The Price of Anarchy in Parallel - Serial Competition with Elastic Demand

John Musacchio

University of California, Santa Cruz, Technology and Information Management Program, 1156 High Street, Santa Cruz, CA 95060, USA

Abstract

We study a game theoretic model of competing network service providers that are connected in parallel and serial combinations and that strategically price their service in the presence of elastic user demand. Demand is elastic in that it diminishes both with higher prices and latency, which in turn grows linearly with a network's usage. To obtain our results, we make an analogy between the game and an electric circuit, in which the slope of each provider's latency function is made to be analogous to electrical resistance. Our bound on efficiency loss depends on the ratio of the conductance of the circuit branch with highest conductance to the combined conductance of all the branches. In terms of the original problem, the bound measures how the worst-case efficiency loss increases as the capability of the system becomes more concentrated in a particular serial combination of players.

Key words: congestion games, oligopoly competition, network competition, selfish routing
JEL: C72

1. Introduction

In many settings, consumers or users have a choice between substitutable combinations of services, where each individual service is offered by a service provider that: i) strategically sets its price, and ii) is subject to a congestion effect, meaning that the more the service is used, the more the value of the service to users degrades. We study the Price of Anarchy (PoA) of such a market, where the PoA is the ratio of the social welfare attained when a social planner chooses prices to maximize social welfare versus the social welfare attained in Nash equilibrium when service providers choose prices strategically. Note that the PoA is just the reciprocal of one minus the efficiency loss of the market. Our results quantify how the worst-case efficiency loss increases as the capability of the system becomes more concentrated in a particular combination of service providers. We reach this result by drawing an analogy to an electric circuit in which the concentration of the market in a particular serial combination of providers is analogous to the ratio of that provider-combination's conductance to the conductance of the overall circuit.

We introduce our model in the context of a market for network services from a single source to a single destination. We suppose that the network service providers in this market are connected in arbitrarily complex parallel and serial groupings. Thus users can choose from several competing paths to the destination, with each path composed of a combination of network providers. Traffic that takes a particular path must pay a price equal to the sum of the prices charged by providers along that path. Each providers network has a latency (degree of conges-

[⋆]This research was supported in part by NSF Grants CNS-0626397 and CNS-0910711.
Email address: johnmm@soe.ucsc.edu (John Musacchio)
URL: http://www.soe.ucsc.edu/∼ johnm/ (John Musacchio)

tion) that grows linearly in the traffic crossing it, and each main parallel "branch" of providers can have an additional fixed latency in a way we describe more precisely later. Given prices from all providers (a price profile), the traffic selfishly distributes itself across the possible paths so that the *disutility* defined as price plus latency – of all used paths must be the same. Moreover, the traffic demand is elastic. In particular there is a function we call a *disutility function* that describes the relationship between the total traffic that chooses to use the network, and the disutility that would induce that total amount of traffic. The disutility function is equivalent to an inverse demand function where disutility plays the role of a "total price" (both monetary and that due to latency) traffic must pay. The payoff to each provider is the revenue - the product of its price and the amount of traffic crossing the provider.

The model we study extends from a model first proposed and studied by Acemoglu and Ozdaglar (2007a) and later extended by Hayrapetyan, Tardos, and Wexler (2007). In the first version of the model studied by Acemoglu and Ozdaglar, the user elasticity is modeled by assuming that all users have a single reservation utility and that if the best available price plus latency exceeds this level, users do not use any service. In this setting, the authors find that the PoA is $(\sqrt{2}+1)/2$ (Acemoglu and Ozdaglar, 2007a). In Acemoglu and Ozdaglar (2007b), the authors extend the model to consider providers in a parallel-serial combination. Traffic chooses between several branches connected in parallel, and then for each branch, the traffic traverses the links of several providers connected serially. Hayrapetyan et al. (2007) extended the model of Acemoglu and Ozdaglar to have elastic demand in the form of a demand function that depends on the sum of price and latency. The topology they consider is just the simple topology of directly competing links between two nodes. They derived the first loose bounds on the price of anarchy for this model. Later Ozdaglar (2008) showed that the bound is actually 1.5, and furthermore that this bound is tight. Ozdaglar's derivation uses techniques of mathematical programming, and is similar to the techniques used in Acemoglu and Ozdaglar (2007a). Musacchio and Wu (2007) provide an independent derivation of the same result using an analogy to an electrical circuit where each branch represents a provider link and the current represents flow.

The present paper generalizes the work of Musacchio and Wu (2007) by considering a more general topology of providers connected in parallel and serial combinations. We derive a bound on the price of anarchy using a circuit analogy similar to the argument used in Musacchio and Wu (2007). However, this previous work only evaluated a topology of providers connected in parallel and also did not derive a bound that depends on market concentration.

To illustrate our results consider the following example. Suppose that a pool of users can choose between n separate paths indexed 1, 2, ..., n between a common source and destination node. Each path, or "branch" has 3 providers connected serially, and the jth provider along path i is indexed ij. Latency across each provider is $a_{ij} f_i$ where f_i is the amount of traffic that takes path i and a_{ij} is the slope of provider ij's latency function. Moreover each branch i has a fixed latency b_i that adds to the providers' latencies. Each provider ij charges a price p_{ij}. As explained previously, a disutility (inverse demand) function $U(\cdot)$ relates the sum of all prices plus latencies (this sum termed "disutility") on the lowest disutility path to the total flow users are willing to offer at this level of disutility. Since traffic is selfish, all used paths in equilibrium have a common disutility level. Note that any one provider can charge a high enough price so as to prevent any traffic from using his branch of the network even if the other providers charged zero. Consequently it would be a Nash equilibrium for all providers in a serially connected branch to charge such a high price. To rule these cases out we define a type a Nash equilibrium called a *zero-flow zero-price equilibrium* which requires that any provider carrying zero flow charge zero price. Our results show that a zero-flow zero-price equilibrium exists for this example (and more generally). Also the worst-case PoA amongst zero-flow zero-price equilibria can be bounded in the following way. Let $\mathcal{S}$ be the set of branches that carry positive flow in any given zero-flow zero-price equilibrium. (Branches with a high enough fixed latency b_i might not carry any flow

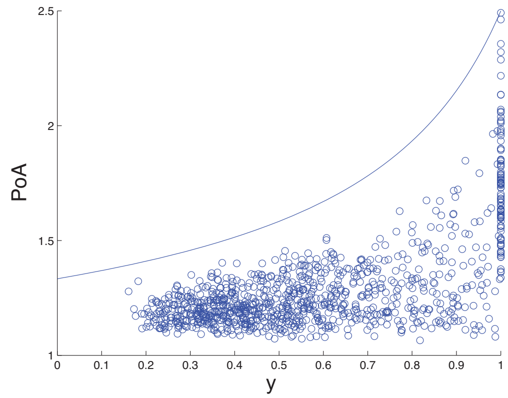

Figure 1: The price of anarchy bound for the example in the introduction. The points marked on the graph are values of the price of anarchy for numerically evaluated examples.

in an equilibrium). Let

$$y = \frac{\max_{i \in \mathcal{S}} \frac{1}{a_{i1}+a_{i2}+a_{i3}}}{\sum_{i \in \mathcal{S}} \frac{1}{a_{i1}+a_{i2}+a_{i3}}}.$$

Note that y describes the rate the "least congestable" branch's latency increases with traffic flow verses the rate the latency of the whole system increases with traffic flow. Our results in this article show that the PoA is no more than

$$\frac{y^2 + 16y - 32}{18y - 24}.$$

This bound along with the PoA of numerically evaluated examples with varying a_{ij}'s and b_i's is given in Figure 1. The figure shows how the PoA increases as the capability of the system becomes more concentrated in one branch of providers.

This kind of bound, that depends on the concentration of the system's capabilities in any one branch is one of the main contributions of this work. Also, the fact that our results bound the PoA for zero-flow zero-price equilibria of the parallel serial case with elastic demand of this family of models started by Acemoglu and Ozdaglar is novel. Finally, our proof technique which makes use of an analogy to an electric circuit is a contribution that is likely to be of use to other researchers analyzing related models of competition. Also note that this example is a topology we call a "simple parallel serial" network. Our results extend to more complicated topologies with arbitrarily complex interconnections of parallel and serial groupings of providers.

1.1. Other related work

The models of Acemoglu and Ozdaglar (2007a), Hayrapetyan et al. (2007), and the present work are part of a stream of recent literature studying the price of anarchy for games of selfish routing. Several works, see for example Roughgarden (2003, 2001, 2002), study games where users are selfish, but the network is passive, the owners of the edges are not strategically choosing their prices. Other work such as Cole et al. (2006) consider the problem of having the network choose prices to induce optimal routing among selfish users, rather than having parts of the network selfishly choose prices to maximize revenue. A recent work by Chawla and Roughgarden (2008) studies a model of Bertrand competition in networks. This work studies a model in which there are capacity constraints, rather than congestion that causes latency as we study here. Another recent work by Babaioff et al. (2009) studies how much malicious players can hurt the utility of others in congestion games.

Another related work is by Johari and Tsitsiklis (2006) where the authors study games where the users request a bit rate from the network, and in turn the network returns to users a price that depends on the sum of the requested rates on each resource. The model of Johari and Tsitsiklis (2006) and the model of this paper study very different situations. In particular in Johari and Tsitsiklis (2006) the main strategic agents are the users and in the model we study the strategic agents are link operators seeking to maximize profit. However, there are many similarities between the two games in terms of the structure of the payoff functions of the players. As in this paper, the authors of Johari and Tsitsiklis (2006) derive bounds on the efficiency loss.

Another very closely related work is by Xi and Yeh (2008). In this work, the authors study a model in which network providers are connected in a directed graph. Roughly, each provider sends a bid – in the form of a function relating price (including congestion costs) to traffic flow – to the providers one hop further away from the destination. The providers get to see the bids of the closer-in neighbors before deciding on a bid to give the further-out neighbors. This setting is quite different from ours because in our model the price setting is simultaneous and also because the congestion traffic sees is an additional cost on top of the prices the traffic pays. Both features make the games strategically different.

1.2. Example applications

One example application of particular interest is in the market for Internet access. In a market for Internet access for a regional pool of users, users select from alternative local Internet Service Providers (ISPs), and these ISPs may in-turn each be connected to separate Network Service Providers (NSPs) that connect the local ISPs to the backbone. Each of these providers may be subject to congestion, and the demand may be elastic as our model supposes. If each ISP is connected to a different NSP, we have a “simple parallel serial” market where users chose between alternative serially connected ISP and NSP options. This scenario fits the model we investigate here. If some ISPs are connected to the same NSP, then the topology is what we call “general parallel serial,” which also falls into the general model we study. Also there is the possibility that some of the ISP and NSPs are merged into a single strategic enterprise while unmerged “stacks” of ISP and NSP compete with this merged stack. This example again fits into our model. One might be interested studying the market for a particular type of application service delivered over the Internet, in which case topologies involving stacks of ISP, NSP, Applications Service Provider (ASP) would be relevant.

Another possible application lies in situations in which firms have choices over competing supply chains to make interchangeable components. Each of the firms in the chain might have a cost structure that exhibits “congestion” effects in that the marginal costs of production increase with quantity. If we suppose that each firm in the supply chain strategically chooses a price margin to charge above its average cost, the situation becomes strategically similar to the model we study here.

In section 2, we make precise our game model, and establish the relationship between the Wardrop equilibrium (equilibrium of selfish traffic selecting routes) and the solution of an analogous electrical circuit. Section 3 proves the main results of this article. Section 4 compares the derived bounds to numerically evaluated examples, and section 5 makes concluding remarks.

2. Model

We consider a network consisting of a single source and destination node, connected by an arbitrary topology of providers connected both in parallel and series as depicted in Figure 2. In order to describe the topology, we create hierarchically nested groups. Groups at an odd numbered level of the hierarchy contain subgroups connected in parallel, and groups at an

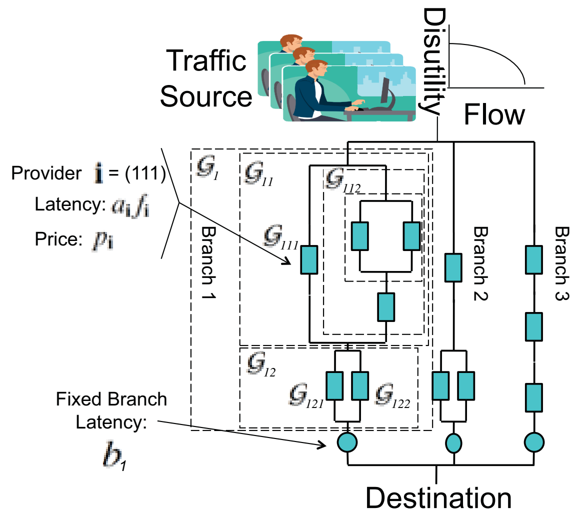

Figure 2: An illustration of the pricing game and the indexing scheme for groups.

even level of the hierarchy contain subgroups connected in series. We make this precise as follows. There are n level 1 groups, each denoted as $\mathcal{G}_i$, $i \in \{1, ..., n\}$ that are connected in parallel. We also refer to the level 1 groups as "branches." Each branch i consists of g_i level 2 groups of providers connected in series, and the set of providers in each of these level 2 groups $j \in \{1, ..., g_i\}$ is denoted as $\mathcal{G}_{ij}$. Each level 2 group $\mathcal{G}_{ij}$ is either a single provider, or consists of g_{ij} level 3 groups connected in parallel, and furthermore, each of these groups is denoted $\mathcal{G}_{ijk}$ where $k \in \{1, ..., g_{ij}\}$. In general each nth level group has indices $i_n, i_{n-1}, i_{n-2}, ..., i_1$ that indicate membership in an $n-1$ level group, an $n-2$ level group, and so on. Each nth level group for n odd, consists of $g_{i_1,i_2...,i_n}$ groups connected in parallel. Likewise each nth level group for n even consists of $g_{i_1,i_2...,i_n}$ groups connected in series. Each provider $\mathbf{i} = (i_1, i_2, ..., i_n)$ charges a price $p_{\mathbf{i}}$ per unit flow and has latency $l_{\mathbf{i}}(f_{\mathbf{i}}) = a_{\mathbf{i}} f_{\mathbf{i}}$ where $f_{\mathbf{i}}$ is the flow through provider $\mathbf{i}$. Users choose the path to take on the basis of the sum of all the prices and latencies encountered a long a path, in a way that we make precise below. We refer to the combination of price plus latency as "disutility," and therefore the disutility encountered on a single provider $\mathbf{i}$'s link is $l_{\mathbf{i}}(f_{\mathbf{i}}) + p_{\mathbf{i}}$.

In addition to the latency of each provider which scales with flow, we allow each branch i to have a fixed latency b_i that does not scale with flow. For each loop free path $\mathcal{P}$ in branch i the disutility incurred by the flow taking that path is $\sum_{\mathbf{i} \in \mathcal{P}} l_{\mathbf{i}}(f_{\mathbf{i}}) + p_{\mathbf{i}} + b_i$. (Loop free means that the path never passes through the same interconnection point more than once.)

Users are nonatomic and are free to choose the path that has the lowest disutility. Therefore in equilibrium, parallel paths that carry positive flow must have a common disutility value, otherwise users would switch paths. This equilibrium of user routing is known as a Wardrop equilibrium (Wardrop, 1952). (We elaborate the detailed conditions for a Wardrop equilibrium for this routing game later in the section.) Each player's i profit or payoff function is the product of their price $p_{\mathbf{i}}$ and the Wardrop equilibrium flow $f_{\mathbf{i}}$, which in-turn is effected by the entire profile of prices from the other players.

As explained in the introduction, demand is elastic. A function $U(f)$ which we call a disutility function relates the total flow f offered by users to the amount of disutility users would be willing to bear in order to offer that total flow. Clearly $U(x)$ is decreasing. We make the additional assumption that $U(x)$ is concave. The assumption that $U(x)$ is concave is a strong assumption, but is necessary to derive the bound in this work. Ozdaglar (2008) and Hayrapetyan et al. (2007) make this assumption as well. We also assume $U(x)$ is continuous and almost everywhere differentiable.

We summarize the components that constitute the pricing game $\mathbf{G}$ in the following definition.

Definition 1. (Game $\mathbf{G}$) The Game $\mathbf{G}$ is characterized by: a set of players (link providers) $\{\mathbf{i}\}$, linear latency functions for each link provider $\mathbf{i}$ of the form $l_{\mathbf{i}}(f_{\mathbf{i}}) = a_{\mathbf{i}} f_{\mathbf{i}}$ with $a_{\mathbf{i}} > 0$, a set of fixed branch latencies $\{b_i\}$ that specify the fixed portion of the latency for each main branch i, and a concave disutility function $U(\cdot)$ that characterizes the demand elasticity. ($U(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ is monotone non-increasing, concave, continuous, piecewise differentiable, and there exists some f_{max} for which $U(f_{max}) = 0$). A strategy profile is a set of prices $\{p_{\mathbf{i}}\}$ for each provider. The payoffs are $\{p_{\mathbf{i}} f_{\mathbf{i}}\}$ where $\{f_{\mathbf{i}}\}$ are the Wardrop equilibrium flows resulting from prices $\{p_{\mathbf{i}}\}$. (It turns out that the Wardrop equilibrium flows are uniquely determined by the price profile, and this is verified in Theorem 1.)

Note that the indexing scheme we have devised is such that the network topology is implicit from the set of player indices $\{\mathbf{i}\}$. The definition below formalizes the notion of a Wardrop equilibrium flow profile in terms of the notation we have defined.

Definition 2. (Wardrop equilibrium) In game $\mathbf{G}$, a flow profile $\{f_{\mathbf{i}}\}$ is a Wardrop equilibrium for price profile $\{p_{\mathbf{i}}\}$ iff the following conditions are satisfied:

- The flow conservation relations – for each group $\mathcal{G}_{\mathbf{i}}$:

$$\begin{aligned} & f_{\mathbf{i}} \geq 0, & \\ & f_{\mathbf{i}} \triangleq f_{\mathbf{j}} = f_{\mathbf{k}} \quad \forall \mathbf{j}, \mathbf{k} \in \mathcal{G}_{\mathbf{i}} & \text{if } \mathcal{G}_{\mathbf{i}} \text{ is a series interconnection,} \\ & f_{\mathbf{i}} \triangleq \sum_{\mathbf{j} \in \mathcal{G}_{\mathbf{i}}} f_{\mathbf{j}} & \text{if } \mathcal{G}_{\mathbf{i}} \text{ is a parallel interconnection.} \end{aligned}$$

- For each group $\mathcal{G}_{\mathbf{i}}$ that consists of a single provider: $D_{\mathbf{i}} = a_{\mathbf{i}} f_{\mathbf{i}} + p_{\mathbf{i}}$.

- For each non-singleton sized group $\mathcal{G}_{\mathbf{i}}$:

$$D_{\mathbf{i}} \triangleq \begin{cases} \sum_{\mathbf{j} \in \mathcal{G}_{\mathbf{i}}} D_{\mathbf{j}} & \text{if } \mathcal{G}_{\mathbf{i}} \text{ is a series interconnection,} \\ \min_{\mathbf{j} \in \mathcal{G}_{\mathbf{i}}} D_{\mathbf{j}} & \text{if } \mathcal{G}_{\mathbf{i}} \text{ is a parallel interconnection.} \end{cases}$$

- For each group $\mathcal{G}_{\mathbf{i}}$ that is a parallel interconnection: $D_{\mathbf{j}} \triangleq D_{\mathbf{i}}$ if $f_{\mathbf{j}} > 0 \quad \forall \mathbf{j} \in \mathcal{G}_{\mathbf{i}}$.

- For all branches (top level groups) $i, j \in 1...n$:

$$\begin{aligned} D_i + b_i &= D_j + b_j & \text{if } f_i,\ f_j > 0, \\ D_i + b_i &\geq D_j + b_j & \text{if } f_i = 0,\ f_j > 0. \end{aligned}$$

The notation $\mathbf{j} \in \mathcal{G}_{\mathbf{i}}$ signifies that group $\mathcal{G}_{\mathbf{j}}$ is a subgroup of $\mathcal{G}_{\mathbf{i}}$ and is exactly one level below in the hierarchy. We are interested in studying the Nash equilibria of the game $\mathbf{G}$ (we will show that at least one Nash equilibrium exists). In this context a Nash equilibrium is a price profile such that no provider can improve his profit by unilaterally changing price. We will compare this to the social optimum pricing of the game $\mathbf{G}$ in which a hypothetical central planner chooses all prices to maximize the social welfare of the whole system – the sum of provider profits plus the welfare of users. Our measure of user welfare is consumer surplus. The classic definition of consumer surplus considers the difference between each user's willingness to pay and what she actually pay, and then sums this difference across all the users. In this model the surplus is found by integrating the difference between each unit of flow's willingness to tolerate disutility, and the disutility the flow actually bears. Therefore the surplus, and hence user welfare, is given by

$$\int_0^f (U(x) - d)dx$$

where f is the total flow carried by the system, and $d = U(f)$ is the equilibrium disutility found on all used branches.

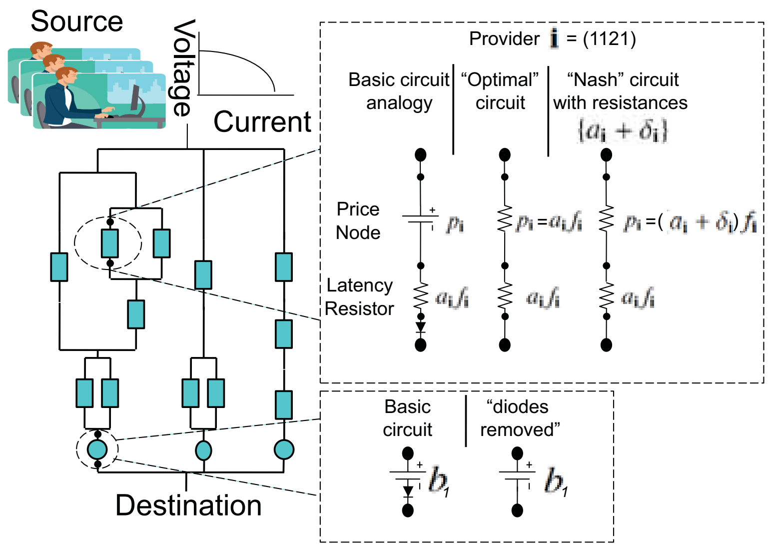

Figure 3: An illustration of the circuit analogy. The notions of "Nash" circuit and "Social Optimal" circuit are introduced in sections 3.1.1 and 3.1.2 respectively.

2.1. Circuit analogy

The Wardrop equilibrium conditions have an exact analogy in the electric circuit pictured in Figure 3 where voltage and current are analogous to disutility and flow. Before we explain the analogy further, we review a few basics of circuit analysis for readers not familiar with the subject. The circuits we consider in our analysis are composed of the following basic components:

- **voltage source:** A circuit element with two terminals that keeps a fixed voltage drop no matter the amount of current (flow) passing through the element. In circuit diagrams, a voltage source is pictured as two parallel lines with a "+" mark on the side where the voltage is positive.
- **resistor:** A resistor with resistance a_i is a two terminal element such that the voltage is always equal $a_i f_i$ where f_i is the current across the resistor. In circuit diagrams, a resistor looks like a "sawtooth" line segment.
- **diode:** An ideal diode is a two-terminal element such that current is allowed in only one direction. A diode is pictured as an arrow pointing in the direction of the allowed direction of the current, and a horizontal line on the other side of the arrow signifying that reverse current is stopped. One can conceptualize a diode as a nonlinear resistor with an infinite resistance when the voltage drop is negative, and zero resistance otherwise.
- **nonlinear power source:** A circuit element whose voltage and current always lie along a specified nonlinear function. The voltage is non increasing with current.

In circuit analysis, one considers interconnections of components like the above and then finds "solutions" (voltages at every interconnection point and currents through each element) by writing equations that represent the following rules:

- The net flow into any interconnection point of elements must be zero. This is known as "Kirchoff's Current Law" (KCL).
- The voltage drop across any series of interconnected elements is the sum of the voltage drops across the individual elements. If there are two paths in the circuit between the

same two nodes, the voltage drop across the paths must be the same. This is known as "Kirchoff's Voltage Law" (KVL).

For the circuits we consider in this paper, the KCL and KVL equations have a unique flow profile solution[1]. For our problem, the KCL rule is analogous to the rule that the flow must be conserved throughout the network. The KVL rule is analogous to the fact that the disutility along any path is the sum of the prices and latencies encountered on that path. The KVL also captures the key property of Wardrop equilibria – that the disutility along competing paths must be the same if they are both being used. In the circuit, parallel branches must have the same voltage drop. In Wardrop equilibria, an unused path can have a disutility higher then competing paths. We model this in the circuit by using diodes to essentially "shut off" paths whose disutility (voltage) is higher than the competing paths.

Definition 3. (Circuit Analogy to Game $\mathbf{G}$) The circuit analogy to game $\mathbf{G}$ consists of:

1. A nonlinear power source with voltage-current characteristic given by disutility function $U(\cdot)$.
2. A series interconnection of circuit components representing each provider $\mathbf{i}$ in game $\mathbf{G}$. Each interconnection contains:
   - a *"price node"* consisting a voltage source $p_{\mathbf{i}}$,
   - a *"latency resistor"* with resistance $a_{\mathbf{i}}$,
   - a perfect diode which prevents the flow from ever becoming negative through provider $\mathbf{i}$,
3. A voltage source of fixed voltage b_i on each branch i,
4. A diode on each branch i preventing the flow from becoming negative through branch i.

As we have said before, the topology is implied from the set of provider indices $\{\mathbf{i}\}$. The lemma below validates that finding a solution to the circuit analogy is equivalent to finding a Wardrop equilibrium. The proof involves building a correspondence between the $D_{i\ldots j}$ in the Wardrop equilibrium definition and voltages in the circuit analogy.

Lemma 1. *A flow profile $\{f_{\mathbf{i}}\}$ is a Wardrop equilibrium to game $\mathbf{G}$ with price profile $\{p_{\mathbf{i}}\}$ if and only if $\{f_{\mathbf{i}}\}$ satisfies the KVL and KCL conditions of the circuit analogy to game $\mathbf{G}$.*

The proof is a relatively simple exercise of comparing circuit equilibrium conditions to Wardrop equilibrium conditions. The detailed proof is in the appendix.

At various points in our analysis, we will make modifications to the basic circuit analogy. For instance, we will replace the "pricing node" of each provider with a resistor to find the socially optimal pricing, and a different, larger, resistor to construct a Nash equilibrium. Also at some points in the analysis, we will need to replace the nonlinear power source with a linear approximation consisting of a fixed voltage source in series with a resistor whose slope matches the slope of the nonlinear power source at some desired operating point.

In circuits, the power an element dissipates is the product of the current times the voltage drop across the element. By analogy, the power dissipated by a provider's price node is the provider's revenue, while the power across a provider's latency resistor is the total utility lost due to the latency of a provider's network.

In circuits it is standard to find the equivalent resistance of a series interconnection of resistors by adding the resistances. Likewise the reciprocal of the sum of reciprocals gives the equivalent

[1]It turns out that assignment of voltages to nodes may not be unique along paths with zero flow in the circuits we consider. For example consider a series interconnection of 2 ideal diodes with a negative voltage across them. The current is zero because the diodes block, but the voltage drop across each particular diode is not uniquely defined.

resistance of a parallel interconnection. These basic facts can be verified directly with the KCL and KVL rules. With that in mind, we define notation for the equivalent resistance of each level of groupings in the circuit.

- If $\mathcal{G}_{i\ldots j}$ is a singleton, then $a_{i\ldots j}$ is the slope of that player's latency function.
- If $\mathcal{G}_{i\ldots j}$ is a group of elements connected serially, then

$$a_{i\ldots j} = \sum_k a_{i\ldots jk}.$$

- If $\mathcal{G}_{i\ldots j}$ is a group of elements connected in parallel, then

$$a_{i\ldots j} = \bigoplus_k a_{i\ldots . jk} \triangleq \left[\sum_k a_{i\ldots jk}^{-1}\right]^{-1}.$$

In the expression above, and subsequently in the paper, we adopt the notation $\bigoplus_k a_{i\ldots . jk} \triangleq \left[\sum_k a_{i\ldots jk}^{-1}\right]^{-1}$, for convenience because we will often have to express the equivalent resistance of parallel interconnections.

2.2. Existence and uniqueness of Wardrop equilibrium for a price profile

In this subsection we use the circuit analogy to verify that given a price profile $\{p_{\mathbf{i}}\}$, which specifies the prices from all the providers (players), that there exists a unique Wardrop equilibrium flow profile $\{f_{\mathbf{i}}\}$. Since provider revenues are found by taking the product of flow and price we need this result to insure that the payoffs of game $\mathbf{G}$ are well defined.

Theorem 1. *In game* $\mathbf{G}$*, for any price profile* $\{p_{\mathbf{i}}\}$ *with* $p_{\mathbf{i}} \geq 0\ \forall \mathbf{i}$*, a Wardrop equilibrium flow profile* $\{f_{\mathbf{i}}\}$ *exists and is unique.*

The proof is by construction of a solution to the a circuit analogy of game $\mathbf{G}$. The detailed proof is in the appendix.

2.3. Thévenin equivalent resistances

In order to evaluate the conditions for Nash equilibrium, we need to understand how a unilateral price change of a provider will change the flow, and in turn profits, of that provider. To address this, we can use a basic result from circuit analysis known as Thévenin's theorem to reduce a complex resistive circuit to a simple one. Thévenin's theorem says that any resistive circuit seen from a "port" or pair of nodes can be reduced to a an open-circuit voltage and a Thévenin equivalent resistance (Chua et al., 1987). In terms of the original problem, this is essentially aggregating the combined effects of elastic demand and competition from other routes into a simple linear model of how a single provider's price change will affect its flow.

Before finding the Thévenin equivalent seen at the price node, we first model the nonlinear voltage source using a linear approximation. In the circuit domain this turns out to be the same as modeling the source as a voltage source of constant voltage in series with a resistor. For a given total flow f, the power source has a voltage-current characteristic with slope $s(f) = U'(f)$. (Note it is possible that $U(\cdot)$ has different left and right hand derivatives at the operating point – we address this possibility later.) Around this operating point, the nonlinear voltage source can be approximated as a constant voltage source of some size V and a serially connected resistor of size s. Note that if the original nonlinear voltage source is connected to a load that draws a flow $f+\epsilon$, the voltage of the source drops to $V - s(f+\epsilon) - O(\epsilon^2)$. (The $-O(\epsilon^2)$ is a consequence of concavity.) Consequently, the voltage from the nonlinear source is never more than that of its linear approximation.

The circuit of Figure 3 is not completely resistive in that it has diodes. For the time being, we will suppose that the circuit is in an operating point for which only diodes at the bottom of each branch are blocking flow. This assumption is equivalent to assuming that providers are not charging a price higher than the disutility of competing paths. We define the set ON to be the set of branches for which the flow is not being blocked by the diodes. Also note that the voltage sources in each provider's "price node" do not affect the Thévenin equivalent resistance seen by other groups. In terms of the original problem, the intuition behind this is that following a small, unilateral change in price by one provider, the change of flow that provider sees will only depend on the latency functions of the other providers that are carrying flow but not the prices of these other providers.

We define the following notation for the Thévenin equivalent resistance seen at different levels of the hierarchy of groups. Let $\delta_{\mathbf{i}}(s, ON)$ denote the Thévenin equivalent resistance seen after removing $\mathcal{G}_i$ from the circuit analogy and measuring from the nodes where the group was connected. Thus

$$\delta_i(s, ON) = \left[\bigoplus_{j \neq i,\ j \in ON} a_j \right] \bigoplus s.$$

This is because after removing group $\mathcal{G}_i$, any current introduced at the nodes where the group used to be connected can travel on paths through each of group $\mathcal{G}_i$'s peers, or it can traverse the path through the power source whose linear approximation we said was a voltage source and resistor of size s. All of these options are in parallel, so the resistance seen is the parallel combination of the resistances of all the paths.

Deeper into the hierarchy, the Thévenin resistance can be found recursively. If $\mathcal{G}_{i\ldots j}$ contains subgroups connected serially, then each subgroup $g_{i\ldots jk}$ "sees" whatever resistance parent group $\mathcal{G}_{i\ldots j}$ saw in parallel with the resistances of the sibling groups. Consequently, the Thévenin resistance is

$$\delta_{i\ldots jk}(s, ON) = \left[\bigoplus_{l \neq k} a_{i\ldots jl} \right] \bigoplus \delta_{i\ldots j}(s, ON).$$

If $\mathcal{G}_{i\ldots jk}(s, ON)$ contains subgroup of elements connected in parallel, then by similar reasoning, we find that

$$\delta_{i\ldots jk} = \delta_{i\ldots j}(s, ON) + \sum_{l \neq k} a_{i\ldots jl}.$$

3. Price of anarchy analysis

Our analysis is organized in the following way. Section 3.1.1 argues why the ratio of price to flow for each provider must be equal to the slope of its latency function plus the Thévenin equivalent resistance seen by that provider in the circuit analogy. This section also introduces the notion of a "Nash circuit," in which each provider's price node is replaced with a resistor. Later we show why a Nash equilibrium must solve the circuit equations of the Nash circuit. Section 3.1.2 argues why the ratio of price to flow for each provider must be equal to the slope of its latency function. This section also introduces the concept of the "optimal circuit," which is similar to the Nash circuit except that each provider's price node is replaced with resistor that is smaller than that used for constructing the Nash circuit. The solution of the optimal circuit gives the social optimum flows and prices. Section 3.2 proves the existence of a Nash equilibrium and shows that the equilibrium is the solution of a Nash circuit. The section obtains these results by addressing the issue of finding a set of branches to be "on" for the purposes of computing Thévenin equivalents that is self consistent. In the course of the analysis, we find how branches that are "off" in Nash equilibrium can still influence the prices charged in the "on" branches. Section 3.3 examines how to compute the equivalent resistances of arbitrary parallel-serial interconnections of providers. Section 3.4 argues why the search for the worst case price of anarchy can be limited to examples with disutility functions that are composed of a flat line segment followed by a downward sloping line segment. Section 3.5 defines some vector and matrix notation and derives a lemma based on the matrix inversion lemma that is a key step in the proof of our main result. Section 3.6 brings all the ideas of the preceding analysis sections together to obtain our final results.

3.1. Price-flow relationships

In this subsection we find relations between the price of each link and the flow on each link with both Nash equilibrium and the social optimum pricing.

3.1.1. Nash equilibrium

Providers connected serially can have very inefficient equilibria. Consider just two providers connected serially, each of which charges a price larger than the highest point on the disutility curve. The flow through these providers will be zero, and furthermore it is a Nash equilibrium for these providers to charge these prices. This is because one player cannot cause the flow to become positive by lowering his price, so each player is playing a best response to the other player's prices. In examples for which there are more than one provider connected serially on each branch, one can construct a Nash equilibrium in which no flow is carried, and therefore the "worst-case" price of anarchy across all Nash equilibria is infinite.

However these infinite price of anarchy Nash equilibria seem unlikely to be played in a real situation. Intuitively, a provider would want to charge a low enough price so that it would be at least possible that he could carry some flow if the other providers on the branch also had a low enough price. With that in mind, we would like to identify a more restricted set of Nash equilibria that seen more "reasonable" and that have a price of anarchy that can be bounded.

In the examples where serial players charge very high prices, the player's best response is not unique, and we can use that fact to chose a class of equilibria that excludes these examples. For instance, we can consider the notion of *strict* Nash equilibrium. Like a conventional Nash equilibrium, a strict Nash equilibrium requires that each player's strategy be a best response to the other players. However a strict Nash equilibrium requires that each player be playing a *unique* best response. In this context, a strict equilibrium requires that every player carry a positive flow, because if otherwise their profit would be zero and there are many prices that lead to a profit of zero. Acemoglu and Ozdaglar (2007b) use the notion of strict equilibrium in order to have a class of equilibria with bounded price of anarchy for parallel-serial competition without elastic demand.

In our model, it is not possible in some cases to find a Nash equilibrium in which all players carry a positive flow, which is required in strict equilibrium. For example, if the fixed latency b_i on a branch i is large enough, it will not carry flow regardless of what the provider prices are.

We would like to define a class of equilibria that has the property that it always exists, and that it is possible to bound the price of anarchy. To that end, we define a notion of a *zero-flow zero-price* equilibrium.

Definition 4. (Zero-Flow Zero-Price Equilibrium) A *zero-flow zero-price equilibrium* is a type of Nash equilibrium of a game $\mathbf{G}$ for which players who are carrying zero-flow are charging a zero-price.

The intuition motivating this definition is that a player who is not attracting any flow will try lowering his price as far as possible in order to attract some flow. We will later show that there always exists a zero-flow zero-price equilibrium, and that the price of anarchy is bounded for such an equilibrium.

At this point, we need to identify the condition for a player to be playing a best response in order to construct Nash equilibria. In the following Theorem we show that a player playing a best response must have a price to flow ratio equal to the Thévenin equivalent resistance he sees from the rest of the circuit, plus the slope of his latency function.

Theorem 2. *Consider a set of prices $\{p_{\mathbf{i}}\}$ and flows $\{f_{\mathbf{i}}\}$ that satisfy the Wardrop equilibrium conditions. Let ON^+ be the set of branches that are on following a small increase in price by player $\mathbf{i}$, and let ON^- be the set of on branches following a small decrease. Likewise let be s^+ and s^- be the magnitude of the right-hand and left-hand derivatives of the disutility function at the operating point. More precisely $s^+ = -\lim_{\epsilon\downarrow 0}(U(f+\epsilon) - U(f))/\epsilon$ and $s^- = -\lim_{\epsilon\uparrow 0}(U(f+\epsilon) - U(f))/\epsilon$.*

A necessary and sufficient condition for a player $\mathbf{i}$ to be playing a best response is that

$$\frac{p_{\mathbf{i}}}{f_{\mathbf{i}}} \in [\delta_{\mathbf{i}}(ON^+, s^-) + a_{\mathbf{i}}, \delta_{\mathbf{i}}(ON^-, s^+) + a_{\mathbf{i}}]. \tag{1}$$

If $ON^+ = ON^- = ON$ and $s^+ = s^- = s$, the condition reduces to

$$\frac{p_{\mathbf{i}}}{f_{\mathbf{i}}} = \delta_{\mathbf{i}}(ON, s) + a_{\mathbf{i}}.$$

Moreover if the above is satisfied, the best response is strict.

Proof. Consider a player $\mathbf{i}$ charging price $p_{\mathbf{i}}$ and carrying flow $f_{\mathbf{i}}$. A unilateral change in price by an amount $+\epsilon$ will change the flow the provider carries by

$$-\frac{\epsilon}{\delta_{\mathbf{i}}(ON^+, s^-) + a_{\mathbf{i}}} - O(\epsilon^2).$$

The $-O(\epsilon^2)$ notation indicates that $\lim_{\epsilon\to 0} -O(\epsilon^2)/\epsilon = 0$. Away from $\epsilon = 0$, $-O(\epsilon^2)$ is non-positive. The term $-O(\epsilon^2)$ is non-positive because convexity of the disutility function makes it such that the actual flow coming from the power source is never more than the "estimate" given buy the source's linear approximation. The profit resulting from this price change is

$$(p_{\mathbf{i}} + \epsilon)\left(f_{\mathbf{i}} - \frac{\epsilon}{\delta_{\mathbf{i}}(ON^+, s^-) + a_{\mathbf{i}}} - O(\epsilon^2)\right) = p_{\mathbf{i}} f_{\mathbf{i}} + \left[-\frac{p_{\mathbf{i}}}{\delta_{\mathbf{i}}(ON^+, s^-) + a_{\mathbf{i}}} + f_{\mathbf{i}}\right]\epsilon - O(\epsilon^2)$$

The above expression is less than the original profit $p_{\mathbf{i}} f_{\mathbf{i}}$ for all nonzero ϵ if and only if

$$\frac{p_{\mathbf{i}}}{f_{\mathbf{i}}} \geq \delta_{\mathbf{i}}(ON^+, s^-) + a_{\mathbf{i}}.$$

A similar analysis for a unilateral price decrease yields

$$\frac{p_{\mathbf{i}}}{f_{\mathbf{i}}} \leq \delta_{\mathbf{i}}(ON^-, s^+) + a_{\mathbf{i}}.$$

□

At this point we introduce the concept of a *Nash circuit*.

Definition 5. (Nash Circuit) The *Nash circuit* with resistances $\{a_{\mathbf{i}} + \delta_{\mathbf{i}}(\mathcal{S}, s)\}$ is the circuit that results by making the following modifications to the basic circuit analogy for game $\mathbf{G}$. For each provider $\mathbf{i}$, replace the price node (a voltage source of size $p_{\mathbf{i}}$) with a resistor of size $a_{\mathbf{i}} + \delta_{\mathbf{i}}(\mathcal{S}, s)$ (see Figure 3) and remove the diode that insures nonnegative flow through provider $\mathbf{i}$. Retain the diodes that insure nonnegative flow through each branch (top level group).

The set $\mathcal{S}$ is in the above definition is an arbitrary subset of branches $\{1...n\}$. Note that the part of the definition about removing diodes does not change the solution of the resulting circuit by the following reasoning. Once each provider's part of the circuit is strictly resistive, and the diodes for each primary branch are retained, it is impossible for the flow through a provider to become negative. The Nash circuit is a useful construction because a solution to it is such that for each provider $\mathbf{i}$, the flow $f_{\mathbf{i}}$ and price $p_{\mathbf{i}} = [a_{\mathbf{i}} + \delta_{\mathbf{i}}(\mathcal{S}, s)]f_{\mathbf{i}}$ satisfy the best response condition (1), provided the set $\mathcal{S}$ equals the set ON. Recall ON is the set of branches that are carrying positive flow.

3.1.2. Social optimum

Now consider the socially optimal pricing. We first give an intuitive argument why the socially optimal price should be $a_{\mathbf{i}}f^*_{\mathbf{i}}$ where $f^*_{\mathbf{i}}$ is the flow of link i in social optimum. The theory of Pigovian taxes says that one way to align the objectives of individuals with that of society as a whole is to make market individuals pay the cost of their externalities on others (Pigou, 1920). Following this notion the socially optimal pricing should price the flow so that each user bears the marginal cost to society of each additional unit of new flow.

The latency on each branch is $a_i f^*_{\mathbf{i}} + b_i$. However, the social cost is the latency times the amount of flow bearing that latency. Therefore the cost is $a_i f^{*2}_{\mathbf{i}} + b_i f^*_{\mathbf{i}}$. Thus the marginal cost is $2a_{\mathbf{i}} f^*_{\mathbf{i}} + b_i$. The latency borne by users is $a_i f^*_{\mathbf{i}} + b_i$. Therefore to make the disutility borne by users reflect marginal social cost, the price should be set to $a_i f^*_{\mathbf{i}}$.

With this observation in mind, we define the following notion of an *optimal circuit* which we will use to conceptualize the social optimum pricing

Definition 6. (Optimal Circuit) The *optimal circuit* for a game $\mathbf{G}$ is the circuit that results from taking the basic circuit analogy for game $\mathbf{G}$ and making the following modifications for each provider $\mathbf{i}$: i) replace the price node with a resistor of size $a_{\mathbf{i}}$, ii) delete the diode in provider $\mathbf{i}$'s part of the circuit, iii) retain the diode on each branch (top level group).

Just as we noted when defining the Nash circuit, the later part of the definition about deleting the diodes in the collection of elements representing each provider is possible because it is impossible for the flow to go backwards once every provider's part of the circuit is strictly resistive, and the diodes for each primary branch are retained (part 3 of Definition 3).

In the lemma that follows, we show that the solution to the optimal flow problem can be found by solving Kirchoff's current and voltage laws for the optimal circuit. To express the optimization problem, we define $\mathcal{T}$ be the set of group indices that refer to groups consisting of a single provider.

Lemma 2. *Consider the game G. A social optimum flow profile $\{f^*_{\mathbf{i}}\}$ routes satisfies Kirchoff's current and voltage laws (KCL and KVL) for the optimal circuit. Therefore, prices satisfying the relation $p_{\mathbf{i}} = a_{\mathbf{i}} f^*_{\mathbf{i}}$ for each provider $\mathbf{i}$ induce optimal flows.*

Proof. A socially optimal assignment of flows to routes solves the following constrained optimization problem:

$$\begin{array}{lll}
\multicolumn{3}{l}{\max \left[\int_0^f U(x)dx\right] - \sum_{\mathbf{i}\in\mathcal{T}} a_{\mathbf{i}} f^2_{\mathbf{i}} - \sum_i b_i f_i} \\
\text{s.t.} & \sum_k f_{i...jk} - f_{i...j} = 0 & \text{for each parallel group } \mathcal{G}_{i...j}, \\
& f_{i...jk} - f_{i...jl} = 0 & \text{for each serial group } \mathcal{G}_{i...j}, \\
& & \text{and } (i...jk), (i...jl) \in \mathcal{G}_{i...j}, \\
& f_{\mathbf{i}} \geq 0 & \forall \mathbf{i}.
\end{array} \tag{2}$$

Note that the objective is concave and Slater's constraint qualification condition holds so there is no duality gap if we use Lagrangian techniques to find the optimal solution (Boyd and Vandenberghe, 2004). We therefore may find optimality conditions for the solution to the above problem by writing the Lagrangian, evaluating the first order conditions as well as the complementary slackness conditions for the Lagrange multipliers associated with the inequality constraints and simplifying.

For groups $\mathcal{G}_{i...j}$ consisting of single providers connected in parallel this procedure reveals that

$$2a_{i...jk}f^*_{i...jk} = V(\mathcal{G}_{i...j}) \quad \forall(i...jk) \in \mathcal{G}_{i...j}$$

where $V(\mathcal{G}_{i...j})$ is the Lagrange multiplier on the flow conservation constraint. This implies that

$$f^*_{i...j} = \frac{V(\mathcal{G}_{i...j})}{2a_{i...j}}.$$

This relation is analogous to the circuit relation that current through resistors connected in parallel is found by dividing the voltage drop by the equivalent resistance of the parallel combination of resistors. For each serial group $\mathcal{G}_{i...j}$ of singletons the fact that $f^*_{i...jk} = f^*_{i...j}$ allows us to write

$$\sum 2a_{i...jk}f^*_{i...j} = \sum_k V(\mathcal{G}_{i...jk}).$$

For convenience, we can define $V(\mathcal{G}_{i...j})$ to be the right side of the above expression. $V(\mathcal{G}_{i...j})$ is analogous to a voltage drop across resistors connected serially. For each parallel grouping $\mathcal{G}_{i...j}$ of arbitrarily sized subgroups, we find that $V(\mathcal{G}_{i...j}) = V(\mathcal{G}_{i...jk})$ for all $(i...jk) \in \mathcal{G}_{i...j}$. In other words, the voltage drops across all branches connected in parallel are the same. This relationship, along with the flow conservation constraint leads to

$$f^*_{i...j} = \frac{V(\mathcal{G}_{i...j})}{2a_{i...j}}.$$

For serial groupings of $\mathcal{G}_{i...j}$ of arbitrary sized subgroups, we can define

$$V(\mathcal{G}_{i...j}) \triangleq \sum_k V(\mathcal{G}_{i...jk})$$

so that

$$f^*_{i...j} = \frac{V(\mathcal{G}_{i...j})}{2a_{i...j}}.$$

For the flow across the N main branches, we have that

$$\begin{cases} D^* - 2a_i f^*_i - b_i = 0 & f^*_i > 0 \\ D^* - b_i < 0 & f^*_i = 0. \end{cases}$$

where $D^* = U(f^*)$ is the disutility of the used branches in the optimal solution.

Thus we have verified the exact parallel between the optimality conditions of the flow assignment problem, and the KCL and KVL equations of the optimal circuit. $\square$

Lemma 3. *Consider the optimal circuit with all diodes removed. (This allows the flow in some branches to possibly be negative.) Then in a solution to circuit equations, the quantity*

$$\left[\int_0^f U(x)dx\right] - \sum_{\mathbf{i}\in\mathcal{T}} a_{\mathbf{i}} f_{\mathbf{i}}^2 - \sum_i b_i f_i$$

is at least as large as the social welfare in social optimum of game **G**.

Proof. Consider the following optimization problem:

$$\begin{array}{ll} \max \left[\int_0^f U(x)dx\right] - \sum_{\mathbf{i}\in\mathcal{T}} a_{\mathbf{i}} f_{\mathbf{i}}^2 - \sum_i b_i f_i & \\ \text{s.t.} \quad \sum_k f_{i...jk} - f_{i...j} = 0 & \text{for each parallel group } \mathcal{G}_{i...j}, \\ \qquad\;\; f_{i...jk} - f_{i...jl} = 0 & \text{for each serial group } \mathcal{G}_{i...j}, \\ & \text{and } (i...jk),(i...jl) \in \mathcal{G}_{i...j}. \end{array}$$

This can be seen to be the same problem as (2) except that the constraint that flows be positive is relaxed. Relaxing the constraints cannot decrease the solution to a maximization problem. It is easy to verify that the first order conditions of the Lagrangian of this problem correspond to the KVL and KCL equations of the optimal circuit with the diodes removed. □

3.2. Nash equilibrium existence

In Theorem 2 we have found conditions for a provider to be playing a strict best response. In this subsection we address whether a Nash equilibrium exists.

To start with, we will suppose that the set ON (the set of branches that are "on.") for purposes of computing Thévenin equivalents is fixed to a particular set $\mathcal{S}$ even if that assignment leads to a different set of branches actually being on in the circuit. Later, we will resolve this by finding a fixed point so that $\mathcal{S} = ON$. We start with the following lemma.

Lemma 4. *Fix a constant $s > 0$ and fix a set $\mathcal{S}$ of branch indices so that $\delta_{\mathbf{i}}(s, \mathcal{S})$ are taken to be constants for all $\mathbf{i}$. The Nash circuit with resistances $\{a_{\mathbf{i}} + \delta_{\mathbf{i}}(s, \mathcal{S})\}$ has a unique solution.*

The proof is just a matter of showing the intuitive result that this simple circuit with fixed resistances and diodes has a unique solution. Since one can construct circuits with diodes that have non unique solutions, a careful argument is needed to show that the circuit we consider has a unique solution. The proof is in the appendix.

The next lemma considers the solution to the circuit when we allow the resistor of size s, which represents demand elasticity, is allowed to vary with the total flow f.

Lemma 5. *Given a set of branch indices $\mathcal{S}$, there is a unique flow profile $\{f_{\mathbf{i}}\}$, total flow f, and slope s such that: i) $s \in [U'_-(f), U'_+(f)]$, and ii) $\{f_{\mathbf{i}}\}$ solves the Nash circuit with resistances $\{a_{\mathbf{i}} + \delta_{\mathbf{i}}(s, \mathcal{S})\}$.*

Proof. For each fixed s Lemma 4 says that there is a unique solution to the Nash circuit. An increase in s increases each $\delta_{\mathbf{i}}(s, \mathcal{S})$ which therefore must decrease the total flow in the unique circuit solution. This relation describes a monotone decreasing function $s \to f(s)$ (note the monotonicity is strict wherever the flow is positive). The range is restricted to $[0, f_{\max}]$ where $f_{\max} = U^{-1}(0)$ and the domain is $[0, \infty)$ The slope of $U(f)$ is monotone increasing by concavity therefore $-U'(f)$ is a monotone increasing function in f mapping $[0, f_{\max}]$ to $[0, s_{\max}]$. The two functions must have a unique intersection point, and this point is the unique solution. The functions may cross at a point of discontinuity of $-U'(f)$. In this case s must lie between the left and right hand derivatives. □

Now we can prove the main result of the section.

Theorem 3. *There exists at least one Nash equilibrium. Furthermore, a Nash equilibrium of one of the following two types exists:*

1. *There is a set of Nash equilibria that satisfies the following two properties: i) In the reduced player game with only players carrying positive flow, the players are in strict Nash equilibrium, ii) The players carrying no flow are in branches such that even if all players in the branch reduced their price to zero, the flow in the branch would still be zero, thus the prices of players in these branches are arbitrary.*

2. *There exists a set of Nash equilibria that satisfies the following properties: i) The players carrying positive flow are playing strict best responses, ii) There is at least one branch i with $b_i = U(f)$ and all players in that branch charge 0 price. iii) Branches with $b_i > U(f)$ carry no flow, and the prices of the players can be chosen arbitrarily.*

Proof. Without loss of generality, number the branches in increasing order of b_i. Let the value of $D(k)$ for each k be the value of $U(f)$ found by analyzing the Nash circuit with resistances $\{a_{\mathbf{i}} + \delta_{\mathbf{i}}(s, \mathcal{S}_k)\}$ where $\mathcal{S}_k = \{1, ..., k\}$ and $s = U'(f)$ or s is in between the left and right hand derivatives of $U(f)$. The solution is well defined and unique by Lemma 5. The function $D(k)$ is monotone non-increasing by the following reasoning. The larger k, the smaller the value of $\delta_{\mathbf{i}}(s, \mathcal{S}_k)$ for each $\mathbf{i}$, and thus the larger f. Since $U(f)$ is monotone non-increasing, it follows that $D(k)$ is monotone non-increasing. A branch i is not on if $b_i \leq D(k)$. Therefore as k increases, the number of on branches will decrease (not-increase). Let $J(k)$ be the number of on branches as a function of k. If there is a fixed point $J(k) = k$ that fixed point corresponds to a Nash equilibrium in which k branches are turned on, and the rest undercut.

If there is no fixed point, then there is a crossing with $J(k) > k$ and $J(k+1) < k+1$. Let $F(\{\delta_{\mathbf{i}}\})$ be the total flow in the Nash circuit with fixed resistances $\{a_{\mathbf{i}} + \delta_{\mathbf{i}}\}$. (Lemma 4 insures this is well defined.) Let f^{k+1} be the unique fixed point flow for set $\mathcal{S}_{k+1}$ that is determined in Lemma 5. In particular, the fixed point satisfies the relations: $f^{k+1} = F(\{\delta_{\mathbf{i}}(s_{k+1}, \mathcal{S}_{k+1})\})$ for a particular $s_{k+1} \in [U'_-(f_{k+1}), U'_+(f_{k+1})]$. Similarly, let f^k be the unique fixed point flow satisfying: $f^k = F(\{\delta_{\mathbf{i}}(s_k, \mathcal{S}_k)\})$, for a particular $s_k \in [U'_-(f_k), U'_+(f_k)]$.

Then in the former case, branch $k+1$ is undercut (i.e. $U(f^{k+1}) \leq b_{k+1}$) but not in the latter case, ($U(f^k) > b_{k+1}$). Let $f = U^{-1}(b_{k+1})$ and note that $f \in [f^k, f^{k+1})$. (Since $f^k < f^{k+1}$ and $U(f^k) > U(f^{k+1})$, $U(\cdot)$ is invertible in the region f lies.)

Note that

$$F(\{\delta_{\mathbf{i}}(U'_+(f), \mathcal{S}_{k+1})\}) \geq f^{k+1}$$

because $U'_+(f) \leq s_{k+1}$. We argue that

$$F(\{\delta_{\mathbf{i}}(U'_-(f), \mathcal{S}_k)\}) < f$$

by the following reasoning. If the flow were larger than f, the fixed point would be larger than f, but we know that the fixed point is f^k. Thus we have that

$$F(\{\delta_{\mathbf{i}}(U'_-(f), \mathcal{S}_k)\}) < f \leq F(\{\delta_{\mathbf{i}}(U'_+(f), \mathcal{S}_{k+1})\}).$$

By continuity, there exists an assignment (not necessarily unique) of resistances

$$\delta^*_{\mathbf{i}} \in [\delta_{\mathbf{i}}(U'_+(f), \mathcal{S}_{k+1}), \delta_{\mathbf{i}}(U'_-(f), \mathcal{S}_k)], \;\; \forall \mathbf{i}$$

for which the flow is $f = F(\{\delta^*_{\mathbf{i}}\})$, and thus $U(f) = b_{k+1}$. Suppose now that each provider $\mathbf{i}$ in branches $\mathcal{S}_k$ charge prices with a price to flow ratio $a_{\mathbf{i}} + \delta^*_{\mathbf{i}}$ and providers in branch $k+1$ charge 0. Branch $k+1$ carries no flow because $U(f) = b_{k+1}$. However, if any provider in branches $1...k$ increase price, they will induce flow in branch $k+1$, so the Thévenin equivalent they (providers in branches $1...k$) see for price increases is $\delta_{\mathbf{i}}(U'_+(f), \mathcal{S}_{k+1})$ and for price decreases it is $\delta_{\mathbf{i}}(U'_-(f), \mathcal{S}_k)$. Thus having the price-flow ratio in the above range satisfies the necessary and sufficient condition for a best response given in Theorem 2.

□

Corollary 1. *There exists a zero-flow zero-price equilibrium.*

Proof. From Theorem 3 one can construct a set of Nash equilibrium of either type 1 or 2 from the statement of Theorem 3. For type 1, the prices of the players carrying no flow are arbitrary, so they can be set to zero. For type 2, each branch i with $b_i > U(f)$ have arbitrary prices, so they can be set to zero. In either case, the resulting Nash equilibrium satisfies the zero-flow zero-price property. □

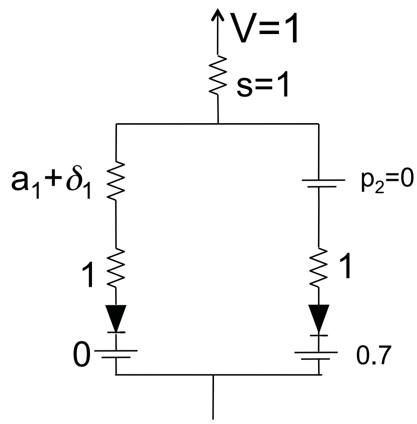

Figure 4: An illustration of Example 1.

Corollary 2. *In any zero-flow zero-price equilibrium, or in any Nash equilibrium of the type described in the statement of Theorem 3, there exist constants* $\{d_{\mathbf{i}}\}$*, with* $d_{\mathbf{i}} \in [0, 1]$ *such that the Nash equilibrium prices and flows satisfy the KVL and KCL equations of a Nash circuit with resistances* $\{a_{\mathbf{i}} + d_{\mathbf{i}}\delta_{\mathbf{i}}(s, ON)\}$ *where* f *is the total flow in Nash equilibrium,* ON *is the set of branches carrying positive flow, and* $s \in [U'_-(f), U'_+(f)]$*.*

Proof. This result follows directly from the proof of Theorem 3. □

The following simple example illustrates the role of the d variables, and how branches that carry no flow can still influence the pricing of the other branches.

Example 1. *Consider a game* $\mathbf{G}$ *with: disutility function* $U(f) = 1 - sf$*, where* $s = 1$*; two providers connected in parallel with latency function slopes* $a_1 = a_2 = 1$*; and fixed branch latencies* $b_1 = 0$*,* $b_2 = 0.7$*. Note that the disutility function can be exactly proxied by a voltage source of voltage 1 and resistance of 1 in series. The example is illustrated in Figure 4.*

The Thévenin equivalent seen by provider 1, assuming for the moment that all branches are "on" is, $\delta_1 = (a_2^{-1} + s^{-1})^{-1} = \frac{1}{2}$*. To try to construct a zero-flow zero-price Nash equilibrium strategy, we consider the circuit analogy with provider 1's price node replaced with a resistor of size* $a_1 + \delta_1 = 1.5$*, and suppose that provider 2's price is 0. The disutility (voltage drop across provider 1) with this substitution will be* $\frac{a_1+a_1+\delta_1}{s+a_1+a_1+\delta_1} = \frac{5}{7}$*. With this level of disutility, no flow would be willing to use provider 2, because the fixed latency on his branch is greater than this. Of course, that contradicts our initial assumption that all the branches were "on" when we computed the Thévenin equivalent provider 1 sees. In other words, we had assumed that provider 1 viewed provider 2 as a competitor that could absorb some flow if provider 1 raised his price slightly, but in fact that does not seem to be the case.*

Next we consider the Thévenin equivalent with branch 2 off. The Thévenin equivalent is $\delta_1 = s = 1$*, which if branch 2 were off, would lead to a voltage drop across provider 1 of* $\frac{a_1+a_1+\delta_1}{s+a_1+a_1+\delta_1} = \frac{3}{4}$*. However, this level of voltage drop (disutulity), would turn provider 2's branch back on again, provided that provider 2 lowered his price sufficiently.*

Now suppose that rather than picking a price such that the price to flow ratio were $\delta_1 + a_1 = 2$*, provider 1 instead picked a price such that the price to flow ratio were* $d\delta_1 + a_1$ *where* $d = \frac{1}{3}$*. Now the voltage drop across provider* 1 *would be* 0.7*, and provider 1's price would be* 2/5*. Moreover, if provider 1 increased price slightly, provider 2 would start carrying some flow. Thus, the Thévenin equivalent provider 1 sees for a small price increase is different than for a small price decrease. Moreover the best response condition (1) is satisfied. We also see that this Nash equilibrium is* *a solution to the Nash circuit with resistances $\{a_i + d_i\delta_i(s(f), ON)\}$ for a particular choice of $\{d_i\}$ (particularly $d_1 = \frac{1}{3}$), just as Corollary 2 says ought to be the case.*

3.3. Equivalent resistances in the Nash circuit

Corollary 2 says that in Nash equilibrium, the prices and flows must solve the Nash circuit with resistances $\{a_\mathbf{i} + d_\mathbf{i}\delta_\mathbf{i}(s, ON)\}$ for some set of non-negative constants $\{d_\mathbf{i}\}$ that are each less than 1. In this subsection we will find expressions for the equivalent resistances of branches of the Nash circuit by composing the resistances corresponding to individual providers. This process of composing resistances will allow us to treat arbitrary parallel-serial topologies by finding an "equivalent" simple parallel serial topology.

Knowing the overall total resistance in each branch of the circuit will allow us to write expressions for the total flow in each branch, which is critical to being able to express the total social welfare achieved in Nash equilibrium. To write expressions for welfare we will also need to express the cost of latency. In terms of the circuit, we need to express the power dissipated by the providers' "latency resistors."

The flow across each provider $\mathbf{i}$ incurs a latency of $a_\mathbf{i} f_\mathbf{i}$ and thus the latency loss if $a_\mathbf{i} f_\mathbf{i}^2$. In the circuit, this corresponds to the power dissipated by the provider $\mathbf{i}$'s latency resistor. Similarly the power dissipated by the resistor of size $a_\mathbf{i} + d_\mathbf{i}\delta_\mathbf{i}$ in provider i's pricing node in the Nash circuit corresponds to the provider's profit. As we compute the effective resistance of groupings, we need to keep track of what fraction of the effective resistance is due to latency cost. In other words if r is the equivalent resistance of a group, when flow across the group is f, the power dissipated is rf^2, some fraction of that will be latency loss. If the power dissipated as latency loss is lf^2, we call l the *equivalent latency resistance* of the group. We make this more precise as follows.

Definition 7. (Equivalent Resistance) Consider a group $\mathcal{G}_{i\ldots j}$ of providers in the Nash circuit with resistances $\{a_\mathbf{i} + d_\mathbf{i}\delta_\mathbf{i}\}$. The *equivalent resistance* r of the group is such that the voltage (latency) across the group is rf whenever the total flow through the group is f. Equivalently the power dissipated by the group is rf^2.

Definition 8. (Equivalent Latency Resistance) Consider a group $\mathcal{G}_{i\ldots j}$ of providers in the Nash circuit with resistances $\{a_\mathbf{i} + d_\mathbf{i}\delta_\mathbf{i}\}$. Suppose that whenever the flow through a group is f, the total power dissipated by the all the provider latency resistors in the group is lf^2. Then the *equivalent latency resistance* of the group is l.

The following lemma states the result we need concerning the equivalent resistance and equivalent latency resistance of groups of providers.

Lemma 6. *Consider the Nash circuit with resistances $\{a_\mathbf{i} + d_\mathbf{i}\delta_\mathbf{i}\}$. Let $m_{i\ldots j}$ be the maximum number of providers a loop free path across group $\mathcal{G}_{i\ldots j}$ crosses. The equivalent resistance $r_{i\ldots j}$ of a group $\mathcal{G}_{i\ldots j}$ is*

$$(1 + \bar{m}_{i\ldots j})a_{i\ldots j} + d_{i\ldots j}\bar{m}_{i\ldots j}\delta_{i\ldots j}$$

form some $\bar{m}_{i\ldots j} \in [0, m_{i\ldots j}]$ and $d_{i\ldots j} \in [0, 1]$. Furthermore, the equivalent latency resistance is

$$(1 + \bar{m}_{i\ldots j})c_{i\ldots j}a_{i\ldots j} + c_{i\ldots j}d_{i\ldots j}\bar{m}_{i\ldots j}\delta_{i\ldots j}$$

for some $c_{i\ldots j} \in [0, \frac{1}{2}]$.

The proof proceeds by using the basic formulas for resistance of parallel and serial interconnections. The fact that the equivalent latency resistance is no more than 1/2 of the overall equivalent resistance is a consequence of each provider's price node resistor is always larger its latency resistor. The detailed proof is in the appendix.

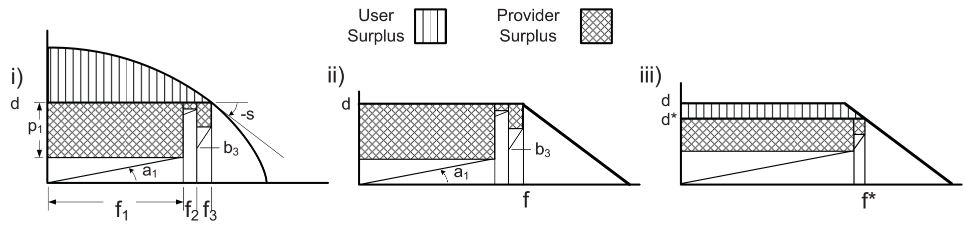

Figure 5: i) The Nash equilibrium of the game $\mathbf{G}$. (In this example three providers compete in parallel, and the shaded areas represent each of their profits.) ii) The Nash equilibrium of the game $\mathbf{G}_t$, where the disutility function of $\mathbf{G}$ has been linearized and "truncated." iii) Game $\mathbf{G}_t$ with social optimum flows. Note that the flow $f^* > f$ and that link 2 is not used when flows are socially optimal.

3.4. Utility function truncation

In this subsection, we argue that given i) a disutility function $U(\cdot)$, ii) a Nash equilibrium of the type characterized by Theorem 3, and Nash equilibrium flow f, iii) an associated price of anarchy, we can create a new example with the disutility function replaced by one that is identically equal to $U(f)$ for flows between 0 and f, and then is a decreasing line of slope s for flows more than s. This new example will have a larger price of anarchy than the old example. Therefore we can restrict our search for the worst case price of anarchy to disutility functions of this shape. We begin the development of the argument with the following lemma.

Lemma 7. *Consider a Nash equilibrium of the game $\mathbf{G}$ with flows $\{f_{\mathbf{i}}\}$ and prices $\{p_{\mathbf{i}}\}$, and total flow f. Let $\mathbf{G}_t$ be a game that is identical to $\mathbf{G}$ except that the disutility curve that is modified from the disutility curve in $\mathbf{G}$ by "truncating" it in the following way:*

$$U(x) = \begin{cases} d & x \leq f \\ d - s^+(x - f) & x > f. \end{cases}$$

Then flows $\{f_{\mathbf{i}}\}$ and prices $\{p_{\mathbf{i}}\}$ describe a Nash equilibrium of game $\mathbf{G}_t$.

Proof. The argument is essentially the same as presented in Hayrapetyan et al. (2007). Any provider $\mathbf{i}$ with positive flow in the Nash equilibrium of $\mathbf{G}$ satisfies condition (1). Moreover any provider with zero flow in the Nash equilibrium of $\mathbf{G}$ must be in a situation such that if that provider unilaterally lowered its to an arbitrarily small positive amount, its flow (revenue) would be zero.

Now consider flows $\{f_{\mathbf{i}}\}$ and prices $\{p_{\mathbf{i}}\}$ in the context of game $\mathbf{G}_l$. These flows and prices satisfy the Wardrop equilibrium conditions. Furthermore, the best response condition for any player $\mathbf{i}$ with positive flow becomes

$$\frac{p_{\mathbf{i}}}{f_{\mathbf{i}}} \in [\delta_{\mathbf{i}}(ON^+, s^+) + a_{\mathbf{i}}, \delta_{\mathbf{i}}(ON^-, \infty) + a_{\mathbf{i}}]$$

which is a weaker condition than the condition that we already know $f_{\mathbf{i}}$ and prices $p_{\mathbf{i}}$ satisfy. Furthermore any player with zero flow cannot induce a positive flow with an arbitrarily small price, for if he could, he would have been able to in game $\mathbf{G}$. We conclude flows $\{f_{\mathbf{i}}\}$ and prices $\{p_{\mathbf{i}}\}$ describe a Nash equilibrium of game $\mathbf{G}_t$. □

Lemma 8. *Consider a Nash equilibrium of the game $\mathbf{G}$ with flows $\{f_{\mathbf{i}}\}$ and prices $\{p_{\mathbf{i}}\}$. By Lemma 7 the same flows and prices describe a Nash equilibrium of game $\mathbf{G}_t$. Let $N(\mathbf{G})$ and $S(\mathbf{G})$ be the social welfare in the Nash equilibrium being considered and in social optimum respectively* *of the game* $\mathbf{G}$*. Likewise let* $N(\mathbf{G}_t)$ *and* $S(\mathbf{G}_t)$ *be the social welfare in the Nash equilibrium and social optimum of game* $\mathbf{G}_t$*. Then*

$$\frac{S(\mathbf{G}_t)}{N(\mathbf{G}_t)} \geq \frac{S(\mathbf{G})}{N(\mathbf{G})}.$$

Proof. The argument is also adapted from Hayrapetyan et al. (2007). Let $U(\mathbf{G})$ be the user welfare in the Nash equilibrium of game $\mathbf{G}$. In the corresponding Nash equilibrium of game $\mathbf{G}_t$, the prices and flows are the same, and thus the provider welfare will be the same in both cases. However, the user welfare will be 0. Thus we have

$$N(\mathbf{G}_t) = N(\mathbf{G}) - U(\mathbf{G}).$$

Consider a new game $\mathbf{G}_{tx}$ with the same latency functions as $\mathbf{G}$ but with a disutility function that has been truncated but not linearized. Therefore

$$U_{tx}(x) = \begin{cases} d & x \leq f \\ U(x) & x > f. \end{cases}$$

where f and $d = U(f)$ are the Nash equilibrium flow and disutility of $\mathbf{G}$, $U_{tx}(\cdot)$ is the disutility function of $\mathbf{G}_{tx}$, and $U(\cdot)$ is the disutility function of $\mathbf{G}$.

From Corollary 2, the Nash equilibrium of $\mathbf{G}$ solves the Nash circuit with appropriately chosen resistances. Similarly, the social optimum solution of game $\mathbf{G}$ solves the optimal circuit by lemma 2. The resistances of the optimal circuit are less than or equal to those of the Nash circuit, so consequently, the total flow f^* carried in social optimum must be no less than than f. Thus the social optimal disutility level $d^* = U(f^*)$ satisfies $d^* \leq d$.

Since we know d^* of game $\mathbf{G}$ is less than d we can view the optimization problem of finding the optimal flow vector for game $\mathbf{G}_{tx}$ as the same problem as finding the optimal flow vector for the game $\mathbf{G}$ but with the constant $U(\mathbf{G}) = \int_0^f (U(x) - d)dx$ subtracted off the objective function. Thus, the social optimum flow vector will be identical for game $\mathbf{G}_{tx}$ as for game $\mathbf{G}_t$. Thus

$$S(\mathbf{G}_{tx}) = S(\mathbf{G}) - U(\mathbf{G}).$$

Now consider the game $\mathbf{G}_t$ where we have both truncated and linearized the disutility function. By convexity, the disutility function of $\mathbf{G}_t$ is never smaller than that of $\mathbf{G}_{tx}$. Thus the solution to the optimization problem of finding the optimal flow vector has an objective function that is never smaller than the objective function that we maximized to find $S(\mathbf{G}_{tx})$. Consequently

$$S(\mathbf{G}_t) \geq S(\mathbf{G}_{tx}) = S(\mathbf{G}) - U(\mathbf{G}).$$

Combining our observations, we have

$$\frac{S(\mathbf{G}_t)}{N(\mathbf{G}_t)} \geq \frac{S(\mathbf{G}) - U(\mathbf{G})}{N(\mathbf{G}) - U(\mathbf{G})} \geq \frac{S(\mathbf{G})}{N(\mathbf{G})}.$$

□

3.5. Vector notation and algebraic identities

We now compare the social optimum and Nash equilibrium welfare. Let n' be the number of providers that carry flow in a zero-flow zero-price equilibrium of the type characterized by Theorem 3. In social optimum, the number of branches that carry no flow cannot decrease. This can be seen by the following reasoning. The zero-flow zero-price equilibrium solves the Nash circuit where the resistances for provider prices are $d_\mathbf{i}\delta_\mathbf{i} + a_\mathbf{i}$ for some set of constants $d_\mathbf{i} \in [0, 1]$. The social optimum circuit has the same structure, but with the resistances on all branches reduced, thus the social optimum has a higher total flow, and thus $U(f^*) \leq U(f)$ where f is the total flow in Nash equilibrium and f^* is the total flow in social optimum. Branches are undercut (carry no flow) only when $U(f^*) \leq b_i$, thus the Nash equilibrium cannot have more branches undercut than the social optimum solution.

We define the following notation. The vectors of flows in Nash equilibrium and social optimum are

$$F = [f_1, f_2, ..., f_n]^T \text{ , and } F^* = [f_1^*, f_2^*, ..., f_n^*]^T$$

respectively. We define

$$\mathbf{V} = [V, V, ...V]^T \text{ , and } b = [b_1, b_2, ...b_n]^T$$

as an n dimensional vector of all V's and the vector of b_i's respectively. The matrices

$$A = \text{diag}(a_1, a_2, ..., a_n) \text{ , } \Delta = \text{diag}(\delta_1, \delta_2, ..., \delta_n) \text{ , } D = \text{diag}(d_1, d_2, ..., d_n)$$

are diagonal matrices of the $a_i's$, δ_i's, and d_i's respectively. It will also be convenient to define

$$J = \begin{bmatrix} 1 & 1 & ... \\ 1 & 1 & ... \\ \vdots & \vdots & \ddots \end{bmatrix}, \quad t \triangleq \text{tr}(A^{-1}), \quad \alpha \triangleq 1/s + t, \text{ and } \quad e \triangleq 1_{n \times 1}.$$

The matrix J above is simply and $n \times n$ matrix of all ones and e is a m dimensional column vector. We state and prove the following lemma.

Lemma 9. *The following identities hold:*

$$(2A + sJ)^{-1} = \frac{1}{2}\left[A^{-1} - \frac{1}{t + 2s^{-1}}A^{-1}JA^{-1}\right],$$
$$\Delta = (\alpha A - I)^{-1}A.$$

Proof. The matrix inversion lemma states that

$$(F + UCV)^{-1} = F^{-1} - F^{-1}U(C^{-1} + VF^{-1}U)^{-1}VF^{-1}$$

where F, U, C, V are arbitrary matrices with only the condition that the matrices are of appropriate dimension and all the inverses in the above expression exist. Let $F = 2A$, $U = e$, $V = e^T$, $C = s/2$ and note that $J = ee^T = 1_{m \times m}$. Then we may apply the matrix inversion lemma in the following way:

$$\begin{aligned}(2A + sJ)^{-1} &= \frac{1}{2}A^{-1} - \frac{1}{4}A^{-1}e\left[\frac{1}{s} + \frac{e^TA^{-1}e}{2}\right]^{-1}e^TA^{-1} \\ &= \frac{1}{2}\left[A^{-1} - A^{-1}e\left[\frac{2}{s} + \text{tr}(A^{-1})\right]^{-1}e^TA^{-1}\right] \\ &= \frac{1}{2}\left[A^{-1} - \frac{1}{2s^{-1} + t}A^{-1}JA^{-1}\right].\end{aligned}$$

To derive the next identity, recall $\delta_i = \left[\sum_{j \neq i} \frac{1}{a_i} + \frac{1}{s}\right]^{-1}$. Therefore

$$\begin{aligned}\Delta^{-1} &= \left[\text{tr}(A^{-1}) + \frac{1}{s}\right] - A^{-1} \\ &= \alpha I - A^{-1}.\end{aligned}$$

Thus $A\Delta^{-1} = \alpha A - I$, and therefore $\Delta A^{-1} = (\alpha A - I)^{-1}$. Note by construction, $\alpha > 1/\min_i a_i$ so that $(\alpha A - I)^{-1}$ exists. Finally we may conclude that $\Delta = (\alpha A - I)^{-1}A$. ◻

3.6. Price of anarchy analysis

In this section, we derive the main results of this paper. We find upper bounds for the price of anarchy of the pricing game, where we take the price of anarchy to mean the worst case ratio of the social welfare in a zero-flow zero-price Nash equilibrium to the social optimum welfare. Our bounds are found to be a function of a parameter y we term the *conductance ratio*. The conductance ratio is the conductance of the most conductive branch divided by the conductance of network as a whole. The conductance ratio is therefore a measure of how concentrated the capabilities of the network are in a single branch. A conductance ratio near 1 means that most of the conductance of the system is concentrated in a single branch. The smaller the conductance ratio, the more that the overall conductance of the system is distributed across multiple branches. Thus, in a sense the conductance ratio reflects the market power or concentration of the system. As one would expect, the PoA bounds that we find increase as the conductance ratio approaches 1.

3.6.1. Simple parallel-serial competition

Our result for the simple parallel-serial topology is given by the following theorem.

Theorem 4. *Consider game* $\mathbf{G}$ *with a simple parallel-serial topology. Consider the following ratio*

$$y = \frac{\max_i 1/a_i}{\sum_i 1/a_i},$$

which is the conductance of the most conductive branch (among those branches that carry positive flow in Nash equilibrium)[2] *divided by the overall conductance. The price of anarchy for zero-flow zero-price Nash equilibria is no more than*

$$\begin{cases} \frac{1}{4}\frac{m^2+2m(1+y)+(y-1)^2}{m} & y \le 1 - m/3 \\ \frac{m^2(2-y)+m(4-y^2-y)+2(y-1)^2}{8m-6my} & y \ge 1-m/3 \end{cases} \tag{3}$$

where m *is the maximum number of providers connected serially. Furthermore, the maximum of the above bound occurs when* $y = 1$, *and consequently the price of anarchy is mo more than*

$$1 + m/2. \tag{4}$$

Proof. We suppose that the system consists of branches indexed by i containing m_i competitors connected serially. We consider a zero-flow zero-price Nash equilibrium of game $\mathbf{G}$ with flow vector F. Using the arguments presented in Section 3.4, the same prices and flows describe a Nash equilibrium for the game $\mathbf{G}_t$ with a truncated disutility function. We suppose that there are n providers that carry positive flow in zero-flow zero-price equilibrium, and therefore we define all vectors to be n dimensional [3]. Using the result of Corollary 2, there exist some set of $\{d_{\mathbf{i}}\}$ such that the zero-flow zero-price equilibrium solves the Nash circuit with resistances $\{a_i + d_{\mathbf{i}}\delta_{\mathbf{i}}\}$.

For each $\mathbf{i} = ij$, δ_{ij} is related to the the Thévenin equivalent seen by the branch i by the relation $\delta_i j = \sum_{j' \neq j} a_{ij'} + \delta_i$. Consequently, for each branch i, the resistances associated with provider prices can be added as follows

$$\sum_{j=1}^{g_i} a_{ij} + d_{ij}\delta_{ij} \ge \sum_{j=1}^{g_i} \left(a_{ij} + \sum_{j' \neq j} a_{ij'} + \delta_i \right) = \sum_{j=1}^{g_i} m_i a_{ij} + \delta_i = m_i a_i + m_i \delta_i$$

[2] Any branch i with fixed latency b_i high enough to prevent it from carrying any flow in zero-flow zero-price equilibrium is not considered when we look for the "most conductive branch."

[3] Even though our vectors only include the branches that carry positive flow, the branches carrying no flow can influence the pricing in the other branches. See section 3.2 and Example 1 in particular to recall how the influence of these "off" branches is modeled by the d parameters.

Consequently, there exists an $\bar{m}_i \in [1, ..., m_i]$ such that the resistance associated with provider price on each branch is $\bar{m}_i a_i + \bar{m}_i d_i \delta_i$. The total resistance of the branch includes this resistance, plus the resistance associated with latency, which is just a_i. Consequently, we may write the the following KVL equations for the Nash circuit:

$$((\bar{M} + I)A + sJ + \bar{M}D\Delta)F = V - b$$

where $\bar{M}$ is a diagonal matrix of $\bar{m}_i$'s and the other vectors and matrices are as was introduced in section 3.5. A flow vector that describes a zero-flow zero-price equilibrium of $\mathbf{G}$ also describes a zero-flow zero-price equilibrium of game $\mathbf{G}_t$. The social welfare W_t in this equilibrium of $\mathbf{G}_t$ is just the provider profit, since the user welfare is zero. This welfare is expressed as the quadratic form

$$W_t = F^T(\bar{M}A + \bar{M}D\Delta)F. \tag{5}$$

Now consider the social optimum welfare. By Lemma 8, the PoA of $\mathbf{G}_t$ is at least as high as that of $\mathbf{G}$. Thus we focus on finding expressions for the optimum social welfare of $\mathbf{G}_t$. As we showed in Lemma 3 an upper bound on the social optimum welfare can be found by evaluating the social optimum circuit with diodes removed.

The KVL equations for the social optimum circuit with diodes removed expressed in vector form are $(2A + sJ)F^* = V - b$, while the provider profit can be seen to be $F^{*T}AF^*$. The consumer surplus for this flow vector F^* and disutility function $U_t(\cdot)$ (recall that this function is constant from 0 to the Nash equilibrium flow level, and then decreases linearly with slope $-s$) is $\frac{s}{2}F^{*T}JF^* - \frac{s}{2}F^TJF$. Combining these relations, and the fact that $F^* = (2A + sJ)^{-1}(V - b)$, an upper bound on social welfare in social optimum is

$$\widetilde{W}_t^* = \frac{1}{2}(V - b)^T(2A + sJ)^{-1}(V - b) - \frac{s}{2}F^TJF.$$

To bound the price of anarchy to be no more than an amount z it is sufficient to show that $zW_t - \widetilde{W}_t^*$ is always positive. Algebra reveals that

$$\begin{aligned}
zW_t - \widetilde{W}_t^* &= F^T\Big(z\bar{M}A + z\bar{M}D\Delta + \frac{1}{2}sJ\Big)F - \\
&\quad \frac{1}{2}F^T\Big((\bar{M} + I)A + sJ + \bar{M}D\Delta\Big)(2A + sJ)^{-1}\Big((\bar{M} + I)A + sJ + \bar{M}D\Delta\Big)F \qquad (6)\\
&= F^T\Big(((z - 1/2)\bar{M} - 1/2I)A + (z - 1/2)\bar{M}D\Delta\Big)F - \\
&\quad \frac{1}{2}F^T\Big((\bar{M} - I)A + \bar{M}D\Delta\Big)(2A + sJ)^{-1}\Big((\bar{M} + I)A + sJ + \bar{M}D\Delta\ \Big)F \\
&= F^T\Big((z - 1)\bar{M}A + (z - 1)\bar{M}D\Delta\Big)F - \\
&\quad \frac{1}{2}F^T\Big((\bar{M} - I)A + \bar{M}D\Delta\Big)(2A + sJ)^{-1}\Big((\bar{M} - I)A + \bar{M}D\Delta\Big)F \\
&= F^T\Big((z - 1)\bar{M}A + (z - 1)\bar{M}D\Delta\Big)F - \\
&\quad \frac{1}{4}F^T\Big[\Big((\bar{M} - I)A + \bar{M}D\Delta\Big)\Big(A^{-1} - \frac{A^{-1}JA^{-1}}{t + 2s^{-1}}\Big)\Big((\bar{M} - I)A + \bar{M}D\Delta\Big)\Big]F
\end{aligned}$$

where $t = \mathrm{tr}(A^{-1})$. The last step results from applying the matrix inversion lemma identity we found in Lemma 9. Consider the matrix in the above expression that is left and right multiplied by F^T and F respectively. To show the above expression is non-negative, we can either show that this matrix is positive definite, or since F always has positive entries, we can simply show that all entries of the matrix are non-negative. We take this approach. It is easy to verify that the off-diagonal elements are positive, therefore we need to focus on finding conditions that make the diagonal elements positive. The diagonal elements are given by

$$\big(\bar{m}(z-1)a_i + \bar{m}(z-1)d_i\delta_i\big) - \frac{1}{2}\big((\bar{m}-1)a_i + \bar{m}d_i\delta_i\big)\left(a_i^{-1} - \frac{a_i^{-2}}{t+2s^{-1}}\right)\big((\bar{m}-1)a_i + \bar{m}d_i\delta_i\big)$$

where we have let $\bar{m} = \bar{m}_i$. Substitution of the identity $\Delta = (\alpha A - I)^{-1}A$ from Lemma 9 followed by some algebra reduces the above expression to

$$\frac{N_0 + N_1 s^{-1} + N_2 s^{-2} + N_3 s^{-3}}{((t+s^{-1})a-1)^2(t+2s^{-1})}$$

where

$$\begin{aligned}
N_0 =& \bar{m}ta_i(ta_i-1)(ta_i-1+d_i)z - \frac{1}{4}(ta_i-1)\times \\
& \big((\bar{m}+1)^2t^2a_i^2 + 2ta_i(\bar{m}^2(d_i-1)+\bar{m}d_i-1) + (\bar{m}(d_i-1)+1)^2\big), \\
N_1 =& \bar{m}\big(4(ta_i-1)^2 + (2+3d_i)(ta_i-1)+d_i\big)z - (\bar{m}+1)^2(ta_i-1)^2 - \\
& \bar{m}(\frac{3}{2}(\bar{m}+1)d_i+2)(ta_i-1) - \bar{m}d_i(\frac{1}{2}\bar{m}d_i+1), \\
N_2 =& \bar{m}(5ta_i-4+2d_i)z - \frac{5}{4}(\bar{m}+1)^2ta_i^2 + \frac{5}{4} + (\frac{5}{4}-d_i)\bar{m}^2 + (-d_i+\frac{3}{2})\bar{m}, \\
N_3 =& 2\bar{m}z - 1/2(\bar{m}-1)^2.
\end{aligned}$$

These expressions are monotone increasing in z. Therefore there exists a sufficiently large z to make each of these expressions positive. For each $i = 0...3$ let z_i be the unique value for which $N_i(z_i) = 0$. The values of z_i can be found to be

$$\begin{aligned}
z_0 &= \frac{1}{4}\frac{(\bar{m}+1)^2x^2 + (2\bar{m}^2d + 4\bar{m} + 2\bar{m}d)x + 4\bar{m}d + \bar{m}^2d^2}{\bar{m}(x+d)(1+x)}, \\
z_1 &= \frac{1}{2}\frac{2(\bar{m}+1)^2x^2 + (3\bar{m}d + 4\bar{m} + 3\bar{m}^2d)x + \bar{m}^2d^2 + 2\bar{m}d}{\bar{m}(4x^2 + (2+3d)x + d)}, \\
z_2 &= \frac{1}{4}\frac{5(\bar{m}+1)^2x + 4\bar{m}^2d + 4\bar{m} + 4\bar{m}d}{\bar{m}(1+5x+2d)}, \qquad z_3 = \frac{1}{4}\frac{(\bar{m}+1)^2}{\bar{m}}
\end{aligned}$$

where $x \triangleq ta_i - 1 \geq 0$, and $d = d_i$. The derivatives w.r.t d of each of the above expressions are all nonnegative. Therefore the highest critical values z_i occur when $d = 1$. Substituting $d = 1$ we have

$$\begin{aligned}
z_0(x,\bar{m}) &= \frac{1}{4}\frac{(\bar{m}+1)^2x^2 + (2\bar{m}^2 + 6\bar{m})x + 4\bar{m} + \bar{m}^2}{\bar{m}(1+x)^2}, \\
z_1(x,\bar{m}) &= \frac{1}{2}\frac{2(\bar{m}+1)^2x^2 + (7\bar{m} + 3\bar{m}^2)x + \bar{m}^2 + 2\bar{m}}{\bar{m}(4x+1)(x+1)}, \\
z_2(x,\bar{m}) &= \frac{1}{4}\frac{5(\bar{m}+1)^2x + 4\bar{m}^2 + 8\bar{m}}{\bar{m}(3+5x)}, \quad z_3(x,\bar{m}) = \frac{1}{4}\frac{(\bar{m}+1)^2}{\bar{m}}.
\end{aligned}$$

The quantities $z_1(x,\bar{m}) - z_2(x,\bar{m})$ and $z_1(x,\bar{m}) - z_3(x,\bar{m})$ reduce to

$$\frac{1}{4}\frac{(\bar{m}+13)(\bar{m}-1)x^2 + (3\bar{m}^2 + 12\bar{m} - 5)x + 4\bar{m} + 2\bar{m}^2}{\bar{m}(4x+1)(x+1)(3+5x)}, \; \frac{1}{4}\frac{(\bar{m}^2 + 4\bar{m} - 5)x + \bar{m}^2 + 2\bar{m} - 1}{\bar{m}(4x+1)(x+1)}$$

respectively which are both positive for $\bar{m} \geq 1$. Thus z_2 and z_3 are dominated by z_1.

Thus the price of anarchy is bounded by the larger of z_1 and z_0. The critical value z_1 minus value z_0 is

$$\frac{1}{4}\frac{((\bar{m}-3)x+\bar{m})((\bar{m}-1)x+\bar{m})}{(1+x)^2(4x+1)\bar{m}}$$

which implies that z_1 is smaller than z_0 whenever $x > \bar{m}/(3-\bar{m})$ or $\bar{m} \geq 3$.

The derivative with respect to x of z_0 and z_1 respectively are

$$-\frac{1}{2}\frac{(\bar{m}-1)x+\bar{m}}{(1+x)^3\bar{m}}, \quad -\frac{1}{2}\frac{(2\bar{m}^2+8\bar{m}-10)x^2+(4\bar{m}^2+8\bar{m}-4)x+2\bar{m}^2+3\bar{m}}{\bar{m}(4x+1)^2(x+1)^2}$$

which are negative. This shows the bound decreases in x. Similarly, taking the derivative of z_0 and z_1 with respect to $\bar{m}$ is positive. Therefore setting $\bar{m} = m$ upper bounds the PoA. Substituting $x = 0$, and $\bar{m} = m$ into z_1 we get $1 + m/2$, which is the upper bound on the price of anarchy in the theorem statement. To get the bound on the price of anarchy in terms of conductance ratio y, we substitute the relation $y = 1/(x+1)$ into the expressions for z_1 and z_0. This yields

$$z_0(y,m) = \frac{1}{4}\frac{m^2+2m(1+y)+(y-1)^2}{m}, \quad z_1(y,m) = \frac{m^2(2-y)+m(4-y^2-y)+2(y-1)^2}{8m-6my}$$

and $z_1(y,m)$ is the larger of the two when $y > 1 - m/3$. □

3.6.2. General topology

In this section, we study the general topology case.

Theorem 5. *The price of anarchy of game* **G** *with a general topology, is no more than 2 times the upper bounds in expressions (3) and (4) from Theorem 4.*

Proof. As we argued for the simple parallel-serial topology, Corollary 2 implies that there exists a set of $\{d_{\mathbf{i}}\}$ such that the zero-flow zero-price equilibrium flows solve the Nash circuit equations with resistances $\{a_{\mathbf{i}} + d_{\mathbf{i}}\delta_{\mathbf{i}}\}$. Moreover, Lemma 6 shows that we can collapse the resistances of the arbitrary groups into equivalent resistances on each branch i. Consequently, for some d_i's (as prescribed by Lemma 6) the following KVL equations must hold

$$((\bar{M}+I)A + sJ + \bar{M}D\Delta)F = V - b.$$

Recall that D collects the d_i's into a diagonal matrix. At this point, we use the other result of Lemma 6 which gives an expression for the maximum "latency resistance" of an arbitrarily complex group of providers. Since the provider profit of a group is equivalent to the total power dissipated by the resistors of the group minus the power dissipated by the latency resistance, we find that the total provider profit in Nash equilibrium is no less than

$$\frac{1}{2}F^T((\bar{M}+I)A + \bar{M}D\Delta)F \geq \frac{1}{2}F^T(\bar{M}A + \bar{M}D\Delta)F. \tag{7}$$

Note that the left expression is just $\frac{1}{2}$ times expression 5 , which recall was the expression we found for provider profit in the simple parallel-serial case.

By the same argument as we used for the simple parallel-serial case, an upper bound on social optimum welfare for a truncated disutility function is

$$\widetilde{W}_t^* = \frac{1}{2}(V-b)^T(2A+sJ)^{-1}(V-b) - \frac{s}{2}F^TJF.$$

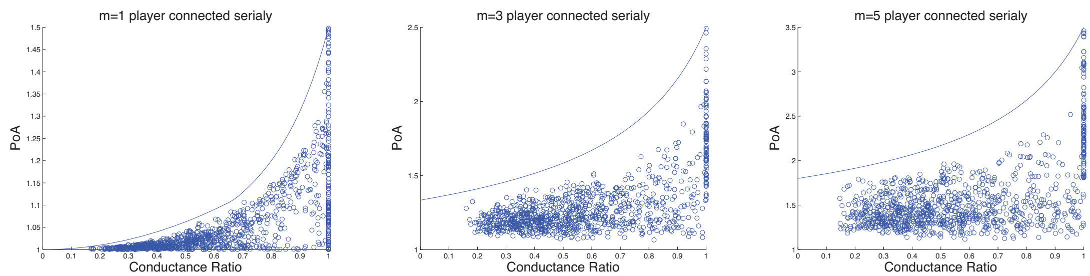

Figure 6: Simple parallel serial PoA bound for the cases where there are either $m = 1$, 3, or 5 providers connected serially. The bounds are plotted as a function of the conductance ratio – the ratio of the conductance of the most conductive branch to the conductance of the whole circuit. The points are the PoA of randomly generated example topologies.

As in the previous proof, we find a bound on the price of anarchy by finding a sufficiently large z to make the below expression non-negative

$$zW_t - \widetilde{W}_t^* = F^T \left(z\frac{1}{2}\bar{M}A + \frac{1}{2}z\bar{M}D\Delta + \frac{1}{2}sJ \right) F - \frac{1}{2}F^T \Big((\bar{M}+I)A + sJ + \bar{M}D\Delta \Big) (2A + sJ)^{-1} \Big((\bar{M}+I)A + sJ + \bar{M}D\Delta \Big) F.$$

This expression is similar to expression (6) studied in Theorem 4. The only difference is that every term that has a factor of z has added factor of $\frac{1}{2}$ in the above expression as compared to (6). Therefore any value of z that made (6) non-negative, twice that value will make the above expression non negative. Therefore two times the upper bounds found in Theorem 4 for the price of anarchy found are upper bounds for the price of anarchy for general parallel serial topologies. □

4. Result characterization and simulated examples

In this section we graph the the PoA bounds found in Theorems 4 and 5. We also compare the bounds to PoA found for randomly generated examples evaluated numerically.

Figure 4 shows the price of anarchy bounds for the simple parallel serial topology as given by Theorem 4. The plots are shown with respect to the ratio of the conductance of the most conductive branch in the circuit to the conductance of the overall circuit. A conductance ratio near 1 means that most of the conductance of the system is concentrated in a single branch. The smaller the conductance ratio, the more that the overall conductance of the system is distributed across multiple branches. In a sense the conductance ratio reflects the market power or concentration of the system.

5. Conclusion

We have seen how a circuit analogy can be used to obtain an upper bound on the price of anarchy for a network routing game with latency, elastic demand, and selfish providers. Furthermore, our bound depends on what may be a useful measure of "market power" in this context: the ratio of the conductance of the most conductive branch of the system to the conductance of the whole system.

The results of this study lead naturally to several lines of future investigation. The conductance ratio metric is an interesting one, and deserves further investigation. The additional factor of 2 between the bounds for the general parallel-serial case and the general parallel-serial case was needed for our proof, but from trying examples it might be that the bound without this factor of 2 still holds. (We found no example of a general parallel-serial topology with PoA outside our bound for the simple parallel-serial topology.) We hope in future work to either prove or disprove this conjecture. Finally, all of our analysis holds only for a single source-destination pair. It might be possible to extend the analysis to multiple sources and destinations. However, electricity flow is inherently a single commodity, so it would take far more than a trivial extension to extend the circuit analogy approach to a multiple source and destination scenario.

References

Acemoglu, D., Ozdaglar, A., February 2007a. Competition and efficiency in congested markets. Mathematics of Operations Research 32 (1), 1 – 31.

Acemoglu, D., Ozdaglar, A., 2007b. Competition in parlallel-serial networks. IEEE Journal on Selected Areas in Communications 25 (6), 1180–1192.

Babaioff, M., Kleinberg, R., Papadimitriou, C., 2009. Congestion games with malicious players. Games and Economic Behavior 67, 22–35.

Boyd, S., Vandenberghe, L., 2004. Convex Optimization. Cambridge University Press, Cambridge, UK.

Chawla, S., Roughgarden, T., April 2008. Bertrand competition in networks. In: International Symposium on Algorithmic Game Theory. Paderborn, Germany.

Chua, L., Desoer, C., Kuh, E., 1987. Linear and Nonlinear Circuits. McGraw-Hill, New York.

Cole, R., Dodis, Y., Roughgarden, T., May 2006. How much can taxes help selfish routing? Journal of Computer and System Sciences 72 (3), 444–467.

Hayrapetyan, A., Tardos, E., Wexler, T., March 2007. A network pricing game for selfish traffic. Distributed Computing 19 (4), 255–266.

Johari, R., Tsitsiklis, J., May 2006. A scalable network resource allocation mechanism with bounded efficiency loss. IEEE Journal on Selected Areas in Communications 24 (5), 992 – 99.

Musacchio, J., Wu, S., October 2007. The price of anarchy in a network pricing game. In: Proceedings of the 45th Annual Allerton Conference on Communication, Control, and Computing. Monticello, IL, pp. 882–891.

Ozdaglar, A., October 2008. Price competition with elastic traffic. Networks 52 (3), 141 – 155.

Pigou, A. C., 1920. The Economics of Welfare. Macmillan.

Roughgarden, T., 2001. Stackelberg scheduling strategies. In: Proceedings of the 33rd Annual Symposium ACM Symposium on Theory of Computing. pp. 104 – 113.

Roughgarden, T., March 2002. How bad is selfish routing? Journal of ACM 49 (2), 236–259.

Roughgarden, T., September 2003. The price of anarchy is independent of the network topology. Journal of Computer and System Sciences 67 (2), 341 – 364.

Wardrop, J., 1952. Some theoretical aspects of road traffic research. Proceedings of the Institute of Civil Engineers, Part II 1 (36), 352 – 362.

Xi, Y., Yeh, E. M., 2008. Pricing, competition, and routing for selfish and strategic nodes in multi-hop networks. In: Proceedings of IEEE INFOCOM. Phoenix, AZ.

6. Appendix

Proof of Lemma 1. $\Leftarrow$ Suppose that $\{f_{\mathbf{i}}\}$, $\{p_{\mathbf{i}}\}$ satisfies the KCL and KVL conditions for the circuit analogy. This requires that across each group $\mathcal{G}_{i...j}$ there is a well defined voltage drop $V_{i...j}$. Moreover If $\mathcal{G}_{i...j}$ contains subgroups in parallel $V_{i...jk} = V_{i...j}$ for all $k \in 1...g_{i..j}$ (KVL) and $f_{i...j} = \sum_{k\in 1...g_{i...j}} f_{i...jk}$ (KCL). If $\mathcal{G}_{i...j}$ contains elements in subgroups in series $V_{i...j} = \sum_{k\in 1...g_{i..j}} V_{i...jk}$ (KVL) and $f_{i...j} = f_{i...jk}$ for all $k \in 1...g_{i...j}$ (KCL).

For each single provider $\mathbf{i}$, let

$$D_{\mathbf{i}} = \begin{cases} V_{\mathbf{i}} & \text{if } f_{\mathbf{i}} > 0, \\ p_{\mathbf{i}} & \text{if } f_{\mathbf{i}} = 0. \end{cases}$$

Note that if $f_{\mathbf{i}} = 0$, KVL requires that $p_{\mathbf{i}} \geq V_{\mathbf{i}}$ and the voltage drop across provider $\mathbf{i}$'s diode in the circuit analogy be nonpositive. (In other words, the diode is preventing the flow from becoming negative.) For each group of providers $\mathcal{G}_{i...j}$ containing parallel subgroups, let

$$D_{i...j} = \begin{cases} V_{i...j} & \text{if } f_{i...j} > 0, \\ \min_{k\in 1...g_{i..j}} D_{i...jk} & \text{if } f_{i...j} = 0. \end{cases}$$

Similarly, for each group $\mathcal{G}_{i...j}$ of providers containing serial subgroups, let

$$D_{i...j} = \sum_{k\in 1...g_{i..j}} D_{i...jk}.$$

With these assignments, it is easy to verify that the Wardrop conditions in Definition 2 are satisfied.

$\Rightarrow$ Now suppose that $\{f_{\mathbf{i}}\}$, $\{p_{\mathbf{i}}\}$ satisfies the Wardrop equilibrium conditions of Definition 2. At the top level of the hierarchy, (branches), let:

$$V_i = U(f) - b_i, \text{ where recall } f = \textstyle\sum_{i=1}^n f_i.$$

To choose the rest of the voltages, we proceed down the hierarchy. Suppose we have defined $V_{i...j}$. Consider 2 cases:

Case 1: Suppose $\mathcal{G}_{i...j}$ contains subgroups in series. Let for each $k \in 1...g_{i...jk}$, $V_{i...jk} = D_{i...jk}$ if $f_{i...jk} > 0$. Otherwise if $f_{i...k} = f_{i...jk} = 0$, then it must be that $D_{i...j} \geq V_{i...j}$ or else the Wardrop conditions would not be satisfied. By the construction of $D_{i...j}$ in Definition 2, it must be that $D_{i...j}$ is equal to

$$\min_{P\in\mathcal{P}_{i...j}} \sum_{\mathbf{i}\in P} p_{\mathbf{i}}$$

where $\mathcal{P}_{i...j}$ is the set of loop free paths (a path P is a sequence of provider indices) that traverses group $\mathcal{G}_{i...j}$. Likewise, for each k, $D_{i...jk}$ is equal to

$$\min_{P\in\mathcal{P}_{i...jk}} \sum_{\mathbf{i}\in P} p_{\mathbf{i}}.$$

Consequently, there exists a set of voltage drops $\{V_{i...jk}\}$ such that

$$\sum_{k\in 1...g_{i...j}} V_{i...jk} = V_{i...j} \text{ and } V_{i...jk} \geq \min_{P\in\mathcal{P}_{i...jk}} \sum_{\mathbf{i}\in P} p_{\mathbf{i}}.$$

With this assignment of voltages, the flows trough groups $\mathcal{G}_{i...jk}$ would be 0 in their circuit analogy because the diodes would be blocking.

Case 2: Suppose $\mathcal{G}_{i...j}$ contains subgroups in parallel. Let for each $k \in 1...g_{i...jk}$, $V_{i...jk} = V_{i...j}$.

By induction, the above procedure defines voltages throughout the circuit analogy. It is then straightforward to verify that these voltages, and flows from the Wardrop equilibrium profile satisfy the KCL and KVL conditions of the circuit analogy. $\square$

Proof of Theorem 1. We construct a unique flow profile that solves the circuit analogy, which by Lemma 1 is equivalent to constructing a Wardrop equilibrium. In this proof we will consider the relationship between flow and voltage drop across groups at each level of the hierarchy. The voltage drop across a group can be thought of as being analogous to the best disutility of competing paths if that group is part of a parallel interconnection. Therefore the voltage drop is the same as "the D" for a group in definition 2 for Wardrop equilibrium whenever that group carries positive flow.

Consider the circuit analogy of a group $\mathcal{G}$ consisting of single serially connected providers. Clearly the flow through these providers is a monotone nondecreasing function of the voltage drop across the group, and this function maps $\mathbb{R}^+$ onto $\mathbb{R}^+$ since the resistors have positive resistance. Therefore the voltage across the group uniquely determines the flow through all providers of the group. Moreover the functional dependence maps $\mathbb{R}^+$ onto $\mathbb{R}^+$ – a property we will refer to simply as "onto" in the rest of the proof. Also note that the voltage across the group is the same as the D for the group in definition 2.

Now consider the circuit analogy of a group $\mathcal{G}$ consisting of single providers connected in parallel, and consider the flow through the group as a function of voltage. When the voltage V is less than $\min_{\mathbf{i}\in\mathcal{G}} p_{\mathbf{i}}$, the flow across all the parallel branches will be zero. As V increases beyond the $\min_{\mathbf{i}\in\mathcal{G}} p_{\mathbf{i}}$ the flow through any providers in the group with $p_{\mathbf{i}} < V$ will carry flow $V/a_{\mathbf{i}}$ and providers with $p_{\mathbf{i}} \geq V$ will carry 0 flow. Consequently the voltage across the group uniquely determines the flow through all providers of the group, and the relation is onto.

We proceed at this point by induction. Suppose that we have shown that: i) a voltage uniquely determines a flow profile of all providers in groups $\mathcal{G}_{i....jk}$, $k = 1...g_{i...j}$, and ii) the flow across each group is a monotone nondecreasing onto function of voltage.

Suppose that $\mathcal{G}_{i....j}$ contains the groups $\mathcal{G}_{i....jk}$, $k = 1...g_{i...j}$ connected in series. For each k, Let $F_{i...jk}(\cdot)$ be the monotone nondecreasing function describing the dependence of the flow through group $\mathcal{G}_{i....jk}$ on voltage. By monotonicity the inverse function $V_{i...jk}(\cdot)$ describing the dependence of voltage on flow exists, and is onto. This function is monotone increasing and may have jump discontinuities. Now consider the function $\sum_k V_{i...jk}(\cdot)$. This function is also monotone increasing with possible jump discontinuities, so its inverse $F_{i...j}(\cdot)$ exists, and moreover the inverse is monotone nondecreasing continuous, and onto. Consequently, given a voltage V a value $f = F_{i...j}(V)$ is uniquely determined. Moreover the voltage of each subgroup $\mathcal{G}_{i....jk}$ by computing $V_{i...jk}(f)$.

Suppose that $\mathcal{G}_{i....j}$ contains the groups $\mathcal{G}_{i....jk}$, $k = 1...g_{i...j}$ connected in parallel. By the inductive hypothesis, we already have monotone, non-decreasing, onto functions $F_{i...jk}(\cdot)$ for each k. Simply take $F_{i...j}(\cdot)$ to be the sum of these functions.

Thus by induction, a voltage across any group uniquely determines the flow through all providers contained in that group. This extends all the way to the top of the hierarchy, so that the voltage at the top of the hierarchy uniquely determines a total flow, and that relation is monotone non-decreasing and onto. Consequently the inverse relation between flow and voltage is monotone increasing with possible jump discontinuities, starts at (0,0) and is onto the positive real numbers. The voltage as a function of flow of the nonlinear power source (representative of the disutility function) is monotone nonincreasing and starts at a positive voltage when flow is 0. These two functions must intersect, giving a unique flow and voltage at the intersection. □

Proof of Lemma 4. Each of the highest level groups $\mathcal{G}_i$ consist of strictly resistive circuit elements and thus have a unique solution for each voltage across each group's connection point, and furthermore each group can be by replaced by a Thévenin equivalent resistance $\tilde{a}_i$ without affecting the solution to the rest of the circuit (Chua et al., 1987).

It remains to show that there is a unique solution of currents traversing the highest level groups. A solution can be constructed by the following procedure. If $\min_i\{b_i\} > U(0)$ then having zero flow on all branches, and all diodes blocking is a valid solution. Otherwise perform the following procedure. i) Let $OFF = \{1, 2, ..., n\}$, $ON = \varnothing$ and $f = 0$. ii) Pick an i^* from $\text{argmin}_{i \in OFF} b_i$. If $U(f) > b_{i*}$ let $OFF \leftarrow OFF - i^*$ and $ON \leftarrow \text{ON} + i^*$. iii) Find a flow assignment assuming that the branches in ON have on diodes and the branches in OFF are replaced with open circuits. Such a solution should be unique because the circuit elements are strictly passive and monotone (Chua et al., 1987)[4]. iv) Let f be the sum of flows across all on branches. Repeat from step (ii) until OFF is exhausted or until $U(f) > b_{i*}$. Flow f is monotone increasing at each repetition because we are always adding branches for more current to flow, and each added branch has a b_i less than the voltage of the terminal the branch is connected to. Because the b_i's of the already on branches are lower, the flow through them never becomes negative.

Uniqueness can be shown as follows. Suppose that there are two solutions to the circuit. Let $\mathbf{f}$ and $\mathbf{f}'$ represent the flow vectors in both solutions. If the same diodes are on and off, the two solutions would represent non unique solutions to the same passive circuit (one where on diodes are replaced with closed circuits and the off ones open.) which is not possible. Therefore the two solutions must have different sets of on and off diodes.

Consider two cases: i) Suppose that the set of ON diodes in each case are not subsets of the other $ON \nsubseteq ON'$ and $ON' \nsubseteq ON$. Then there exists $b_i < d' < b_j$ and $b_j < d < b_i$ which is impossible. ii) Suppose $ON' \subset ON$, then $f > f'$ because more branches are open, but then $U(f) < U(f')$ which implies that any diode that is on in solution $\mathbf{f}'$ ought to be on in solution $\mathbf{f}$. □

Proof of Lemma 6. We build a general argument by first considering interconnections of single providers, and then later interconnections of groups of providers.

CASE 1: Group $\mathcal{G}_{i...j}$ consists of single providers connected serially

The resistance of the group $\mathcal{G}_{i...j}$ is just the sum across the resistances of each provider $(i...jk)$, which includes both the resistance $a_{i...jk} + d_{i...jk}\delta_{i...jk}$ representing price and the resistance $a_{i...jk}$ representing latency. Thus the total resistance is given by

$$r_{i...j} = \sum_k 2a_{i...jk} + d_{i...jk}\delta_{i...jk}$$

Substituting the formula for $\delta_{i...jk}$ when $(i...jk)$ is part of a serial group, we have

$$r_{i...j} = \sum_k 2a_{i...jk} + \sum_{k' \neq k} d_{i...jk}\delta_{i...j} + d_{i...jk}a_{i...jk'}.$$

Since $d_{i...jk} \leq 1$ and the sumation has $m_{i...j} - 1$ terms, there exists an $\bar{m}_{i...j} \in [1, m_{i...j}]$ and $d_{i...j} \in [0, 1]$ satisfying

$$r_{i...j} = (1 + \bar{m}_{i...j})a_{i...j} + \bar{m}_{i...j}d_{i...j}\delta_{i...j}.$$

Since the group is connected serially, all the flow passes through each resistor. Therefore the equivalent resistance is the sum of the latency resistors $\sum a_{i...jk}$ which by definition is $a_{i...j}$.

CASE 2: Group $\mathcal{G}_{i...j}$ consists of single providers connected in parallel

[4] A strictly passive, monotone circuit element is such that for any different voltages v_1, v_2, the corresponding flows f_1 and f_2 satisfy the relation $(v_1 - v_2)(f_1 - f_2) > 0$ (Chua et al., 1987).

The equivalent resistance of the group $\mathcal{G}_{i...j}$ is

$$\bigoplus_k (2a_{i...jk} + d_{i...jk}\delta_{i...jk}) = \bigoplus_k (2a_{i...jk} + \delta_{i...j} \bigoplus_{k' \neq k} d_{i...jk} a_{i..jk'})$$
$$= 2a_{i...j} + d_{i...j}\delta_{i...j}$$

for some $d_{i...j} \in [0, 1]$. A portion of the power dissipated in the group corresponds to provider profits, with the remainder corresponding to latency loss. Since more than half the resistance of each branch corresponds to the provider's price, the latency loss must be less than half the power dissipated. Therefore the effective latency resistance is $c_{i...j}2a_{i...j} + c_{i...j}d_{i...j}\delta_{i...j}$ for some $c_{i...j} \in [0, .5]$.

CASE 3: Group $\mathcal{G}_{i...j}$ consists of groups of providers

The proof proceeds by induction. Suppose that it has been shown that for each subgroup $\mathcal{G}_{i...jk} \subset \mathcal{G}_{i...j}$, the equivalent resistance is

$$(1 + \bar{m}_{i...jk})a_{i...jk} + d_{i...jk}\bar{m}_{i...jk}\delta_{i...jk}$$

for some $\bar{m}_{i...jk} \in [0, m_{i...jk}]$ and $d_{i...jk} \in [0, 1]$. Also suppose that it has been shown that the equivalent latency resistance is

$$c_{i...jk}(1 + \bar{m}_{i...jk})a_{i...jk} + c_{i...jk}d_{i...jk}\bar{m}_{i...jk}\delta_{i...jk}$$

for some $c_{i...jk} \in [0, \frac{1}{2}]$. Note that both of these facts are true if group $\mathcal{G}_{i...jk}$ consists of single providers by either the analysis for Case 1 or Case 2.

Case 3a: Suppose group $\mathcal{G}_{i...j}$ contains subgroups connected in parallel:

We claim that the equivalent resistance is

$$(1 + \bar{m}_{i...j})a_{i...j} + d_{i...j}\bar{m}_{i...j}\delta_{i...j}$$

for some $\bar{m}_{i..j} \in [0, m_{i...j}]$ and $d_{i...j} \in [0, 1]$ by the following reasoning. Consider a hypothetical collection of resistors of size $\max_{k'}(1 + \bar{m}_{i...jk'})a_{i...jk}$ $k = 1, 2, ...$ connected in parallel. The equivalent resistance of this interconnection is $\max_{k'}(1 + \bar{m}_{i...jk'})a_{i...j}$. Now add a resistor of size $m_{i...j}\delta_{i...j}$ placed in series on each of the branches. This increases the resistance by an amount not more than $m_{i...j}\delta_{i...j}$. Consequently the resistance of this hypothetical interconnection is now no more than $(1 + m_{i...j})a_{i...j} + m_{i...j}\delta_{i...j}$.

Now decrease the resistance on each branch k to be $(1 + \bar{m}_{i...jk})a_{i....jk}$. This will decrease the overall resistance of the interconnection, and also make the hypothetical interconnection's resistance match that of $\mathcal{G}_{i...j}$. It should now be possible to find suitable constants $\bar{m}_{i...j}$ and $d_{i...j}$.

The equivalent latency resistance is

$$(1 + \bar{m}_{i...j})c_{i...j}a_{i...j} + c_{i...j}d_{i...j}\bar{m}_{i...j}\delta_{i...j}$$

for some $c_{i...j} \in [0, \frac{1}{2}]$, which can be seen by the following reasoning. The latency resistance of each subgroup is no more than half the total resistance. Therefore no more than half the total power dissipated group $\mathcal{G}_{i...j}$ can be due to latency. Thus the equivalent latency resistance is no more than half the overall equivalent resistance.

Case 3b: Suppose group $\mathcal{G}_{i...j}$ contains subgroups connected in series:

Then the equivalent resistance is

$$r_{i...j} = \sum_k (1 + \bar{m}_{i...jk})a_{i...jk} + d_{i...jk}\bar{m}_{i...jk}\delta_{i...jk}$$

Substuting the definition of $\delta_{i...jk}$ we have the following relatons

$$\begin{aligned}
r_{i...j} &\leq \sum_k (1+m_{i...jk})a_{i...jk} + m_{i...jk}\delta_{i...jk} \\
&\leq \sum_k \left[(1+m_{i...jk})a_{i...jk} + m_{i...jk}\left(\delta_{i...j} + \sum_{k'\neq k} a_{i...jk'}\right)\right] \\
&\leq \sum_k \left[a_{i...jk} + m_{i...jk}\sum_{k'} a_{i...jk'} + m_{i...jk}\delta_{i...j}\right] \\
&\leq (1+m_{i...j})a_{i...j} + m_{i...j}\delta_{i...j}.
\end{aligned}$$

Also it is clear that $r_{i...j} \geq a_{i...j}$. Thus there exists an $\bar{m}_{i...j} \in [1, m_{i...j}]$ and $d_{i...j} \in [0,1]$ such that

$$r_{i..j} = (1+\bar{m}_{i...j})a_{i...j} + d_{i...j}\bar{m}_{i...j}\delta_{i...j}.$$

The equivalent latency resistance is no more than half of the above quantity, because the equivalent latency resistance of each subgroup is no more than half the equivalent resistance of each subgroup. $\square$