Braess's Paradox in Large Sparse Graphs

Fan Chung[*] and Stephen J. Young

Department of Mathematics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093
`{fan,s7young}@math.ucsd.edu`

Abstract. Braess's paradox, in its original context, is the counter-intuitive observation that, without lessening demand, closing roads can improve traffic flow. With the explosion of distributed (selfish) routing situations understanding this paradox has become an important concern in a broad range of network design situations. However, the previous theoretical work on Braess's paradox has focused on "designer" graphs or dense graphs, which are unrealistic in practical situations. In this work, we exploit the expansion properties of Erdős-Rényi random graphs to show that Braess's paradox occurs when $np \geq c \log(n)$ for some $c > 1$.

Keywords: Braess's paradox, price of anarchy, random graphs, selfish routing

1 Introduction

In 1968 Dietrich Braess observed that there were road networks such that if the travellers were behaving selfishly it was possible to improve everyone's travel time by *removing* roads, even roads with extremely fast travel times. Specifically, he considered the situation illustrated in Figure 1 in the case of routing one unit of flow from s to t. As we can see in Figure 1(a) when the users behave selfishly all of the flow passes through the zero latency central edge and the total latency is 2. However, as we can see in Figure 1(b), by removing the central zero latency edge, the selfish routing will spread the flow uniformly over the paths between s and t, resulting in an overall latency of $\frac{3}{2}$.

Since its discovery Braess's paradox has spawned a significant amount of work aimed at understanding the full implications of the paradox, both theoretically [3, 5, 11, 12] and via anecdotal observations [2, 7]. In many ways the recent trend towards studying the "Price of Anarchy" [8, 10] has its roots in the discovery of Braess's paradox. However, these "worst case" analyses via designer instances give little insight into the practical consequences of Braess's paradox. Although the anecdotal evidence indicates that Braess's paradox can occur in real world, it gives little to no feeling for how prevalent or severe the paradox can be in real world networks. Recently, Valiant and Roughgarden [16] began

to answer this question by providing the first proof that Braess's paradox could occur in a large class of non-designer graphs. Specifically, they showed that in sufficiently dense instances of Erdős-Rényi random graphs with affine latency functions satisfying certain mild conditions, Braess's paradox occurs with high probability (that is, with probability $1 - \mathcal{O}(n^{-c})$ for some $c > 0$). In this work we extend their results to almost all conected Erdős-Rényi random graphs. That is, with similar mild conditions on the latency function, Braess's paradox occurs with high probability in Erdős-Rényi random graphs with expected degree at least a $c' \log(n)$ for some $c' > 1$.

(a) Original Network with equilibrium latency 2
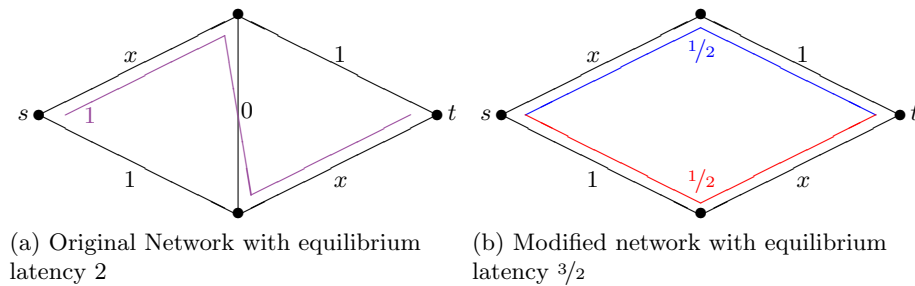
(b) Modified network with equilibrium latency $3/2$

Fig. 1: Braess's Paradox

We consider a single commodity flow on an undirected Erdős-Rényi random graph G with a designated sink s and source t with latency functions $\ell = \{\ell_e\}_{e \in E(G)}$ associated to each edge. Letting $\mathcal{P}$ be the set of simple s-t paths in G, a *flow* is a function $f \colon \mathcal{P} \longrightarrow \mathbb{R}^{\geq 0}$. The flow on an edge e with respect to a flow f is $f_e = \sum_{P \in \mathcal{P}_e} f(P)$, where $\mathcal{P}_e$ is the set of paths in $\mathcal{P}$ containing the edge e. We will say an edge is *flow carrying* if $f_e > 0$, similarly, we will say a vertex v is flow carrying if it is incident to some flow carrying edge. The *traffic rate*, or value, of a flow f, is $R = \sum_{P \in \mathcal{P}} f(P)$. The *latency of a path* P with respect to a flow f, which we will denote as $\ell_P(f)$, is $\sum_{e \in E(P)} \ell_e(f_e)$. Notice that the latency of a path depends on the value of the flow on an edge, not the value of the flow on the particular path under consideration.

We concern ourselves with the case of routing an infinitely divisible flow from s to t such that each "unit" of flow behaves selfishly. That is, given a fixed traffic rate R we are interested in the properties of those flows such that there is no incentive for any "unit" of flow to change the path it is on. Alternatively, we may say that such a flow is at Nash equilibrium. One way of characterizing such flows is that for every pair $P_1, P_2 \in \mathcal{P}$, if $f(P_1) > 0$, then $\ell_{P_1}(f) \leq \ell_{P_2}(f)$ [17]. In particular, this implies that all flow carrying paths have the same latency at Nash equilibrium. It has been shown that every selfish routing network has a Nash equilibrium flow and further, all flow carrying paths in all possible Nash equilibrium flows have the same latency [1]. Thus, we may define $\mathcal{L}(G, \ell, R)$ as

the common latency of all flow carrying paths at Nash equilibrium in the graph G with the latency functions ℓ at the traffic rate R. Given this notation, Braess's paradox may be restated as the observation that there exists an 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, define the *Braess ratio* of the instance, denoted $\rho(G, \ell, R)$, as

$$\max_{G' \subseteq G} \frac{\mathcal{L}(G, \ell, R)}{\mathcal{L}(G', \ell, R)}. \tag{1}$$

Note that the Braess ratio is specification of the price of anarchy to the context of this context.

1.1   Previous Work

Recently, driven by the obvious practical applications there has been some work attempting to answer the question of whether networks can be designed to avoid Braess's paradox. A, perhaps, more important question is whether Braess's paradox can be exploited to improve the performance of already existent real world networks. Roughgarden emphatically answers these questions in the negative in [13], by showing that unless P = NP there is not an $\frac{n}{2}$ approximation algorithm to determining a subnetwork which achieves a Braess ratio larger than 1. Further, even if the latency functions are restricted to be affine, there is still not a $\frac{4}{3}$ approximation algorithm, and this result is tight [14]. This leads naturally to the following important practical question: Is Braess's paradox a prevalent phenomenon or, like the exponential examples for the simplex method [6, 15], is it an academic curiosity that can be ignored in practice? The recent work of Valiant and Roughgarden [16] has begun to address this fundamental question. In order to state and understand their results, we need the following definition. A pair of distributions $\mathcal{A}$ and $\mathcal{B}$ is *reasonable* if

- $\mathcal{A}$ has bounded support $[A_{\min}, A_{\max}]$ with $A_{\min} > 0$,
- there is some closed interval I_A of positive length, such that for every nontrivial subinterval $J \subseteq I_A$, $\mathbb{P}(\mathcal{A} \in J) > 0$, and
- there is some interval $I_B = [0, \eta]$, with $\eta > 0$, such that for every nontrivial subinterval $J \subseteq I_B$, $\mathbb{P}(\mathcal{B} \in J) > 0$.

Within this context they were able to show the following theorem.

Theorem 1 ([16]). *Let $p = \Omega\left(n^{-\frac{1}{2}+\zeta}\right)$ be an edge sampling probability with $\zeta > 0$ and let $\mathcal{A}$ and $\mathcal{B}$ be reasonable distributions. Let G be an Erdős-Rényi random graph with edge probability p and let $\ell_e = a_e f_e + b_e$ for all edges, where (a_e, b_e) is distributed as $\mathcal{A} \times \mathcal{B}$. There is a constant $\rho = \rho(\zeta, \mathcal{A}, \mathcal{B}) > 1$ such that, with high probability the instance (G, ℓ), admits a choice of traffic rate R such that the Braess ratio of the instance is at least ρ.*

Given the importance and self evident "correctness" of Nash equilibrium flows in a practical context (especially in the context of automobile traffic) it is unsurprising that there has been significant previous work on the properties of such flows. We collect a few of the more useful results in the following lemma.

Lemma 2. *Given an instance (G, ℓ, R) and an Nash equilibrium flow f for each vertex v define $d_s(v)$ as the shortest path from v to s with respect to the latencies $\ell(f)$. Define $d_t(f)$ 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 $d_s(v) + d_s(t) \geq \mathcal{L}(G, \ell, R)$ with equality if v is a flow carrying vertex.*
2. *[13] If f is a 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 if and only if equality holds whenever $f_e > 0$.*
3. *[4, 9] For every network G and strictly increasing latency functions ℓ, $\mathcal{L}(G, \ell, R)$ is continuous and strictly increasing function of R.*
4. *[13] 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.*

Now by part (3) of this lemma, if the latency functions are all strictly increasing, there is a function $R_\ell^G(\mathcal{L})$ which gives the unique value R so that the latency of a Nash equilibrium flow on the network G with latencies ℓ is precisely $\mathcal{L}$. When the underlying graph is clear, we will denote this simply as $R_\ell(\mathcal{L})$.

1.2   Our Contribution

In the work of Valiant and Roughgarden the critical structural property they use in their proofs is that if $p \gg n^{-\frac{1}{2}+\zeta}$ for some $\zeta > 0$, then $\mathcal{G}(n, p)$ has polynomially many disjoint paths of length two between any two vertices. Whereas if p is $\mathcal{O}\left(n^{-\frac{1}{2}+o(1)}\right)$ there are clearly not even polynomially many paths of length two between any two vertices. More importantly, very few, if any real world networks share this property. The primary contribution of our work is to generalize their methodology to rely on a more prevalent real world property, expansion.

In Section 2 we analyze an idealized network and reveal two key properties of Nash equilibria that will be used in our proof. Specifically, we observe that at particular traffic rate in this idealized network every internal vertex is at the same latency distance from s and further, every internal vertex is equidistant from s and t. In the more general case, this breaks down into two claims. The first, which we prove in Section 3, is that for any two internal vertices u and v, their latency distance to s differs by at most a fixed constant δ. Valiant and Roughgarden prove a similar result for dense Erdős-Rényi random graphs in their δ-lemma using a specialized case of our expansion based argument. The second claim is that the internal vertices are equidistant from s and t. Clearly, this cannot hold if the internal vertices do not all have the same latency distance to s, however, as we show in Section 4 the latency distance for internal vertices is roughly balanced between exiting s and entering t. We note that Valiant and Roughgarden's proof of the balance lemma would suffice at this point, however, their proof depends on a somewhat unnatural and cumbersome discretization argument which we are able to completely avoid. Finally, in Section 5 we show,

to our knowledge, the first proof that Braess's paradox can occur in sparse, non-designer graphs with the following theorem.

Theorem 3. *Let G be an Erdős-Rényi random graph on n vertices with edge probability p. Let $\mathcal{A}$ and $\mathcal{B}$ be reasonable distributions and let all latency functions have the form $\ell_e(f_e) = a_e f_e + b_e$ where (a_e, b_e) is distributed according to $\mathcal{A} \times \mathcal{B}$. There are constants $\delta > 0$, $c > 1$, and $\rho > 1$ such that, if $\mathbb{P}\left(\mathcal{B} \leq \frac{\delta}{\log(n)}\right) pn \geq c\log(n)$, then there is a flow rate R such that the instance (G, ℓ, R) has Braess's ratio at least ρ with high probability.*

2   Sketch of Ideas via a Motivating Example

In many ways the ideas behind the our proofs are motivated by the following example. Let G be a graph consisting of s, t, and a complete bipartite graph $K_{n,n}$ such that s and t are each adjacent to every vertex on opposite sides of the bipartition. That is, $(\Gamma(s), \Gamma(t))$ forms the bipartition of $K_{n,n}$. Define the latency function such that every edge in the $K_{n,n}$ has latency 0 and every edge adjacent to s or t has latency function $ax + b$ where (a, b) is distributed as $\mathcal{A} \times \mathcal{B}$. See Figure 2(a). Since the latencies in $K_{n,n}$ are all 0, we can explicitly calculate the flow given two values, c and $\mathcal{L}$, where c is the value of $d_s(v)$ for any vertex in the $K_{n,n}$ and $\mathcal{L}$ is the overall latency of the network. Specifically, the total flow is

$$R = \sum_{\substack{v \in \Gamma(s) \\ b_{\{s,v\}} \leq c}} \frac{c - b_{\{v,s\}}}{a_{\{v,s\}}} = \sum_{\substack{u \in \Gamma(t) \\ b_{\{s,u\}} \leq \mathcal{L}-c}} \frac{\mathcal{L} - c - b_{\{u,s\}}}{a_{\{u,s\}}}. \tag{2}$$

We note that this implies that, up to lower order terms, $c = \mathcal{L} - c$ with high probability.

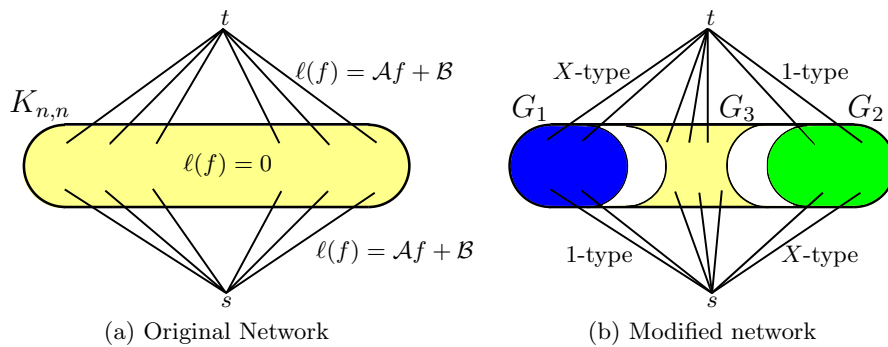(a) Original Network   (b) Modified network

Fig. 2: Motivating Example

Now, if $\mathcal{A}$ and $\mathcal{B}$ are reasonable, then let $I_{\mathcal{A}}$ and $I_{\mathcal{B}}$ be the two intervals witnessing their reasonableness. As in [16] choose $A_1 < A_2$ arbitrarily in the interior $I_{\mathcal{A}}$ and B in the interior of $I_{\mathcal{B}}$. Letting $\mathcal{B}_{\leq B}$ be $\mathcal{B}$ restricted to be at most B, we then have that in our motivating example

$$R_{\ell}(2B) = \frac{B - \mathbb{E}\left[\mathcal{B}_{\leq B}\right]}{\mathbb{E}\left[\mathcal{A}\right]} \mathbb{P}(\mathcal{B} \leq B)\, n \qquad (3)$$

up to low order terms. Now we want to partition $K_{n,n}$ into three parts, yielding graphs G_1, G_2, and G_3 so that $R_{\ell}^{G_1}(2B) + R_{\ell}^{G_2}(2B) + R_{\ell}^{G_3}(2B) > R_{\ell}^{G}(2B)$. If we succeed in this, then the graph $G' = G_1 \cup G_2 \cup G_3$ is a proper subgraph of G which can route more flow than G at the same latency. Then, since the latency function is strictly increasing by Lemma 2, this implies that Braess's paradox occurs. In order to do that we will again mimic the work of Valiant and Roughgarden [16] and partition the vertices into three classes based on the latency function of the edge adjacent with s or t. Specifically, fix $\epsilon > 0$ so that $(1 - \epsilon)A_2 > A_1$ and let E_1 be the edges with latency function $ax + b$ such that $a \leq A_1$ and $b \in (B, (1 + \epsilon)B)$. These are the 1-*type* edges. Similarly, define the X-*type* edges as those where the latency function $ax + b$ is such that $a \in ((1 - \epsilon)A_2, A_2)$ and $b \leq \epsilon B$ and denote the set of such edges by E_X. We will then choose the partition of the $K_{n,n}$ to force the use of 1-type edges which are underutilized in the routing in G. To that end, define G_1 as the graph induced by $\{s, t\} \cup \{v \in \Gamma(s) \mid \{v, s\} \in E_1\} \cup \{u \in \Gamma(t) \mid \{u, t\} \in E_X\}$. Further define G_2 as the graph induced by $\{s, t\} \cup \{v \in \Gamma(s) \mid \{v, s\} \in E_X\} \cup \{u \in \Gamma(t) \mid \{u, t\} \in E_1\}$. Finally, G_3 is the graph induced by $\{s, t\}$ and the vertices not in G_1 or G_2. In other words, G_1 contains all the 1-type edges adjacent to s and all the X-type adjacent to t, G_2 contains all the X-type edges adjacent to s and all the 1-type adjacent to t, and the remaining edges are in G_3. See Figure 2(b). Now if the probability of being a 1-type edge is greater than the probability of being a X-type edge, randomly move 1-type edges (and their associated vertices) to G_3 so that the expected degrees of s and t are the same in G_1 and G_2. By performing the analogous operation if the probability of being an X-type edge is greater than the probability of being a 1-type edge, we may assume that the expected degrees of s and t in the same in each of G_1, G_2, and G_3. That is, letting D be the expected degree of s in G_1, we have $D = \mathbb{E}\left[\Gamma_{G_1}(s)\right] = \mathbb{E}\left[\Gamma_{G_2}(s)\right] = \mathbb{E}\left[\Gamma_{G_1}(t)\right] = \mathbb{E}\left[\Gamma_{G_2}(t)\right]$.

Now consider $R_{\ell}^{G_3}(2B)$. Since the distribution of latencies are the same at s and t in G_3, for every $v \in \Gamma_{G_3}(\{s, t\})$, $d_s(v) = B$ up to lower order terms. Similarly, when considering $R_{\ell}^{G}(2B)$ we have for every vertex $v \in \Gamma_G(\{s, t\})$, $d_s(v) = B$ up to lower order terms as well. Thus, the flow on edges in present in G_3 is the same (again, up to low order terms) as the flow on the corresponding edges in G. Furthermore, this also implies that the 1-type edges in G have asymptotically zero flow. Hence, we have that asymptotically $R_{\ell}^{G}(2B) - R_{\ell}^{G_3}(2B) \leq \frac{B}{(1-\epsilon)A_2} D$. Now we return to G_1. By similar arguments as above there exists some constant c such that for every 1-type edge e in G_1, $a_e f_e + b_e = c$ and for every X-type edge e', $a_{e'} f_{e'} + b_{e'} = 2B - c$. Summing over all 1-type and X-type edges in G_1

we have, asymptotically,

$$2BD = \sum_{v \in \Gamma_{G_1}(s)} a_{\{v,s\}} f_{\{v,s\}} + b_{\{v,s\}} + \sum_{u \in \Gamma_{G_1}(t)} a_{\{u,t\}} f_{\{u,t\}} + b_{\{u,t\}} \qquad (4)$$

$$\leq \sum_{v \in \Gamma_{G_1}(s)} A_1 f_{\{v,s\}} + (1+\epsilon)B + \sum_{u \in \Gamma_{G_1}(t)} A_2 f_{\{u,t\}} + \epsilon B \qquad (5)$$

$$= (A_1 + A_2) R_\ell^{G_1}(2B) + (1+2\epsilon)BD. \qquad (6)$$

Solving, we have that $R_\ell^{G_1}(2B) \geq \frac{(1-2\epsilon)B}{A_1+A_2}D$. Similarly, $R_\ell^{G_2}(2B) \geq \frac{(1-2\epsilon)B}{A_1+A_2}D$ and thus, if

$$\frac{2(1-2\epsilon)B}{A_1+A_2}D > \frac{B}{(1-\epsilon)A_2}D. \qquad (7)$$

then Braess's paradox occurs with high probability in this example. By rearranging, we get that this condition is equivalent to

$$2(1-2\epsilon)(1-\epsilon) > 1 + \frac{A_1}{A_2}. \qquad (8)$$

Since $A_1 < A_2$, there is some $\epsilon > 0$ which makes this inequality true and thus Braess's paradox occurs.

We notice that there are two key observations that allow Braess's paradox to occur in this example. The first is that for any internal vertices u and v, $d_s(u) = d_s(v)$. This observation is mimicked by the δ-lemma, which will be proved in Section 3, which says that $|d_s(u) - d_s(v)| \leq \delta$ for some small $\delta > 0$. The next key observation is that $\mathcal{L} - c = c$ and so the latency is balanced exiting s and entering t, which corresponds naturally to the balance lemma which we prove in Section 4.

3   Small Latency Separates Interior Vertices

The basis for Valiant and Roughgarden's proof of the δ-lemma is that if p is $\Omega\left(n^{-\frac{1}{2}+\zeta}\right)$ with $\zeta > 0$, then there are many paths of length two between any pair of vertices. This structural property allows the flow to be spread out among many internally disjoint paths, yielding a relatively small increase in latency between any two internal vertices. Although no similar property holds for sparser Erdős-Rényi graphs, the following expansion property will provide a sufficient analog.

Lemma 4. *Let G be an Erdős-Rényi random graph on n vertices with edge probability p. There is some $c > 1$ such that, if $np > c \log(n)$, then with high probability every subset U of the vertices is such that $|\Gamma(U)| \geq (\frac{e-1}{e} - o(1)) \min \{np|U|, n\}$*

Roughly speaking, this lemma allows us to view the graph through the lens of a series of s-t cuts, each of which has a relatively large number of edges crossing the cut. This large number of edges crossing the cut allows for the flow from s to

t to be spread out and guarantees that there are some low latency edges crossing the cut. Then, by moving the endpoints of these low latency edges across the cut, we get a new s-t cut and can repeat this procedure. However, there are two primary difficulties in applying this methodology. The first is that the initial cut can have a relatively small number of edges compared to the subsequent cuts, and the second is that at the some point the number of edges crossing the cut begins to decline. We deal with first by allowing the first cut to use edges with higher latency in order to get an initial "core" of vertices with which to start the method. The second difficulty is dealt with by building cuts from s and from t, and recognizing that when those cuts are close to colliding there are a large number of short (in fact, length at most two) internally disjoint paths between the two sets. This last step is the only step needed in the denser case dealt with in [16]

Lemma 5. *Let G be an Erdős-Rényi random graph on n vertices with edge probability p. Let $\mathcal{A}$ and $\mathcal{B}$ be reasonable distributions and let all latency functions have the form $\ell_e(f_e) = a_e f_e + b_e$ where (a_e, b_e) is distributed according to $\mathcal{A} \times \mathcal{B}$. For any sufficiently small fixed $\delta > 0$, there are some constants $c > 1$ and $n_0 > 0$ such that, if $n > n_0$, $np > c\log(n)$ and $\mathbb{P}\left(\mathcal{B} \leq \frac{\delta}{\log(n)}\right) pn \geq 4$, then for any two flow carrying vertices u, v other than s and t in the instance $(G, \ell, R_\ell(2B))$, we have $|d_s(u) - d_s(v)| \leq 7\delta$ and $|d_t(u) - d_t(v)| \leq 7\delta$ with high probability.*

Proof. Since the case where p is a constant was resolved by Valiant and Rough-garden in [16] we may assume without loss of generality that $p \to 0$. Let v_s be a flow carrying vertex that minimizes $d_s(v_s)$ and let v_t be a flow carrying vertex that maximizes $d_s(v_t)$. Since, for flow carrying vertices v, $d_s(v) + d_t(v) = 2B$, in order to show the lemma it suffices to show that $d_s(v_t) - d_s(v_s) \leq 7\delta$ for sufficiently large n. We note that the amount of flow entering v_s is at most $\frac{d_s(v_s)}{A_{\min}} \leq \frac{2B}{A_{\min}}$. Further, since by Chernoff bounds the $\deg(s) \leq \frac{3}{2}np$ with high probability, $R_\ell(2B) \leq \frac{3Bnp}{A_{\min}}$. Additionally, by Chernoff bounds, with high probability there are least $\frac{2}{3}\mathbb{P}(\mathcal{B} \leq \delta) np$ edges e adjacent to v_s with $b_e \leq \delta$. At most half of these have flow more that $2\frac{2B}{A_{\min}} \frac{1}{\frac{2}{3}\mathbb{P}(\mathcal{B} \leq \delta)np}$, thus there are at least $\frac{1}{3}\mathbb{P}(\mathcal{B} \leq \delta) np$ vertices v such that $d_s(v) \leq d_s(v_s) + A_{\max} \frac{6B}{\mathbb{P}(\mathcal{B} \leq \delta)np} + \delta$. With this in mind, let $c_0 = d_s(v_s) + A_{\max} \frac{6B}{\mathbb{P}(\mathcal{B} \leq \delta)np} + \delta$ and let $U_0 = \{v \in V(G) \mid d_s(v) \leq c_0\}$ and note that $|U_0| \geq \frac{1}{3}\mathbb{P}(\mathcal{B} \leq \delta) np$.

We now will inductively define a sequence U_i and c_i such that $U_i \subset U_{i+1}$ and $c_i < c_{i+1}$, stopping when $|\Gamma(U_i)| \geq \frac{3n}{5}$. Suppose then that U_i and c_i are defined and that $|\Gamma(U_i)| < \frac{3n}{5}$ and let $\gamma = \frac{\delta}{\log(n)}$. Then, by Lemma 4 and noting that $\frac{3}{5} < 1 - \frac{1}{e}$, with high probability $|\Gamma(U_i) \setminus U_i| \geq \frac{3}{5}np|U_i|$ for sufficiently large n. Furthermore, (again by Chernoff bounds) with probability at least $1 - e^{-\frac{5}{24}\mathbb{P}(\mathcal{B} \leq \gamma)np|U_i|} > 1 - e^{-\frac{\mathbb{P}(\mathcal{B} \leq \gamma)np|U_i|}{5}}$ there are at least $\frac{1}{2}\mathbb{P}(\mathcal{B} \leq \gamma) np|U_i|$ vertices in $\overline{U_i}$ (the compliment of U_i) that are connected to a vertex in U_i by 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 witnesses for membership in U_i'. That is, for every vertex $v \in U_i'$ there is a unique edge

$e \in E_i$ so that $e \in U_i \times \{v\}$ and $b_e \leq \gamma$. Now since $(U_i, \overline{U_i})$ is a cut and we may assume that the Nash equilibrium flow is cycle free, there is $R_\ell(2B) \leq \frac{3Bnp}{A_{\min}}$ flow crossing the cut from U_i to $\overline{U_i}$. But then at most half of the edges in E_i have flow greater than $2\frac{3Bnp}{A_{\min}|U_i'|}$ and in particular at least half of the vertices $v \in U_i'$ have

$$d_s(v) \leq c_i + \gamma + A_{\max}\frac{6Bnp}{A_{\min}\,|U_i'|} \leq c_i + \gamma + \frac{12A_{\max}Bnp}{A_{\min}np\,|U_i|} = c_i + \gamma + \frac{12A_{\max}B}{A_{\min}\,|U_i|}. \quad (9)$$

Thus we define $c_{i+1} = c_i + \gamma + \frac{12A_{\max}B}{A_{\min}|U_i|}$ and $U_{i+1} = \{v \in V(G) \mid d_s(v) \leq c_{i+1}\}$. By the above we have that

$$|U_{i+1}| \geq \left(\frac{1}{4}\mathbb{P}(\mathcal{B} \leq \gamma)\,np + 1\right)|U_i| \geq \left(\frac{1}{4}\mathbb{P}(\mathcal{B} \leq \gamma)\,np + 1\right)^i |U_0|. \quad (10)$$

If i^* is the first i such that $|\Gamma(U_i)| \geq \frac{3n}{5}$, then this implies that

$$i^* \leq \frac{\log(\frac{3n}{5|U_0|})}{\log(\frac{1}{4}\mathbb{P}(\mathcal{B} \leq \gamma)\,np + 1)} \leq \frac{\log(\frac{9}{5\mathbb{P}(\mathcal{B} \leq \delta)p})}{\log(\frac{1}{4}\mathbb{P}(\mathcal{B} \leq \gamma)\,np + 1)}. \quad (11)$$

Since, by assumption $\mathbb{P}(\mathcal{B} \leq \gamma)\,np \geq c\log(n)$ for some $c > 1$, we have $\frac{1}{p} \leq \frac{n}{\log(n)}$, and thus $i^* \leq \log(n)$ for sufficiently large n. As a consequence, we have that

$$c_{i^*} \leq c_0 + \gamma i^* + \sum_{i=0}^{i^*} \frac{12A_{\max}B}{A_{\min}\,|U_i|} \quad (12)$$

$$\leq c_0 + \frac{\delta}{\log(n)}\log(n) + \sum_{i=0}^{\log(n)} \frac{12A_{\max}B}{A_{\min}\left(\frac{1}{4}\mathbb{P}(\mathcal{B} \leq \gamma)\,np + 1\right)^i |U_0|} \quad (13)$$

$$= c_0 + \delta + \frac{12A_{\max}B}{A_{\min}\,|U_0|}\sum_0^{\log(n)}\left(\frac{1}{4}\mathbb{P}(\mathcal{B} \leq \gamma)\,np + 1\right)^{-i} \quad (14)$$

$$\leq \left(d_s(v_s) + \frac{6BA_{\max}}{\mathbb{P}(\mathcal{B} \leq \delta)\,np} + \delta\right) + \delta + \frac{36A_{\max}B}{A_{\min}\mathbb{P}(\mathcal{B} \leq \delta)\,np}\sum_{i=0}^{\infty}2^{-i} \quad (15)$$

$$= d_s(v_s) + 2\delta + \frac{BA_{\max}\left(6A_{\min} + 72\right)}{A_{\min}\mathbb{P}(\mathcal{B} \leq \delta)\,np}. \quad (16)$$

At this point it worth noting that the failure probability in the recursive construction of U_{i^*} is at most

$$\sum_{i=0}^{i^*} e^{-\frac{\mathbb{P}(\mathcal{B} \leq \gamma)\,np|U_i|}{5}} = \sum_{i=0}^{i^*} e^{-\frac{\mathbb{P}(\mathcal{B} \leq \gamma)\,np\left(\frac{1}{4}\mathbb{P}(\mathcal{B} \leq \gamma)\,np+1\right)^i|U_0|}{5}} \leq \sum_{i=0}^{i^*} e^{-\frac{2^{i+2}|U_0|}{5}} \quad (17)$$

$$\leq \sum_{i=0}^{i^*} e^{-\frac{4}{15}2^i\mathbb{P}(\mathcal{B} \leq \delta)\,np} \quad (18)$$

$$\leq \log(n)e^{-\frac{4}{15}\mathbb{P}(\mathcal{B} \leq \delta)\,np}. \quad (19)$$

Thus, with high probability, this construction method succeeds.

In a similar way, we can define $c'_{j^*} = d_s(v_t) - 2\delta - \frac{BA_{\max}(6A_{\min}+72)}{A_{\min}\mathbb{P}(\mathcal{B}\leq\delta)np}$ and $V_{j^*} = \{v \mid d_s(v) \geq c'_{j^*}\}$ and have that $|\Gamma(V_{j^*})| \geq \frac{3n}{5}$. Without loss of generality, we may assume that $c'_{j^*} - c_{i^*} > 0$ and $V_{j^*} \cap U_{i^*} = \varnothing$. Now since $|\Gamma(U_{i^*})| + |\Gamma(V_{j^*})| \geq \frac{6n}{5}$ there are at least $\frac{n}{10}$ edge disjoint paths of length at most 2 between U_{i^*} and V_{j^*}. Furthermore, by Chernoff bounds, this implies that with high probability there are at least $\frac{1}{12}\mathbb{P}(\mathcal{B}\leq\delta)^2 n$ such paths where all edges e on the path have $b_e \leq \delta$. Now, at most half of those paths have flow more that $2\frac{12R_\ell(2B)}{\mathbb{P}(\mathcal{B}\leq\delta)^2 n}$ and thus

$$c'_{j^*} - c_{i^*} \leq 2\delta + \frac{48A_{\max}R_\ell(2B)}{\mathbb{P}(\mathcal{B}\leq\delta)^2 n} \leq 2\delta + \frac{144A_{\max}Bp}{A_{\min}\mathbb{P}(\mathcal{B}\leq\delta)^2}. \tag{20}$$

Putting all the pieces together we have that $d_s(v_s) - d_s(v_t) \leq 6\delta + \frac{144BA_{\max}}{\mathbb{P}(\mathcal{B}\leq\delta)^2}p + \frac{BA_{\max}(12A_{\min}+144)}{A_{\min}\mathbb{P}(\mathcal{B}\leq\delta)np}$, which, for sufficiently large n, is at most 7δ. $\square$

This control on the spread of flow carrying vertices leads to the following control on the spread of all vertices other than s and t.

Corollary 6. *Let G be an Erdős-Rényi random graph on n vertices with edge probability p. Let $\mathcal{A}$ and $\mathcal{B}$ be reasonable distributions and let all latency functions have the form $\ell_e(f_e) = a_e f_e + b_e$ where (a_e, b_e) is distributed according to $\mathcal{A} \times \mathcal{B}$. For any sufficiently small fixed $\delta > 0$ there are some constants $c > 1$ and $n_0 > 0$ such that, if $n > n_0$ and $\mathbb{P}\left(\mathcal{B} \leq \frac{\delta}{\log(n)}\right)np \geq c\log(n)$, then for any two vertices u, v other than s and t in the instance $(G, \ell, R_\ell(2B))$, we have $|d_s(u) - d_s(v)| \leq 8\delta$ and $|d_t(u) - d_t(v)| \leq 8\delta$ with high probability.*

Proof. Let u and v be arbitrary vertices in $\mathcal{G}(n,p) - \{s,t\}$. Now if c is large enough then $\mathcal{G}(n,p) - \{s,t\}$ restricted to those edges e where $b_e \leq \frac{\delta}{\log(n)}$ is connected and has diameter at most $\log(n)$ with high probability. Let P be a path from u to v in the restricted graph and suppose without loss of generality that $d_s(u) \leq d_s(v)$. If the path P contains at most one flow carrying vertex, then none of the edges along the path carry any flow and thus $d_s(v) \leq d_s(u) + \log(n)\frac{\delta}{\log(n)} = d_s(u) + \delta$ by Lemma 2, and the claim follows. Thus there are at least two flow carrying vertices on the path P. Let u_f be the closest (in terms of the path) flow carrying vertex to u and similarly for v_f. Let g_u be the number of edges on P between u and u_f and similarly for g_v. Since none of the first g_u edges carry flow there is a path from u to t of length at most $g_u\frac{\delta}{\log(n)} + d_t(u_f) = g_u\frac{\delta}{\log(n)} + 2B - d_s(u_f)$. Further, since $d_s(u) + d_t(u) \geq 2B$ we have $2B - d_s(u) \leq d_t(u) \leq g_u\frac{\delta}{\log(n)} + 2B - d_s(u_f)$. Arguing analogously $d_s(v) \leq g_v\frac{\delta}{\log(n)} + d_s(v_f)$. Combining the inequalities we have

$$2B + d_s(v) - d_s(u) \leq (g_u + g_v)\frac{\delta}{\log(n)} + d_s(v_f) - d_s(u_f). \tag{21}$$

By assumption $d_s(u) \leq d_s(v)$ and by Lemma 5, $|d_s(v_f) - d_s(u_f)| \leq 7\delta$, thus $|d_s(v) - d_s(u)| \leq (g_u + g_v)\frac{\delta}{\log(n)} + 7\delta \leq 8\delta$. A similar argument shows that $|d_t(v) - d_t(u)| \leq 8\delta$. $\square$

At this point it is worth noticing that the condition that $\mathbb{P}\left(\mathcal{B} \leq \frac{\delta}{\log(n)}\right) pn$ is likely necessary for this method to work as otherwise $\mathcal{L}(G, \ell, R) \to \infty$ even when $R = 0$.

4   Equal Distance Separates the Interior from s and t

In order to prove the balance lemma, Valiant and Roughgarden discretize the space of latency functions $a_e f_e + b_e$ into collections where $a_e \in [i\tau, (i+1)\tau)$ and $b_e \in [j\tau, (j+1)\tau)$ for some pair (i, j) and a fixed small constant τ. They then show that if the latency cost for leaving s is significantly more than the latency cost for entering t, this implies that the flow on an edge in collection (i, j) leaving s is at least a constant factor larger (independent of i and j) than the flow entering t on an edge in collection (i, j). As this implies that more flow leaves s than enters t, this clearly is a contradiction yielding the balance lemma. We proceed in a similar manner, except we note that the random variables under consideration are all bounded and thus by applying the law of large numbers, we may side step the discretization argument in favor of a more direct argument.

Lemma 7. *Let G be an Erdős-Rényi random graph on n vertices with edge probability p. Let $\mathcal{A}$ and $\mathcal{B}$ be reasonable distributions and let all latency functions have the form $\ell_e(f_e) = a_e f_e + b_e$ where (a_e, b_e) is distributed according to $\mathcal{A} \times \mathcal{B}$. For any sufficiently small fixed $\delta > 0$ there are some constants $c > 1$ and $n_0 > 0$ such that, if $n > n_0$ and $\mathbb{P}\left(\mathcal{B} \leq \frac{\delta}{\log(n)}\right) np \geq c \log(n)$, then for any vertex v other than s and t in the instance $(G, \ell, R_\ell(2B))$, we have $d_s(v) \leq B + 10\delta$ with high probability.*

Proof. We proceed by contradiction. Suppose that v is a vertex, other than s and t, and that $d_s(v) > B + 10\delta$. By Corollary 6, we have that for every vertex u other then s and t, $d_s(u) > B + 2\delta$. Furthermore, since for every flow carrying vertex w, $d_s(w) + d_t(w) = 2B$, this implies that every flow carrying vertex has $d_t(w) < B - 2\delta$. Now consider the flow leaving s. Let e an edge adjacent to s, then we have that $a_e f_e + b_e > B + 2\delta$. Further, this implies that if $b_e \leq B + 2\delta$, then $f_e > \frac{B + 2\delta - b_e}{a_e} > 0$. Let E_s be the set of such edges. Now for any constant $\epsilon' > 0$, we have by Chernoff bounds that $|E_S| \geq (1 - \epsilon')\mathbb{P}(\mathcal{B} \leq B + 2\delta)\, pn$ with high probability (since $B + 2\delta$ is a constant and $pn \geq c \log(n)$). Furthermore, since the range of $\mathcal{A}$ is bounded and we are concerned only with a bounded range of $\mathcal{B}$, we may apply the law of large numbers to get convergence. In particular, if we let $\mathcal{B}_{[x,y]}$ be the random variable $\mathcal{B}$ conditioned on $x \leq \mathcal{B} \leq y$, we have that

$$R_\ell(2B) \geq \sum_{e \in E_s} \frac{B + 2\delta - b_e}{a_e} \tag{22}$$

$$= \left(B + 2\delta - \mathbb{E}\left[\mathcal{B}_{[0, B+2\delta]}\right] - o(1)\right)\left(\mathbb{E}\left[\frac{1}{\mathcal{A}}\right] + o(1)\right)|E_s| \tag{23}$$

$$\geq (1 - \epsilon')\frac{B + 2\delta - \mathbb{E}\left[\mathcal{B}_{[0, B+2\delta]}\right] - o(1)}{\mathbb{E}\left[\mathcal{A}\right] + o(1)}\mathbb{P}(\mathcal{B} \leq B + 2\delta)\, pn. \tag{24}$$

Applying a similar argument to the edges adjacent to t we have that with high probability,

$$R_\ell(2B) \le (1 + \epsilon') \frac{B - 2\delta - \mathbb{E}\left[\mathcal{B}_{[0,B-2\delta]}\right] + o(1)}{\mathbb{E}\left[\mathcal{A}\right] - o(1)} \mathbb{P}(\mathcal{B} \le B - 2\delta)\, pn. \qquad (25)$$

Thus in order to provide the contradiction, it suffices to show that for sufficiently large n

$$\frac{(1 - \epsilon') \frac{B + 2\delta - \mathbb{E}\left[\mathcal{B}_{[0,B+2\delta]}\right] - o(1)}{\mathbb{E}[\mathcal{A}] + o(1)} \mathbb{P}(\mathcal{B} \le B + 2\delta)\, pn}{(1 + \epsilon') \frac{B - 2\delta - \mathbb{E}\left[\mathcal{B}_{[0,B-2\delta]}\right] + o(1)}{\mathbb{E}[\mathcal{A}] - o(1)} \mathbb{P}(\mathcal{B} \le B - 2\delta)\, pn} > 1. \qquad (26)$$

Observing that ϵ' is arbitrary, this is equivalent to showing

$$\frac{\left(B + 2\delta - \mathbb{E}\left[\mathcal{B}_{[0,B+2\delta]}\right]\right) \mathbb{P}(\mathcal{B} \le B + 2\delta)}{\left(B - 2\delta - \mathbb{E}\left[\mathcal{B}_{[0,B-2\delta]}\right]\right) \mathbb{P}(\mathcal{B} \le B - 2\delta)} > 1. \qquad (27)$$

We note that, for positive a, b, c, $\frac{a+c}{b+c} > 1$ if and only if $\frac{a}{c} > 1$. Thus, by adding $\mathbb{E}\left[\mathcal{B}_{[0,B-2\delta]}\right] \mathbb{P}(\mathcal{B} \le B - 2\delta)$ to the numerator and denominator, we have that demonstrating Equation 27, is equivalent to showing that

$$\frac{(B + 2\delta)\, \mathbb{P}(\mathcal{B} \le B + 2\delta) - \mathbb{E}\left[\mathcal{B}_{[B-2\delta,B+2\delta]}\right] \mathbb{P}(|\mathcal{B} - B| \le 2\delta)}{(B - 2\delta)\, \mathbb{P}(\mathcal{B} \le B - 2\delta)} > 1 \qquad (28)$$

Observing that $(B + 2\delta)\, \mathbb{P}(\mathcal{B} \le B + 2\delta) \ge (B - 2\delta)\, \mathbb{P}(B - 2\delta)$, we can rearrange this to

$$0 < \left(B + 2\delta - \mathbb{E}\left[\mathcal{B}_{[B-2\delta,B+2\delta]}\right]\right) \mathbb{P}(|\mathcal{B} - B| \le 2\delta) + 4\delta \mathbb{P}(\mathcal{B} \le B - 2\delta). \qquad (29)$$

But this follows immediately from the choice of B and δ, and the reasonableness of $\mathcal{B}$. $\square$

5   Braess's Paradox Occurs in Erdős-Rényi Graphs

As in Section 2 and in Valiant and Roughgarden's work [16], in order to form the more efficient subnetwork we will classify edges adjacent to s and t as either 1-type, X-type or unclassified. We will then use these classifications to create a graph G' so that $R_\ell^{G'}(2B(1 - \mu)) > R_\ell^G(2B)$ for some μ. In the construction that follows it will be convenient to suppose that $p \ll \frac{1}{\sqrt{n}}$, and in particular, that $\Gamma(s) \cap \Gamma(t) = \varnothing$ with high probability. We will now construct G' be partitioning the internal vertices of G into three sets and considering only edges that are induced by one of the sets or are incident to one of s and t.

To that end, let p_1 be the probability that an edge is 1-type, that is, the latency function $a_e f_e + b_e$ satisfies that $b_e \in (B, (1+\epsilon)B)$ and $a_e \le A_1$. Similarly, define p_X as the probability that an edge is X-type, in other words, $b_e \le \epsilon B$ and $a_e \in ((1 - \epsilon)A_2, A_2)$. Now define $p^* = \min\{p_1, p_X\}$. With this notation we create three collections of vertices as follows:

- For each $v \in \Gamma(s) - \{t\}$ if the edge $\{s, v\}$ is 1-type assign v to the set V_{1X} with probability $\frac{p^*}{p_1}$. If it is X-type assign v to the set V_{X1} with probability $\frac{p^*}{p_X}$. Otherwise assign v to V_U.
- For each $u \in \Gamma(t) - \{s\}$, if the edge $\{u, t\}$ is 1-type assign u to the set V_{X1} with probability $\frac{p^*}{p_1}$. If it is X-type assign u to the set V_{1X} with probability $\frac{p^*}{p_X}$. Otherwise assign u to V_U.
- For each $v \notin \Gamma(s) \cup \Gamma(t) - \{s, t\}$ assign v uniformly at random to one of V_{1X}, V_{X1}, or V_U.

Define the graph G_{1X} as the subgraph of G induced by $V_{1X} \cup \{s, t\}$ excluding the edge $\{s, t\}$. Similarly define G_{X1} and G_U with the edge $\{s, t\}$ allowed to be present in G_U. It is important to note that each edge adjacent to s or t appears in precisely one of the three graphs. Thus, letting $G' = G_{1X} \cup G_{X1} \cup G_U$, we have $\deg_{G'}(s) = \deg_G(s)$ and $\deg_{G'}(t) = \deg_G(t)$.

Now in a similar manner as Lemma 5, Lemma 7 and Corollary 6 we have the following results

Lemma 8. *Let G be an Erdős-Rényi random graph on n vertices with edge probability p. Let $\mathcal{A}$ and $\mathcal{B}$ be reasonable distributions and let all latency functions have the form $\ell_e(f_e) = a_e f_e + b_e$ where (a_e, b_e) is distributed according to $\mathcal{A} \times \mathcal{B}$. For any sufficiently small fixed $\delta > 0$ there are some constants $c > 1$ and $n_0 > 0$ such that if $n > n_0$, $\mathbb{P}\left(\mathcal{B} \leq \frac{\delta}{\log(n)}\right) np \geq c\log(n)$, and G_{1X}, G_{X1}, G_U are defined as above, then the instance $\left(G', \ell, R_\ell^{G'}(2B(1-\mu))\right)$ satisfies that:*

1. *For any vertices v and u other than s and t, both in one of G_{1X}, G_{X1}, or G_U, $|d_s(v) - d_s(u)| \leq 8\delta$, and*
2. *for any vertex v other than s and t in G_U, $B - 10\delta \leq d_s(v)$,*

with high probability.

With this lemma in hand we can proceed to the proof of our main theorem.

Proof (Theorem 3). Our goal will be to show that there is some $\delta > 0$ and $\mu \in (0, 1)$ such that $R_\ell^{G'}(2B(1-\mu)) > R_\ell(2B)$. Since the latency of a Nash equilibrium flow is strictly increasing in the flow if all the latency functions are affine by Lemma 2, this implies that $\mathcal{L}\left(G', \ell, R_\ell^{G'}(2B)\right) \leq (1-\mu)\mathcal{L}\left(G, \ell, R_\ell^G(2B)\right)$ and thus the Braess's ratio is at least $(1-\mu)^{-1}$ at the flow rate $R_\ell^G(2B)$. First, we will consider the difference in flow between G and G' over the subgraph G_U and then we will consider the differences in flow over G_{1X} and G_{X1}. For a given edge e (adjacent to s or t), denote by f_e the flow on the edge in G and denote by f'_e the flow on the edge in G'.

Since, for ϵ sufficiently small p^* is a constant that is strictly less than $\frac{1}{2}$, there is some constant c^* such that $p^* pn$ and $(1 - 2p^*)pn$ are at least $c^* \log(n)$. Thus, we have that for any fixed constant $\epsilon^* > 0$, if N is the neighbor of s or T in G_{1X} or G_{X1}, then $||N| - p^* pn| \leq \epsilon p^* pn$. Similarly, if N is the neighborhood of s or t in G_U, then $||N| - (1 - 2p^*)pn| \leq \epsilon^*(1 - 2p^*)pn$ with high probability.

Suppose e is adjacent to s in G_u. Then by the Lemma 7, we have that $a_e f_e + b_e \leq B + 10\delta$, and in particular $f_e \leq \frac{B+10\delta-b_e}{a_e}$. Similarly, by Lemma 8 we have $a_e f'_e + b_e \geq B - 10\delta$ and $f'_e \geq \frac{B-10\delta-b_e}{a_e}$ and thus

$$f_e - f'_e \leq \frac{B + 10\delta - b_e}{a_e} - \frac{B - 10\delta - b_e}{a_e} = \frac{20\delta}{a_e} \leq \frac{20\delta}{A_{\min}}. \tag{30}$$

Summing over the neighbors of s in G_U we have that there is at most $(1+\epsilon^*)(1-2p^*)pn\frac{20\delta}{A_{\min}}$ more flow along those edges in G than in G'.

Now let e_s be adjacent to s in G_{1X} and let e_t be adjacent to t in G_{1X}. Then we have that

$$2B(1 - \mu) \leq a_{e_s} f'_{e_s} + b_{e_s} + 8\delta + a_{e_t} f'_{e_t} + b_{e_t} \tag{31}$$
$$\leq A_1 f'_{e_s} + (1 + \epsilon)B + 8\delta + A_2 f'_{e_t} + \epsilon B. \tag{32}$$

Thus $B - 2\mu B - 2\epsilon B - 8\delta \leq A_1 f'_{e_s} + A_2 f'_{e_t}$. Summing over all choices of e_s and e_t we get

$$\frac{|\Gamma_{G_{1X}}(s)|\,|\Gamma_{G_{1X}}(t)|\,(B - 2\epsilon B - 2\mu B - 8\delta)}{A_1\,|\Gamma_{G_{1X}}(t)| + A_2\,|\Gamma_{G_{1X}}(s)|} \leq R_\ell^{G_{1X}}(2B(1 - \mu)). \tag{33}$$

In particular, $\frac{(1-\epsilon^*)^2}{1+\epsilon^*}\frac{B-2\mu B-2\epsilon B-8\delta}{A_1+A_2}p^*pn \leq R_\ell^{G_{1X}}(2B(1 - \mu))$. Similarly, we have that $\frac{(1-\epsilon^*)^2}{1+\epsilon^*}\frac{B-2\mu B-2\epsilon B-8\delta}{A_1+A_2}p^*pn \leq R_\ell^{G_{X1}}(2B(1 - \mu))$.

Finally, consider the flow in G on the edges adjacent to s that appear in either G_{1X} or G_{X1}. For the edges in that appear in G_{1X} the flow is at most $\frac{10\delta}{A_{\min}}$ by Lemma 7. Similarly, the flow in G for an edge that appears in G_{X1} is at most $\frac{B+10\delta}{(1-\epsilon)A_2}$.

Thus. letting $\epsilon, \epsilon^*, \delta, \mu \to 0$, we have that in this limit

$$\frac{R_\ell^{G'}(2B(1 - \mu)) - R_\ell^G(2B)}{pn} \geq \frac{2B}{A_1 + A_2} - \frac{B}{A_2} > 0. \tag{34}$$

where the last inequality follows from the fact that $A_1 < A_2$. Thus, by continuity, for $\epsilon, \epsilon^*, \mu, \delta$ sufficiently small, $R_\ell^{G'}(2B(1 - \mu)) > R_\ell^G(2B)$ and Braess's paradox occurs with Braess ratio $(1 - \mu)^{-1}$. □

Now the proof of Valiant and Roughgarden [16] deals with the case where $p \gg n^{-\frac{1}{2}+\zeta}$ for some $\zeta > 0$ and our proof deals with the case where $p \ll n^{-\frac{1}{2}}$, leaving a small gap between these results. The only difficulty in extending our results to cover the gap is dealing with that fact that within this range of p there is a positive probability of $\Gamma(s) \cap \Gamma(t) \neq \varnothing$. However, this gap can be closed by using the more complicated partitioning scheme used by Valiant and Roughgarden which appropriately deals with vertices in $\Gamma(s) \cap \Gamma(t)$. Thus, with an appropriate choice of distribution for the latency function, there is a flow for which Braess's paradox occurs with high probability in $\mathcal{G}(n, p)$ almost down to the connectivity threshold. In fact, it is plausible that this work could be expanded to a broader class of expanders than $\mathcal{G}(n, p)$ under certain degree conditions.

References

1. M. Beckmann, C. B. McGuire, and C. B. Winsten. *Studies in the economics of transportation.* Yale University Press, 1956.
2. C. Fisk and S. Pallottino. Empirical evidence for equilibrium paradoxes with implications for optimal planning strategies. *Transportation Research Part A: General*, 15(3):245 – 248, 1981.
3. M. Frank. The Braess paradox. *Math. Programming*, 20(3):283–302, 1981.
4. M. A. Hall. Properties of the equilibrium state in transportation networks. *Transportation Science*, 12(3):208 – 216, August 1978.
5. 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*, volume 1, pages 437 – 445 vol.1, 2002.
6. V. Klee and G. J. Minty. How good is the simplex algorithm? In *Inequalities, III (Proc. Third Sympos., Univ. California, Los Angeles, Calif., 1969; dedicated to the memory of Theodore S. Motzkin)*, pages 159–175. Academic Press, New York, 1972.
7. G. Kolata. What if they closed 42d street and nobody noticed? *The New York Times*, December 25, 1990.
8. E. Koutsoupias and C. Papadimitriou. Worst-case equilibria. In *16th Annual Symposium on Theoretical Aspects of Computer Science*, volume 1563 of *Lecture Notes in Comput. Sci.*, pages 404–413. Springer Berlin, 1999.
9. H. Lin, T. Roughgarden, and É. Tardos. A stronger bound on Braess's paradox. In *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 340–341 (electronic), New York, 2004. ACM.
10. C. Papadimitriou. Algorithms, games, and the internet. In *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*, pages 749–753 (electronic), New York, 2001. ACM.
11. E. I. Pas and S. L. Principio. Braess' paradox: Some new insights. *Transportation Research Part B: Methodological*, 31(3):265 – 276, 1997.
12. C. M. Penchina. Braess paradox: Maximum penalty in a minimal critical network. *Transportation Research Part A: Policy and Practice*, 31(5):379 – 388, 1997.
13. T. Roughgarden. On the severity of Braess's Paradox: designing networks for selfish users is hard. *J. Comput. System Sci.*, 72(5):922–953, 2006.
14. T. Roughgarden and É. Tardos. How bad is selfish routing? *J. ACM*, 49(2):236–259 (electronic), 2002.
15. D. A. Spielman and S.-H. Teng. Smoothed analysis of algorithms: why the simplex algorithm usually takes polynomial time. *J. ACM*, 51(3):385–463 (electronic), 2004.
16. G. Valiant and T. Roughgarden. Braess's paradox in large random graphs. *Random Structures Algorithms*, to appear.
17. J. Wardrop. Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers, Part II*, 1(36):352–362, 1952.