an (α, β) -expander on n vertices, let s and t be a source sink pair, and let $(\mathcal{H}, \mathcal{B})$ be reasonable distributions. For any fixed $\delta > 0$, if (G, ℓ) is a δ -reasonable instance and \mathcal{L} is constant, then for any two flow carrying vertices u, v other than s and t , we have that $|d_s(u) - d_s(v)| \leq 7\delta$ and $|d_t(u) - d_t(v)| \leq 7\delta$ with probability at least $1 - e^{-\Omega(\kappa)}$, where $\kappa = \min_{v \in \Gamma(s) \cup \Gamma(t) - \{s, t\}} \deg v - 1$.*

Proof. Without loss of generality we may assume that $\deg(s) \leq \deg(t)$. Now by the properties of the Nash equilibrium flow from Lemma 2, for every flow carrying vertex v , $d_s(v) + d_t(v) = \mathcal{L}$ and thus it suffices to prove that $|d_s(u) - d_s(v)| \leq 7\delta$ for any pair of flow carrying vertices. Let v_s be the flow carrying vertex (other than s) which minimizes $d_s(v_s)$ and similarly let v_t be the flow carrying vertex (other than t) which maximizes $d_s(v_t)$ (and hence minimizes $d_t(v_t)$). Since the latency functions on a flow carrying edge are strictly positive, $\{s, v_s\}, \{t, v_t\} \in E(G)$. Furthermore, since $d_s(v_s) \leq \mathcal{L}$, the flow on the edge $\{s, v_s\}$ is at most $H_{\min}^{-1}(\mathcal{L})$ and thus the total flow R is at most $\deg(s)H_{\min}^{-1}(\mathcal{L})$ and the flow out of v_s is at most $H_{\min}^{-1}(d_s(v_s)) \leq H_{\min}^{-1}(\mathcal{L})$.

By Chernoff bounds, with probability at least $1 - e^{-\frac{\kappa \mathbb{P}(\mathcal{B} \leq \delta)}{8}}$ there are at least $\frac{\kappa \mathbb{P}(\mathcal{B} \leq \delta)}{2}$ vertices $v \in V(G) - \{s\}$ such that $b_{\{v, v_s\}} \leq \delta$. By the pigeonhole principle at most half of these edges have flow more than $\frac{4H_{\min}^{-1}(\mathcal{L})}{\mathbb{P}(\mathcal{B} \leq \delta)\kappa}$. Thus there are at least $\frac{\kappa}{4\mathbb{P}(\mathcal{B} \leq \delta)} + 1$ vertices v such that $d_s(v) \leq d_s(v_s) + H_{\max} \left(\frac{4H_{\min}^{-1}(\mathcal{L})}{\mathbb{P}(\mathcal{B} \leq \delta)\kappa} \right) + \delta$. Thus define $c_0 = d_s(v_s) + H_{\max} \left(\frac{4H_{\min}^{-1}(\mathcal{L})}{\mathbb{P}(\mathcal{B} \leq \delta)\kappa} \right) + \delta$ and $U_0 = \{v \mid d_s(v) \leq c_0\}$.

We will now inductively define a sequence U_i and c_i , such that $U_i \subset U_{i+1}$, $G[U_i]$ is connected, and $c_i < c_{i+1}$, stopping when $|\Gamma(U_i) \cup U_i| \geq (1 + \beta)n/2$. Furthermore, each U_i will be defined as $\{v \in V(G) \mid d_s(v) \leq c_i\}$. We note that since each U_i will be connected, $|\Gamma(U_i) \cup U_i| = |\Gamma(U_i)|$. Thus, suppose c_i and U_i are defined and $|\Gamma(U_i)| < (1 + \beta)n/2$ and let $\gamma = \delta/\log(n)$. Now consider the vertices in $\Gamma(U_i) \setminus U_i$. With probability at least $1 - e^{-\frac{\kappa \mathbb{P}(\mathcal{B} \leq \gamma) \alpha |U_i|}{8}}$ there are at least $\frac{\alpha \mathbb{P}(\mathcal{B} \leq \gamma)}{2} |U_i|$ vertices in $V(G) \setminus U_i$ that are adjacent to U_i across an edge e with $b_e \leq \gamma$. Let U'_i be the set of such vertices and let E_i be a set of edges witnessing membership in U'_i . That is, for each $v \in U'_i$ there is a unique edge $e \in E_i$ such that $b_e \leq \gamma$ and $e \in \{v\} \times U_i$. Now since U_i and its complement form an $s - t$ cut in G and all the latency functions are zero only if there is zero flow on the edge, there are precisely R units of flow crossing the cut. But then, at most half of the edges in E_i have flow greater than $\frac{4R}{\mathbb{P}(\mathcal{B} \leq \gamma) \alpha |U_i|}$, and thus at least half of the vertices $v \in U'_i$ are such that

$$d_s(v) \leq c_i + H_{\max} \left(\frac{4R}{\mathbb{P}(\mathcal{B} \leq \gamma) \alpha |U_i|} \right) + \gamma.$$

Thus define $c_{i+1} = c_i + H_{\max} \left(\frac{4R}{\mathbb{P}(\mathcal{B} \leq \gamma) \alpha |U_i|} \right) + \gamma$ and $U_{i+1} = \{v \in V(G) \mid d_s(v) \leq c_{i+1}\}$. By the above, we have that

$$|U_{i+1}| \geq |U_i| + \frac{1}{4} \mathbb{P}(\mathcal{B} \leq \gamma) \alpha |U_i| = \left(1 + \frac{1}{4} \mathbb{P}(\mathcal{B} \leq \gamma) \alpha \right) |U_i| \geq \left(1 + \frac{1}{4} \mathbb{P}(\mathcal{B} \leq \gamma) \alpha \right)^{i+1} |U_0|.$$

Now let i^* be the first i such that $|\Gamma(U_i)| \geq (1 + \beta)n/2$. Thus

$$i^* \leq \frac{\log \left(\frac{(1+\beta)n}{2|U_0|} \right)}{\log \left(1 + \frac{1}{4} \mathbb{P}(\mathcal{B} \leq \gamma) \alpha \right)} \leq \frac{\log \left(\frac{(1+\beta)n}{\kappa \mathbb{P}(\mathcal{B} \leq \delta)} \right)}{\log \left(1 + \frac{1}{4} \mathbb{P}(\mathcal{B} \leq \gamma) \alpha \right)}.$$

For sufficiently large n , since $\kappa \rightarrow \infty$ and $\alpha\mathbb{P}(\mathcal{B} \leq \gamma) > 4$, $i^* \leq \log(n)$. Thus

$$\begin{aligned}
c_{i^*} &\leq c_0 + \gamma i^* + \sum_{i=0}^{i^*-1} H_{\max} \left(\frac{4R}{\mathbb{P}(\mathcal{B} \leq \gamma) \alpha |U_i|} \right) \\
&\leq c_0 + \frac{\delta}{\log(n)} \log(n) + \sum_{i=0}^{i^*-1} H_{\max} \left(\frac{4R}{\mathbb{P}(\mathcal{B} \leq \gamma) \alpha (1 + \frac{1}{4}\mathbb{P}(\mathcal{B} \leq \gamma) \alpha)^i |U_0|} \right) \\
&\leq c_0 + \delta + \sum_{i=0}^{i^*-1} \left(1 + \frac{1}{4}\mathbb{P}(\mathcal{B} \leq \gamma) \alpha \right)^{-i} H_{\max} \left(\frac{4R}{\mathbb{P}(\mathcal{B} \leq \gamma) \alpha |U_0|} \right) \\
&\leq c_0 + \delta + H_{\max} \left(\frac{4R}{\mathbb{P}(\mathcal{B} \leq \gamma) \alpha |U_0|} \right) \sum_{i=0}^{\infty} 2^{-i} \\
&\leq d_s(v_s) + H_{\max} \left(\frac{4H_{\min}^{-1}(\mathcal{L})}{\mathbb{P}(\mathcal{B} \leq \delta) \kappa} \right) + 2\delta + 2H_{\max} \left(\frac{16R}{\mathbb{P}(\mathcal{B} \leq \gamma) \alpha \mathbb{P}(\mathcal{B} \leq \delta) \kappa} \right),
\end{aligned}$$

where the third inequality follows from observing that for $\lambda > 0$, $H_{\max}((1 + \lambda)x) \leq (1 + \lambda)H_{\max}(x)$ by the convexity of H_{\max} and $H_{\max}(0) = 0$. Now since $\mathbb{P}(\mathcal{B} \leq \gamma) \alpha \kappa / \deg(s) \rightarrow \infty$ and $\kappa \rightarrow \infty$, for sufficiently large n this is at most $d_s(v_s) + 7/3\delta$.

Noting that

$$\begin{aligned}
\sum_{i=0}^{i^*-1} e^{-\frac{\mathbb{P}(\mathcal{B} \leq \gamma) \alpha |U_i|}{8}} &\leq \sum_{i=0}^{i^*-1} e^{-\frac{\mathbb{P}(\mathcal{B} \leq \gamma) \alpha (1 + \frac{1}{4}\mathbb{P}(\mathcal{B} \leq \gamma) \alpha)^i \mathbb{P}(\mathcal{B} \leq \delta) \kappa}{32}} \leq \sum_{i=0}^{i^*-1} e^{-\frac{\mathbb{P}(\mathcal{B} \leq \gamma) \alpha 2^i \mathbb{P}(\mathcal{B} \leq \delta) \kappa}{32}} \\
&\leq \left(\sum_{i=0}^{i^*-1} e^{-2^i} \right)^{\frac{\mathbb{P}(\mathcal{B} \leq \delta) \mathbb{P}(\mathcal{B} \leq \gamma) \kappa \alpha}{32}} \leq \left(\sum_{i=0}^{\infty} e^{-2^i} \right)^{\frac{\mathbb{P}(\mathcal{B} \leq \delta) \mathbb{P}(\mathcal{B} \leq \gamma) \kappa \alpha}{32}} \leq \left(\frac{3}{5} \right)^{\frac{\mathbb{P}(\mathcal{B} \leq \delta) \kappa}{8}}
\end{aligned}$$

And thus, the probability of this procedure failing is at most $e^{-\Omega(\kappa)}$. So this construction procedure succeeds with probability at least $1 - e^{-\Omega(\kappa)} \rightarrow 1$, since $\kappa \gg 1$.

In a similar way we can define $c_j^* = d_s(v_t) - H_{\max} \left(\frac{4H_{\min}^{-1}(\mathcal{L})}{\mathbb{P}(\mathcal{B} \leq \delta) \kappa} \right) - 2\delta - 2H_{\max} \left(\frac{16R}{\mathbb{P}(\mathcal{B} \leq \gamma) \alpha \mathbb{P}(\mathcal{B} \leq \delta) \kappa} \right) \geq d_s(v_t) - 7/3\delta$ and $V_{j^*} = \{v \in V(G) \mid d_s(v) \geq c_{j^*}^*\}$ and have that $|\Gamma(V_{j^*})| \geq (1 + \beta) \frac{n}{2}$. Without loss of generality we may assume that $U_{i^*} \cap V_{j^*} = \emptyset$ and thus, as $|\Gamma(U_{i^*})| + |\Gamma(V_{j^*})| \geq (1 + \beta)n$, there are at least $\frac{\beta}{2}n$ edge disjoint paths of length at most 2 between U^* and V^* . Hence, with probability at least $1 - e^{-\frac{\mathbb{P}(\mathcal{B} \leq \delta)^2 \beta n}{16}}$, there are at least $\beta/4n$ of these paths such that all edges e on the path have $b_e \leq \delta$. But then at most half of these paths carry more than $8R/\beta n$ units of flow, implying that $c_{j^*}^* - c_{i^*} \leq 2\delta + H_{\max} \left(\frac{8R}{\beta n} \right)$, which for sufficiently large n is at most $7/3\delta$. Thus, $d_s(v_t) - d_s(v_s) \leq 7\delta$, as desired. \square

In order to understand the flow structure of the Nash equilibrium flow for a δ -reasonable instance, it is also necessary to understand the range of potential latencies distances of the non-flow carrying vertices. The following corollary shows that the distinction between flow carrying and non-flow carrying vertices is essentially meaningless for $d_s(\cdot)$. Again, thinking in terms of $d_s(v)$ as a random variable, this is saying that in some sense this random variable is independent of whether the vertex v is flow carrying (which can also be viewed as a random variable). As a result we have the following corollary.

Corollary 6. *Let G be an (α, β) -expander on n vertices, let s and t be a source sink pair, and let $(\mathcal{H}, \mathcal{B})$ be a pair of reasonable distributions. For any fixed $\delta > 0$, if (G, ℓ) is a δ -reasonable instance and \mathcal{L} is constant, then for any two vertices u, v , other than s and t , we have that $|d_s(u) - d_s(v)| \leq 12\delta$ and $|d_t(u) - d_t(v)| \leq 12\delta$ with probability at least $1 - e^{-\Omega(\kappa)}$.*

Proof. Again, we assume without loss of generality that $\deg(s) \leq \deg(t)$ and let v'_s and v'_t be the vertices (other than s and t) which minimize and maximize $d_s(\cdot)$, respectively. Now by a slight modification of the argument used in the proof of Lemma 5 there is a path of latency at most 5δ from v'_s to the set of flow carrying vertices, and similarly for v'_t . Let u_s be the end of the path from v'_s which is flow carrying, and similarly let u_t be the end of the path from v'_t which is flow carrying. Thus $d_t(v'_s) \leq 5\delta + d_t(u_s) = 5\delta + \mathcal{L} - d_s(u_s)$. Furthermore, since $d_s(v'_s) + d_t(v'_s) \geq \mathcal{L}$, we have $\mathcal{L} - d_s(v'_s) \leq 5\delta + \mathcal{L} - d_s(u_s)$. Now $d_s(v'_t) \leq \mathcal{L} + d_s(u_t)$ and thus combining the inequalities we have

$$\mathcal{L} + d_s(v'_t) - d_s(v'_s) \leq 5\delta + \mathcal{L} + d_s(u_t) - d_s(u_s).$$

But by assumption $d_s(v'_t) \geq d_s(v'_s)$, and so since u_s and u_t are flow carrying $|d_s(v'_t) - d_s(v'_s)| \leq 12\delta$, completing the proof. \square

We also control the behavior of the latency distances in $(G'_{1X}(c, \epsilon), \ell, R(\mathcal{L}))$, $(G'_{X1}(c, \epsilon), \ell, R(\mathcal{L}))$, and $(G'_U(c, \epsilon), \ell, R(\mathcal{L}))$ by the following corollary.

Corollary 7. *Let G be an (α, β) -expander on n vertices, let s and t be a source sink pair, and let $(\mathcal{H}, \mathcal{B})$ be a pair of reasonable distributions. For any fixed $\epsilon, \delta > 0$, if (G, ℓ) is a δ -reasonable instance, \mathcal{L} is constant, and $p = \min\{p_1 p_X, 1 - 2p_1 P_X\}$, then for any two vertices u, v in any of the instances $(G'_{1X}(c, \epsilon), \ell, R(\mathcal{L}))$, $(G'_{X1}(c, \epsilon), \ell, R(\mathcal{L}))$, or $(G'_U(c, \epsilon), \ell, R(\mathcal{L}))$, other than s and t , we have that $|d_s(u) - d_s(v)| \leq \left(8 + \frac{4}{\log(1+p)}\right) \delta$ and $|d_t(u) - d_t(v)| \leq \left(8 + \frac{4}{\log(1+p)}\right) \delta$ with probability at least $1 - e^{-\Omega(\kappa)}$.*

This result follows immediately by the combined proof of Lemma 5 and Corollary 6 and the observation that every vertex in any of the three graphs is present independently with probability at least p . Specifically, in all of the growth stages instead of considering only the vertices which are adjacent via an edge with small constant term, we also limit to those that are in the correct graph. At this point it is worth noting that the condition that $\alpha \mathbb{P}\left(\mathcal{B} \leq \frac{\delta}{\log(n)}\right) > 4$ is not tight, and can be replaced by $\alpha \mathbb{P}\left(\mathcal{B} \leq \frac{\delta}{\log(n)}\right) > c > 0$ for some fixed constant c .

4. “EXPECTATION” OF LATENCY DISTANCE

Now if $d_s(v)$ is viewed as a random variable which, by the above, concentrates about some value then the natural question is what value? That is, what is serving the role of “ $\mathbb{E}[d_s(v)]$ ”? As we will see below, this is a more difficult question than usual when analyzing a random variable. In fact, in order to find this value even in a limited case, we need to following two technical lemmas.

Lemma 8. *If \mathcal{H} is part of a reasonable distribution pair, then for any fixed constant T , the function $\mathbb{E}[\mathcal{H}^{-1}](x) = \int h^{-1}(x) d\mathcal{H}(h)$ exists and is continuous, concave, and strictly increasing on $[0, T]$.*

Proof. First we note that since for any fixed $x \in [0, T]$, $H_{\max}^{-1}(x) \leq h^{-1}(x) \leq H_{\min}^{-1}(x)$, the integral converges and thus there is a limit function. Now consider a sequence of functions h_1, \dots, h_k chosen from \mathcal{H} , and let $H_k^{-1} = -\frac{1}{k} \sum_{i=1}^k h_k^{-1}$. Each H_k^{-1} is a convex function, and furthermore, by the law of large numbers for all $x \in [0, T]$, $H_k^{-1}(x) \rightarrow -\mathbb{E}[\mathcal{H}^{-1}](x)$. That is, H_k^{-1} converges point-wise to $-\mathbb{E}[\mathcal{H}^{-1}]$ and thus $-\mathbb{E}[\mathcal{H}^{-1}]$ is convex [14]. Furthermore, this implies that the convergence is uniform. But since each of the H_k^{-1} are continuous, the uniform convergence implies that $\mathbb{E}[\mathcal{H}^{-1}]$ is also continuous.

Let $x < y$, then

$$\mathbb{E} [\mathcal{H}^{-1}](y) - \mathbb{E} [\mathcal{H}^{-1}](x) = \int h^{-1}(y) - h^{-1}(x) d\mathcal{H}(h) \geq \int_{(1-\epsilon)H_2}^{H_2} h^{-1}(y) - h^{-1}(x) d\mathcal{H}(h).$$

But now since $(1 - \epsilon)H_2 \leq h \leq H_2$ this implies that $H_2^{-1}(x) \leq h^{-1}(x) \leq H_2^{-1}(\frac{x}{1-\epsilon})$. Thus, for sufficiently small $\epsilon > 0$, $\frac{x}{1-\epsilon} \leq \frac{x+y}{2}$, and thus $h^{-1}(x) \leq \mathcal{H}_2^{-1}(\frac{x+y}{2})$. Hence $\mathbb{E} [\mathcal{H}^{-1}](y) - \mathbb{E} [\mathcal{H}^{-1}](x) \geq \mathbb{P}((1 - \epsilon)H_2 \leq \mathcal{H} \leq H_2) (H_2^{-1}(y) - H_2^{-1}(\frac{x+y}{2})) > 0$. \square

Lemma 9. *If $(\mathcal{H}, \mathcal{B})$ is a reasonable pair of distributions where \mathcal{B} has a continuous distribution on $[0, \nu]$ except possibly at 0, then for any $\deg(s)$, $\deg(t)$, and $\mathcal{L} \leq \nu$, there is a unique value $c_s = c_s(\mathcal{H}, \mathcal{B}, \deg(s), \deg(t), \mathcal{L})$ such that*

$$\deg(s) \int \int_0^{c_s} h^{-1}(c_s - b) d\mathcal{B}(b) d\mathcal{H}(h) = \deg(t) \int \int_0^{\mathcal{L} - c_s} h^{-1}(\mathcal{L} - c_s - b) d\mathcal{B}(b) d\mathcal{H}(h).$$

Proof. First note by Lemma 8, the function $\mathbb{E} [\mathcal{H}^{-1}]$ exists, and is concave and continuous. Thus by rearranging the order of integration it suffices to show that there is a solution to

$$\deg(s) \int_0^{c_s} \mathbb{E} [\mathcal{H}^{-1}](c_s - b) d\mathcal{B}(b) = \deg(t) \int_0^{\mathcal{L} - c_s} \mathbb{E} [\mathcal{H}^{-1}](\mathcal{L} - c_s - b) d\mathcal{B}(b).$$

Now since \mathcal{B} has a continuous distribution except potentially at 0 and $\mathbb{E} [\mathcal{H}^{-1}]$ is continuous, both sides of this equation are continuous functions of c_s on $(0, \mathcal{L})$. Furthermore, the left hand side is clearly an increasing function of c_s while the right hand side is a decreasing function of c_s . Observing that $\int_0^0 \mathbb{E} [\mathcal{H}^{-1}](-b) d\mathcal{B}(b) = 0$ and that $\int_0^{\mathcal{L}} \mathbb{E} [\mathcal{H}^{-1}](\mathcal{L} - b) d\mathcal{B}(b) > 0$, completes the proof. \square

Given a graph with a designated source sink pair and a pair of reasonable distributions $(\mathcal{H}, \mathcal{B})$, we will refer to c_s as the balance point of the instance when it is defined. As we see in the following lemma the balance point serves as the role of “ $\mathbb{E} [d_s(v)]$ ”.

Lemma 10. *Let G be an (α, β) -expander on n vertices, let s and t be a source sink pair, and let $(\mathcal{H}, \mathcal{B})$ be a pair of reasonable distributions where \mathcal{B} has a continuous distribution function except potentially at 0. For any fixed $\delta > 0$, if (G, ℓ, R) is a δ -reasonable instance and $c_s, \mathcal{L} - c_s < \nu$, then for any vertex v , other than s and t , we have that $d_s(v) \leq c_s + 13\delta$ with probability at least $1 - e^{-\Omega(\kappa)}$.*

Proof. We proceed by contradiction. Suppose v is a vertex such that $d_s(v) > c_s + 13\delta$, then by Corollary 7, every other vertex u , other than s and t , has $d_s(u) \geq c_s + \delta$. In particular, every flow carrying vertex u has $d_s(u) \geq c_s + \delta$ and $d_t(u) \leq \mathcal{L} - c_s - \delta$. This implies that every edge e adjacent to s with $b_e \leq c_s + \delta$, carries a positive amount of flow, thus the expected amount of flow on an edge is at least $\int \int_0^{c_s + \delta} h^{-1}(c_s + \delta - b) d\mathcal{B}(b) d\mathcal{H}(h)$. Similarly the expected amount of flow on an edge adjacent to t is at most $\int \int_0^{\mathcal{L} - c_s - \delta} h^{-1}(\mathcal{L} - c_s - \delta - b) d\mathcal{B}(b) d\mathcal{H}(h)$. Thus, for each edge e adjacent to s , we define S_e as the random variable that is 0 if $b_e > c_s + \delta$ and $h_e^{-1}(c_s - \delta - b_e)$ otherwise, and let $S = \sum_e S_e$. We define similarly T_e for each edge adjacent to t and let $T = \sum_e T_e$. Then

$$R_s = \mathbb{E} [S] = \deg(s) \int \int_0^{c_s + \delta} h^{-1}(c_s + \delta - b) d\mathcal{B}(b) d\mathcal{H}(h), \quad \text{and}$$

$$R_t = \mathbb{E} [T] = \deg(t) \int \int_0^{\mathcal{L} - c_s - \delta} h^{-1}(\mathcal{L} - c_s - \delta - b) d\mathcal{B}(b) d\mathcal{H}(h).$$

Now, since \mathcal{H} is bounded below by H_{\min} and all functions in \mathcal{H} are convex, $0 \leq S_e, T_e \leq H_{\min}^{-1}(\mathcal{L})$. Thus, assuming that $R_s \geq R_t$, by Hoeffding bounds the probability that $T \geq S$ is at most

$2 \exp \left[-\frac{(R_s - R_t)^2}{(\sqrt{\deg(s)} + \sqrt{\deg(t)})^2 H_{\min}^{-1}(\mathcal{L})^2} \right]$, by Hoeffding bounds. Thus in order to prove the contradiction, it suffices to show that $R_s - R_t$ is large enough. However we have that

$$\begin{aligned}
R_s - R_t &= \deg(s) \int \int_0^{c_s + \delta} h^{-1}(c_s + \delta - b) d\mathcal{B}(b) d\mathcal{H}(h) \\
&\quad - \deg(t) \int \int_0^{\mathcal{L} - c_s - \delta} h^{-1}(\mathcal{L} - c_s - \delta - b) d\mathcal{B}(b) d\mathcal{H}(h) \\
&= \deg(s) \int_0^{c_s + \delta} \mathbb{E}[\mathcal{H}^{-1}](c_s + \delta - b) d\mathcal{B}(b) \\
&\quad - \deg(t) \int_0^{\mathcal{L} - c_s - \delta - b} \mathbb{E}[\mathcal{H}^{-1}](\mathcal{L} - c_s - \delta - b) d\mathcal{B}(b) \\
&> \deg(s) \int_0^{c_s} \mathbb{E}[\mathcal{H}^{-1}](c_s - b) d\mathcal{B}(b) - \deg(t) \int_0^{\mathcal{L} - c_s} \mathbb{E}[\mathcal{H}^{-1}](\mathcal{L} - c_s - b) d\mathcal{B}(b) \\
&\quad + \deg(s) \int_{c_s}^{c_s + \delta} \mathbb{E}[\mathcal{H}^{-1}](c_s + \delta - b) d\mathcal{B}(b) \\
&\quad + \deg(t) \int_{\mathcal{L} - c_s - \delta}^{\mathcal{L} - c_s} \mathbb{E}[\mathcal{H}^{-1}](\mathcal{L} - c_s - \delta - b) d\mathcal{B}(b) \\
&= \deg(t) \int_{\mathcal{L} - c_s - \delta}^{\mathcal{L} - c_s} \mathbb{E}[\mathcal{H}^{-1}](\mathcal{L} - c_s - \delta - b) d\mathcal{B}(b).
\end{aligned}$$

But since \mathcal{B} is locally dense around c_s and $\mathcal{L} - c_s$, by the definition of c_s and Lemma 9 this implies that for any fixed δ , $R_s - R_t$ is $\Omega(\deg(s) + \deg(t))$, completing the proof. \square

Now consider the situation in $G'_U(c, \epsilon)$. We first consider the effect of the sub-sampling on the balance point calculation, specifically let

$$\mathcal{E}(K, \epsilon, c) = \iint_0^K (p_X \mathbb{1}_{[c, (1+\epsilon)c]}(b) \mathbb{1}_{[(1-\epsilon)H_1, H_1]}(h) + p_1 \mathbb{1}_{[0, \epsilon c]}(b) \mathbb{1}_{[(1-\epsilon)H_2, H_2]}(h)) h^{-1}(K - b) d\mathcal{H}(h) d\mathcal{B}(b).$$

Thus we have that if c'_s is the balance point for $G'_U(c, \epsilon)$ then

$$\begin{aligned}
&\deg(s) \left(\int_0^{c'_s} \mathbb{E}[\mathcal{H}^{-1}](c'_s - b) d\mathcal{B}(b) - \mathcal{E}(c'_s, \epsilon, c) \right) \\
&= \deg(t) \left(\int_0^{\mathcal{L} - c'_s} \mathbb{E}[\mathcal{H}^{-1}](\mathcal{L} - c'_s - b) d\mathcal{B}(b) - \mathcal{E}(\mathcal{L} - c'_s, \epsilon, c) \right).
\end{aligned}$$

Of particular interest, is the case where $\deg(t) = (1 + o(1))\deg(s)$, and in this case we have $c'_s = (1 + o(1))c_s$ for any fixed $\epsilon > 0$.

5. BRAESS'S PARADOX

Returning to the viewpoint of $d_s(v)$ being a random variable, the previous two sections say that in a δ -reasonable instance with $\deg(t) = (1 + o(1))\deg(s)$, we have that $d_s(v)$ concentrates around its “expected value” and further we can identify that expected value as $\mathcal{L}/2$ if \mathcal{L} is sufficiently small. Additionally, these properties are preserved when passing to the subgraph G'_{1X} , G'_{X1} , and G'_U and moreover, the “expectation” for vertices in G'_U does not change too much. This control on $d_s(t)$ allows us to explicitly compare $R_G(\mathcal{L})$ and $R_{G'(\mathcal{L}/2, \epsilon)}(\mathcal{L}(1 - \mu))$, in order to prove that

$\mathcal{L}(G'(\mathcal{L}/2, \epsilon), \ell, R_G(\mathcal{L})) < \mathcal{L}$. As a consequence, we have a Braess ratio of at least $\frac{1}{1-\mu}$ and thus we are able to prove Theorem 1.

Proof of Theorem 1. Without loss of generality we will assume that $\deg(s) \leq \deg(t)$. We first fix $B < \nu/2$. In order to show Braess's paradox occurs we will show that for some sufficiently small ϵ , μ , and δ , we have $R < R'$ where $\mathcal{L}(G, \ell, R) = 2B$ and $\mathcal{L}(G'(B, \epsilon), \ell, R') = 2B(1 - \mu)$. Specifically, we show that the flow to achieve latency $2B(1 - \mu)$ in G'_U and G'_{1X} is not too much smaller than the flow on those edges in G . Then we analyze the flow in G'_{X1} and show that the flow gained in this subgraph is larger than the flow previously lost. Thus to achieve latency $2B(1 - \mu)$ in G' it is necessary to send more flow than to achieve latency $2B$ in G , and thus by the monotonicity of the latency, Braess's paradox occurs.

Let $\epsilon > 0$ be fixed with a specific value determined later in the proof. As ϵ is fixed $p = \min\{p_1 p_x, 1 - 2p_1 p_x\}$, is also fixed. We first consider the flow in G'_U relative to those edges in G . First we note that since $\deg(t) = (1 + o(1)) \deg(s)$, we have that $c_s = B + o(1)$. Then by the argument in Section 4 this implies that $c'_s = B + o(1)$. For each edge e in G'_U adjacent to s , let f_e be the flow in G and f'_e be the flow in G'_U . Then, by Corollary 7, $h_e(f_e) + b_e \leq B + 13\delta + o(1)$ and $h_e(f'_e) + b_e \geq B - \left(9 + \frac{4}{\log(1+p)}\right) \delta + o(1)$ and hence $f_e - f'_e \leq h_e^{-1}(B - b_e + 13\delta + o(1)) - h_e^{-1}(B - b_e - \left(9 + \frac{4}{\log(1+p)}\right) \delta + o(1))$. But for sufficiently large n , this is at most $h_e^{-1}(B - b_e + 14\delta) - h_e^{-1}(B - b_e - \left(10 + \frac{4}{\log(1+p)}\right) \delta)$. Now by Chernoff bounds the degree of s in G'_U is at least $(1 - \epsilon') \deg(s)$ with probability at least $1 - e^{-\Omega(\deg s)}$. Thus, since $h_e^{-1}(B - b_e + 14\delta) - h_e^{-1}(B - b_e - \left(10 + \frac{4}{\log(1+p)}\right) \delta)$ is a bounded random variable it has finite expectation. Thus by Hoeffding bounds,

$$\sum_{e \in \{s\} \times U'} h_e^{-1}(B + 14\delta - b_e) - h_e^{-1}\left(B - \left(10 + \frac{4}{\log(1+p)}\right) \delta - b_e\right) \leq C(\delta, \epsilon) \deg(s)$$

for some $C(\delta, \epsilon)$ with probability at least $1 - e^{-\Omega(\deg(s))}$. We now wish to show that $\lim_{\delta \rightarrow 0} C(\delta, \epsilon) = 0$ for any fixed ϵ . In order to do this it suffices to show that

$$\lim_{\delta \rightarrow 0} \mathbb{E} \left[h_e^{-1}(B + 14\delta - b_e) - h_e^{-1}\left(B - \left(10 + \frac{4}{\log(1+p)}\right) \delta - b_e\right) \right] = 0.$$

But an easy calculation shows that the expectation is at most

$$\mathbb{P}(B \leq \mathcal{B} \leq B + 14\delta) \mathbb{E}[\mathcal{H}^{-1}](14\delta) + \int_0^B \mathbb{E}[\mathcal{H}^{-1}](B + 14\delta - b) - \mathbb{E}[\mathcal{H}^{-1}]\left(B - \left(10 + \frac{4}{\log(1+p)}\right) \delta - b\right) d\mathcal{B}(b).$$

Now since $\mathbb{E}[\mathcal{H}^{-1}](B + 14\delta - b) - \mathbb{E}[\mathcal{H}^{-1}]\left(B - \left(10 + \frac{4}{\log(1+p)}\right) \delta - b\right) \leq \mathbb{E}[\mathcal{H}^{-1}](B + 14\delta)$, by the dominated convergence theorem, as $\delta \rightarrow 0$,

$$\int_0^B \mathbb{E}[\mathcal{H}^{-1}](B + 14\delta - b) - \mathbb{E}[\mathcal{H}^{-1}]\left(B - \left(10 + \frac{4}{\log(1+p)}\right) \delta - b\right) d\mathcal{B}(b) \rightarrow 0.$$

Further, since $\mathbb{E}[\mathcal{H}^{-1}](14\delta) \rightarrow 0$ as $\delta \rightarrow 0$, this expectation necessarily approaches 0. Thus, by controlling δ , the amount of flow "lost" in G'_U can be controlled.

Now we consider the amount of flow "gained" by G'_{1X} and G'_{X1} relative to those edges in G . Let R_{1X} and R'_{1X} be the flow leaving s on edges in G'_{1X} in G and G'_{1X} , respectively. Define R_{X1} and R'_{X1} analogously. Similarly to above let f_e and f'_e denote the flow on an edge in G and G'_{1X}

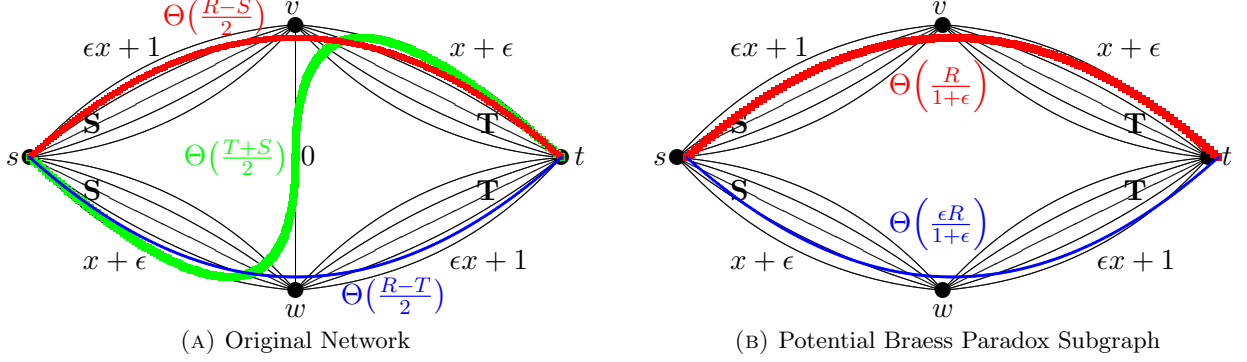


FIGURE 3. Modified Canonical Braess's Paradox Example

respectively. Now if e is adjacent to s and e' is adjacent to t , then we have that

$$H_1(f'_e) + (1 + \epsilon)B + \left(8 + \frac{4}{\log(1 + p)}\right) \delta + H_2(f'_{e'}) + \epsilon B \geq 2B(1 - \mu).$$

But then this implies that

$$H_1\left(\frac{R'_{1X}}{(1 - \epsilon')p_1p_X \deg(s)}\right) + H_2\left(\frac{R'_{1X}}{(1 - \epsilon')p_1p_X \deg(t)}\right) \geq (1 - 2\mu - 2\epsilon)B - \left(8 + \frac{4}{\log(1 + p)}\right) \delta.$$

Using the fact that $\deg(t) \geq \deg(s)$, this implies that $(H_1 + H_2)\left(\frac{R'_{1X}}{(1 - \epsilon')p_1p_X \deg(s)}\right) \geq (1 - 2\mu - 2\epsilon)B - 13\delta$. Combining this with the similar statement for R'_{X1} , we have

$$R'_{1X} + R'_{X1} \geq 2(1 - \epsilon')p_1p_X \deg(s)(H_1 + H_2)^{-1}((1 - 2\mu - 2\epsilon)B - \left(8 + \frac{4}{\log(1 + p)}\right) \delta).$$

Now consider the behavior of R_{1X} , for each edge e in G_{1X} we have that $(1 - \epsilon)H_1(f_e) + B \leq B + 14\delta$, and thus $f_e \leq H_1^{-1}\left(\frac{14\delta}{1 - \epsilon}\right)$ and thus $R_{1X} \leq (1 + \epsilon')p_1p_X \deg(s)H_1^{-1}\left(\frac{14\delta}{1 - \epsilon}\right)$. Similarly, $R_{X1} \leq (1 + \epsilon')p_1p_X \deg(s)H_2^{-1}\left(\frac{B + 14\delta}{1 - \epsilon}\right)$. Thus, by letting $\epsilon', \delta \rightarrow 0$, it suffices to show that there is some $\epsilon, \mu > 0$ such that

$$2p_1p_X \deg(s)(H_1 + H_2)^{-1}((1 - 2\mu - 2\epsilon)B) - p_1p_X \deg(s)H_2^{-1}\left(\frac{B}{1 - \epsilon}\right) > 0.$$

It then suffices to show that $2(H_1 + H_2)^{-1}(B) - H_2^{-1}(B) > 0$, as then if μ, ϵ are sufficiently small then there is a "gain" of flow on $G'_{1X} \cup G'_{X1}$. Since the loss of flow on G'_U can be bounded in terms of δ for a fixed μ and ϵ , a sufficiently small choice of δ yields the result. But now we note that

$$2(H_1 + H_2)^{-1}(B) > 2(2H_2)^{-1}(B) = 2H_2^{-1}(B/2) \geq H_2^{-1}(B). \quad \square$$

At this point it is worth noting that the condition $\deg(t) = (1 + o(1))\deg(s)$ is necessary. Consider the following modification of the canonical example of Braess's paradox. The vertex s is attached to the vertex v with S edges each having latency function $x \mapsto \epsilon x + 1$ and to vertex w with S edges having latency function $x \mapsto x + \epsilon$. Similarly, t is attached to v with T edges having latency function $x \mapsto x + \epsilon$ and to w with T edges having latency function $\epsilon x + 1$. Finally, there is a 0 latency edge between v and w (see Figure 3).

A quick calculation shows that if $S \leq T$

$$\mathcal{L} = \begin{cases} 2\epsilon + \frac{R}{S} + \frac{R}{T} & R \leq (1 - \epsilon)S \\ \frac{\epsilon}{1 + \epsilon} \frac{R}{S} + \frac{2\epsilon^2 + \epsilon + 1}{\epsilon + 1} + \frac{R}{T} & (1 - \epsilon)S < R \leq (1 - \epsilon)T \\ \frac{\epsilon}{1 + \epsilon} \left(\frac{R}{S} + \frac{R}{T}\right) + 2\frac{\epsilon^2 + 1}{\epsilon + 1} & (1 - \epsilon)T < R \end{cases}$$

Now in the case of deleting the edge between v and w , the latency is

$$\mathcal{L}' = 1 + \epsilon + \frac{(\epsilon S + T)(S + \epsilon T)}{(1 + \epsilon)(S + T)ST} R.$$

Letting $\delta = \frac{S}{T}$ and taking the difference we have

$$\mathcal{L} - \mathcal{L}' = \begin{cases} \epsilon - 1 + \frac{1+2\epsilon-\epsilon^2+\delta}{(1+\delta)(1-\epsilon)} \frac{\delta R}{S} & R \leq (1 - \epsilon)S \\ \frac{\epsilon(\epsilon-1)}{1+\epsilon} + \frac{2\epsilon+\delta-\epsilon^2}{(1+\epsilon)(1+\delta)} \frac{\delta R}{S} & (1 - \epsilon)S < R \leq (1 - \epsilon)T \\ \frac{(1-\epsilon)^2}{(1+\epsilon)^2} - \frac{(1-\epsilon)^2}{(1+\epsilon)(1+\delta)} \frac{\delta R}{S} & (1 - \epsilon)T < R \end{cases}.$$

Fixing ϵ , and letting δ be arbitrarily small can see that the in order for $\mathcal{L} - \mathcal{L}' > 0$, $\frac{\delta R}{S}$ can not go to 0. That is, R is on the order of $\frac{S}{\delta} = T$. But in this case the \mathcal{L} is on the order of $\frac{T}{S}$ which is not bounded by a constant. Hence, when the graph is unbalanced we should not expect Braess's paradox to occur if the latency is $\Theta(1)$.

REFERENCES

- [1] NOGA ALON AND JOEL H. SPENCER, *The probabilistic method*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience [John Wiley & Sons], New York, second ed., 2000. With an appendix on the life and work of Paul Erdős.
- [2] MARTIN BECKMANN, C. B. MCGUIRE, AND CHRISTOPHER B. WINSTEN, *Studies in the economics of transportation*, Yale University Press, 1956.
- [3] D. BRAESS, *Über ein Paradoxon aus der Verkehrsplanung*, Unternehmensforschung, 12 (1968), pp. 258–268.
- [4] DIETRICH BRAESS, ANNA NAGURNY, AND TINA WAKOLBINGER, *On a paradox of traffic planning*, Transportation Science, 39 (2005), pp. 446 – 450.
- [5] FAN CHUNG AND STEPHEN J. YOUNG, *Braess's paradox in large sparse random graphs*, in 6th Workshop on Internet and Network Economics, 2010.
- [6] MICHALIS FALOUTSOS, PETROS FALOUTSOS, AND CHRISTOS FALOUTSOS, *On power-law relationships of the internet topology*, in SIGCOMM '99: Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication, New York, NY, USA, 1999, ACM, pp. 251–262.
- [7] CAROLINE FISK AND STEFANO PALLOTTINO, *Empirical evidence for equilibrium paradoxes with implications for optimal planning strategies*, Transportation Research Part A: General, 15 (1981), pp. 245 – 248.
- [8] DIMITRIS FOTAKIS, ALEXIS C. KAPORIS, AND PAUL G. SPIRAKIS, *Efficient methods for selfish network design*, in Proceedings of the 36th International Colloquium on Automata, Languages and Programming: Part II, ICALP '09, Berlin, Heidelberg, 2009, Springer-Verlag, pp. 459–471.
- [9] MARGUERITE FRANK, *The Braess paradox*, Math. Programming, 20 (1981), pp. 283–302.
- [10] ALAN M. FRIEZE, *Edge-disjoint paths in expander graphs*, SIAM J. Comput., 30 (2001), pp. 1790–1801 (electronic).
- [11] CHRISTOS GKANTSIDIS, MILENA MIHAIL, AND AMIN SABERI, *Conductance and congestion in power law graphs*, SIGMETRICS Perform. Eval. Rev., 31 (2003), pp. 148–159.
- [12] JERROLD W. GROSSMAN AND PATRICK D. F. ION, *On a portion of the well-known collaboration graph*, in Proceedings of the Twenty-sixth Southeastern International Conference on Combinatorics, Graph Theory and Computing (Boca Raton, FL, 1995), vol. 108, 1995, pp. 129–131.
- [13] MICHAEL A. HALL, *Properties of the equilibrium state in transportation networks*, Transportation Science, 12 (1978), pp. 208 – 216.
- [14] J.-B. HIRIART-URRUTY AND C. LEMARÉCHAL, *Fundamentals of Convex Analysis*, Springer Verlag, Heidelberg, 2001.
- [15] WASSILY Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc., 58 (1963), pp. 13–30.
- [16] SHLOMO HOORY, NATHAN LINIAL, AND AVI WIGDERSON, *Expander graphs and their applications*, Bull. Amer. Math. Soc. (N.S.), 43 (2006), pp. 439–561 (electronic).
- [17] H. KAMEDA, *How harmful the paradox can be in the braess/cohen-kelly-jeffries networks*, in INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE, vol. 1, 2002, pp. 437 – 445 vol.1.

- [18] JON KLEINBERG AND RONITT RUBINFELD, *Short paths in expander graphs*, in 37th Annual Symposium on Foundations of Computer Science (Burlington, VT, 1996), IEEE Comput. Soc. Press, Los Alamitos, CA, 1996, pp. 86–95.
- [19] GINA KOLATA, *What if they closed 42d street and nobody noticed?*, The New York Times, December 25 (1990).
- [20] ELIAS KOUTSOPIAS AND CHRISTOS PAPADIMITRIOU, *Worst-case equilibria*, in 16th Annual Symposium on Theoretical Aspects of Computer Science, vol. 1563 of Lecture Notes in Comput. Sci., Springer Berlin, 1999, pp. 404–413.
- [21] TOM LEIGHTON AND SATISH RAO, *Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms*, J. ACM, 46 (1999), pp. 787–832.
- [22] HENRY LIN, TIM ROUGHGARDEN, AND ÉVA TARDOS, *A stronger bound on Braess’s paradox*, in Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms, New York, 2004, ACM, pp. 340–341 (electronic).
- [23] NATHAN LINIAL, ERAN LONDON, AND YURI RABINOVICH, *The geometry of graphs and some of its algorithmic applications*, Combinatorica, 15 (1995), pp. 215–245.
- [24] ALEXANDER LUBOTZKY, *Expander graphs in pure and applied mathematics*. Notes for the Colloquium Lectures at the 2011 Joint Mathematics Meetings.
- [25] MILENA MIHAIL, CHRISTOS PAPADIMITRIOU, AND AMIN SABERI, *On certain connectivity properties of the internet topology*, J. Comput. System Sci., 72 (2006), pp. 239–251.
- [26] IGAL MILCHTAICH, *Network topology and the efficiency of equilibrium*, Games Econom. Behav., 57 (2006), pp. 321–346.
- [27] M. E. J. NEWMAN, *The structure and function of complex networks*, SIAM Rev., 45 (2003), pp. 167–256 (electronic).
- [28] CHRISTOS PAPADIMITRIOU, *Algorithms, games, and the internet*, in Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing, New York, 2001, ACM, pp. 749–753 (electronic).
- [29] ERIC I. PAS AND SHARI L. PRINCIPIO, *Braess’ paradox: Some new insights*, Transportation Research Part B: Methodological, 31 (1997), pp. 265 – 276.
- [30] CLAUDE M. PENCHINA, *Braess paradox: Maximum penalty in a minimal critical network*, Transportation Research Part A: Policy and Practice, 31 (1997), pp. 379 – 388.
- [31] TIM ROUGHGARDEN, *On the severity of Braess’s Paradox: designing networks for selfish users is hard*, J. Comput. System Sci., 72 (2006), pp. 922–953.
- [32] TIM ROUGHGARDEN AND ÉVA TARDOS, *How bad is selfish routing?*, J. ACM, 49 (2002), pp. 236–259 (electronic).
- [33] GREGORY VALIANT AND TIM ROUGHGARDEN, *Braess’s paradox in large random graphs*, Random Structures & Algorithms, 37 (2010), pp. 495–515.
- [34] VIJAY V. VAZIRANI, *Approximation algorithms*, Springer-Verlag, Berlin, 2001.
- [35] DUNCAN J. WATTS AND STEVEN H. STROGATZ, *Collective dynamics of ‘small-world’ networks*, Nature, 393 (1998), pp. 440–442.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, SAN DIEGO, 9500 GILMAN DRIVE # 0112, LA JOLLA, CA 92093

E-mail address, Fan Chung: fan@math.ucsd.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, SAN DIEGO, 9500 GILMAN DRIVE # 0112, LA JOLLA, CA 92093

E-mail address, Stephen J. Young: s7young@math.ucsd.edu

(Wenbo Zhao) DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, UNIVERSITY OF CALIFORNIA, SAN DIEGO, 9500 GILMAN DRIVE # 0404, LA JOLLA, CA 92093

E-mail address, Wenbo Zhao: w3zhao@ucsd.edu