

Spanning Trees With Edge Conflicts and Wireless Connectivity

Magnús M. Halldórsson

School of Computer Science, Reykjavik University, Iceland
mmh@ru.is

Guy Kortsarz

Rutgers University, Camden, NJ, USA
guyk@camden.rutgers.edu

Pradipta Mitra

Google Research, New York, USA
ppmitra@gmail.com

Tigran Tonoyan

School of Computer Science, Reykjavik University, Iceland
ttonoyan@gmail.com

Abstract

We introduce the problem of finding a spanning tree along with a partition of the tree edges into fewest number of feasible sets, where constraints on the edges define feasibility. The motivation comes from wireless networking, where we seek to model the irregularities seen in actual wireless environments. Not all node pairs may be able to communicate, even if geographically close – thus, the available pairs are specified with a link graph $\mathcal{L} = (V, E)$. Also, signal attenuation need not follow a nice geometric formula – hence, interference is modeled by a conflict (hyper)graph $\mathcal{C} = (E, F)$ on the links. The objective is to maximize the efficiency of the communication, or equivalently, to minimize the length of a schedule of the tree edges in the form of a coloring.

We find that in spite of all this generality, the problem can be approximated linearly in terms of a versatile parameter, the inductive independence of the interference graph. Specifically, we give a simple algorithm that attains a $O(\rho \log n)$ -approximation, where n is the number of nodes and ρ is the inductive independence, and show that near-linear dependence on ρ is also necessary. We also treat an extension to Steiner trees, modeling multicasting, and obtain a comparable result.

Our results suggest that several canonical assumptions of geometry, regularity and “niceness” in wireless settings can sometimes be relaxed without a significant hit in algorithm performance.

2012 ACM Subject Classification Theory of computation → Routing and network design problems, Networks → Network design and planning algorithms

Keywords and phrases spanning trees, wireless capacity, aggregation, approximation algorithms

Digital Object Identifier 10.4230/LIPIcs.ICALP.2018.158

Related Version A full version of the paper is available at <https://arxiv.org/abs/1803.04578>.

Funding The first and last authors are supported by grants nos. 152679-05 and 174484-05 from the Icelandic Research Fund. The second author is supported by NSF grant 1540547.



© Magnús M. Halldórsson, Guy Kortsarz, Pradipta Mitra, and Tigran Tonoyan; licensed under Creative Commons License CC-BY

45th International Colloquium on Automata, Languages, and Programming (ICALP 2018).
Editors: Ioannis Chatzigiannakis, Christos Kaklamani, Dániel Marx, and Donald Sannella;
Article No. 158; pp. 158:1–158:15



Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



1 Introduction

We introduce the problem of finding a spanning tree along with a partition of the tree edges into fewest number of feasible sets, which are independent sets in a given conflict (hyper)graph. The motivation comes from wireless networking, where we seek a basic communication structure while capturing the irregularities seen in actual wireless environments.

A spanning tree is the minimal structure for connecting the given set of nodes into a mutually communicable network. The *cost* of a communication spanning tree is the *time* required to *schedule* all the tree edges – the transmission *links* – while obeying the *interference* caused by simultaneous transmissions. The *scheduling complexity* of the tree represents its throughput capacity: how much communication can be sustained in the long run. The task might be to *aggregate* the data measured at the sensor nodes, or to *broadcast* using one-to-one communication to all nodes of the network.

Algorithmic studies of wireless connectivity to date have generally involved strong “niceness” assumptions. One core assumption is that points are located in the Euclidean plane and *all (close enough) pairs of nodes are available as links* for use in the spanning tree. Interference modeling has become progressively more realistic, starting with range-based graph models to the fractional *SINR* model of interference, but the common thread is that *interference is a direct function of the geometry*. While natural, these assumptions depend on a simplified view of the nature of wireless communication.

Wireless networking in the real world behaves quite different from these theoretical models [10, 32, 38] and typically displays a high degree of irregularity. This manifests in how the strength of signals (and the corresponding interference) often varies greatly within the same region, and is often poorly correlated with distance [2]. This behavior holds even in simple outdoor environments, but is magnified inside buildings. It is also evidenced by fluctuations, sensitivity to environmental changes (even levels of humidity), and hard-to-explain unreliability.

There has been increased emphasis for greater robustness in the design and analysis of wireless algorithms to address the observed irregularities. In the world of communications engineering, the default is to introduce stochastic distributions, e.g., on signal strengths. The algorithms world prefers more adversarial effects, but that can easily lead to intractability.

The objective of this work is to embrace this irregularity in connectivity problems. We replace the previous assumptions by the opposite premises:

A link may not be usable even if it should be.

and

Interference need not follow (or even relate to) the underlying geometry.

Technically, the former premise means that the set of usable or *available links* is now given as a link graph $\mathcal{L} = (V, L)$. We place no restrictions on the structure of this graph. The second premise implies another graph, this time on top of the *links*. Namely, the conflict (hyper)graph $\mathcal{C} = (L, F)$ specifies whether a given pair of links in L can coexist in the same color (of a spanning tree). In the CONNECTIVITY SCHEDULING problem, we seek a spanning tree T of \mathcal{L} and a coloring of the links of T minimizing the number of colors used.

These formulations naturally raise a number of questions: Can arbitrary sets of available/usable links actually be handled effectively? Can we disconnect the conflicts/interference from the geometry? Since the ugly specter of intractability is bound to raise its head somewhere, what are minimal restrictions that keep these problems well-approximable?

Our Results. Given the generality of the CONNECTIVITY SCHEDULING problem, it is unsurprising that it is very hard even to approximate. We show that strong $n^{1-\epsilon}$ -approximation hardness holds, even for the natural special case of 2-hop interference. Instead, we aim to obtain approximations in terms of natural instance parameters.

We show that the problem is approximable within $O(\rho \log n)$ -factor, where ρ is the *inductive independence* of the (fractional) conflict graph. This is particularly relevant since ρ is known to be constant in both of the predominant interference models: the physical (or SINR) model, and the protocol model. This is attained by a simple greedy algorithm that can be viewed as a combination of Kruskal’s MST algorithm and a link scheduling algorithm for the physical model. In contrast, we find that the (perhaps more natural) approach of selecting and coloring an MST fails badly.

We also generalize the problem to Steiner trees and obtain a similar logarithmic approximation. The results carry over to the SINR model, where we obtain the first results that hold for general metric spaces. We also give a supplementary result in the full version (for space reasons), involving natural geometric interference assumptions, namely that all links of length smaller than a threshold are available.

Definitions. In line with a modern view of wireless interference, we represent the interference conflicts by a fractional conflict graph $\mathcal{C} = (L, W)$. Here L is the set of communication links and $W : L \times L \rightarrow \mathbb{R}^+$ is a function on ordered pairs of links, where $W(e, f)$ represents (or approximates) the degree to which a transmission on link e interferes with a transmission on link f . Of particular interest are functions W in terms of geometric relationships involving link lengths and distances between links. Note that W may be asymmetric. For convenience, let $W(e, e) = 0$. We shall write $W(S, e) = \sum_{f \in S} W(f, e)$ and $W(f, S) = \sum_{e \in S} W(f, e)$. Let $\mathcal{C}[Y] = (Y, W|_Y)$ denotes the subgraph induced by a given subset $Y \subseteq L$.

A set S of links is an independent or a feasible set if $W(S, e) \leq 1$, for all $e \in S$. A coloring of $\mathcal{C} = (L, W)$ is a partition of L into independent sets. Observe that when W is a 0-1 function, we have the usual independent sets and colorings of graphs. Also, the formulation with fractional conflicts corresponds to hypergraphs that contain a hyperedge for each minimal set S' where $W(S', e) \geq 1$ holds for some $e \in S'$.

We can now state our CONNECTIVITY SCHEDULING problem formally:

Given a link graph $\mathcal{L} = (V, L)$ and a fractional conflict graph $\mathcal{C} = (L, W)$, we seek a spanning tree T of \mathcal{L} and a coloring of $\mathcal{C}[T]$, using the fewest number of colors.

A fractional conflict graph $\mathcal{C} = (L, W)$ is said to be ρ -inductive independent, w.r.t. an ordering \prec of the links, if for every link e and every feasible set I with $e \prec I$, $W(I, e) + W(e, I) \leq \rho$, where $e \prec I$ means that e precedes each link in I . Here, “inductive” refers to how the interference is measured only towards later links, and “independence” that it is towards independent sets. In geometric settings (including range-based and SINR models), \prec corresponds to a non-decreasing ordering by link length.

For a fractional conflict graph $\mathcal{C} = (L, W)$, let $\chi(\mathcal{C})$ denote the smallest number of independent sets into which L can be partitioned; when \mathcal{C} is an ordinary graph, $\chi(\mathcal{C})$ is the chromatic number of \mathcal{C} . We view a coloring of \mathcal{C} also as a schedule and refer to the colors also as slots (which could be time slots or frequency bands).

Notable Instantiations. CONNECTIVITY SCHEDULING has a number of special cases of independent interest, both graph-based and geometric:

- A well-studied setting is where two links conflict if they are incident on a common link, i.e., when \mathcal{C} is the square of the line graph of the link graph \mathcal{L} . This case corresponds to bidirectional version of the classic *radio network* model. The directed version of CONNECTIVITY SCHEDULING was treated in [9] as the *radio aggregation scheduling problem*.
- In *range-based or disk models*, nodes are embedded in the plane and two links are adjacent if the distance between (the closest points on) them is less than K times the length of the longer link, where K is some fixed constant. In a variant, the condition is on distances between particular nodes on the links. Also, in the related *protocol* model, adjacency occurs if the distance is less than K_1 times the length of the longer link plus K_2 times the length of the shorter link, for some constants K_1, K_2 .
- The original driving motivation is when nodes and links are embedded in a metric space and the fractional conflicts follow the *geometric SINR model* of interference in terms of the lengths and distances between links. Before this work, only the case when \mathcal{L} is the complete graph over a set of points in a Euclidean metric was considered.
- A different geometric version is when we view that no signal gets transmitted between nodes on unavailable links, perhaps due to an obstacle. The links are then unavailable, but the nodes also don't interfere with each other. We refer to this as the Missing Links version.
- A natural special case occurs when link unreliability is restricted by link length, so that only reasonably long links are unavailable or attenuated, but short links follow the normal SINR laws (short links are reliable). This is treated in the full version of this paper.
- Finally, when the conflict graph \mathcal{C} is the *line graph* of the link graph \mathcal{L} , i.e., $\mathcal{C} = L(\mathcal{L})$, we obtain the well-known *minimum degree spanning tree* (MDST) problem, where given a graph \mathcal{L} , the goal is to find a spanning tree of smallest maximum degree. By König's theorem, the chromatic number of the line graph of a tree (in fact, of any bipartite graph) is equal to the maximum degree of the tree. This problem has more structure that allows for better solution: while it is NP-hard, it can be approximated within an additive one [8]. In particular, $L(\mathcal{L})$ is claw-free (does not contain an induced star graph $K_{1,3}$), which is stronger than being 2-inductive independent), and is intimately related to \mathcal{L} .

Related Work. The connectivity problem in the geometric SINR model was first considered by Moscibroda and Wattenhofer [35]. It was, in fact, the first work on worst-case analysis in the SINR model. They show that unlike in random networks, the worst-case connectivity depends crucially on the use of power control, and with optimal power control, $O(\log^4 n)$ slots suffice to connect the nodes. They soon improved this to $O(\log^2 n)$ [36, 34]. Currently, the best upper bounds known are $O(\log n)$ [17] and $O(\log^* \Lambda)$ [23], where Λ is the ratio between the longest to the shortest length of a link in a minimum spanning tree (MST), a structural parameter that is independent of n . Both of these results hold for the MST of the pointset; there are pointsets where $\Omega(\log^* \Lambda)$ slots are necessary for scheduling an MST [21].

The scheduling complexity of connectivity relates closely to the efficiency of *aggregation*, a key primitive for wireless sensor networks. We refer the reader to [26] for bibliography on aggregation/collection problems.

There are many approaches that have been proposed to model irregularity in wireless networks. We first examine static cases, or the modeling of non-geometric behavior. The basic SINR model allows the pathloss constant α to be adjusted [13], giving a first-order approximation of the signal gain. In the engineering community, it is most common to assume that the deviations are drawn from a particular stochastic distribution, typically

assuming independence of events. In the TCS camp, the prevailing approach is to view the variations as conforming the plane into a non-Euclidean metric space [7, 16], while retaining some tractable characteristics. This can also entail identifying appropriate parameters [4].

For frequent temporal changes, the standard engineering assumption is Rayleigh fading. Dams et al. [6] (see also [22]) showed that link scheduling algorithms are not significantly affected by such variation, assuming independence across time.

For unpredictably changing behavior, there is much research on adapting to new conditions, particularly with exponential backoff. A theoretic model proposed to specifically capture unreliability is the *dual graph model* [33], which extends the radio network model to a pair of graphs, the reliable and the unreliable links, where the latter are under adversarial control. The focus there is on distributed algorithms for one-shot problems, like global and local broadcast problems, where the nodes do not know which links are reliable. As far as we know, it has not been considered in settings involving a long-term communication structure.

Inductive independence was first defined by [1] and studied by [37] in the graph setting, while the weighted version was introduced by Hoefer and Kesselheim [25]. It has been used as a performance measure for various problems related to wireless networks, including admission control [11], dynamic packet scheduling [31, 15], and spectrum auctions [25, 24, 15].

Outline of the paper. We first examine, in Sec. 2, how the standard approach – finding a minimum spanning tree – fares for our problem, and show that it can give poor solutions in every known interference model when there are missing or unreliable links. We then give in Sec. 3 a greedy algorithm for CONNECTIVITY SCHEDULING achieving $O(\rho \log n)$ -approximation, where ρ is the inductive independence number of the conflict graph. This dependence on ρ is shown to be essentially tight in Sec. 5. We also obtain a similar approximation of a *Steiner or multicast* version of the problem in Sec. 4. Implications of our results to the SINR (or physical) model are given in Sec. 6. The rest of the paper can safely be read without any background in that model. We then close with open problems. Missing proofs, as well as a brief primer on SINR concepts, are given in the full version of the paper.

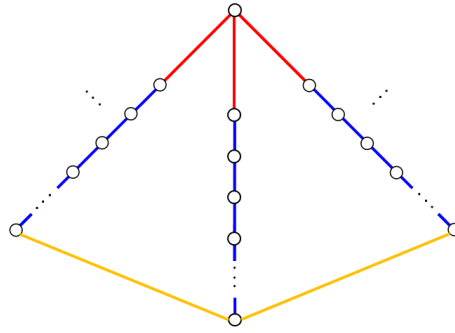
2 MST Fails

In a basic setting, the nodes are located on the plane, and the interference between two links is a function of the lengths of links (distance between the two end-nodes), and the distance between the (endpoints of) links. For instance, in the SINR model, the interference between two links is a decreasing function of their distance, and an increasing function of the length of the interfered link. In this setting, the Euclidean minimum spanning tree (MST) over the set of nodes is a natural candidate for connectivity, since it favors short links and has low degree (or, more generally, contains few links in the vicinity of any node). Indeed, the MST of n nodes can be scheduled in $O(\log n)$ slots in the Euclidean SINR model [17].

Somewhat surprisingly, we find that when the set of possible links is restricted, the MST can actually fail quite badly. This holds in every reasonable model of interference.

► **Interference Assumption 1.** *We say that an interference model is reasonable if: a) incident links cannot be scheduled together, while b) sparse instances of equal length links can be scheduled in $O(1)$ slots, where a set of length ℓ links is sparse if any ball of radius ℓ contains $O(1)$ endpoints of those links.*

Every geometrically-defined wireless interference model known satisfies this reasonableness property. In particular, this holds in the protocol and Euclidean SINR models.



■ **Figure 1** The construction from Thm. 1

► **Theorem 1.** *For any n , there is an instance of n nodes embedded on the plane together with a spanning tree that is schedulable in $O(1)$ slots while scheduling the MST requires $\Omega(n^{1/3})$ slots, in every reasonable interference model.*

Proof. Let $k \geq 1$ be a number and $K = 2k^2$. Let $V = \{o\} \cup \{v_{i,j} : i = 0, 1, \dots, k-1, j = 0, 1, \dots, K-1\}$ denote the set of $n = kK + 1 = 2k^3 + 1$ nodes. We position the nodes in the plane using polar coordinates, with the node o as the origin. For node $v_{i,j}$, angular coordinate $r_{i,j}$ is $2\pi \cdot i/k$, while its radial coordinate is $k + j$.

The links are given by $L = O \cup T \cup Y$, where $O = \{(o, v_{i,1}) : i = 0, \dots, k-1\}$, $T = \{(v_{i,j}, v_{i,j+1}) : i = 0, \dots, k-1, j = 0, \dots, K-2\}$, $Y = \{(v_{i,K-1}, v_{i+1 \bmod k, K-1}) : i = 0, \dots, k-1\}$, or the *ordinary*, the *tiny* and the *yuge* links. That is, the link graph is in the form of a wheel, centered at origin, with k spokes, and K nodes on each spoke (see Fig. 1). Ordinary links are incident with the origin, while the yuge links form the tire of the wheel.

We observe that $d(v_{i,K-1}, v_{i+1 \bmod k, K-1}) > k = d(o, v_{i,1})$, for any i, i' . Thus, the MST consists of the ordinary and tiny links, $S \cup L$. Since all the ordinary links have an endpoint in the origin, they must all be scheduled in different slots, implying that the MST requires $k = \Theta(n^{1/3})$ slots. On the other hand, a more efficient solution is to use the set Q , consisting of T, Y and one (arbitrary) link from O . This set Q is a union of three sparse subsets, and therefore can be scheduled in $O(1)$ slots. ◀

This same example shows why the known results for Euclidean SINR do not carry over to general metric spaces (even without missing links). Namely, one could simply form a metric space on the n nodes by shortest-path distances in the link graph.

3 Greedy Algorithm

A natural greedy approach is to find a large feasible subset of edges, assign it a fresh color, contract it, and iterate on the contracted graph. The key step is obtaining a constant-approximation for a maximum feasible subset. A logarithmic approximation then follows from a set cover argument.

We assume in this section that \mathcal{L} can have parallel edges but no loops. We assume that the conflict graph \mathcal{C} is ρ -inductive independent for a number $\rho > 0$, and the corresponding conflict function W and ordering of edges \prec are given. In the maximum feasible forest problem, the goal is to find a maximum cardinality subset of edges of \mathcal{L} , which is both independent in \mathcal{C} and acyclic in \mathcal{L} .

The algorithm, given as Alg. 1, is a greedy Kruskal-like algorithm that mixes the edge selection criteria of wireless capacity algorithms [16, 29] with the classic MST algorithm of

Kruskal, thus the name CAPKRUSKAL. It processes the edges in order of precedence \prec and adds an edge to the forest if: a) the interference on that edge from previously selected edges is small, and b) the edge does not induce a cycle (as per Kruskal). We state it in terms of the classic union-find operations of MAKESET, CONNECTED, and UNION.

Algorithm 1 CAPKRUSKAL(\mathcal{L}, \mathcal{C})

```

1: MAKESET( $v$ ), for each  $v \in V(\mathcal{L})$ 
2:  $S \leftarrow \emptyset$ 
3: for  $e = (u, v)$  in  $L$  in  $\prec$  order do
4:   if  $W(S, e) + W(e, S) \leq 1/2$  and not
      CONNECTED( $u, v$ ) then
5:      $S \leftarrow S \cup \{e\}$ 
6:     UNION( $u, v$ )
7:   end if
8: end for
9: return  $S' = \{e \in S : W(S, e) \leq 1\}$ 

```

Algorithm 2 CONN(\mathcal{L}, \mathcal{C})

```

1:  $i \leftarrow 0$ 
2:  $\mathcal{L}_0 \leftarrow \mathcal{L}$ 
3: while  $\mathcal{L}_i$  has an edge do
4:    $S_i \leftarrow$  CAPKRUSKAL( $\mathcal{L}_i, \mathcal{C}[\mathcal{L}_i]$ )
5:    $\mathcal{L}_{i+1} \leftarrow$  Contract( $\mathcal{L}_i, S_i$ )
6:    $i \leftarrow i + 1$ 
7: end while
8: return  $S_0, S_1, \dots, S_{i-1}$ 

```

Recall that a subset S of edges in \mathcal{L} is feasible if $W(S, e) = \sum_{f \in S} W(f, e) \leq 1$, for all $e \in S$. Define the ordered weight function W^+ as $W^+(e, f) = W(e, f)$ if $e \prec f$, and $W^+(e, f) = 0$, otherwise. Similarly, define W^- as $W^-(e, f) = W(e, f)$ if $f \prec e$, and $W^-(e, f) = 0$, otherwise. Also define the cumulative versions $W^+(S, e)$, $W^-(e, S)$ as before.

We say that a set S is semi-feasible if for each $e \in S$, $W^+(S, e) + W^-(e, S) \leq 1/2$. Namely, the weighted indegree from shorter nodes and to longer nodes is bounded, but the total indegree of e may not be. By an averaging argument, a semi-feasible set I contains a feasible subset of at least half its size. Indeed, using semi-feasibility and sum rearrangements, we have,

$$\sum_{e \in S} W(S, e) = \sum_{e \in S} (W^+(S, e) + W^-(e, S)) \leq \frac{|S|}{2} \quad (1)$$

so for at least half of the links $e \in S$ it holds that $W(S, e) \leq 1$.

► **Theorem 2.** *Let F be a maximum feasible forest of \mathcal{L} . Then CAPKRUSKAL(\mathcal{L}, \mathcal{C}) outputs a feasible forest of size $\Omega(|F|/\rho)$.*

Proof. Let S and S' be the sets computed in CAPKRUSKAL(\mathcal{L}, \mathcal{C}). By definition, S' is feasible. To argue that S' is large, we examine an arbitrary feasible forest, break it into three parts, and show that none of the parts can be too large compared to S' . This will hold, in particular, for the optimal feasible forest. By (1), we can focus on bounding $|S|$, as $|S'| \geq |S|/2$.

Let I be a feasible forest. Observe that the selection condition of the algorithm is equivalent to $W^+(S, e) + W^-(e, S) \leq 1/2$, since the edges are considered in the order of \prec . Let I_R be those edges e in I that failed the degree condition ($W^+(S, e) + W^-(e, S) > 1/2$), and I_T those edges $e = (u, v)$ in I that failed the connectivity condition (CONNECTED(u, v)). The rest, $I_S = I \setminus (I_R \cup I_T)$ are contained in S . We bound these sets in terms of S .

Since I_T contains only edges inside components that S also connects (recalling that I induces a forest), $|I_T| \leq |S|$. Also, clearly $I_S \subseteq I \cap S \subseteq S$, so $|I_S| \leq |S|$. To bound the size of I_R , observe first that by the definition of ρ -inductive independence, $W^-(I_R, f) + W^+(f, I_R) \leq$

ρ , for every edge $f \in S$. This implies that

$$W^-(I_R, S) + W^+(S, I_R) = \sum_{f \in S} [W^-(I_R, f) + W^+(f, I_R)] \leq \rho \cdot |S|.$$

On the other hand, by the selection criteria,

$$W^+(S, I_R) + W^-(I_R, S) = \sum_{e \in I_R} [W^+(S, e) + W^-(e, S)] > \sum_{e \in I_R} \frac{1}{2} = \frac{|I_R|}{2}.$$

Thus, $|I_R| \leq 2\rho \cdot |S|$ and $|I| \leq (2\rho + 2)|S| \leq 4(\rho + 1)|S'|$. \blacktriangleleft

Connectivity Scheduling Algorithm. The algorithm CONN repeatedly calls CAPKRUSKAL to obtain a large independent set of links and assigns it to a new color class. These links are then contracted and the process repeated until we have obtained a spanning tree.

The contraction of an edge is defined in the standard way, except we discard loops. Note that *contraction leaves the conflict graph \mathcal{C} intact*. The operation $\text{Contract}(\mathcal{L}, S)$ contracts all edges in S of a link graph \mathcal{L} and outputs the resulting graph.

The pseudocode of the algorithm is given in Alg. 2. The proof of the following theorem, which is relegated to the full version, follows the classic set cover argument [27].

► **Theorem 3.** CONN terminates in $O(\rho \log n) \cdot \chi$ rounds, where χ is the number of colors needed for coloring an optimum spanning tree.

4 Multicast Tree Schedules

A natural generalization of CONNECTIVITY SCHEDULING is to allow for a set of optional nodes that can be used in the tree construction but need not. Formally, the node set V contains a subset X of terminals and we seek a Steiner tree that spans all the terminals. As before, we ask also for the shortest schedule of the tree links. We refer to this as the STEINER CONNECTIVITY SCHEDULING.

It is not hard to construct examples for which optimal multicast trees are arbitrarily better than trees that use only the terminals, even in a geometric setting. One instance can be obtained from the example of Sec. 2 by restricting the terminals to only the origin and the nodes incident on yuge links.

We give an algorithm for STEINER CONNECTIVITY SCHEDULING with *unweighted* conflict graph \mathcal{C} , and analyse it in terms of a parameter similar to ρ but involving clique covers rather than independence (this setting is also applicable to the SINR model, see Sec. 6). An unweighted graph $\mathcal{C} = (L, E)$ is η -simplicial if there is an ordering \prec of L such that for each link $v \in L$, the subgraph induced by v 's neighbors that are later in the ordering can be covered with η cliques. We refer to neighbors later in the ordering as post-neighbors. As before, in the geometric setting, the ordering is given by link length. Observe that $\rho \leq \eta$, while the best bound in the other direction is $\eta \leq \rho \log n$.

Our algorithm is a reduction to a multi-dimensional version of the Steiner tree (MMST) problem, recently treated by Bilò et al. [3]. In MMST, each edge of the input graph has an associated d -dimensional weight vector, where the weight of edge e along dimension i indicates how much of the i -th *resource* is required by e . The objective is to find a tree that minimizes the ℓ_p -norm of its load vector, where the load vector of a Steiner tree is the sum of the weight vectors of its edges. We use the ℓ_∞ -norm, as we want to minimize the maximum use of a resource. They give a greedy $O(\log d)$ -approximation algorithm for that case.

Given an instance of STEINER CONNECTIVITY SCHEDULING with link graph \mathcal{L} and conflict graph \mathcal{C} , our reduction is as follows. Each link e in \mathcal{L} is itself (or corresponds to) a resource, so there are n (=number of edges) resources. The weight of link f along dimension e is 1 if f is a post-neighbor of e in the conflict graph \mathcal{C} , and 0 otherwise.

Suppose now that the MMST algorithm of [3] returns a tree T with ℓ_∞ -norm Z . Then, the sum of the tree edges along each dimension is at most Z , namely, each link (whether in T or not) has at most Z post-neighbors in T . In particular, $\mathcal{C}[T]$ is Z -inductive, and can then be colored greedily using $Z + 1$ colors.

On the other hand, consider an optimal tree T^* and let Z^* denote the infinity norm of its load vector. From [3], we know that $Z = O(\log n) \cdot Z^*$. Let f be a link with Z^* post-neighbors in T^* , and let N_f be its set of post-neighbors in T^* . By assumption, $\mathcal{C}[N_f]$ can be covered with η cliques, and thus N_f contains a clique of size at least $|N_f|/\eta = Z^*/\eta$. It follows that the length of the schedule of the optimal tree is at least the chromatic number of $\mathcal{C}[N_f]$, which is at least Z^*/η . Thus, our solution yields a $O(\eta \log n)$ -approximation.

► **Theorem 4.** *There is a $O(\eta \log n)$ -approximation algorithm for STEINER CONNECTIVITY SCHEDULING, where conflicts are given by a η -simplicial graph.*

It is a folklore that $\eta \leq 6$ in disk graphs.

► **Corollary 5.** *There is a $O(\log n)$ -approximation algorithm for STEINER CONNECTIVITY SCHEDULING, where conflicts are given by a disk graph.*

5 Hardness of Approximation

It is easy to see that with an arbitrary conflict graph \mathcal{C} , the problem is hard to approximate. For instance, if the link graph \mathcal{L} is already a spanning tree, CONNECTIVITY SCHEDULING becomes simply the classical graph coloring problem (of \mathcal{C}). We show below that the hardness extends to other more restricted settings. These results also show that near-linear dependence on ρ , the inductive independence, is unavoidable.

We first show that hardness holds when \mathcal{C} is the square of the line graph of \mathcal{L} (for general \mathcal{L}), $\mathcal{C} = L^2(\mathcal{L})$, then extend the construction to the case when \mathcal{L} is a complete graph and \mathcal{C} is general (Thm. 7), and to signal strength models (Sec. 6). This corresponds to (bidirectional) 2-hop interferences: two transmission links conflict if they are incident on a common edge. The reduction is from the DISTANCE-2 EDGE COLORING problem in general graphs, also known as STRONG EDGE COLORING: Given a graph \mathcal{L} , find a partition of the edge set into induced matchings, i.e., induced subgraphs where every vertex is of degree 1.

► **Theorem 6.** *The CONNECTIVITY SCHEDULING problem is hard to approximate within a $n^{1-\epsilon}$ -factor, for any $\epsilon > 0$, even when $\mathcal{C} = L^2(\mathcal{L})$.*

Proof. Given an instance of STRONG EDGE COLORING with graph $G' = (V', E')$, we construct an instance of CONNECTIVITY SCHEDULING problem with the graph \mathcal{L} constructed as follows. Consider a bipartite graph $G'' = (V_1, V_2, E)$, as follows. For each vertex v in V' , there are two vertices v_1, v_2 in $V = V_1 \cup V_2$, where $v_i \in V_i$, $i = 1, 2$. If $uv \in E'$ then v_1u_2 and v_2u_1 are in E . Link graph \mathcal{L} is obtained from G'' by taking a complete binary tree with $|V_2|$ leaves and identifying each leaf with a vertex of V_2 . The conflict graph is given by a simple graph \mathcal{C} with vertex set E , where $e_1, e_2 \in E$ are adjacent in \mathcal{C} if and only if they form an induced matching in G' , i.e., there is no edge in G' connecting an endpoint of e_1 to an endpoint of e_2 . This completes the construction.

First, let us show that a strong edge coloring of G' can be used to construct a spanning tree in \mathcal{L} with a similar coloring number. Consider a strong coloring that partitions the edges of G' into c color classes E_1, E_2, \dots, E_c . Each class E_i induces a pair of feasible slots S_i, S'_i in \mathcal{L} , where $S_i = \{v_1u_2 : uv \in E_i\}$ and $S'_i = \{u_1v_2 : uv \in E_i\}$. Indeed, since E_i is an induced matching in G' , each of the slots S_i, S'_i is also an induced matching in \mathcal{L} (and hence independent in \mathcal{C}). Note that the edges in these slots cover all vertices of \mathcal{L} , except for the binary tree. We also add $O(\log n)$ slots to the schedule, two for each layer in the binary tree. The number of slots used then is $O(c + \log n)$. This gives us a connected subgraph of \mathcal{L} that can be scheduled in $O(c + \log n)$ slots.

Next, consider a spanning tree of \mathcal{L} with a corresponding schedule of the edges in slots S_1, S_2, \dots, S_t . Ignoring all edges within the binary tree, we obtain a partition of the edges of the bipartite graph G'' between V_1 and V_2 . We claim that each class corresponds to an induced matching in G' , leading to a strong edge coloring of G' with t colors.

Consider a pair of edges v_1u_2 and w_1x_2 in the same feasible slot. Since they are feasible, there are no edges v_1x_2 nor w_1u_2 in \mathcal{L} , and thus no edges vx nor wu in E . Then vu and wx form an induced matching in G' .

Hence, the optimum number of colors in strong edge coloring of G' is within a constant factor plus a logarithmic term of the optimal number of slots needed for scheduling a spanning tree in \mathcal{L} . Since the former is hard to approximate within $n^{1-\epsilon}$ -factor [5], so is the latter. ◀

► **Theorem 7.** *For general graphs \mathcal{C} , the CONNECTIVITY SCHEDULING problem is hard to approximate within a $n^{1-\epsilon}$ -factor, for any $\epsilon > 0$, even if the link graph \mathcal{L} is complete.*

Proof. We modify the instance of Thm. 6, by adding to \mathcal{L} all edges that were not there and make them adjacent (in \mathcal{C}) to *all* other edges in the graph. If these new edges are used in a spanning tree, they have to be scheduled separately in individual time slots. Thus, using them can only increase the length of any schedule. ◀

6 Implications to Signal Strength Models

We consider in this section the implementation and implication of our results to signal strength models, most importantly metric SINR model.

SINR-feasibility, besides the underlying metric, also depends on the *transmission power control* regime. Different power control methods give different notions of feasibility. Nevertheless, it is known that for most interesting cases, SINR-feasibility has constant-inductive independence property. In particular, power control is usually split into two modes: fixed monotone power schemes, where links use only local information, such as the link length, to define the power level, and global power control, where all power levels are controlled simultaneously to give larger independent sets. The former includes the uniform power mode, where all links use equal power. Another technical issue is *directionality* of links, which is not explicitly addressed by our general results, but will be addressed below.

Let us start the discussion from Euclidean metrics (or more generally doubling metrics). For the global power control mode, [29] introduced a weight function W and proved that with this function, the conflict graph of any set of links is constant-inductive independent (see [29, Thm. 1]), so our results apply here directly (except for directionality issues, addressed below). Similarly, for fixed monotone power schemes (excluding uniform power), [15] showed that in order to get constant-inductive independence, one may take the natural weight function, affectance (also called *relative* or normalized interference) [15, Thm.3.3]. In all cases, the ordering \prec corresponds to a non-decreasing order of links by length.

For general metric spaces, a slightly more technical definition of inductive independence is used, where a fractional conflict graph $\mathcal{C} = (L, W)$ is (ρ, γ) -inductive independent, w.r.t. an ordering \prec of the links, if for every link e and every feasible set $I \in \mathcal{F}$ with $e \prec I$, there is a subset $I' \subseteq I$ of size $|I'| \geq |I|/\gamma$, such that $W(I, e) + W(e, I) \leq \rho$. The old definition corresponds to the setting $\gamma = 1$. It is easily verified that Thms. 2 and 3 extend to cover this new definition, with approximation ratios multiplied by a factor of γ . Now, the counterparts of the results from the previous paragraph in general metrics can be found in [18, Lemmas 2,4] and [30, Thm. 1, Lemma 3], where it is shown that with appropriate weight functions, feasibility for any fixed monotone power scheme (including uniform power), as well as feasibility with global power control, can be expressed by a fractional conflict graph, which is $(O(1), O(1))$ -inductive independent.

The claims above concern settings where the links have fixed directions. In particular, if we apply Thm. 3 to the weighted functions from the previous paragraph, then we should add “there exists a direction of links, such that...” to the claim. This issue is easily resolved for the global power control mode, where the weight function of [29] does not depend on directions. Namely, it gives a schedule, such that whatever direction is assigned to the links, one can find a power assignment that makes it work (the power assignment could be different for different orientations of links).

For oblivious powers, the following trick applies. It is known that for a set of links with some direction and an oblivious power assignment, and with the weight function W defined in terms of the affectances, if $W(e, S) \leq 1/2$ for all $e \in S$ (call this dual-feasibility), then there is another oblivious power assignment (called the *dual* of the original one) that makes S feasible with the reversed directions of links [28]. Thus, we would like to have schedules with slots S being also dual-feasible. To this end, it is enough to modify CAPKRUSKAL, so that the threshold $1/2$ in the acceptance condition is replaced with $1/4$, and the output set S' is given by $S' = \{e \in S : (W(S, e) \leq 1) \wedge (W(e, S) \leq 1/2)\}$. Very similar methods then show that this again gives an $O(\rho)$ -approximation to the maximum feasible forest problem. The rest of the analysis is left intact, so we obtain an $O(\log n)$ -approximation as before, but with schedule slots that are both feasible and dual-feasible. Then we can replace each slot with its two copies and revert the directions of links in one of the copies. Every link thus gets scheduled in both directions, while the schedule length increases by a factor of two.

Summarizing the observations above, we state the following theorem.

► **Theorem 8.** *There is an $O(\log n)$ -approximation to CONNECTIVITY SCHEDULING problem in the SINR model in arbitrary metric spaces. This holds both in the case of fixed monotone power assignments, and for arbitrary power control. It holds even when only a subset of the node-pairs are available as links (but interferences follow the metric SINR definitions).*

These are the first results that hold in general metrics. They are necessarily relative approximations, since in general metric spaces, there is no good upper bound on the connectivity number, even for complete graphs. Two simple examples are the metric induced by the star $K_{1,t}$ with unit-length edges, and the unit metric formed by distances on the unit-length clique metric.

For the case of points in the plane (i.e., a complete link graph with conflicts induced by distances), connectivity can be achieved in $O(\log n)$ slots [17]. Since it is not known if $O(1)$ slots always suffice, this result is not directly implied by Thm. 3. However, it was also shown in [17] that the MST contains a feasible forest of $\Omega(n)$ edges. The rest of our analysis (using constant-inductive independence) then implies a result matching [17].

► **Corollary 9.** *Let P be a set of points in the plane. Then, CONN finds and schedules a spanning tree of P in $O(\log n)$ slots.*

Steiner trees. In the geometric SINR model with a fixed monotone power scheme (with not all links available), we reduce the problem to a graph question as follows. It was observed in [14] that links of the same length class behave approximately like unit-disk graphs, where a length class refers to links whose lengths differ by at most a factor of 2. Namely, there are constants c_1 and c_2 such that for a set S of links of length approximately ℓ , if all links are of mutual distance greater than $c_2\ell$, then they form a feasible set, whereas any pair of links in S of distance at most $c_1\ell$ must be scheduled separately.

We modify the reduction to MMST to that of the graph construction so that weight of link f along dimension e is 1 only if f is a post-neighbor of e in \mathcal{C} and f and e are of the same length class. We then take the resulting tree and schedule the length classes separately, at an extra cost of $O(\log \Lambda)$ (the number of length classes).

► **Corollary 10.** *There is a $O(\log \Lambda \log n)$ -approximation algorithm for STEINER CONNECTIVITY SCHEDULING in the geometric SINR model, under any fixed monotone power scheme.*

Using power control, we can do considerably better. The main result of [19] shows that for any set L of links, there is an unweighted conflict graph $\mathcal{C}(L)$, such that every independent set in \mathcal{C} is feasible, and the chromatic number of \mathcal{C} is at most $O(\log^* \Delta)$ factor away from the optimum schedule length of L (using global power control). Moreover, \mathcal{C} is constant-simplicial [19, Prop. 1].

► **Corollary 11.** *There is a $O(\log n \log^* \Lambda)$ -approximation algorithm for STEINER CONNECTIVITY SCHEDULING in the geometric SINR model with global power control.*

A similar result with $O(\log \log \Lambda)$ -factor holds also for certain monotone power schemes (but not, for instance, uniform power) [20].

Hardness. A special Missing Links variant of the geometric case is where the nodes/links are embedded in the plane and all interferences are either zero or follow the SINR model (with either fixed power or global power control).

► **Theorem 12.** *The geometric Missing Links variant is $n^{1-\epsilon}$ -hard to approximate, for any $\epsilon > 0$. It is also $\Lambda^{2-\epsilon}$ -hard, where Λ is the ratio between the longest to the shortest node distance. This holds even if all unavailable links are missing links.*

Proof. We embed the instance of Thm. 6 in the plane. The nodes of V_1 are located in a unit square in a mesh pattern, $1/\sqrt{n}$ apart in \sqrt{n} columns \sqrt{n} abreast. At a unit distance, a similar unit square holds the nodes of V_2 . The length of an edge in \mathcal{L} (in distance in the plane) is then between 1 and 4.

An induced matching in \mathcal{L} corresponds to a set of links with no mutual interference. On the other hand, a pair of links that are incident on a common edge or share a vertex, will receive interference from each other according to the SINR formula (using the shared edge or each other). Given that distances along available edges vary only by a constant factor, the interference between the links is a constant (specifically, at least $1/4^\alpha$, where α is the “pathloss” constant of the SINR model). Thus, in the setting where the SINR threshold is at least the reciprocal of that constant (i.e., $\beta \geq 4^\alpha$), feasible sets are necessarily induced matchings in \mathcal{L} . We can then conclude by recalling a “signal-strengthening” result [12] that shows that varying the threshold by a constant factor only affects the schedule length by a constant factor.

The longest node distance is at most $\log n$, which is from the root of the binary tree to its leaves, while the shortest distance is $1/\sqrt{n}$. Thus, $\Lambda \leq 4\sqrt{n} \log n$, and $n^{1-\epsilon} \geq \Lambda^{2-\epsilon'}$, for some $\epsilon' \geq \epsilon/3$.

We can restrict the available edges incident to (non-leaf) nodes on the binary tree to the tree edges alone. Thus, non-leaf nodes in the tree must be connected via the tree edges. Then, all unavailable edges are missing edges. ◀

7 Open Issues

Many related problems are left addressing; we list the most prominent ones.

- Latency minimization: Bounding the time it takes for a packet to filter through the tree from a leaf to a root (and back). This requires optimizing both the height of the tree as well as the ordering of the links in the schedule.
- Directed case: Finding an arborescence. This requires new techniques, as our argument crucially depends on the graph being undirected.
- Distributed algorithms: This relates also to the issue of detecting or learning whether a link is usable/reliable or not.

References

- 1 Karhan Akcoglu, James Aspnes, Bhaskar DasGupta, and Ming-Yang Kao. Opportunity cost algorithms for combinatorial auctions. In *Computational Methods in Decision-Making, Economics and Finance*, pages 455–479. Springer, 2002.
- 2 Nouha Baccour, Anis Koubaa, Luca Mottola, Marco Antonio Zuniga, Habib Youssef, Carlo Alberto Boano, and Mario Alves. Radio link quality estimation in wireless sensor networks: a survey. *ACM Trans. Sensor Netw.*, 8(4):34, 2012.
- 3 Vittorio Bilò, Ioannis Caragiannis, Angelo Fanelli, Michele Flammini, and Gianpiero Monaco. Simple greedy algorithms for fundamental multidimensional graph problems. In *44th International Colloquium on Automata, Languages, and Programming, ICALP 2017, July 10-14, 2017, Warsaw, Poland*, pages 125:1–125:13, 2017.
- 4 Marijke H. L. Bodlaender and Magnús M. Halldórsson. Beyond geometry: towards fully realistic wireless models. In *ACM Symposium on Principles of Distributed Computing, PODC '14, Paris, France, July 15-18, 2014*, pages 347–356, 2014.
- 5 Parinya Chalermsook, Bundit Laekhanukit, and Danupon Nanongkai. Graph products revisited: Tight approximation hardness of induced matching, poset dimension and more. In *SODA*, pages 1557–1576. SIAM, 2013.
- 6 Johannes Dams, Martin Hoefer, and Thomas Kesselheim. Scheduling in wireless networks with Rayleigh-fading interference. *IEEE Transactions on Mobile Computing*, 14(7):1503–1514, 2015.
- 7 A. Fanghänel, T. Kesselheim, H. Räcke, and B. Vöcking. Oblivious interference scheduling. In *PODC*, pages 220–229, 2009.
- 8 Martin Fürer and Balaji Raghavachari. Approximating the minimum-degree steiner tree to within one of optimal. *Journal of Algorithms*, 17(3):409–423, 1994.
- 9 Rajiv Gandhi, Magnús M. Halldórsson, Christian Konrad, Guy Kortsarz, and Hoon Oh. Radio aggregation scheduling. In *Algosensors*, pages 169–182, 2015.
- 10 Deepak Ganesan, Bhaskar Krishnamachari, Alec Woo, David Culler, Deborah Estrin, and Stephen Wicker. Complex behavior at scale: An experimental study of low-power wireless sensor networks. Technical report, UCLA/CSD-TR 02, 2002.
- 11 Oliver Göbel, Martin Hoefer, Thomas Kesselheim, Thomas Schleiden, and Berthold Vöcking. Online independent set beyond the worst-case: Secretaries, prophets, and periods. In *ICALP*, pages 508–519, 2014.
- 12 Olga Goussevskaia, Magnús M Halldórsson, and Roger Wattenhofer. Algorithms for wireless capacity. *IEEE/ACM Transactions on Networking (TON)*, 22(3):745–755, 2014.

- 13 P. Gupta and P. R. Kumar. The Capacity of Wireless Networks. *IEEE Trans. Information Theory*, 46(2):388–404, 2000.
- 14 M. M. Halldórsson. Wireless scheduling with power control. *ACM Transactions on Algorithms*, 9(1):7, December 2012.
- 15 Magnús M. Halldórsson, Stephan Holzer, Pradipta Mitra, and Roger Wattenhofer. The power of oblivious wireless power. *SIAM J. Comput.*, 46(3):1062–1086, 2017.
- 16 Magnús M. Halldórsson and Pradipta Mitra. Wireless capacity with oblivious power in general metrics. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2011, San Francisco, California, USA, January 23-25, 2011*, pages 1538–1548, 2011.
- 17 Magnús M. Halldórsson and Pradipta Mitra. Wireless connectivity and capacity. In *Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2012, Kyoto, Japan, January 17-19, 2012*, pages 516–526, 2012.
- 18 Magnús M. Halldórsson and Pradipta Mitra. Nearly optimal bounds for distributed wireless scheduling in the SINR model. *Distributed Computing*, 29(2):77–88, 2016. doi:10.1007/s00446-014-0222-7.
- 19 Magnús M. Halldórsson and Tigran Tonoyan. How well can graphs represent wireless interference? In *STOC*, pages 635–644, 2015.
- 20 Magnús M. Halldórsson and Tigran Tonoyan. The price of local power control in wireless scheduling. In *35th IARCS Annual Conference on Foundation of Software Technology and Theoretical Computer Science, FSTTCS 2015, December 16-18, 2015, Bangalore, India*, pages 529–542, 2015.
- 21 Magnús M. Halldórsson and Tigran Tonoyan. Optimal aggregation throughput is nearly constant. In *MobiHoc*, 2017. Poster abstract.
- 22 Magnús M. Halldórsson and Tigran Tonoyan. Wireless link capacity under shadowing and fading. In *MobiHoc*, pages 27:1–27:10, 2017.
- 23 Magnús M. Halldórsson and Tigran Tonoyan. Optimal aggregation throughput is nearly constant. In *38th IEEE International Conference on Distributed Computing Systems, July 2–5, 2018, Vienna, Austria*, 2018.
- 24 Martin Hoefer and Thomas Kesselheim. Secondary spectrum auctions for symmetric and submodular bidders. *ACM Transactions on Economics and Computation*, 3(2):9, 2015.
- 25 Martin Hoefer, Thomas Kesselheim, and Berthold Vöcking. Approximation algorithms for secondary spectrum auctions. *ACM Transactions on Internet Technology (TOIT)*, 14(2-3):16, 2014.
- 26 Özlem Durmaz Incel, Amitabha Ghosh, and Bhaskar Krishnamachari. Scheduling algorithms for tree-based data collection in wireless sensor networks. In *DCOSS*, pages 407–445. Springer, 2011.
- 27 David S. Johnson. Approximation algorithms for combinatorial problems. *J. Comput. Syst. Sci.*, 9:256–278, 1974. doi:10.1016/S0022-0000(74)80044-9.
- 28 T. Kesselheim and B. Vöcking. Distributed contention resolution in wireless networks. In *DISC*, pages 163–178, August 2010.
- 29 Thomas Kesselheim. A constant-factor approximation for wireless capacity maximization with power control in the SINR model. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2011, San Francisco, California, USA, January 23-25, 2011*, pages 1549–1559, 2011.
- 30 Thomas Kesselheim. Approximation algorithms for wireless link scheduling with flexible data rates. In *ESA*, pages 659–670, 2012.
- 31 Thomas Kesselheim. Dynamic packet scheduling in wireless networks. In *PODC*, pages 281–290, 2012.

- 32 David Kotz, Calvin Newport, Robert S Gray, Jason Liu, Yougu Yuan, and Chip Elliott. Experimental evaluation of wireless simulation assumptions. In *Proceedings of the 7th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems*, pages 78–82. ACM, 2004.
- 33 Fabian Kuhn, Nancy Lynch, Calvin Newport, Rotem Oshman, and Andrea Richa. Broadcasting in unreliable radio networks. In *PODC*, pages 336–345. ACM, 2010.
- 34 Thomas Moscibroda. The worst-case capacity of wireless sensor networks. In *IPSN*, pages 1–10, 2007.
- 35 Thomas Moscibroda and Roger Wattenhofer. The complexity of connectivity in wireless networks. In *INFOCOM 2006. 25th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies, 23-29 April 2006, Barcelona, Catalunya, Spain, 2006*.
- 36 Thomas Moscibroda, Roger Wattenhofer, and Aaron Zollinger. Topology control meets SINR: The scheduling complexity of arbitrary topologies. In *MOBIHOC*, pages 310–321, 2006.
- 37 Yuli Ye and Allan Borodin. Elimination graphs. *ACM Trans. Algorithms*, 8(2):14:1–14:23, 2012.
- 38 Marco Zúñiga Zamalloa and Bhaskar Krishnamachari. An analysis of unreliability and asymmetry in low-power wireless links. *ACM Transactions on Sensor Networks (TOSN)*, 3(2):7, 2007.