The price of imperfect competition for a spanning network

Hervé Moulin[a] and Rodrigo A. Velez[b]

[a]Department of Economics, Rice University, Houston, TX 77251

[b]Department of Economics, Texas A&M University, College Station, TX 77843

January 18, 2010

Abstract

A buyer procures a network to span a given set of nodes; each seller bids to supply certain edges, then the buyer purchases a minimal cost spanning tree.

Irrespective of the pattern of bidding licenses and costs, an efficient tree is constructed in any equilibrium of the Bertrand game.

If each seller can only bid for a single edge, or for a set of mutually disconnected edges, we evaluate the *price of imperfect competition* (PIC), namely the ratio of the total cost that could be charged to the buyer in some equilibrium, to the true minimal cost. If costs satisfy the triangle inequality we show that the PIC is at most 2 for an odd number of nodes, and at most $2\frac{n-1}{n-2}$ for an even number n of nodes. Surprisingly, this worst case ratio does not improve when the cost pattern is ultrametric (a much more demanding substitutability requirement), although the overhead is much lower on average under typical probabilistc assumptions.

Acknowledgments: Several comments by Tim Roughgarden, Jay Sethuraman and Vijay Vazirani have been very helpful.

1 Introduction

Bertrand competition for differentiated commodities typically implies some welfare losses, as in the Hotelling model. We consider a special case where it does not, and where the surplus that the competing firms are able to extract admits a simple upper bound.

A buyer procures a network spanning a given set of nodes, while the sellers bid for different edges of the network. Efficiency requires to build a minimal cost spanning tree. This well known optimization problem has a variety of applications (including rail infrastructure, the internet's backbone, water distribution, etc..; see [15] for a survey).

If one seller is the sole bidder for a certain edge, or a group of edges, competition is imperfect and sellers can typically bid higher than their true cost.

But this effect is mitigated by the fact that the edges are partial substitutes: for any edge e, there are several alternative paths ensuring the connection of e's two end-nodes.

We observe first that irrespective of the ownership structure and of the cost pattern, in equilibrium an efficient (minimal cost) spanning tree is constructed (Proposition 1). Our main result (Theorem 1) is that if each seller bids for a single edge (sections 4 and 5), or for several mutually disconnected edges (section 6), and true costs satisfy the triangular inequality, the buyer's surcharge essentially cannot exceed 100% of the true minimal cost.

We define the *price of imperfect competition* as the ratio of the total cost that *could* be charged to the buyer in *some* equilibrium, to the minimal cost. Inspired by the recent literature on the *price of anarchy* ([8], [13], [11], [10]), we compute first the worst case (largest) charge in equilibrium for a given set of costs (Proposition 2), next the worst possible ratio under two familiar assumptions on the cost pattern.

Write the cost of connecting two nodes a, b as c_{ab}. The triangular inequality ($c_{ab} \leq c_{ad} + c_{db}$ for all a, b, d) is realistic in most network connection models. We show (Theorem 1) that under this assumption the price of imperfect competition is at most 2 for an odd number of nodes, and at most $2\frac{n-1}{n-2}$ for an even number n of nodes.

The cost pattern is ultrametric if $c_{ab} \leq \max\{c_{ad}, c_{db}\}$ for all a, b, d. This is stronger, and implies much closer substitutability, than the triangular inequality. A typical ultrametric distance[1] comes from a genealogical tree of which the leafs are the objects to be compared (in our case, the nodes $a, b, \cdots$), so the mapping of leafs to objects is one-to-one. Assume also that all paths from the root to a leaf are of equal length. Then the distance between two objects is the number of steps they must go up in the tree to reach a common ancestor (their degree of kinship). Ultrametric costs play an important role in the design of rules sharing the cost of an optimal spanning tree between the different nodes (see [12], [5], [2], [3]).

Surprisingly, the price of imperfect competition does not improve when the cost pattern is ultrametric; however, the tree structures for which the worst ratio is achieved are much fewer (Theorem 2); also the overhead is much lower on average under typical probabilistic assumptions (subsection 5.4).

2 Basic definitions and notation

Let V be the set of nodes with cardinality n and generic elements $a, b, \cdots$. Let the set of edges, i.e., non oriented pairs in V, be E with generic elements $e, f, \cdots$, and cardinality $\frac{n(n-1)}{2}$. We let $\Gamma \subseteq 2^E$, with generic element γ, be the set of trees spanning V. The path between two nodes a, b on the spanning tree γ is denoted $[a, b]_\gamma$. We write $\Delta(e, \gamma)$ for the set of edges f across e on γ, i.e., such that $f \neq e$ and $\gamma + f - e$ is still a spanning tree. Finally, $Ad(e, \gamma)$ is the set of edges adjacent to e in γ (i.e., they are not e but share a node with e).

[1] Actually it captures the general structure of ultrametric distances on finite sets. See [1].

The cost "matrix" is $c \in \mathbb{R}_+^E$. For each $\gamma \in \Gamma$ let $c(\gamma) \equiv \sum_{e\in\gamma} c_e$ be the cost of γ; the minimal cost of a tree spanning V is $\lambda(c) \equiv \min_{\gamma\in\Gamma} c(\gamma)$ and $\Gamma(c)$ is the set of minimal cost spanning trees for c. The set $\Gamma(c)$ obtains by the well known Kruskal Algorithm (KA): (i) starting from the empty graph, add one edge that has minimum cost among the edges not yet chosen and whose addition does not induce a cycle; (ii) repeat as many times as possible ([9]).

3 The Bertrand game for general costs and bidding rights

A single buyer requests bids for each edge in E from a set of sellers S (with generic elements $i, j, \cdots$) in order to construct a spanning tree. Each seller i is allowed (licenced) to bid for a subset of edges denoted $l(i)$. The sets $l(i)$ may overlap, and we assume $\cup_{i\in N} l(i) = E$. If only one seller is requested to bid for a given edge, we can think of him as the owner of the edge.

For each $e \in l(i)$, seller i can build e at a cost c_e^i. The efficient cost of edge e is $c_e^* = \min_{\{i:e\in l(i)\}} c_e^i$.

A strategy of seller i is $p^i \equiv (p_e^i)_{e\in l(i)} \in \mathbb{R}_+^{l(i)}$, specifying his bid for each of "his" edges.

Given a strategy profile $p \equiv (p^i)_{i\in S}$, the buyer purchases a cheapest spanning tree, namely $\gamma(p) \in \Gamma(\pi)$, where $\pi_e \equiv \min_{\{i:e\in l(i)\}} p_e^i$ is the lowest bid for edge e. The Bertrand game is well defined once we choose a tie-breaking rule for the case where $\Gamma(\pi)$ contains several cheapest trees, and for selecting a seller if there are several optimal bids for an edge $e \in \gamma(p)$. The net profit of seller i is then $\sum(p_e^i - c_e^i)$, where the sum bears over all edges e in $\gamma(p)$ for which i's bid wins.

We keep the tie-breaking rule entirely general, and omit it for simplicity in the notation. It will be clear that our results are independent of this rule (that could even be probabilistic).

In the Bertrand game with pure strategies just described, it is well known that the concept of Nash equilibria is not the right one: because the payoff functions are discontinuous around a tie, there may be no Nash equilibrium at all. The simplest example has two nodes and one edge, for which two sellers with different costs $c^1 < c^2$ compete (and ties broken in favor of seller 2, or randomly); see e.g. [16], Chapter 5.

This difficulty can be resolved in several convincing, equivalent ways. From the handful of possible approaches[2], the most convenient for our purpose turns out to be the notion of ε-equilibrium (see [6]), namely strategy profiles ${}^{\varepsilon}p$ from which no seller can change his bid and increase his profit by more than ε. We call p a *limit equilibrium* if it is the limit of a sequence of ε-equilibria where ε goes

[2] Other solutions include: to extend the definition of equilibria and make the tie breaker an outcome of agents' strategic interaction ([14]); to eliminate weakly dominated strategies ([4]); to assume integer-valued bids; to use mixed strategies ([7], Chapter 10).

to zero. In the single edge example, this implies identical bids $p^1 = p^2 \in [c^1, c^2]$ and that seller 1 wins the contract.

Note that in an ε-equilibrium ${}^{\varepsilon}p$, the buyer selects a minimal cost tree $\gamma({}^{\varepsilon}p) \in \Gamma({}^{\varepsilon}\pi)$, and because of ties the sequence $\gamma({}^{\varepsilon}p)$ may have several limit points; thus in a given limit equilibrium p several trees can be built, but of course they all cost $\lambda(\pi)$; similarly we could have more than one possible winner for each edge. These complications are totally inconsequential in the results below.

Our first result states that in our model competition in prices ensures that an efficient network is built (all proofs omitted in the body of the paper are gathered in section 7).

Proposition 1 *In each limit equilibrium of the Bertrand game, the buyer purchases a minimal cost spanning tree $\gamma \in \Gamma(c^*)$ (though he typically pays more than $\lambda(c^*)$).*

We stress that Proposition 1 requires no assumption about the number of edges for which a seller can bid. This means that even if a seller is the unique provider of a subset of edges, he will not manipulate the market to construct an inefficient tree.

4 The price of imperfect competition

We now investigate what is the price that the buyer pays for these efficient equilibrium trees under the assumption that each seller bids for exactly one edge.

For $c \in \mathbb{R}_+^E$, $b \in \mathbb{R}_+$, and $e \in E$, replacing the e-coordinate of c by b yields the cost matrix (c_{-e}, b). Let $E(c)$ be the set of edges that belong to some minimal cost spanning tree at c.

Consider an edge $e \in E(c)$. For our purpose a key quantity is how high the cost of e can be, ceteris paribus, while e remains in at least one minimal cost spanning tree:

$$\mu_e(c) \equiv \max\left\{b \in \mathbb{R}_+ : e \in E(c_{-e}, b)\right\}. \tag{1}$$

We show now that $\mu_e(c)$ is the minimum cost among the substitutes of e.

Lemma 1 *Let $e \in E(c)$. Then for each $\gamma \in \Gamma(c)$ such that $e \in \gamma$, $\mu_e(c) = \min\{c_f : f \in \Delta(e, \gamma)\}$.*

Proof We choose $c, e \in E(c)$, and $\gamma \in \Gamma(c)$ such that $e \in \gamma$.

Step 1. $\mu_e(c) \leq \min\{c_f : f \in \Delta(e, \gamma)\}$. Pick $b \in \mathbb{R}_+$ such that $b > \min\{c_f : f \in \Delta(e, \gamma)\}$ and $f \in \Delta(e, \gamma)$ such that $c_f < b$. Thus, $\lambda(c_{-e}, b) \leq \lambda(c) - c_e + c_f$. We claim that $e \notin E(c_{-e}, b)$. Suppose on the contrary that there is $\gamma' \in \Gamma(c_{-e}, b)$ such that $e \in \gamma'$. Clearly $\gamma' \in \Gamma(c)$ because from (c_{-e}, b) to c we reduce the cost of a single edge of γ' (so $c(\gamma'') < c(\gamma')$ would imply $(c_{-e}, b)(\gamma'') < (c_{-e}, b)(\gamma')$). Therefore, $\lambda(c_{-e}, b) = \lambda(c) - c_e + b$, implying $b \leq c_f$, a contradiction.

Step 2. Set now $b = \min\{c_f : f \in \Delta(e, \gamma)\}$ and check that $\gamma \in \Gamma(c_{-e}, b)$. Recall a well known fact about minimal cost spanning trees, equivalent to property (10) in the proof of Proposition 1: for any cost c' and tree $\gamma' \in \Gamma$

$$\gamma' \in \Gamma(c') \Leftrightarrow \{\text{for all } f = ab \notin \gamma',\ c_f \geq \max\{c_e | e \in [a, b]_{\gamma'}\}\}. \tag{2}$$

This implies at once $\gamma \in \Gamma(c_{-e}, b)$.■

We say that an edge $e \in E(c)$ is *essential* when it belongs to each minimal cost spanning tree at c. This property informs the comparison between c_e and $\mu_e(c)$.

Lemma 2 *Let $e \in E(c)$. If e is essential, then $\mu_e(c) > c_e$; if e is not essential, then $\mu_e(c) = c_e$.*

Proof The first statement is an easy consequence of Lemma 1. Now, let $e \in E(c)$ be non essential, i.e., there is a spanning tree $\gamma \in \Gamma(c)$ such that $e \notin \gamma$. Let $b > c_e$. We check that $e \notin E(c_{-e}, b)$. Suppose on the contrary that there is $\gamma' \in \Gamma(c_{-e}, b)$ such that $e \in \gamma'$. Thus, $b + \sum_{f \in \gamma' \setminus \{e\}} c_f \leq \sum_{f \in \gamma} c_f$. As in step 1 of the previous proof, $\gamma' \in \Gamma(c)$ as well. Therefore $\lambda(c) = c_e + \sum_{f \in \gamma' \setminus \{e\}} c_f < b + \sum_{f \in \gamma' \setminus \{e\}} c_f \leq \sum_{f \in \gamma} c_f$, in contradiction of $\gamma \in \Gamma(c)$.
■

The following lemma states that the sum of the costs $\mu_e(c)$ along a minimal cost spanning tree is an invariant of the cost matrix c.

Lemma 3 *For any $\gamma, \gamma' \in \Gamma(c)$ we have*

$$\sum_{e \in \gamma} \mu_e(c) = \sum_{e' \in \gamma'} \mu_{e'}(c),$$

so this sum can be written as $\mu(c)$. Moreover, $\mu(c)$ is weakly increasing in c

Proof From $\gamma, \gamma' \in \Gamma(c)$ we get $\sum_{e \in \gamma \setminus \gamma'} c_e = \sum_{e' \in \gamma' \setminus \gamma} c_{e'}$. By Lemma 2 we have $c_e = \mu_e(c)$ for all $e \in \gamma \setminus \gamma'$, and $c_{e'} = \mu_{e'}(c)$ for all $e' \in \gamma' \setminus \gamma$. Combining these equalities

$$\sum_{e \in \gamma} \mu_e(c) - \sum_{e \in \gamma'} \mu_e(c) = \sum_{e \in \gamma \setminus \gamma'} \mu_e(c) - \sum_{e' \in \gamma' \setminus \gamma} \mu_{e'}(c) = \sum_{e \in \gamma \setminus \gamma'} c_e - \sum_{e' \in \gamma' \setminus \gamma} c_{e'} = 0,$$

as stated.

For the second statement, we fix an edge $e \in E$, $b \in \mathbb{R}_+$, and two costs c, c' such that $c \leq c'$. An easy consequence of the KA is $e \in E(c_{-e}, b) \Rightarrow e \in E(c'_{-e}, b)$. By (1) this implies that $c \to \mu_e(c)$ is weakly increasing for e fixed. Thus $b \to \mu(c_{-e}, b)$ is weakly increasing on any interval where a given tree remains optimal. At the boundary between two such intervals, we just proved that $\mu(c_{-e}, b)$ is the same when computed for either tree. This implies the desired monotonicity.■

We can now describe the range of the aggregate prices that the buyer may be charged in the limit equilibria of the Bertrand game.

Proposition 2
i) Suppose that each seller bids for exactly one edge. Then in any limit equilibrium of the Bertrand game the buyer pays at least $\lambda(c^*)$ and at most $\mu(c^*)$ (recall $c^*_e = \min_{\{i: e = l(i)\}} c^i_e$).
ii) Suppose that each seller is the exclusive bidder for exactly one edge, so we identify a player with his edge e and c_e is this bidder's cost. For any $c \in \mathbb{R}^E_+$

and $\gamma \in \Gamma(c)$, the profile

$$p^e = \mu_e(c) \text{ if } e \in \gamma; \, p^e = c_e \text{ if } e \notin \gamma, \tag{3}$$

is a limit equilibrium of the Bertrand game where the buyer pays $\mu(c)$.

Note that if we refine the Nash equilibrium concept by one round of elimination of weakly dominated strategies, we find that in all remaining limit equilibria, the buyer pays exactly $\mu(c)$. This is because bidding below one's cost c_e^i is a weakly dominated strategy for bidder i, so that all equilibria are described by (3), where γ varies in $\Gamma(c)$.

We define the *price of imperfect competition (PIC)* for a cost matrix c as the ratio of the maximal aggregate price payed by the buyer in an equilibrium of the Bertrand game, to the true minimal cost $\lambda(c)$.[3] By Proposition 2, whenever each seller is the exclusive bidder for exactly one edge,

$$PIC(c) = \frac{\mu(c)}{\lambda(c)}.$$

In contrast to the Price of Anarchy (see Introduction), the PIC measures the effect to the buyer's welfare of decentralizing the procurement of a spanning network, as opposed to measuring it for the whole economy. In our model imperfect competition induces no aggregate welfare loss.

5 Bounding the price of imperfect competition

In this section we study the price of imperfect competition whenever each seller is the exclusive bidder for exactly one edge. First, we characterize the conditions under which the buyer will pay no overhead above the true minimal cost. Then we investigate the PIC under two familiar assumptions on the cost pattern.

The overhead that the buyer can expect to pay above the true minimal cost depends on the degree of substitutability of the edges. For instance, suppose that for some cost matrix c there is a unique minimal cost spanning tree γ. If costs outside γ are unboundedly large, then there is no substitute for any edge in γ. Hence the overhead payed by the buyer is infinite. At the other extreme, if *all* edges in E have the same cost, Lemma 1 gives $\mu(c) = \lambda(c)$ so the overhead is zero (despite the fact that edges are *not* perfect substitutes).

5.1 Zero overhead

Proposition 2 implies that the buyer pays no overhead at all if $\lambda(c^*) = \mu(c^*)$, and by Lemmas 2 and 3, this happens if and only if every edge in $E(c^*)$ is non essential. We can reformulate these conditions in a more intuitive way.

Lemma 4 *If each seller is the exclusive bidder for exactly one edge, the overhead is zero in all limit equilibria ($\lambda(c) = \mu(c)$) if and only if at each stage of each instance of KA for c, we can select at least two edges.*

[3] We adopt the convention $0/0 = 1$ and for each $b > 0$, $b/0 = +\infty$.

Proof Suppose $\lambda(c) = \mu(c)$, and for some stage of an instance of KA for c there is only one edge, say $e \in E$, that can be selected. Let γ be the tree constructed in such an instance of KA. Since for each $f \in \Delta(e,\gamma)$, $\gamma + f - e$ is a tree, then $c_e < \min\{c_f : f \in \Delta(e,\gamma)\}$. Thus, $\mu_e(c) > c_e$ and $\mu(c) > \lambda(c)$. Conversely, suppose that each stage of each instance of KA leaves two or more choices, and pick $e \in E(c)$. Then one can choose an instance of KA that always select an edge different from e, so there is $\gamma \in \Gamma(c)$ such that $e \notin \gamma$.■

5.2 Metric costs

A cost matrix $c \in \mathbb{R}_+^E$ satisfies the triangular inequality if for all edges e, f, and g forming a triangle, we have $c_e \leq c_f + c_g$. We speak of a *metric cost c,* and let $T(V)$ be the set of metric costs.

Given $c \in T(V)$ and $\gamma \in \Gamma(c)$, there is in $T(V)$ a largest d such that for each $e \in \gamma$, $d_e = c_e$. For each edge $e = ab \notin \gamma$, it is given by $d_e \equiv \sum_{f\in[a,b]_\gamma} c_f$; moreover, $\gamma \in \Gamma(d)$. By Lemma 1, for $e \in \gamma$ we have $\mu_e(d) = c_e + \min\{c_f : f \in Ad(e,\gamma)\}$. From $c \leq d$ and the fact that $\mu(c)$ is weakly increasing in c (Lemma 3), we get $\mu(c) \leq \mu(d)$. Therefore,

$$\mu(c) \leq \lambda(c) + \sum_{e\in\gamma} \min\{c_f : f \in Ad(e,\gamma)\}. \tag{4}$$

Equivalently,

$$PIC(c) \leq 1 + \frac{\sum_{e\in\gamma} \min\{c_f : f \in Ad(e,\gamma)\}}{\lambda(c)}. \tag{5}$$

This inequality is an equality for the cost d just described. In particular if $c_e = 1$ for all $e \in \gamma$, we have $\min\{c_f : f \in Ad(e,\gamma)\} = 1$. Therefore, $\mu(d) = 2\lambda(d)$ and consequently, $PIC(d) = 2$.[4]

Our main result is that this is essentially the worst possible PIC over all metric costs.

Theorem 1 *Suppose that each seller is the exclusive bidder for exactly one edge. Then,*

$$\max_{c\in T(V)} PIC(c) = \begin{cases} 2 & \text{if } n \text{ is odd,} \\ 2\frac{n-1}{n-2} & \text{if } n \text{ is even.} \end{cases}$$

The proof that no metric cost can have a higher PIC is in the Appendix. We check here that these two bounds are reached.

We observed just before stating Theorem 1, that for any spanning tree γ, there is a cost matrix c such that $\gamma \in \Gamma(c)$ and $PIC(c) = 2$. Interestingly for n even, choosing c such that $PIC(c) = 2\frac{n-1}{n-2}$ places many more constraints on the structure of the optimal tree. We start with an example.

Label the nodes as $V = \{1, \cdots, n\}$ and consider the line-tree γ with edges $i(i+1)$ for $1 \leq i \leq n-1$. Set $c_{i(i+1)} = 0$ if i is odd, $= 1$ if i is even, so we have

[4] Many other metric costs give the same PIC, for instance $c_e = 1$ if $e \in \gamma$, $c_f = 2$ if $f \notin \gamma$.

$\frac{n}{2}$ free edges, including the two end edges. Set $c_{i(i+2)} = 1$ for $1 \leq i \leq n-2$, and $c_e \geq 1$ for other edges (we do not need to specify their costs as long as they remain metric). Lemma 1 gives $\mu_e(c) = 1$ for all $e \in \gamma$, hence $\mu(c) = n-1$ while $\lambda(c) = \frac{n}{2} - 1$.

Lemma 5
i) If n is odd, every spanning tree $\gamma \in \Gamma$ admits a cost matrix $c \in T(V)$ such that $\gamma \in \Gamma(c)$ and $\frac{\mu(c)}{\lambda(c)} = 2$.
ii) If n is even, the spanning tree $\gamma \in \Gamma$ admits a cost matrix $c \in T(V)$ such that $\gamma \in \Gamma(c)$ and $\frac{\mu(c)}{\lambda(c)} = 2\frac{n-1}{n-2}$ if and only if the edges of γ contain a perfect matching of V (i.e., we can find $\frac{n}{2}$ edges in γ such that each node in V is the end point of exactly one such edge).

5.3 Ultrametric costs

Let $U(V) \subset T(V)$ be the set of ultrametric cost matrices, i.e., $c_e \leq \max\{c_f, c_g\}$ for all edges e, f, and g forming a triangle. For any $c \in U(V)$, the tree γ is optimal ($\gamma \in \Gamma(c)$) if and only if

$$\text{for each } f = ab \notin \gamma,\ c_f = \max_{e \in [a,b]_\gamma} c_e, \tag{6}$$

(recall that for an arbitrary c, optimality of γ implies $c_f \geq \max_{e \in [a,b]_\gamma} c_e$).

Pick $e \in \gamma$ and $f = ab \in \Delta(e, \gamma)$: there is at least one edge $g \in Ad(e, \gamma) \cap [a, b]_\gamma$, and (6) implies $c_f \geq \max\{c_e, c_g\}$; conversely, for each $g \in Ad(e, \gamma)$ the edge f forming a triangle with e and g is in $\Delta(e, \gamma)$ and $c_f = \max\{c_e, c_g\}$. Therefore Lemma 1 gives

$$\mu_e(c) = \max\{c_e, \min\{c_f : f \in Ad(e, \gamma)\}\}. \tag{7}$$

For each $c \in U(V)$ we can write $\mu(c)$ as follows:

$$\mu(c) = \lambda(c) + \sum_{e \in \gamma} \min\{(c_f - c_e)_+ : f \in Ad(e, \gamma)\}.$$

Thus,

$$PIC(c) = 1 + \frac{\sum_{e \in \gamma} \min\{(c_f - c_e)_+ : f \in Ad(e, \gamma)\}}{\lambda(c)}. \tag{8}$$

Comparing (7) with (4), we see that the overhead on c_e is $\min\{(c_f - c_e)_+ : f \in Ad(e, \gamma)\}$ for an ultrametric cost, and (at most) $\min\{c_f : f \in Ad(e, \gamma)\}$ for a triangular one.

An example where the difference is extreme is for a line-tree γ with nodes $1, \cdots, n$, and $c_{12} < c_{23} < \cdots < c_{(n-1)n}$. Let c^u be the cost given by (6) outside γ. It is the unique ultrametric cost such that $\gamma \in \Gamma(c^u)$. For edge 12, the overhead is $c_{23} - c_{12}$. For $i \geq 2$, the cheapest edge in $Ad(i(i+1), \gamma)$ is $(i-1)i$, implying that the overhead on $c_{i(i+1)}$ is zero for c^u. Therefore, $\mu(c^u) - \lambda(c^u) = c_{23} - c_{12}$. By contrast, let c^t be the largest metric cost compatible with the given costs on γ, i.e., $c^t_{ij} \equiv \sum_{k=i}^{j-1} c_{k(k+1)}$ for all $i < j$. It is easy to compute $\mu(c^t) - \lambda(c^t) = c_{12} + 2c_{23} + c_{34} + \cdots + c_{(n-2)(n-1)}$. For instance if $c_{i(i+1)} = i$ for $1 \leq i \leq n-1$, we find $PIC(c^u) = 1 + \frac{2}{n(n-1)}$ while $PIC(c^t) \geq 1.8$ for all n.

However, the worst case $\max_{c \in U(V)} \frac{\mu(c)}{\lambda(c)}$ is *the same* for triangular and ultrametric costs. To see this, it is enough to check that these bounds can be reached in $U(V)$, because $U(V) \subset T(V)$. Consider the line tree γ with alternating costs 0 and 1, and 0 at both end edges if n is even, and set a cost of 1 for any edge outside γ. This cost c is clearly ultrametric, and (8) gives the $PIC(c) = 2$ for odd n and $PIC(c) = 2\frac{n-1}{n-2}$ for n even.

Theorem 2
i) Suppose that each seller is the exclusive bidder for exactly one edge. Then, $\max_{c \in T(V)} PIC(c) = \max_{c \in U(V)} PIC(c)$.
ii) If n is odd, the spanning tree $\gamma \in \Gamma$ admits a cost matrix $c \in U(V)$ such that $\gamma \in \Gamma(c)$ and $\frac{\mu(c)}{\lambda(c)} = 2$ if and only if γ has at least one leaf edge of which the inner end point is of degree two.
iii) If n is even, the spanning tree $\gamma \in \Gamma$ admits a cost matrix $c \in U(V)$ such that $\gamma \in \Gamma(c)$ and $\frac{\mu(c)}{\lambda(c)} = 2\frac{n-1}{n-2}$ if and only if the edges of γ contain a perfect matching of V.

Thus we can reach a PIC of 2 with a metric cost for any tree, whereas for a tree not covered by statements *ii*), *iii*) above, the PIC one can reach with an ultrametric cost is smaller, and sometimes much smaller. The most extreme example is the simple star, where all nodes but one are of degree one: the worst PIC of an ultrametric cost for which this star is optimal is $1 + \frac{1}{n-2}$.[5]

5.4 Expected overhead

In the past two subsections we showed that the worst case scenario for the PIC in the metric and ultrametric domains coincide. We also showed that for a given tree, say γ, the maximum PIC that can be achieved for a matrix c such that $\gamma \in \Gamma(c)$ is generally lower in the later domain. This suggests that the "expected" PIC is lower in the ultrametric domain than in the metric domain. We confirm this intuition by calculating the expected PIC under several probabilistic assumptions.

It is well known that to each triangular cost matrix, say $c \in T(V)$, there is an associated "irreducible cost matrix," $c^u \in U(V)$, defined as the smallest cost matrix which is bounded above by c and such that $\lambda(c^u) = \lambda(c)$.

This motivates the following probabilistic model. Let $C \subset T(V)$ be a set of triangular matrices (endowed with an appropriate σ-algebra) and p a probability measure on C. We associate with (C, p) the ultrametric probability space (C^u, p^u) defined by: $C^u \equiv \{c^u : c \in C\}$ and for each $D \subset C^u$, $p^u(D) \equiv p(\{c : c^u \in D\})$. The two spaces, (C, p) and (C^u, p^u) are comparable

[5]The worst PIC in $U(V)$ is always attained by a matrix c with $\gamma \in \Gamma(c)$ such that for each $e \in \gamma$, $c_e \in \{0, 1\}$. See proof of statement (*ii*) in Theorem 2.

in the sense that for each cost matrix $c \in C$ there is a cost matrix $c^u \in C^u$ with equivalent cost structure, i.e., $\Gamma(c) \subset \Gamma(c^u)$ and $\lambda(c) = \lambda(c^u)$.

We calculate first the expected PIC, for the triangular spaces and their associated ultrametric spaces, when the nodes $\{1, \ldots n\}$ are i.i.d uniformly in the unit cube $[0,1]^L$, and costs correspond to the euclidian distance. Formally

$$C \equiv \left\{ c \in \mathbb{R}^E : \text{for each } e = ab \in E,\, c_e = ||a - b||,\, a, b \in [0,1]^L \right\}, \tag{9}$$

and p is the Lebesgue measure. In contrast to the the maximum PIC (which is essentially constant as a function of the number of nodes), the expected PIC, both in the triangular and ultrametric spaces, is decreasing with respect to the number of nodes. Figure 1 shows the ratio of the expected PIC in the two domains for $L \in \{1, \ldots, 6\}$ and $n \in \{1, \ldots, 20\}$.

Consider now the ratio of the expected overhead $PIC\,(c) - 1$ in the triangular domain, to that in the ultrametric domain. It turns out this ratio is increasing in n, with an initial value of at least 2 for $n = 3$. See Figure 1 (c).

Second, we perform two exercises that shed some light on the difference between the expected overhead in the triangular and ultrametric domains conditional on the optimal tree being (i) a line-tree, and (ii) a simple star.

We calculate first the expected PIC when nodes $\{1, \ldots n\}$ are located in a line following the order $1, 2, \ldots, n$ and the distance between two consecutive nodes is i.i.d. uniformly on $[0,1]$. We compare the expected overhead between the triangular costs (9) and their ultrametric counterparts for $n \in \{1, \ldots, 20\}$. See Figure 2 (a).

Finally, we perform the parallel exercise for a simple star γ with center 1 and in which the costs of the edges $1i, 2 \leq i \leq n$, are i.i.d. uniformly on $[0,1]$. We compare the metric costs $c_e = c_{1i} + c_{1j}$ for $e = ij, i, j \geq 2$, and its ultrametric counterpart $c_e^u = \max\{c_{1i}, c_{1j}\}$. The ratio of the two expected overheads metric over ultrametric, grows approximately linearly in n with slope 1, starting with 2.6 for $n = 3$. See Figure 2 (b).

The computations above show that for our typical probabilistic assumptions the expected overhead for metric costs triangular is at least twice that for ultrametric costs.

6 The case of multiple ownership

If each seller bids for a set of mutually disconnected edges (no two edges share a node), our two main results, Theorems 1 and 2, still hold. Indeed it is straightforward to adapt the proof of Proposition 2 to show the following. If the buyer purchases γ and edge $e \in \gamma$ is sold by seller i, then the maximum price the buyer can pay for edge e in equilibrium is $\min\{c_f | f \in \Delta(e, \gamma),\, f \notin l(i)\}$. Because edges in $l(i)$ are mutually disconnected, $Ad(e, \gamma) \cap l(i) = \emptyset$, therefore the maximum equilibrium overhead for e is at most $\min\{c_f | f \in Ad(e, \gamma)\}$, and the two bounds for the maximal PIC, 2 and $2\frac{n-1}{n-2}$, still hold.

If sellers may bid for sets of possibly connected edges, these bounds are not preserved.

Proposition 1 still holds: an optimal tree is purchased in any (limit) equilibrium. However if a seller owns several edges forming a *cut* (a set of edges intersecting every spanning tree), he has monopsony power so in our model he can extract infinite surplus.

Assume now that no seller owns a cut. Then the PIC can increase quite significantly. Here is an example.

Suppose $n = kp, V = \{1, \cdots, n\}$, and p sellers can bid each for a different k-clique of edges. That is, seller 1 bids for all edges connecting nodes $1, \cdots, k$, seller 2 for all edges connecting nodes $k+1, \cdots, 2k$, and so on until seller p who bids for edges inside $(p-1)k+1, \cdots, pk$. Other edges are covered by other sellers. Assume that the cost of any edge inside any clique is zero, and the cost of all other edges is 1. Then the profile of bids where all edges carry a price of 1 (so only the p clique-owners overcharge), and the line-tree $1 - 2 - \cdots - n$ is purchased, is a limit equilibrium where the PIC is $\frac{n-1}{p-1}$. If $k = 2$ a clique is a single edge so we are in the situation of section 4, and indeed $\frac{n-1}{p-1} = 2\frac{n-1}{n-2}$. For larger k we have $\frac{n-1}{p-1} = \frac{kp-1}{p-1} \geq k$, so the PIC increases linearly in the clique-size.

We conjecture that when the number n of nodes becomes large, and each seller is the exclusive bidder for at most a k-clique, then the maximum PIC converges to k.

7 Appendix A: Proofs

7.1 Proposition 1

Step 1. We prove a simple fact about minimal cost spanning trees: the cost of an inefficient spanning tree can always be reduced by replacing a single edge. Formally, for any $c \in \mathbb{R}_+^E$ we have

$$\{\gamma \in \Gamma \setminus \Gamma(c)\} \Rightarrow \text{ there exists } g \in \gamma \text{ and } f \in \Delta(g,\gamma) \text{ such that } c_f < c_g, \quad (10)$$

(hence $\gamma + f - g$ is a cheaper spanning tree).

Let $\gamma \in \Gamma \setminus \Gamma(c)$. Consider the instance of KA in which one selects edges in γ whenever possible. Let f be the first edge not in γ that is added. Then $\gamma + f$ has a cycle. Thus, there must be an edge in the cycle, say g, such that $g \in \gamma$ and $c_f < c_g$, for otherwise f would have not been added. Since g and f belong to the same cycle at $\gamma + f$, then $f \in \Delta(g,\gamma)$.

Step 2. We prove the following claim. Fix a Bertrand game $(V; E; S; c^i, i \in S)$ as above, an ε-equilibrium ${}^{\varepsilon}p$, and an edge $e \in \gamma({}^{\varepsilon}p)$ that is built by seller i, so ${}^{\varepsilon}p_e^i = {}^{\varepsilon}\pi_e$. Then ${}^{\varepsilon}p_e^i \geq c_e^* - 2\varepsilon$ (recall $c_e^* = \min_{\{j:e\in l(j)\}} c_e^j$).

Assume to the contrary ${}^{\varepsilon}p_e^i < c_e^* - 2\varepsilon$ and consider the following alternative bidding strategy $\widetilde{p}^i$:

- $\widetilde{p}_e^i \equiv c_e^i$;
- $\widetilde{p}_f^i \equiv \max\{c_f^i, {}^{\varepsilon}p_f^i\}$, for every $f \notin \gamma({}^{\varepsilon}p)$, and for every $f \in \gamma({}^{\varepsilon}p)$ that i does not build at ${}^{\varepsilon}p$;

- $\widetilde{p}^i_g \equiv^\varepsilon p^i_g - \frac{\varepsilon}{n-1}$ for every edge $g \in \gamma(^\varepsilon p)\diagdown\{e\}$ that i builds at $^\varepsilon p$.

Let $F = \{g \in \gamma(^\varepsilon p)\diagdown\{e\}|\ i$ builds g at $^\varepsilon p\}$. Irrespective of the tie-breaking rule, seller i still builds at $(^\varepsilon p_{-i}, \widetilde{p}^i)$ every edge in F: indeed his bids decrease strictly for those edges, and increase at least weakly for other edges, so the KA picks all of F. His profit from edges in F decreases by less than ε, while he saves more than 2ε on e, whether or not he still builds e at $(^\varepsilon p_{-i}, \widetilde{p}^i)$. Finally he does not lose anything on other edges he may be called to build. Thus his net profit increases by more than ε, contradiction.

Step 3. Fix the game $(V; E; S; c^i, i \in S)$ as above, an inefficient spanning tree $\gamma \notin \Gamma(c^*)$, and a limit equilibrium p such that γ is built in a sequence of ε-equilibria $^\varepsilon p$ where ε goes to zero. We derive a contradiction. By Step 1 there exist $g \in \gamma$ and $f \in \Delta(g, \gamma)$ such that $c^*_f < c^*_g$ and $\gamma + f - g$ is a cheaper spanning tree. Set $\delta = c^*_g - c^*_f$ and choose for some $\varepsilon < \frac{\delta}{5}$ an ε-equilibrium $^\varepsilon p$ where γ is built.

Let j be the seller chosen to build g at $^\varepsilon p$, and i be such that $c^i_f = c^*_f$. Distinguish two cases.

Case 1: $i \neq j$. We construct an alternative strategy $\widetilde{p}^i$ as follows:

- $\widetilde{p}^i_f \equiv^\varepsilon p^j_g - \varepsilon$;
- $\widetilde{p}^i_e \equiv \max\{c^i_e, ^\varepsilon p^i_e\}$, for every $e \notin \gamma$, and for every $e \in \gamma$ that i does not build at $^\varepsilon p$;
- $\widetilde{p}^i_e \equiv^\varepsilon p^i_e - \frac{\varepsilon}{n-1}$ for every edge $e \in \gamma$ that i builds at $^\varepsilon p$.

As above, seller i still builds at $(^\varepsilon p_{-i}, \widetilde{p}^i)$ every edge he was building at $^\varepsilon p$, because their price has decreased strictly so they are part of all optimal trees at $(^\varepsilon p_{-i}, \widetilde{p}^i)$. His profit decreases by at most ε for those edges. Moreover he now builds f as well, with a profit of more than 2ε: indeed by Step 2 and the definition of i we have

$$^\varepsilon p^j_g - \varepsilon \geq c^j_g - 3\varepsilon \geq c^*_g - 3\varepsilon = c^i_f + \delta - 3\varepsilon > c^i_f + 2\varepsilon.$$

Finally he does not lose money for edges other than f that he was not building and may be called to build. Contradiction.

Case 2: $i = j$. The alternative strategy of seller i is now:

- $\widetilde{p}^i_f \equiv^\varepsilon p^i_g - \varepsilon$; $\widetilde{p}^i_g =^\varepsilon p^i_g$;
- $\widetilde{p}^i_e \equiv \max\{c^i_e, ^\varepsilon p^i_e\}$, for every $e \notin \gamma$, and for every $e \in \gamma$ that i does not build at $^\varepsilon p$;
- $\widetilde{p}^i_e \equiv^\varepsilon p^i_e - \frac{\varepsilon}{n-1}$ for every edge $e \in \gamma\diagdown\{g\}$ that i builds at $^\varepsilon p$.

Once again, any optimal tree at $(^\varepsilon p_{-i}, \widetilde{p}^i)$ contains f, and includes all other edges different from g that i builds at $^\varepsilon p$. So i still builds the latter edges, and f as well. Distinguish two cases. If the tree purchased at $(^\varepsilon p_{-i}, \widetilde{p}^i)$ contains g, and i still builds g (for the same price), the same accounting argument shows that i's profit increases by more than ε, and yields a contradiction. If i does not build g any more (whether or not g is built at all), then his new profit on f exceeds his old profit (or loss) on g by more than 4ε:

$$\{({}^{\varepsilon}p_g^i - \varepsilon) - c_f^i\} - \{{}^{\varepsilon}p_g^i - c_g^i\} = c_g^i - c_f^i - \varepsilon \geq c_g^* - c_f^* > 4\varepsilon$$

and the proof is complete.

7.2 Proposition 2

Statement i) Fix a limit equilibrium p and for any $\varepsilon > 0$, an ε-equilibrium ${}^{\varepsilon}p$ converging to p. By Step 2 in the proof of Proposition 1, if edge e is selected at ${}^{\varepsilon}p$ ($e \in \gamma({}^{\varepsilon}p)$) and is built by agent i, then ${}^{\varepsilon}p_e^i \geq c_e^* - 2\varepsilon$. Therefore if e is selected at p (i.e., $e \in \gamma({}^{\varepsilon}p)$ for a sequence ε going to zero and such that $\gamma({}^{\varepsilon}p)$ is constant) and built by i we have $p_e^i \geq c_e^*$. This proves that the buyer pays at least $\lambda(c^*)$.

Next we pick an arbitrary $\varepsilon > 0$, an ε-equilibrium ${}^{\varepsilon}p$, and $e \in \gamma({}^{\varepsilon}p)$. We prove by contradiction that ${}^{\varepsilon}\pi_e \geq \mu_e(c^*) + 3\varepsilon$ is impossible. This will imply $\pi_e \leq \mu_e(c^*)$ in the limit, and conclude the proof.

Let $f \in E$ and $i \in N$ be such that $\mu_e(c_e^*) = c_f^* = c_f^i$, where $f \in \Delta(e, \gamma({}^{\varepsilon}p))$. Distinguish two cases.

Case 1: i does not build e at ${}^{\varepsilon}p$. Consider the switch by this player to the following strategy $\widetilde{p}^i$:

- $\widetilde{p}_f^i \equiv c_f^i + 2\varepsilon$;
- $\widetilde{p}_g^i \equiv \max\{c_g^i, {}^{\varepsilon}p_g^i\}$, for every $g \neq f$ and such that $g \notin \gamma({}^{\varepsilon}p)$, and for every $g \in \gamma({}^{\varepsilon}p)$ that i does not build at ${}^{\varepsilon}p$;
- $\widetilde{p}_g^i \equiv {}^{\varepsilon}p_g^i - \frac{\varepsilon}{n-1}$ for every edge $g \in \gamma({}^{\varepsilon}p)$ that i builds at ${}^{\varepsilon}p$.

The price to the buyer of edge f is now strictly below that of e; and that of any edge in $\gamma({}^{\varepsilon}p)$ that i builds at ${}^{\varepsilon}p$ goes down strictly, so those edges as well as f are essential at $({}^{\varepsilon}p_{-i}, \widetilde{p}^i)$. Seller i's profit goes up by more than ε because he was building at most $n-2$ edges at ${}^{\varepsilon}p$.

Case 2: i builds e at ${}^{\varepsilon}p$. then we define the deviation $\widetilde{p}^i$ as follows:

- $\widetilde{p}_f^i \equiv {}^{\varepsilon}p_e^i - \varepsilon$; $\widetilde{p}_e^i = \max\{c_e^i, {}^{\varepsilon}p_e^i\}$;
- $\widetilde{p}_g^i \equiv \max\{c_g^i, {}^{\varepsilon}p_g^i\}$, for every $g \neq f$ and such that $g \notin \gamma({}^{\varepsilon}p)$, and for every $g \in \gamma({}^{\varepsilon}p)$ that i does not build at ${}^{\varepsilon}p$;
- $\widetilde{p}_g^i \equiv {}^{\varepsilon}p_g^i - \frac{\varepsilon}{n-1}$ for every edge $g \in \gamma({}^{\varepsilon}p) \diagdown \{e\}$ that i builds at ${}^{\varepsilon}p$.

The usual accounting argument shows that i's profit grows by more than ε.
Statement ii). Fix c, γ, and p as in (3), and $\varepsilon > 0$. We check that the following profile of bids ${}^{\varepsilon}p$:

$${}^{\varepsilon}p^e \equiv \mu_e(c) - \varepsilon \text{ if } e \in \gamma; {}^{\varepsilon}p^e \equiv c_e \text{ if } e \notin \gamma,$$

is an ε-equilibrium. The characterization (2) of $\Gamma(c')$ (Step 2 in the proof of Lemma 1) can be refined as follows:

$$\Gamma(c') \equiv \{\gamma'\} \Leftrightarrow \{\text{for all } f \equiv ab \notin \gamma' \colon\ c_f > \max\{c_e : e \in [a,b]_{\gamma'}\};$$

(we omit the straightforward proof). This implies $\Gamma({}^{\varepsilon}p) = \{\gamma\}$. Each seller $e \notin \gamma$ has zero profit, and cannot get positive profit by changing his bid. A seller $e \in \gamma$ has the net profit $\mu_e(c) - c_e - \varepsilon$, possibly negative but no less than $-\varepsilon$. If he raises his price by more than ε, he is no longer part of the optimal tree (or trees). We conclude that no deviation raises his profit by more than ε, as desired.■

7.3 Theorem 1 and Lemma 5

By inequality (8) it will be enough to prove, for any $\gamma \in \Gamma$ and $c \in \mathbb{R}_+^{\gamma}$

$$\sum_{e\in\gamma} \min\{c_f : f \in Ad(e,\gamma)\} \leq \sum_{e\in\gamma} c_e \text{ if } n \text{ is odd,} \tag{11}$$

$$\sum_{e\in\gamma} \min\{c_f : f \in Ad(e,\gamma)\} \leq \frac{n}{n-2} \sum_{e\in\gamma} c_e \text{ if } n \text{ is even.} \tag{12}$$

We use the following notation. Let $G \subseteq E$ and $\sigma : \{1,\ldots,|G|\} \to G$ be a bijection. The convex cone of costs for G ordered by σ is $K_\sigma(G) \equiv \{c \in \mathbb{R}_+^G : c_{\sigma(1)} \leq c_{\sigma(2)} \leq \cdots \leq c_{\sigma(|G|)}\}$. The canonical basis of $K_\sigma(G)$ is the set $\{b^k \in \{0,1\}^G : 1 \leq k \leq |G|, b^k_{\sigma(j)} = 1 \Leftrightarrow j \geq k\}$

Let σ be a bijection such that $c \in K_\sigma(\gamma)$. The expressions on the left and right of (11) and (12) are positively linear (commutes with linear combinations with positive coefficients) on $K_\sigma(\gamma)$. Therefore it is enough to prove (11) and (12) for the profiles b^k in the canonical basis of $K_\sigma(\gamma)$. As σ is arbitrary, this amounts to prove (11) and (12) for costs c such that $c_e \in \{0,1\}$ for all $e \in \gamma$. It will be convenient to describe each such c by the subset B of its "free" edges: $e \in B \Leftrightarrow c_e = 0; e \in \gamma \setminus B \Leftrightarrow c_e = 1$. With the notation $Ad(B,\gamma) = \cup_{e\in B} Ad(e,\gamma)$, we have then $\sum_{e\in\gamma} c_e = n - |B| - 1$ and

$$\min\{c_f : f \in Ad(e,\gamma)\} = \begin{cases} 0 & \text{if } e \in Ad(B,\gamma), \\ 1 & \text{otherwise.} \end{cases}$$

Thus $\sum_{e\in\gamma} \min\{c_f | f \in Ad(e,\gamma)\} = n - 1 - |Ad(B,\gamma)|$. Now (11) and (12) are rewritten as

$$|Ad(B,\gamma)| \geq |B| \text{ if } n \text{ is odd; } (n-2)|Ad(B,\gamma)| + 2(n-1) \geq n|B| \text{ if } n \text{ is even.} \tag{13}$$

Property (13) is clear if $B = \emptyset$ so we assume from now on that B is non empty. We say that B, a non empty subset of γ, is of "type 1" if $|Ad(B,\gamma)| \geq |B|$, and of "type 2" if $|Ad(B,\gamma)| = |B| - 1$.

We prove by induction on $n = |V|$ the property $\mathcal{P}_n$, combining statements (14), (15), and (16):

$$|Ad(B,\gamma)| \geq |B| - 1 \text{ for any } B, \emptyset \neq\ B \subseteq \gamma \quad \text{(there is no "type 3")}, \tag{14}$$

$$B \text{ is of type 1 if all the edges incident to some node of } \gamma \text{ are outside } B, \tag{15}$$

$$B \text{ is of type 1 if t contains at least two adjacent edges of } \gamma. \tag{16}$$

The statement $\mathcal{P}_3$ is clear because $|Ad(B,\gamma)| = |B|$, whether B contains a single edge or $B = \gamma$. We assume $\mathcal{P}_3, \ldots, \mathcal{P}_{n-1}$ and fix an arbitrary subset B of γ. We prove (15) first. Let i be a node of which all incident edges are outside B.

Case 1: Suppose first i is a leaf,[6] so its unique incident edge e is also a leaf, and is not in B. Distinguish two subcases.

If $Ad(e,\gamma)$ contains at least one edge f in B, the inductive assumption (14) implies $|Ad(B,\gamma - e)| \geq |B| - 1$; on the other hand $Ad(B,\gamma) = Ad(B,\gamma - e) + e$, so B is of type 1. If $Ad(e,\gamma) \cap B = \emptyset$, pick one leaf f in $Ad(e,\gamma)$ and consider the subtree $\gamma[f;e]$ away from e and with f as a leaf, i.e., generated by f and all the nodes i such that f is between i and e. By the inductive assumption (15), if B intersects $\gamma[f;e]$ we have $|Ad(B \cap \gamma[f;e], \gamma[f;e])| \geq |B \cap \gamma[f;e]|$. Summing up these inequalities over f in $Ad(e,\gamma)$ such that $B \cap \gamma[f;e] \neq \emptyset$, gives (15) because the sets $Ad(B \cap \gamma[f;e], \gamma[f;e])$ are pairwise disjoint and cover $Ad(B,\gamma)$ when f varies in $Ad(e,\gamma)$; similarly the sets $B \cap \gamma[f;e]$ cover B.

Case 2: Suppose i is of degree two or more. For each edge e incident to i, hence outside B, apply as above the inductive assumption (15) to the subtree $\gamma[e;i]$ away from i and with e as a leaf (if $B \cap \gamma[e;i] \neq \emptyset$), then sum up as above over all edges incident to i.

We prove (16) next. If B contains at least two adjacent edges of γ, pick a non trivial maximal subtree δ of γ, entirely in B and containing these two edges. If $\delta = \gamma$ property (16) is clear. Otherwise every edge f between δ and $B \setminus \delta$ is outside B, so we can again apply (15) to $\gamma[f;\delta]$ and sum up over f, because the sets $Ad(B \cap \gamma[f;\delta], \gamma[f;\delta])$ are pairwise disjoint. Thus $Ad(B,\gamma)$ contains at least $|B \cap \gamma \setminus \delta|$ edges outside δ. Moreover it also contains all edges in δ. This implies $|Ad(B,\gamma)| \geq |B|$ as desired.

Finally we prove (14). Suppose B is such that $|Ad(B,\gamma)| < |B|$. By (16) no two edges in B are adjacent, and by (15) every node is the endpoint of an edge in B. Thus every node is the end point of a unique edge in B, in other words the edges of B form a perfect matching of V. In particular n is even and $|B| = \frac{n}{2}$. Moreover $Ad(B,\gamma) = \gamma \setminus B$ hence $|Ad(B,\gamma)| = \frac{n}{2} - 1$, completing the proof of (14) and of $\mathcal{P}_n$.

This argument also shows that if B is of type 2 then n is even, and in this case $|B| = \frac{n}{2}$, so that the right hand side inequality in (13) is an equality. The latter inequality holds if B is of type 1 as well, so the proof of (13) and of Theorem 1 is complete.

[6] A node of degree 1 (a single edge in γ has this node as an end); we also call edge $e = ab$ a leaf of γ if either a or b is a leaf.

It remains to prove statement $ii)$ in Lemma 5. If n is even and the edges of γ contain a perfect matching B, then $|B| = \frac{n}{2}, |Ad(B,\gamma)| = \frac{n}{2} - 1$, and the above argument shows that for the metric cost $c_e = 0$ if $e \in B$, $= 1$ if $e \in E \diagdown B$, (12) holds as an equality, or equivalently $PIC(c) = 2\frac{n-1}{n-2}$.

Conversely, suppose for some $c \in T(V)$ we have $PIC(c) = 2\frac{n-1}{n-2}$. Then the restriction of c to γ is non zero, because $c_e = 0$ for all $e \in \gamma$ implies $c \equiv 0$ and $\lambda(c) = \mu(c) = 0$, so $PIC(c) = 1$. Moreover (12) holds as an equality. Say $c \in K_\sigma(\gamma)$ then c is a positive linear combination of the cost vectors b^k in the canonical basis of this cone. Inequality (12) holds for each b^k , therefore it must be an equality for at least one non zero b^k (one with a strictly positive coefficient in the decomposition of c). The cost vector b^k takes only the values $0, 1$, so we can represent it as above by the set B of its free edges; now equality in (12) amounts to

$$(n-2)|Ad(B,\gamma)| + 2(n-1) = n|B|.$$

As $b^k \neq 0$, B is a strict subset of γ, $|B| < n-1$. This implies $|Ad(B,\gamma)| < |B|$, i.e., B is of type 2. We can repeat the argument three paragraphs earlier showing that B is a perfect matching of V.

7.4 Theorem 2

Statement iii): n is even. As $U(V) \subset T(V)$, the "only if" statement follows from statement $ii)$ in Lemma 5. Conversely if γ contains a perfect matching B, the cost function $c_e = 0$ for all $e \in B; c_e = 1$ for all $e \in \gamma \setminus B$ is clearly ultrametric, and as we checked immediately after stating Theorem 1, its PIC is $2\frac{n-1}{n-2}$.

Statement ii): n is odd. For "if,"suppose node a is a leaf of γ, and the leaf edge $f = ab$ is adjacent (in γ) to a single edge $g = bd$. It is easy to check that the cost function: $c_e = 1$ if e connects $\{a, b\}$ to $V \diagdown \{a, b\}$, $c_e = 0$ otherwise, is ultrametric. Clearly $\lambda(c) = 1$, and (7) implies $\mu_f(c) = \mu_g(c) = 1$, therefore $PIC(c) = 2$.

For "only if", assume that for some $c \in U(V)$ we have $PIC(c) = 2$. First we mimick the argument in the proof of Lemma 5 showing that such a c can be chosen with $c_e \in \{0, 1\}$ for all $e \in \gamma$. By (8) and Theorem 1 we have for all $c \in U(V)$

$$\sum_{e\in\gamma} \min\{(c_f - c_e)_+ : f \in Ad(e,\gamma)\} \leq \sum_{e\in\gamma} c_e, \tag{17}$$

with equality if and only if $PIC(c) = 2$. The latter implies that the restriction of c to γ, denoted c as well, is non zero, and if $c \in K_\sigma(\gamma)$, it is a positive linear combination of cost vectors b^k also in $K_\sigma(\gamma)$ with all coordinates 0 or 1. Moreover inequality (17) is linear in this cone, so for at least one of the vectors b^k (17) is an equality. And b^k can be extended (uniquely) to an ultrametric cost on E.

Next, as in the previous proof, we represent our $0, 1$- cost function c by the set B of its free edges. For any $e \in \gamma$, we have $\min\{(c_f - c_e)_+ : f \in Ad(e,\gamma)\} = 1$

if e is an isolated edge in B, i.e., it is not adjacent to any other edge in B; and $\min\{(c_f - c_e)_+ : f \in Ad(e, \gamma)\} = 0$ for any other edge in γ. Let B^* be the subset of isolated edges in B: equality in (17) amounts to $|B^*| = |\gamma \diagdown B|$. We prove by contradiction below that at least one $e \in B^*$ is adjacent to exactly one edge in $\gamma \diagdown B$: this edge is the announced leaf edge of which the inner end point has degree two.

Clearly each $e \in B^*$ is adjacent to at least one edge in $\gamma \diagdown B$. Suppose each $e \in B^*$ is adjacent to at least two edges in $\gamma \diagdown B$. Notice that an edge in $\gamma \diagdown B$ is adjacent to at most two edges in B^*. Consider the bipartite graph with nodes the elements of B^* on the left side, and those of $\gamma \diagdown B$ on the right side (the same number of nodes on each side). It is a simple matter to check that if each left node is connected to at least two right nodes, and each right node to at most two left nodes, the bipartitie graph contains a cycle, which in turn gives a cycle in γ, contradiction.

8 Appendix B: Figures

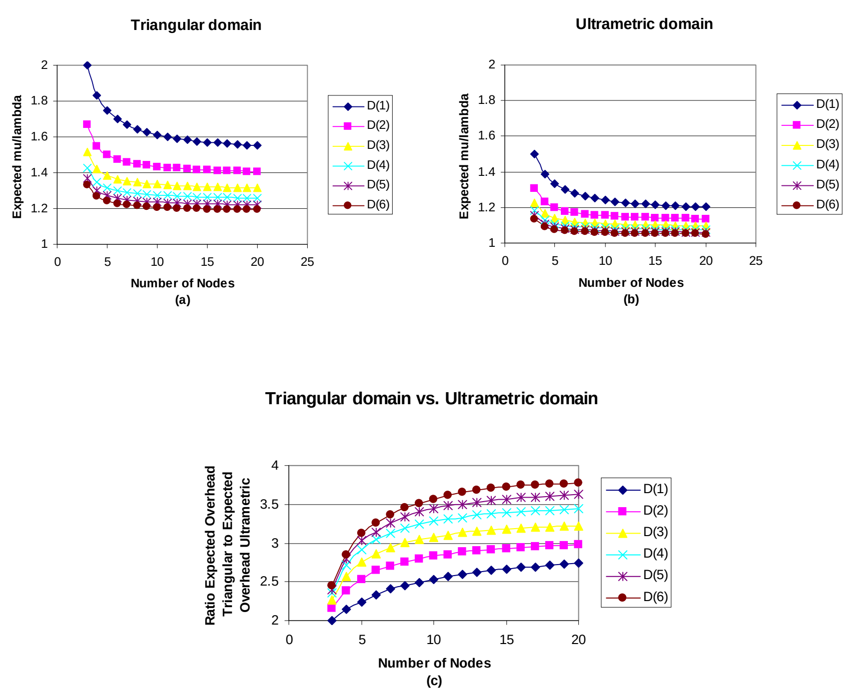

Figure 1: PIC and expected overhead unit cube model.

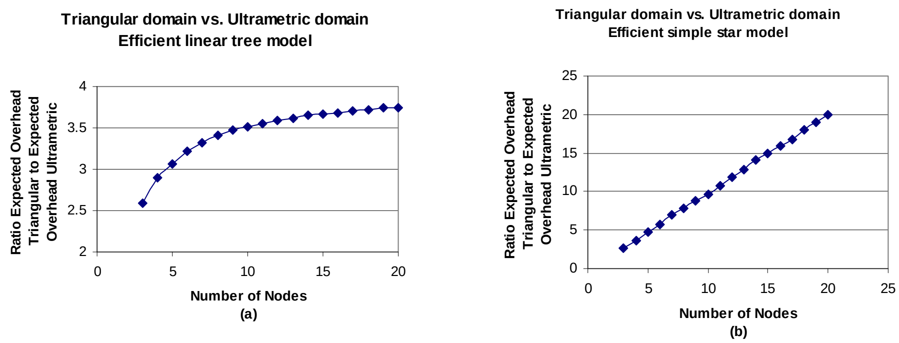

Figure 2: Expected overhead linear and star tree models.

References

[1] N. Ailon, M. Charikar, Fitting tree metrics: Hierarchical Clustering and Phylogeny, *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, 2005.

[2] G. Bergantiños, J. Vidal-Puga, A fair rule in minimum cost spanning tree problems, *J. Econ. Theory* 137, 2007, 326-352.

[3] A. Bogomolnaia, H. Moulin, Sharing a minimal cost spanning tree: beyond the folk solution, *Games Econ. Behav.* 69, 2010, 238-248.

[4] T. Borgers, Iterated elimination of dominated strategies in a Bertrand-Edgeworth model, *Review of Economic Studies* 59, 1992, 163-176.

[5] R. Brânzei, S. Moretti, H. Norde, S. Tijs, The P-value for cost sharing in minimum cost spanning tree situations, *Theory and Decision, 56, 2004, 47-61.*

[6] D. Fudenberg, D. Levine, Limit games and limit equilibria, *J. Econ. Theory* 38, 1986, 261-279.

[7] J. Hirschleifer, J. Riley, *The analytics of uncertainty and information,* Cambridge Surveys of Economic Literature, Cambridge University Press, 1992.

[8] E. Koutsoupias, C. Papadimitriou, Worst case equilibria, *Proceedings of the 16th Symposium on Theoretical Aspects of Computer Science*, 1999, pp. 404-413.

[9] J. Kruskal, On the shortest spanning subtree of a graph and the traveling salesman problem, *Proc. Am. Math. Soc.* 7, 1956, 48-50.

[10] H. Moulin, The price of anarchy of serial, average, and incremental cost sharing, *Econ. Theory*, 36, 3, 2008, 379-405.

[11] N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani, (Editors), *Algorithmic Game Theory*, Cambridge University Press, 2007.

[12] H. Norde, S. Moretti, S. Tijs, Minimum cost spanning tree games and population monotonic allocation schemes, *European Journal of Operations Research 154, 1, 2001, 84-97.*

[13] T. Roughgarden, E. Tardos, How bad is selfish routing? *J. ACM* 49 (2), 2002, pp.236-259.

[14] L. Simon, W. Zame, Discontinuous games and endogenous sharing rules, *Econometrica* 58, 1990, 861-872.

[15] W. Sharkey, Network models in economics. In: Ball et al. (eds.) *Handbooks of Operations Research and Management Science*, pp. 713-765. New York: Elsevier,1995.

[16] J. Tirole, *The Theory of Industrial Organization*, The MIT Press, Cambridge, London, 1988.