

Optimized Scheduled Multiple Access Control for Wireless Sensor Networks

Ioannis Ch. Paschalidis,[†] Wei Lai,[‡] and Xiangdong Song[§]

Abstract—We consider wireless sensor networks with multiple sensor modalities that capture data to be transported over multiple frequency channels to potentially multiple gateways. We study a general problem of maximizing a utility function of achievable transmission rates between communicating nodes. Decisions involve routing, transmission scheduling, power control, and channel selection, while constraints include physical communication constraints, interference constraints, and fairness constraints. Due to its structure the formulation grows exponentially with the size of the network. Drawing upon large-scale decomposition ideas in mathematical programming, we develop a cutting-plane algorithm and show that it terminates in a finite number of iterations. Every iteration requires the solution of a subproblem which is NP-hard. To solve the subproblem we (i) devise a particular relaxation that is solvable in polynomial time and (ii) leverage polynomial-time approximation schemes. A combination of both approaches enables an improved decomposition algorithm which is efficient for solving large problem instances.

Index Terms—Mathematical programming/optimization, multiple frequency channels, routing, transmission scheduling, wireless sensor networks.

I. INTRODUCTION

TRANSMISSION scheduling, which amounts to deciding when a node should send information and to which other node, has attracted much attention in the context of large-scale low-power *Wireless Sensor Networks (WSNETs)* [1], [2]. These networks, operate under stringent resource limitations due to wireless communications and scant energy resources provided by batteries. Commercial wireless sensor networking devices primarily use a random multiple access (MAC) mechanism to transmit information. These devices are capable of switching between multiple frequency channels, but currently this is mostly done for robustness purposes. With the flexibility of operating neighboring sensors in non-overlapping frequencies, and with the use of transmission scheduling rather than a random MAC, we can substantially reduce interference and eliminate packet collisions, thus, achieving much higher radio resource utilization as well as greater energy preservation. This can only benefit a plethora of sensor networking applications, including indoor location detection [3], [4], habitat monitoring [5], surveillance, etc. For surveillance and habitat

monitoring applications in particular, cameras are extremely useful and they start to appear in several WSNET devices [6]. Cameras, though, require higher throughput from the network and accentuate the case for efficient transport.

To utilize the scarce resources (i.e., wireless medium and energy) in the most efficient manner, transmission scheduling can not be seen in isolation but needs to be coordinated with: (i) *channel selection*, that is, determining over which frequency channel a certain transmission between two nodes should take place; (ii) *power control*, which consists of selecting the appropriate power level for a given transmission between two nodes; and (iii) *routing*, that is, determining how should data be routed from a given source to the intended destination. To that end, one must relax the standard (OSI) networking layering architecture as the set of decisions we outlined are typically dealt with in different layers. Making these decisions jointly is often referred to as *cross-layer design* [7], [8].

Scheduling in wireless networks has received a lot of attention. The majority of the literature, starting with the seminal work of [9], has considered a problem where packets arrive to various nodes according to a stochastic process and have to be transported over the network to specific destinations. Every node can queue packets while they are awaiting transmission and the objective is to schedule packet transmissions so as to maximize throughput while maintaining the stability of all queues. In this setting, it is of interest to characterize the maximum throughput region – the set of arrival rate vectors for which the network is stabilized under some scheduling policy – and to devise such stable policies. Subsequent work (see, e.g., [10], [11], [12], [13], [14], [15], [16]) has devised a host of policies that differ in their complexity and the fraction of the maximum throughput region they can achieve. Optimization techniques have also been used extensively in solving networked control problems, e.g., [17]. For a collection of work on applications of optimization in network control we refer to [18].

The problem we consider is different in that we assume there is an infinite supply of packets at given sources that need to be transported over the network to specific destinations. The objective is to maximize throughput or, more generally, functions of the achievable transmission rates. Essentially, we consider a WSNET in a congestion period when tons of data are generated and ask the question how to efficiently “empty” the network. Some earlier works have made attempts towards this direction for different problem settings. For instance, a framework is proposed in [1] to maximize the network lifetime with interference considerations, assuming that the actual transmission rates achieve the Shannon capacity. The

Research partially supported by the NSF under grants EFRI-0735974, DMI-0330171, ECS-0426453 and by the DOE under grant DE-FG52-06NA27490.

[†] Corresponding author. Center for Information & Systems Eng., Dept. of Electrical & Computer Eng., and Division of Systems Eng., Boston University, 15 St. Mary’s St., Brookline, MA 02446, e-mail: yannisp@bu.edu, url: <http://ionia.bu.edu/>.

[‡] Center for Information & Systems Eng., Boston University, e-mail: wei@alum.bu.edu.

[§] Center for Information & Systems Eng., Boston University, e-mail: xsong@bu.edu.

joint routing and power control problem is considered in [19] without frequency diversity.

To solve the joint transmission scheduling, channel selection, power control and routing optimization problem we propose an exact and efficient solution approach inspired by our earlier work in [2], [20], [21]. More specifically, our approach amounts to an iterative algorithm that decomposes the optimization problem into a *master problem* and a *subproblem*. We employ a dual cutting-plane strategy where in every iteration we are required to solve the subproblem which is NP-hard. A similar strategy was also used in [2] but there we assumed a Signal-to-Interference-plus-Noise (SINR)-based physical layer. Here, as well in the preliminary work in [20], [21] we use a protocol-based physical layer according to which interfering transmissions “collide.” Such a model, and our assumptions on transmission achievable rates as well as interference effects between nodes, are in accordance with our experiments with prevailing wireless sensor networking devices. In this paper, we significantly improve upon [20], [21] by developing new strategies for dealing with the complexity of the subproblem. Specifically, to solve the subproblem we (i) devise a new more parsimonious integer linear programming formulation, (ii) develop a particular relaxation that is solvable in polynomial time and (iii) leverage polynomial time approximation schemes. As confirmed by our numerical results, an appropriate combination of these approaches enables an improved decomposition algorithm which is quite efficient for solving large problem instances. The contributions of our work to the existing literature can be summarized as follows.

- (1) Our formulation of the utility maximization problem is close to the one we proposed in our prior work (cf. [2], [20], [21]), and is distinct from those in others’ work in the literature. More specifically, we explicitly enumerate the extreme points (transmission rate vectors) in the utility maximization problem, and search for the optimal transmission scheduling, channel selection, power control and routing policy that achieves the optimal utility. Compared to the existing work in the literature (e.g., [8], [19], [22], [23], [24], [25]), our model does not assume any prespecified routing policy and considers frequency diversity in an integrated framework.
- (2) The interference model we consider in this work is different from what has been considered in the literature (e.g., [26]) and is more practical in wireless sensor networks. Our subproblem is a result of the interference model and has not received much attention in the existing work. To solve the subproblem, we propose a new relaxation, i.e., the matching relaxation, that can be solved in polynomial time, and we leverage the polynomial time approximation scheme. Combining these elements we devise a series of new algorithms.
- (3) We examine and compare the computational efficiency of the proposed algorithms in solving the subproblem. To the best of our knowledge, our work is the first one that reports this comparison.

The rest of the paper is organized as follows. In Sec. II we present the system model and formulate the utility maximiza-

tion problem. Sec. III presents our decomposition algorithm and establishes its convergence. In Sec. IV we develop strategies for generating dual “cuts” that involve solving the subproblem efficiently. These lead to an improved decomposition algorithm. We present illustrative numerical results in Sec. V. Conclusions are in Sec. VI.

Notational Conventions: Throughout the paper all vectors are assumed to be column vectors. We use lower case boldface letters to denote vectors and for economy of space we write $\mathbf{x} = (x_1, \dots, x_R)$ for the column vector \mathbf{x} . \mathbf{x}' denotes the transpose of \mathbf{x} and $\mathbf{0}$ the vector of all zeroes. We use upper case boldface letters to denote matrices. We use script letters to define sets and denote by $\text{Conv}(\mathcal{A})$ the convex hull of a set \mathcal{A} , and by $|\mathcal{A}|$ its cardinality. We denote by $1_{\mathcal{A}}(\mathbf{x})$ the indicator function of $\mathbf{x} \in \mathcal{A}$. When \mathcal{A} is described by a simple condition, say $\mathbf{x} \geq \mathbf{0}$, we simply write $1(\mathbf{x} \geq \mathbf{0})$.

II. NETWORK MODEL AND PROBLEM FORMULATION

We consider a WSNET with N nodes each of which can receive, transmit and relay information with a single port/antenna that it carries. These nodes can be either *sensors* or *data sinks*. We assume that nodes do not multicast information, so each transmission has a specific receiver and is from one node to another. Since they carry a single antenna, nodes cannot receive and transmit simultaneously. Furthermore, receiving nodes cannot receive information from multiple nodes simultaneously. We also assume that there are C frequency channels available for all the nodes in the network. At any point in time each node may only transmit or receive over one of these channels, or, otherwise, remain silent. However, nodes may keep switching between different channels in order to reduce interference.

Sensors in the WSNET collect different types of data depending on the physical system or process they monitor (e.g., temperature, pressure, levels of harmful agents, etc.) and want to relay them to other (sensor or data sinks) nodes. As a result, the WSNET carries multiple types of traffic, differing in information content and utility associated with their successful transmission. A *traffic class* refers to a certain type of traffic with a particular origin and destination; suppose we have K traffic classes. We denote by $s(k)$ and $d(k)$ the source and destination of class k , for $k = 1, \dots, K$. Note that in our model the destinations of the traffic can be either data sinks that collect information, or sensor nodes that need data from other sensors, for instance, to enable in-network processing.

Let p_{ijkc} denote the power used by node i to transmit class k traffic to node j over channel c , for $i, j = 1, \dots, N$, $k = 1, \dots, K$, $c = 1, \dots, C$. We will refer to such a transmission as the (i, j, k, c) transmission. Let us write \mathbf{p} for the N^2KC -dimensional vector of powers and denote by p_{ijkc} its component corresponding to the (i, j, k, c) transmission. For any (i, j, k, c) transmission we assume the transmission power p_{ijkc} may be adjusted (continuously or discretely) between 0 and \bar{p}_i , the maximum power available at node i , and we denote by \mathcal{P} the set of all \mathbf{p} 's that are technically feasible with the available hardware.

We assume that the adjacent-channel interference is negligible due to proper filtering and channel tuning, and thus the

transmissions in different frequency channels do not interfere with each other. On the other hand, concurrent transmissions in the same channel may cause interference to each other and render them unsuccessful. In accordance with our experiments with commercial devices such as the MICA2 motes manufactured by Crossbow Inc., we assume that for any (i, j, k, c) transmission, the transmission is successful if the *Received Signal Strength Indication (RSSI)* at node j is above a certain threshold; the maximum transmission rate is a constant γ if the transmission is successful and 0 otherwise. Apparently, the existing hardware imposes severe limitations (processor clock speed, limitations of the radio chip, speed of the bus, etc.) that prohibit the transmission rate to reach the Shannon capacity of the channel, thus, resulting in a constant maximum transmission rate for all SINR levels above a certain threshold.

For $i = 1, \dots, N$, $0 \leq p \leq \bar{p}_i$ and $c = 1, \dots, C$, let $\mathcal{Q}_t(i, p, c)$ be the set of nodes to which node i can successfully transmit with power level p over channel c when not being interfered by other transmissions. Let also $\mathcal{Q}_l(i, p, c)$ the set of nodes in the neighborhood of the sender i which cannot receive any other transmission over channel c due to the interference caused by node i 's transmission. The nodes in $\mathcal{Q}_l(i, p, c)$ may, however, establish connections to other nodes using different frequency channels if these are interference free. We will be making the following assumption, which is natural for wireless communications. It simply says that higher power allows communicating with more nodes but also causes more interference.

Assumption 1 (Monotonicity) For any node $i = 1, \dots, N$, channel $c = 1, \dots, C$, and power levels $p_1 \leq p_2$ we have $\mathcal{Q}_t(i, p_1, c) \subseteq \mathcal{Q}_t(i, p_2, c)$ and $\mathcal{Q}_l(i, p_1, c) \subseteq \mathcal{Q}_l(i, p_2, c)$.

Denote by r_{ijkc} the transmission rate for the (i, j, k, c) transmission. We assume that the transmitting node i transmits with the maximum possible rate, which is equal to γ when the transmission is successful. Throughout this work we will adopt the convention that $r_{ijkc} = -r_{jikc}$ for any (i, j, k, c) . Thus, if an (i, j, k, c) transmission is in progress we have $r_{ijkc} = \gamma$ and $r_{jikc} = -\gamma$ for $i, j = 1, \dots, N$, $k = 1, \dots, K$, and $c = 1, \dots, C$. We can interpret r_{ijkc} as the *net flow rate* of traffic class k information from node i to node j over channel c . We write \mathbf{r} for the N^2KC -dimensional vector of r_{ijkc} 's and denote by r_{ijkc} its component that corresponds to the net flow rate for an (i, j, k, c) or a (j, i, k, c) transmission.

Let $s_{ijkc} \in \{0, 1\}$ be the indicator of an (i, j, k, c) transmission, namely, $s_{ijkc} = 1$ if the (i, j, k, c) transmission is in progress and 0 otherwise. Let \mathbf{s} be the vector of s_{ijkc} 's. Then

$$r_{ijkc} = \gamma(s_{ijkc} - s_{jikc}), \quad \forall i, j, k, c. \quad (1)$$

Furthermore, the transmission restrictions introduced thus far translate into the following set of conditions

$$s_{ijkc} + s_{uiwl} \leq 1, \quad \forall i, j, k, c, u, w, l, \quad (2)$$

$$s_{ijkc} + s_{ivwl} \leq 1, \quad \forall (j, k, c) \neq (v, w, l), \forall i, \quad (3)$$

$$s_{ijkc} + s_{ujwl} \leq 1, \quad \forall (i, k, c) \neq (u, w, l), \forall j, \quad (4)$$

$$s_{ijkc} = 0, \quad \forall i, k, c, \forall j \notin \mathcal{Q}_t(i, \bar{p}_i, c), \quad (5)$$

$$0 \leq p_{ijkc} \leq \bar{p}_i s_{ijkc}, \quad \forall i, j, k, c, \quad (6)$$

$$s_{ijkc} \in \{0, 1\}, \quad \forall i, j, k, c, \quad (7)$$

$$s_{ijkc} + s_{uvwc} \leq 1, \quad \forall c, \forall (i, j, k) \neq (u, v, w) \text{ such that } v \in \mathcal{Q}_l(i, p_{ijkc}, c). \quad (8)$$

Condition (2) states that nodes cannot transmit and receive simultaneously, condition (3) states that nodes can only transmit traffic of a single class to a single other node over a single channel, condition (4) states that nodes can receive only a single traffic class from a single other node over a single channel, and condition (5) states that node j should be "reachable" from i in order to receive information. Finally, condition (8) states that while i is transmitting no node v in i 's interference set $\mathcal{Q}_l(i, p_{ijkc}, c)$ can be receiving from another node over the same channel c because packets will "collide" at node v . We will be referring to this condition as the *interference constraint*. We denote by \mathcal{S} the set of all (\mathbf{s}, \mathbf{p}) complying with the transmission restrictions (2)–(8). We will refer to a $(\mathbf{s}, \mathbf{p}) \in \mathcal{S}$ as a *valid* transmission scheme; note that it specifies who talks to whom along with the corresponding power level and frequency channel. We will call \mathbf{r} a *valid* transmission rate vector if it is associated with a transmission scheme $(\mathbf{s}, \mathbf{p}) \in \mathcal{S}$, and we let \mathcal{R} be the set of all valid transmission rate vectors. Clearly \mathcal{R} is a finite set, as $r_{ijkc} \in \{0, \gamma, -\gamma\}$, $\forall i, j, k, c$.

Denote by \tilde{L} the cardinality of set \mathcal{R} , and consider the complete collection of transmission rate vectors $\mathbf{r}^1, \dots, \mathbf{r}^{\tilde{L}} \in \mathcal{R}$. A long-term average transmission rate vector of $\mathbf{r} = \sum_{n=1}^{\tilde{L}} \alpha_n \mathbf{r}^n$, where $\alpha_n \geq 0$ for all $n = 1, \dots, \tilde{L}$ and $\sum_{n=1}^{\tilde{L}} \alpha_n = 1$, can be achieved by the WSNET if we use a time-sharing strategy and operate the network a fraction α_n of time according to a transmission scheme associated with \mathbf{r}^n . It follows that any point in the convex hull $\text{Conv}(\mathcal{R})$ of \mathcal{R} is achievable by such a time-sharing strategy. Note also, that $\text{Conv}(\mathcal{R})$ is a polytope (i.e., bounded polyhedron) as the convex hull of the finite set \mathcal{R} .

Over the long run, the WSNET should obey flow conservation laws, i.e., the traffic of each class should not accumulate in any node other than its destination. Hence,

$$\sum_{j=1}^N \sum_{c=1}^C r_{ijkc} = 0, \quad \forall i \neq s(k), d(k), \forall k,$$

that is, class k traffic flow into i equals class k traffic outflow from node i over all the channels.

We seek to maximize the overall utility of transmissions in the WSNET, expressed as a function $F(\mathbf{r})$ of the long-term average transmission rate vector \mathbf{r} . We assume that $F(\mathbf{r})$ is continuous, concave, and bounded in $\text{Conv}(\mathcal{R})$. In most practical scenarios, $F(\mathbf{r})$ takes the form of some function $\tilde{F}(\tilde{\mathbf{r}})$ where $\tilde{\mathbf{r}}$ is the vector of $\tilde{r}_{ijk} = \sum_{c=1}^C r_{ijkc}$, for all i, j, k . Nonetheless, in the sequel we will work with the more general function $F(\mathbf{r})$; clearly a function $\tilde{F}(\tilde{\mathbf{r}})$ can be handled as a special case. Note that by considering system utility, we cover a large variety of objectives studied in the literature, including weighted throughput which is a linear function of \mathbf{r} . Moreover, $F(\mathbf{r})$ needs not to be a sum of individual utilities associated with each traffic class. Rather, it can represent quite general performance metrics of interest that model interdependent behavior of the various sensors, e.g.,

when, for instance, clusters of sensors collaborate towards a common goal.

We are interested in utility maximization subject to fairness constraints. We model fairness considerations as a set of R linear inequalities $\mathbf{A}\mathbf{r} \leq \mathbf{b}$, where $\mathbf{A} \in \mathbb{R}^{R \times N^2 KC}$ and $\mathbf{b} \in \mathbb{R}^R$ are given. For example, these constraints can impose equality among all transmission rates, or force the destinations of various classes of traffic to never send out any traffic of their own classes. Let \mathcal{F} be the set of rates that satisfy fairness constraints and flow conservation, i.e.,

$$\mathcal{F} \triangleq \left\{ \mathbf{r} \mid \mathbf{A}\mathbf{r} \leq \mathbf{b}, \sum_{j=1}^N \sum_{c=1}^C r_{ijkc} = 0, \forall i \neq s(k), d(k), \forall k \right\},$$

and to exclude trivial cases assume $\text{Conv}(\mathcal{R}) \cap \mathcal{F} \neq \emptyset$.

We can formulate the utility optimization problem as

$$\begin{aligned} \max \quad & F(\mathbf{r}) \\ \text{s.t.} \quad & \mathbf{r} \in \text{Conv}(\mathcal{R}) \cap \mathcal{F}, \end{aligned} \quad (9)$$

which is a convex optimization problem. An important observation is that we seek to maximize utility over the convex hull of \mathcal{R} rather than \mathcal{R} itself (as for example in earlier work, e.g., [19], [26]). This is bound to yield higher system utility and as we have seen the WSNET operates by time-sharing among different transmission schemes.

Let $\mathbf{r}^1, \dots, \mathbf{r}^L$ denote the extreme points of $\text{Conv}(\mathcal{R})$. Any $\mathbf{r} \in \text{Conv}(\mathcal{R})$ can be expressed as a convex combination of those. Incorporating the definition of \mathcal{F} , (9) becomes

$$\begin{aligned} \min \quad & -F(\mathbf{r}) \\ \text{s.t.} \quad & \mathbf{r} - \sum_{n=1}^L \alpha_n \mathbf{r}^n = 0, \\ & \sum_{n=1}^L \alpha_n = 1, \\ & \mathbf{A}\mathbf{r} \leq \mathbf{b}, \\ & \sum_{j=1}^N \sum_{c=1}^C r_{ijkc} = 0, \forall i \neq s(k), d(k), \forall k, \\ & \alpha_n \geq 0, n = 1, \dots, L. \end{aligned} \quad (10)$$

The problem above maximizes a concave function over a polyhedron. It can be solved using, for example, the *conditional gradient method*. If $F(\mathbf{r})$ is linear, then it is a linear programming problem.

The challenge with problem (10) is that $\text{Conv}(\mathcal{R})$ can have a humongous number of extreme points – in general a number which grows exponentially with the size N of the network. One could simply enumerate all extreme points but this approach can only solve very small instances (on the order of 5-6 nodes). Instead, we will develop a decomposition algorithm that will enable us to solve sizable instances. This algorithm does not need to know $\mathbf{r}^1, \dots, \mathbf{r}^L$ in advance. It generates them as needed and identifies the ones that should be used to achieve optimality.

III. A DECOMPOSITION METHOD

In this section we propose a decomposition method for solving (10). For linear utilities the method is a *column generation* method for solving large-scale linear programming problems. To handle the nonlinear objective we present it as a *cutting plane* method for the dual problem.

To develop the decomposition approach consider the problem (10), to which we will be referring as the *master problem*. Let $(\boldsymbol{\lambda}, \mu, \boldsymbol{\sigma}, \boldsymbol{\nu})$ be the dual vectors. The dual function $G(\boldsymbol{\lambda}, \mu, \boldsymbol{\nu}, \boldsymbol{\sigma})$ is given by

$$\begin{aligned} G(\boldsymbol{\lambda}, \mu, \boldsymbol{\nu}, \boldsymbol{\sigma}) &= \inf_{\boldsymbol{\alpha} \geq \mathbf{0}, \mathbf{r}} \left\{ -F(\mathbf{r}) + \boldsymbol{\lambda}'(\mathbf{r} - \sum_n \alpha_n \mathbf{r}^n) \right. \\ &\quad + \mu(\sum_n \alpha_n - 1) + \boldsymbol{\sigma}'(\mathbf{A}\mathbf{r} - \mathbf{b}) \\ &\quad \left. + \sum_k \sum_{i \neq s(k), d(k)} \nu_{ik} \sum_j \sum_c r_{ijkc} \right\} \\ &= G_1(\boldsymbol{\lambda}, \boldsymbol{\nu}, \boldsymbol{\sigma}) + G_2(\boldsymbol{\lambda}, \mu) - \mu - \boldsymbol{\sigma}'\mathbf{b}, \end{aligned}$$

where

$$\begin{aligned} G_1(\boldsymbol{\lambda}, \boldsymbol{\nu}, \boldsymbol{\sigma}) &= \inf_{\mathbf{r}} \left\{ -F(\mathbf{r}) + (\boldsymbol{\lambda}' + \boldsymbol{\sigma}'\mathbf{A})\mathbf{r} \right. \\ &\quad \left. + \sum_k \sum_{i \neq s(k), d(k)} \nu_{ik} \sum_j \sum_c r_{ijkc} \right\}, \\ G_2(\boldsymbol{\lambda}, \mu) &= \inf_{\boldsymbol{\alpha} \geq \mathbf{0}} \sum_n (\mu - \boldsymbol{\lambda}'\mathbf{r}^n) \alpha_n. \end{aligned}$$

Let

$$\begin{aligned} \mathcal{D}_1 &= \{(\boldsymbol{\lambda}, \boldsymbol{\nu}, \boldsymbol{\sigma}) \mid G_1(\boldsymbol{\lambda}, \boldsymbol{\nu}, \boldsymbol{\sigma}) > -\infty\} \\ \mathcal{D}_2 &= \{(\boldsymbol{\lambda}, \mu) \mid G_2(\boldsymbol{\lambda}, \mu) > -\infty\} \end{aligned}$$

and note that

$$\mathcal{D}_2 = \{(\boldsymbol{\lambda}, \mu) \mid \mu - \boldsymbol{\lambda}'\mathbf{r}^n \geq 0, n = 1, \dots, L\},$$

$$G_2(\boldsymbol{\lambda}, \mu) = \begin{cases} 0, & \text{if } (\boldsymbol{\lambda}, \mu) \in \mathcal{D}_2, \\ -\infty, & \text{otherwise,} \end{cases}$$

and \mathcal{D}_1 is independent of $\mathbf{r}^1, \dots, \mathbf{r}^L$. Then the dual of the master problem (10) is

$$\begin{aligned} \max \quad & G_1(\boldsymbol{\lambda}, \boldsymbol{\nu}, \boldsymbol{\sigma}) - \mu - \boldsymbol{\sigma}'\mathbf{b} \\ \text{s.t.} \quad & (\boldsymbol{\lambda}, \boldsymbol{\nu}, \boldsymbol{\sigma}) \in \mathcal{D}_1, \\ & \mu - \boldsymbol{\lambda}'\mathbf{r}^n \geq 0, \quad n = 1, \dots, L, \\ & \boldsymbol{\sigma} \geq \mathbf{0}. \end{aligned} \quad (11)$$

Since the master problem is a convex optimization problem there is no duality gap.

Suppose now we have an extreme point of $\text{Conv}(\mathcal{R})$, say \mathbf{r}^1 , which belongs to \mathcal{F} . Let $m \in \{1, \dots, L\}$, and consider

$$\begin{aligned} \min \quad & -F(\mathbf{r}) \\ \text{s.t.} \quad & \mathbf{r} - \sum_{n=1}^m \alpha_n \mathbf{r}^n = 0, \\ & \sum_{n=1}^m \alpha_n = 1, \\ & \mathbf{A}\mathbf{r} \leq \mathbf{b}, \\ & \sum_{j=1}^N \sum_{c=1}^C r_{ijkc} = 0, \forall i \neq s(k), d(k), \forall k, \\ & \alpha_n \geq 0, n = 1, \dots, m, \end{aligned} \quad (12)$$

which we call the *restricted master problem* at the m th iteration. Suppose we solve this problem to optimality; its dual is identical to (11) with the exception that only constraints $\mu - \boldsymbol{\lambda}'\mathbf{r}^n \geq 0$, for $n = 1, \dots, m$, appear. We refer to this latter problem as the *restricted dual problem* at the m th iteration. Let $(\mathbf{r}^{(m)}, \boldsymbol{\alpha}^{(m)}; \boldsymbol{\lambda}^{(m)}, \mu^{(m)}, \boldsymbol{\nu}^{(m)}, \boldsymbol{\sigma}^{(m)})$ be an optimal primal-dual pair for the restricted master problem. The dual variables are dual feasible and satisfy $(\boldsymbol{\lambda}^{(m)}, \boldsymbol{\nu}^{(m)}, \boldsymbol{\sigma}^{(m)}) \in \mathcal{D}_1$, $\boldsymbol{\sigma}^{(m)} \geq \mathbf{0}$, and $\mu^{(m)} - \boldsymbol{\lambda}^{(m)'}\mathbf{r}^n \geq 0$, for all $n = 1, \dots, m$. If $\mu^{(m)} - \boldsymbol{\lambda}^{(m)'}\mathbf{r}^n \geq 0$ for all $n = 1, \dots, L$ then we have

a primal-dual pair for (10) and we are done. Otherwise, we need to generate an extreme point, say \mathbf{r}^{m+1} , of $\text{Conv}(\mathcal{R})$ that violates dual feasibility, solve the $m + 1$ st restricted master problem, and continue iterating in this fashion. We next examine how to produce “cuts” in the dual, i.e., how to generate an extreme point that violates dual feasibility.

A. The subproblem

At the m th iteration we seek an extreme point \mathbf{r}^{m+1} of $\text{Conv}(\mathcal{R})$ satisfying $\mu^{(m)} - \lambda^{(m)'} \mathbf{r}^{m+1} < 0$. As the extreme points of $\text{Conv}(\mathcal{R})$ are also in \mathcal{R} , we might as well generate a point \mathbf{r} that minimizes $\mu^{(m)} - \lambda^{(m)'} \mathbf{r}$ over \mathcal{R} . This suggests the subproblem

$$\begin{aligned} \max \quad & \lambda' \mathbf{r} \\ \text{s.t.} \quad & \mathbf{r} \in \mathcal{R}, \end{aligned} \quad (13)$$

with cost vector $\lambda = \lambda^{(m)}$.

We now establish some properties of (13). $\lambda \in \mathbb{R}^{N^2KC}$ is the dual vector corresponding to the first constraint of (10). Denote by λ_{ijkc} the element of λ corresponding to r_{ijkc} and let $\tilde{\pi}_{ijkc} = \gamma(\lambda_{ijkc} - \lambda_{jikc})$. Recall that s_{ijkc} is the indicator of the (i, j, k, c) transmission. Then

$$\begin{aligned} \lambda' \mathbf{r} &= \sum_{k=1}^K \sum_{i=1}^N \sum_{j=1}^N \sum_{c=1}^C \gamma \lambda_{ijkc} (s_{ijkc} - s_{jikc}) \\ &= \sum_{k=1}^K \sum_{i=1}^N \sum_{j=1}^N \sum_{c=1}^C \tilde{\pi}_{ijkc} s_{ijkc}, \end{aligned}$$

and consequently problem (13) is equivalent to

$$\begin{aligned} \max \quad & \sum_{k=1}^K \sum_{i=1}^N \sum_{j=1}^N \sum_{c=1}^C \tilde{\pi}_{ijkc} s_{ijkc} \\ \text{s.t.} \quad & (\mathbf{s}, \mathbf{p}) \in \mathcal{S}, \end{aligned} \quad (14)$$

where \mathcal{S} is the set of transmission schemes (\mathbf{s}, \mathbf{p}) satisfying constraints (2)–(8). For all i, j, k, c define

$$\pi_{ijkc} \triangleq \begin{cases} \tilde{\pi}_{ijkc}, & \text{if } j \in \mathcal{Q}_t(i, \bar{p}_i, c), \\ 0, & \text{otherwise,} \end{cases}$$

$\hat{p}_{ijkc} \triangleq \min\{p \mid 0 \leq p \leq \bar{p}_i, j \in \mathcal{Q}_t(i, p, c)\}$, and let $\hat{\mathbf{p}}$ denote the vector of \hat{p}_{ijkc} 's. Moreover, for all $i, j = 1, \dots, N$ and $c = 1, \dots, C$ let

$$\omega_{ijc} = \max_{k=1, \dots, K} \pi_{ijkc}$$

and construct a set \mathcal{H} as follows: for each $1 \leq i, j \leq N$ and channel $c = 1, \dots, C$ select only one k such that $k = \arg\max_{1 \leq t \leq K} \pi_{ijt c}$, and let (i, j, k, c) be an element of \mathcal{H} .

Remark : In practice, the minimum power level \hat{p}_{ijkc} depends on the locations of nodes i and j as well as the channel c , but is not dependent on the traffic class k . As such, with a slight abuse of notation we will denote by \hat{p}_{ijc} the minimum power needed for node i to reach node j over channel c .

The next theorem shows that solving (14) amounts to solving an *Integer Linear Programming* (ILP) problem; the proof is in Appendix VII-A.

Theorem III.1 Suppose \mathbf{x}^* is an optimal solution to the ILP

$$\begin{aligned} \max \quad & \sum_{(i,j,c) \in \mathcal{H}} \omega_{ijc} x_{ijc} \\ \text{s.t.} \quad & \sum_{j=1}^N \sum_{c=1}^C x_{ijc} + \sum_{j=1}^N \sum_{c=1}^C x_{jic} \leq 1, \quad \forall i, \\ & x_{ijc} + \sum_{u=1}^N x_{uvc} \leq 1, \quad \forall c, \text{ such that} \\ & \quad \quad \quad \forall v \neq j, v \in \mathcal{Q}_l(i, \hat{p}_{ijc}, c), \\ & x_{ijc} \in \{0, 1\}, \quad \forall i, j, c. \end{aligned} \quad (15)$$

Let

$$s_{ijkc}^* = 1_{\mathcal{H}}(i, j, k, c) x_{ijc}^*, \quad \forall i, j, k, c, \quad (16)$$

and set $s_{ijkc}^* = 0$ whenever $\pi_{ijkc} = 0$. Define $p_{ijkc}^* = \hat{p}_{ijc} s_{ijkc}^*$ and let \mathbf{p}^* be the vector of the p_{ijkc}^* 's. Then $(\mathbf{s}^*, \mathbf{p}^*)$ is an optimal solution to problem (14).

We summarize the discussion on the subproblem as follows: to compute an optimal solution \mathbf{r}^* of (13) we first follow the procedure described in Thm. III.1 to obtain an \mathbf{s}^* , then compute \mathbf{r}^* as in (1). To achieve this \mathbf{r}^* we operate the WSNET as follows: (i, j, k, c) transmissions occur only if $s_{ijkc}^* = 1$ and if so at the minimum power \hat{p}_{ijc} needed for node i to reach node j over channel c .

B. The decomposition algorithm

We now have all the ingredients to present the decomposition algorithm and establish its convergence. The algorithm is in Fig. 1; we assume that (10) is feasible and we will discuss at the end of this Section how this assumption can be relaxed.

-
- 1) **Initialization:** Let $\mathbf{r}^1 \in \text{Conv}(\mathcal{R}) \cap \mathcal{F}$ and set $m = 1$.
 - 2) **m -th iteration:**
 - a) Solve the restricted master problem (12) with $\mathbf{r}^1, \dots, \mathbf{r}^m$ to obtain an optimal primal-dual pair $(\mathbf{r}^{(m)}, \boldsymbol{\alpha}^{(m)}; \boldsymbol{\lambda}^{(m)}, \mu^{(m)}, \boldsymbol{\nu}^{(m)}, \boldsymbol{\sigma}^{(m)})$.
 - b) Solve the subproblem (13) with cost vector $\boldsymbol{\lambda}^{(m)}$ as outlined in Section III-A. Let \mathbf{r}^{m+1} be the optimal solution obtained.
 - c) If $\mu^{(m)} - \boldsymbol{\lambda}^{(m)'} \mathbf{r}^{m+1} \geq 0$ stop; $(\mathbf{r}^{(m)}, \boldsymbol{\alpha}^{(m)})$ is an optimal solution of (10). Otherwise, set $m := m + 1$ and go to step 2a.
-

Fig. 1. The decomposition algorithm.

The following theorem, whose proof is in Appendix VII-B, establishes the convergence of the algorithm. The key idea is that at every iteration we generate a new transmission rate vector and there is a finite number of those.

Theorem III.2 Assume that (10) is feasible. Then the decomposition algorithm of Fig. 1 terminates with an optimal solution of (10) in a finite number of iterations.

C. Initialization

We conclude this section by outlining how to initialize the algorithm of Fig. 1. We require an initial vector $\mathbf{r}^1 \in \text{Conv}(\mathcal{R}) \cap \mathcal{F}$. In many cases of practical interest $\mathbf{r}^1 = \mathbf{0}$ would be feasible, which is the case when $\mathbf{b} \geq \mathbf{0}$. This includes $\mathbf{b} = \mathbf{0}$ which can be interpreted to mean that fairness is *relative*. Arguably, this covers the majority of practical cases. If $\mathbf{b} \neq \mathbf{0}$, then it might still be possible to reformulate the fairness constraints so that $\mathbf{b} \geq \mathbf{0}$. Otherwise, some extra work needs to be done to discover an initial feasible solution. To this end, consider the following *auxiliary master problem*

$$\begin{aligned} \text{s.t. } \quad & \mathbf{r} - \sum_{n=1}^L \alpha_n \mathbf{r}^n = \mathbf{0}, \\ & \sum_{n=1}^L \alpha_n = 1, \\ & \mathbf{A}\mathbf{r} + \mathbf{y} = \mathbf{b}, \\ & \sum_{j=1}^N \sum_{c=1}^C r_{ijkc} = 0, \quad \forall i \neq s(k), d(k), \forall k, \\ & \alpha_n \geq 0, \quad n = 1, \dots, L, \end{aligned} \quad (17)$$

where we introduce the vector of auxiliary variables \mathbf{y} . This problem can be solved using a similar decomposition algorithm as in Fig. 1. We start with $m = 1$, $\mathbf{r}^1 = \mathbf{0}$, and note that $\mathbf{r} = \mathbf{0}$, $\alpha_1 = 1$, $\mathbf{y} = \mathbf{b}$ form a feasible solution. The dual of (17) is almost identical to (11) with a modified definition of $G_1(\boldsymbol{\lambda}, \boldsymbol{\nu}, \boldsymbol{\sigma})$. The subproblem remains the same as before and the decomposition approach applies. If the optimal solution of (17) satisfies $\mathbf{y} \geq \mathbf{0}$ then we are done as we have a feasible solution of (10) to initialize the algorithm in Fig. 1 (this could involve time-sharing between several transmission schemes). Otherwise, (10) is infeasible.

IV. DECOMPOSITION ALGORITHM REVISITED

The efficiency of the algorithm of Fig. 1 critically depends on how efficiently we can solve the subproblem. As outlined in Section III-A, solving the subproblem amounts to solving an ILP. General ILPs are hard to solve (they are NP-complete); solvers invariably use branch-and-bound methods which, depending on the problem and its size, can take a long time. In this section we devise a relaxation of (15) and leverage *Polynomial Time Approximation Schemes (PTAS)* to arrive at more efficient algorithms for solving problem (10) based on the algorithm of Fig. 1.

A. A relaxation based on maximum weighted matching

We first relax the interference constraints from problem (15). The resulting problem is

$$\begin{aligned} \max \quad & \sum_{(i,j,c)} \omega_{ijc} x_{ijc} \\ \text{s.t. } \quad & \sum_{j=1}^N \sum_{c=1}^C x_{ijc} + \sum_{j=1}^N \sum_{c=1}^C x_{jic} \leq 1, \quad \forall i, \\ & x_{ijc} \in \{0, 1\}, \quad \forall i, j, c. \end{aligned} \quad (18)$$

In this section we show that it is equivalent to a *Maximum Weighted Matching (MWM) problem* and can actually be solved in polynomial time. Let us call problem (18) the *MWM relaxation* of problem (15). We will next proceed to solve problem (18).

Let $\mathcal{V} = \{1, \dots, N\}$ where each element of \mathcal{V} corresponds to a sensor node of the WSNET. Consider the undirected graph

$\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where \mathcal{E} is the complete set of edges between nodes in \mathcal{V} . To each edge $(i, j) \in \mathcal{E}$ we associate a weight $\hat{\omega}_{ij}$ such that

$$\hat{\omega}_{ij} = \max_{1 \leq c \leq C} \max\{\omega_{ijc}, \omega_{jic}\}, \quad \forall i, j \in \mathcal{V}. \quad (19)$$

Note that $\hat{\omega}_{ij} = \hat{\omega}_{ji} \geq 0$, $\forall i, j$. Let us also construct a set \mathcal{C} as follows: for each $1 \leq i, j \leq N$, we select only one c satisfying

$$c = \text{argmax}_{1 \leq q \leq C} \max\{\omega_{ijq}, \omega_{jiq}\},$$

and let (i, j, c) be an element of \mathcal{C} . The next theorem establishes that solving the MWM relaxation of the subproblem amounts to solving a maximum weighted matching problem for graph \mathcal{G} where the edge weights are given in (19). The proof is in Appendix VII-C.

Theorem IV.1 *Suppose \mathbf{y}^* is an optimal solution to the maximum weighted matching problem*

$$\begin{aligned} \max \quad & \sum_{(i,j) \in \mathcal{E}} \hat{\omega}_{ij} y_{ij} \\ \text{s.t. } \quad & \sum_{j|(i,j) \in \mathcal{E}} y_{ij} \leq 1, \quad \forall i \\ & y_{ij} = y_{ji}, \quad \forall i, j, \\ & y_{ij} \in \{0, 1\}, \quad \forall i, j. \end{aligned} \quad (20)$$

Then, an optimal solution \mathbf{x}^ to problem (18) satisfies*

$$x_{ijc}^* = 1_{\mathcal{C}}(i, j, c) y_{ij}^*, \quad \forall i, j, c.$$

Remark : It should be noted that (20) is always feasible ($\mathbf{y} = \mathbf{0}$ is a feasible solution), thus, it is always possible to obtain an optimal solution to problem (18).

The maximum weighted matching problem is a well studied problem in graph theory. Many algorithms and heuristics for different matching variants have been proposed and it has been shown that (20) can be solved in $O(|\mathcal{V}|^3)$ amount of time [27], that is, polynomial in the size of the input. In our case, $|\mathcal{V}| = N$ and it takes $O(CN^2)$ additional time to calculate the weights and obtain \mathbf{x}^* from \mathbf{y}^* , thus, the complexity of problem (18) is $O(CN^2 + N^3)$.

B. MWM-based cuts

The polynomial solvability of the MWM relaxation of the subproblem (15) enables a more efficient algorithm based on the decomposition algorithm shown in Fig. 1. More specifically, for each iteration of the decomposition algorithm we seek a transmission rate vector \mathbf{r} that violates the dual feasibility (a ‘‘cut’’ of the dual feasible set); we either supply \mathbf{r} to the restricted master problem, or declare success when such an \mathbf{r} does not exist. We can produce cuts by solving the MWM relaxation, thus, avoiding solving the subproblem to optimality at each iteration. The proposed algorithm, to be referred to as MWM, is shown in Fig. 2. Note that the algorithm employs the MWM relaxation to produce a (potentially infeasible) solution to the subproblem and then uses a greedy approach to ‘‘prune’’ that solution and make it feasible. Specifically, we first attempt (cf. Step 2) to switch high-valued interfering transmissions to interference-free channels and then we silence (cf. Step 4) low-valued transmissions that cause interference.

-
- 1) **Initialization:** Let \mathbf{y} be an optimal solution to problem (18) obtained as described in Thm. IV.1. Set \mathbf{r} according to (1) and (16), and let $\mathcal{A} = \{(i, j, k, c) \mid s_{ijkc} = 1\}$.
 - 2) **Iteration (channel selection):** Select an $(i, j, k, c) \in \mathcal{A}$ with the largest weight π_{ijkc} . Check all interference constraints (8) to which s_{ijkc} participates.

If these constraints are not satisfied:

 - a) If there exists another channel \hat{c} so that switching the (i, j, k, c) transmission to \hat{c} (i.e., setting $s_{ijkc} = 0$ and $s_{ijk\hat{c}} = 1$) satisfies the interference constraints then switch the transmission to \hat{c} and adjust \mathbf{r} accordingly.
 - b) Else skip.

Remove (i, j, k, c) from \mathcal{A} . If $\mathcal{A} = \emptyset$ break; else go to the beginning of 2).
 - 3) Let $\mathcal{A} = \{(i, j, k, c) \mid s_{ijkc} = 1\}$.
 - 4) **Iteration (collision avoidance):** Select an $(i, j, k, c) \in \mathcal{A}$ with the smallest weight π_{ijkc} . Check all interference constraints (8) to which s_{ijkc} participates. If these constraints are not satisfied set $s_{ijkc} = 0$ and adjust \mathbf{r} accordingly. Remove (i, j, k, c) from \mathcal{A} . If $\mathcal{A} = \emptyset$ exit and output \mathbf{r} ; else go to the beginning of 4).
-

Fig. 2. MWM algorithm that constructs a feasible solution to the subproblem (13).

Note that the transmission rate vector \mathbf{r} obtained from the algorithm in Fig. 2 is a feasible solution to the subproblem (13). Let $\lambda^{(m)}$ and $\mu^{(m)}$ be the optimal dual variables corresponding to the first and second constraints of (12) at the m -th iteration, respectively. Then if $\mu^{(m)} < \lambda^{(m)'} \mathbf{r}$, we have generated a cut in the dual feasible set, we can set $\mathbf{r}^{(m+1)} = \mathbf{r}$, and continue the iteration in the decomposition algorithm.

However, we cannot terminate the algorithm based on \mathbf{r} . In particular, let \mathbf{r}^* be the optimal solution to the subproblem (13). Clearly $\lambda^{(m)'} \mathbf{r}^* \geq \lambda^{(m)'} \mathbf{r}$. Therefore, even if $\mu^{(m)} < \lambda^{(m)'} \mathbf{r}^*$ we may have $\mu^{(m)} \geq \lambda^{(m)'} \mathbf{r}$, and thus the stopping criterion is not met. In this case we have to solve the subproblem to optimality (by solving (15)) in order to check the stopping criterion.

C. PTAS-based cuts

Next we seek solutions to problem (15) via an alternative route – PTAS (Polynomial Time Approximation Scheme) for geometric graphs. Related approaches have been proposed in the literature to solve other NP-hard problems, e.g., maximum independent set [28] and maximum weighted matching problem with K -hop interference constraints [26]. We combine this approach with MWM-based cuts to solve the problem more efficiently. Furthermore, we propose a randomized approximation scheme.

Denote by \mathcal{L} the two-dimensional region in which all the sensor nodes reside and the position of each sensor node is

determined by its geographical location.¹ Similar with [28], [26], let us first specify an error level ϵ for the PTAS and find the smallest Θ such that $(\frac{\Theta}{\Theta+1})^2 \geq 1 - \epsilon$.

We divide the region \mathcal{L} into horizontal strips and from top to bottom label the strips as $\mathcal{B}_1, \dots, \mathcal{B}_{M_0}$. The region \mathcal{L} is divided into as many as possible horizontal strips so that the interference caused by any transmitting node in \mathcal{B}_m does not affect any node (in any channel) in a strip $\mathcal{B}_{m'}$ if strip \mathcal{B}_m is not neighboring to $\mathcal{B}_{m'}$, $\forall m, m' \in \{1, \dots, M_0\}$. Now, for a given integer $\theta_0 \in \{0, \dots, \Theta\}$, let us remove from region \mathcal{L} every strip whose label number $m \in \{1, \dots, M_0\}$ satisfies the condition: $m \bmod (\Theta + 1) = \theta_0$. Note that by removing a strip we delete all nodes inside the strip. Next, we organize the strips remaining into subregions by putting in the same subregion strips that are neighboring each other. Suppose that this yields a number $T(\theta_0)$ of non-neighboring subregions, each of width $O(\Theta)$; we denote them by $\mathcal{B}_{(\theta_0,1)}, \dots, \mathcal{B}_{(\theta_0,T(\theta_0))}$. Next, consider an arbitrary subregion $\mathcal{B}_{(\theta_0,t)}$ ($1 \leq t \leq T(\theta_0)$) and partition it into as many as possible vertical strips so that the interference caused by any transmitting node in one strip does not affect any node (in any channel) in a non-neighboring strip. In this fashion we generate a number of rectangles which from left to right we label as $\mathcal{B}_{(\theta_0,t)}^1, \dots, \mathcal{B}_{(\theta_0,t)}^{M_t}$. For some fixed integer $\theta_t \in \{0, \dots, \Theta\}$ we remove from this collection of rectangles the ones with label number $m \in \{1, \dots, M_t\}$ satisfying the condition: $m \bmod (\Theta + 1) = \theta_t$. As before, we organize the remaining rectangles into rectangular subregions by putting in the same subregion rectangles that are neighboring each other. Suppose that this yields a number $\hat{T}(\theta_0, \theta_t)$ of non-neighboring rectangular subregions, each of width $O(\Theta^2)$; we denote them by $\mathcal{B}_{(\theta_0,t)}^{(\theta_t,1)}, \dots, \mathcal{B}_{(\theta_0,t)}^{(\theta_t,\hat{T}(\theta_0,\theta_t))}$. We repeat this process for every horizontal subregion $\mathcal{B}_{(\theta_0,t)}$, $t = 1, \dots, T(\theta_0)$, thus, partitioning region \mathcal{L} into a collection of non-neighboring rectangular subregions $\mathcal{B}_{(\theta_0,t)}^{(\theta_t,\tau)}$, $t = 1, \dots, T(\theta_0)$, $\tau = 1, \dots, \hat{T}(\theta_0, \theta_t)$. This collection is parametrized by $\theta_0, \theta_1, \dots, \theta_{T(\theta_0)}$; a different parameter set leads to a different collection of non-neighboring rectangular subregions, each of size $O(\Theta^2)$.

Note that due to the construction any transmitting node in a rectangular subregion $\mathcal{B}_{(\theta_0,t)}^{(\theta_t,\tau)}$ does not interfere with any nodes in any other rectangular subregion. Therefore, we can focus on each subregion individually, and for each subregion obtain a part of the transmission rate vector by solving problem (15) to optimality and computing rates implied by the \mathbf{s}^* of Thm. III.1. The resulting global transmission rate vector \mathbf{r} is a feasible solution to the subproblem (13). Since the parameters $\theta_0, \theta_1, \dots, \theta_{T(\theta_0)}$ specify the way \mathcal{L} is partitioned we can tune them in order to obtain a “good” solution. More specifically, let $\Gamma(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)}) = \lambda^{(m)'} \mathbf{r}(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)})$ where $\lambda^{(m)}$ is the optimal dual vector corresponding to the first constraints of (12) at the m -th iteration, and $\mathbf{r}(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)})$ is the solution we obtain using the method described above. The algorithm described in Fig. 3, to be referred to as PTAS, yields a feasible solution to (13) with maximum $\Gamma(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)})$.

¹We assume \mathcal{L} is two-dimensional in this paper, but the extension to the three-dimensional case is straightforward.

-
- 1) **Initialization:** Specify an error level ϵ . Let $\mathbf{r} = \mathbf{0}$ and Θ equal to the smallest integer such that $(\frac{\Theta}{\Theta+1})^2 \geq 1 - \epsilon$.
 - 2) For each combination $(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)})$ such that $0 \leq \theta_0, \theta_1, \dots, \theta_{T(\theta_0)} \leq \Theta$ partition \mathcal{L} into rectangular subregions, solve problem (15) for each individual subregion and calculate $\Gamma(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)})$.
 - 3) Let

$$(\theta_0^*; \theta_1^*, \dots, \theta_{T(\theta_0)}^*) = \operatorname{argmax}_{\theta} \Gamma(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)}).$$

Output $\mathbf{r}(\theta_0^*; \theta_1^*, \dots, \theta_{T(\theta_0)}^*)$.

Fig. 3. PTAS algorithm that constructs a feasible solution to the subproblem (13).

PTAS is actually “asymptotically optimal” in the sense that when ϵ goes to 0, $\mathbf{r}(\theta_0^*; \theta_1^*, \dots, \theta_{T(\theta_0)}^*)$ converges to an optimal solution of (13). More specifically, let Γ^* be the optimal value of the subproblem (13); similar with [26] we have the following theorem which provides the guarantee; the proof is in Appendix VII-D.

Theorem IV.2 $\Gamma(\theta_0^*; \theta_1^*, \dots, \theta_{T(\theta_0)}^*) \geq (\frac{\Theta}{\Theta+1})^2 \Gamma^*$.

Note that with PTAS we still solve an ILP for each of the subregions. However, as the sizes of the ILP’s are much smaller than the one for the whole region \mathcal{L} , they usually take much less time to solve. On the other hand, we might as well utilize the algorithm in Fig. 2 to construct a feasible solution by solving the corresponding MWM relaxation for each rectangular subregion. This yields a third algorithm, to be referred to as PTAS+MWM, which is described in Fig. 4.

-
- 1) **Initialization:** Specify an error level ϵ . Let $\mathbf{r} = \mathbf{0}$ and Θ equal to the smallest integer such that $(\frac{\Theta}{\Theta+1})^2 \geq 1 - \epsilon$.
 - 2) For each combination $(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)})$ such that $0 \leq \theta_0, \theta_1, \dots, \theta_{T(\theta_0)} \leq \Theta$ reorganize \mathcal{L} into into rectangular subregions, solve the MWM relaxation and construct a feasible solution for each subregion according to the algorithm in Fig. 2, and calculate $\Gamma(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)})$.
 - 3) Let

$$(\theta_0^*; \theta_1^*, \dots, \theta_{T(\theta_0)}^*) = \operatorname{argmax}_{\theta} \Gamma(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)}).$$

Output $\mathbf{r}(\theta_0^*; \theta_1^*, \dots, \theta_{T(\theta_0)}^*)$.

Fig. 4. PTAS+MWM algorithm that constructs a feasible solution to the subproblem (13).

In some instances in order to solve problem (10) with the decomposition approach, it may be computationally expensive to search exhaustively for the optimal combination $(\theta_0^*; \theta_1^*, \dots, \theta_{T(\theta_0)}^*)$ at each iteration of the decomposition algorithm. Rather, in the same spirit of searching for a “good” solution for the subproblem, at the m -th iteration we only need a combination of $(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)})$ such that $\lambda^{(m)'} \mathbf{r}(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)}) > \mu^{(m)}$, where $\lambda^{(m)}$ and $\mu^{(m)}$ are

the optimal dual variables corresponding to the first and second constraints of problem (12) at the m -th iteration, respectively. As such, we propose another algorithm – RAS (Randomized Approximation Scheme)+MWM – based on PTAS+MWM to construct a feasible solution of the subproblem at the m -th iteration, as shown in Fig. 5. More specifically, we randomly choose a set of parameters $(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)})$ and test if they induce a valid transmission vector; we continue picking the parameters until we find one or the maximum number of trials is exceeded. (A similar idea was used in [16] but in a different context.) As will be clear in Sec. V, RAS+MWM cuts may considerably reduce the running times of the decomposition algorithm with a reasonable maximum number of trials.

-
- 1) **Initialization:** Specify an error level ϵ . Let $\mathbf{r} = \mathbf{0}$ and Θ equal to the smallest integer such that $(\frac{\Theta}{\Theta+1})^2 \geq 1 - \epsilon$. Specify the number of trials W and set $w = 1$. Let $\mathcal{F} = \emptyset$.
 - 2) **Iteration:** With a uniform probability distribution select $\theta_0 \in \{0, \dots, \Theta\}$ and then $\theta_i \in \{0, \dots, \Theta\}, \forall i \in \{1, \dots, T(\theta_0)\}$.
 - a) If $(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)}) \in \mathcal{F}$, go to the beginning of step 2).
 - b) Else add $(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)})$ to \mathcal{F} . According to this combination reorganize \mathcal{L} into into rectangular subregions, solve the MWM relaxation and construct a feasible solution for each subregion with the algorithm in Fig. 2, and calculate $\mathbf{r}(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)})$.
 - c) If $\lambda^{(m)'} \mathbf{r}(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)}) > \mu^{(m)}$ or $w = W$ exit and output $\mathbf{r}(\theta_0; \theta_1, \dots, \theta_{T(\theta_0)})$; else set $w := w + 1$ and go to step 2).
-

Fig. 5. RAS+MWM algorithm that constructs a feasible solution to the subproblem (13).

We have proposed four algorithms – MWM, PTAS, PTAS+MWM, RAS+MWM – to generate feasible solutions to the subproblem that can produce cuts in the course of the decomposition algorithm. In terms of computational cost and for large enough networks and Θ the algorithms are ordered as MWM, RAS+MWM, PTAS+MWM, PTAS, from least to most expensive. Although only PTAS can produce an arbitrarily close to optimal solution of the subproblem, our numerical experience shows that putting PTAS in the framework of the decomposition algorithm significantly slows down the computation due to the overhead of solving ILPs. As a result we only incorporate MWM, PTAS+MWM and RAS+MWM; the final algorithm for problem (10) is in Fig. 6 with MWM and PTAS+MWM cuts. To incorporate RAS+MWM in the framework, we can simply replace step (d) in the algorithm in Fig. 6 by the algorithm described in Fig. 5.

V. NUMERICAL RESULTS

In this section we present some illustrative numerical results to assess the efficiency of the proposed approach.

- 1) **Initialization:** Let $\mathbf{r}^1 \in \text{Conv}(\mathcal{R}) \cap \mathcal{F}$ and set $m = 1$.
- 2) **m -th iteration:**
 - a) Solve the restricted master problem (12) with $\mathbf{r}^1, \dots, \mathbf{r}^m$ to obtain an optimal primal-dual pair $(\mathbf{r}^{(m)}, \boldsymbol{\alpha}^{(m)}; \boldsymbol{\lambda}^{(m)}, \mu^{(m)}, \boldsymbol{\nu}^{(m)}, \boldsymbol{\sigma}^{(m)})$.
 - b) Use MWM (cf. Fig. 2) to obtain a feasible solution \mathbf{r} to the subproblem (13).
 - c) If $\mu^{(m)} - \boldsymbol{\lambda}^{(m)\prime} \mathbf{r} < 0$, let $\mathbf{r}^{(m+1)} = \mathbf{r}$, set $m := m + 1$ and go to step 2a).
 - d) Obtain a feasible solution \mathbf{r} to the subproblem (13) with PTAS+MWM (cf. Fig. 4).
 - e) If $\mu^{(m)} - \boldsymbol{\lambda}^{(m)\prime} \mathbf{r} < 0$, let $\mathbf{r}^{(m+1)} = \mathbf{r}$, set $m := m + 1$ and go to step 2a).
 - f) Solve the subproblem (13) to optimality by solving (15). Let \mathbf{r}^{m+1} be the optimal solution obtained.
 - g) If $\mu^{(m)} - \boldsymbol{\lambda}^{(m)\prime} \mathbf{r}^{m+1} \geq 0$ stop; $(\mathbf{r}^{(m)}, \boldsymbol{\alpha}^{(m)})$ is an optimal solution of (10). Else set $m := m + 1$ and go to step 2a).

Fig. 6. The improved decomposition algorithm with MWM and PTAS+MWM cuts.

The example we consider is a WSNET with 85 nodes uniformly distributed over a square region of 185×185 (feet)². There are 400 classes of traffic, generated by 80 source nodes and collected by 5 data sinks. Each node has a maximum transmission rate γ of 4.8 Kbits per second (Kbps). The objective is to maximize throughput, namely,

$$F(\mathbf{r}) = \sum_{i \in \mathcal{A}} \sum_{j=1}^N \sum_{k=1}^K \sum_{c=1}^C r_{ijkc},$$

where \mathcal{A} is the set of all the source nodes. We impose the fairness constraint requiring all traffic classes to have equal rate.

To construct the sets $\mathcal{Q}_t(i, p, c)$ and $\mathcal{Q}_l(i, p, c)$ we assume that the transmit power P_t decreases with the distance d as $P_t d^{-2}$. Accounting for the fact that commercial WSNET nodes are not sensitive enough to measure small differences in RSSI (cf. second branch of (21)) we will be using

$$RSSI = \begin{cases} R_t - 20 \log_{10} d, & d > 1, \\ R_t, & d \leq 1, \end{cases} \quad (21)$$

where R_t is P_t measured in dBm, RSSI is in dBm, and d is in feet. Using this expression, for any transmitting node i we let $\mathcal{Q}_t(i, p, c)$ include all nodes with RSSI above -83.0 dBm; such nodes can reliably receive packets from i . In $\mathcal{Q}_l(i, p, c)$ we include all nodes with RSSI above -85.0 dBm, these nodes can not receive any other transmission when i is transmitting.

A. The value of multiple frequency channels

We next evaluate the benefit of using multiple frequency channels. We use the decomposition algorithm in Fig. 6 to obtain the maximum achievable throughput. Table I lists the maximum achievable throughput for the cases of using 1 and 2 channels. Notice that with a maximum transmission rate γ of 4.8 Kbps and 5 data sinks the maximum throughput is at most 24 Kbps which we achieve with 2 frequency channels.

TABLE I
MAXIMUM THROUGHPUT AS A FUNCTION OF THE NUMBER OF CHANNELS AVAILABLE IN THE NETWORK.

# of Channels	Throughput (Kbps)
1	16.07
2	24

B. The value of cuts

To assess the benefit of utilizing the MWM-based and PTAS-based cuts, we compare (i) the algorithm in Fig. 1, which only utilizes the cuts generated by solving the ILP (13) (let us call this type of cuts INTEGER), and the following: (ii) the algorithm in Fig. 6 but without steps (d)–(e) (i.e., only MWM-based cuts are being used), (iii) the algorithm in Fig. 6, that is, with both types of cuts (MWM and PTAS+MWM), and (iv) the algorithm in Fig. 6 with step (d) replaced by the algorithm in Fig. 5, namely, with RAS+MWM type cuts where the maximum number of trials in RAS+MWM is set to 5. In Fig. 7 and Fig. 8, and for the case of one and two frequency channels, respectively, we plot the throughput achieved by the policies obtained during the course of the algorithms mentioned above.

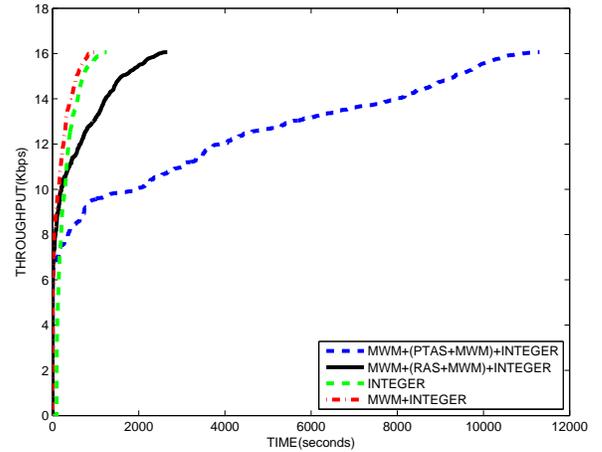


Fig. 7. Comparison between various cuts in the case of a single available channel.

Table II records the running time (in seconds) for each algorithm to achieve 90% of the optimal objective value², namely, 14.46Kbps for one frequency channel and 19.57Kbps for two, respectively.

A few observations are in order. First, MWM-based cuts offer significant improvement in terms of the computational efficiency. More specifically, in the cases of one and two frequency channels, the running times of the algorithm with only MWM-based cuts are 23.51% and 39.68% less than those of the algorithm in Fig. 1, respectively. Second, for smaller instances (with one frequency channel) adding RAS+MWM

²All the programs were run on a computer with a CPU running at 3.06 GHz, and 3.6 Gbytes of main memory; background processes were minimal.

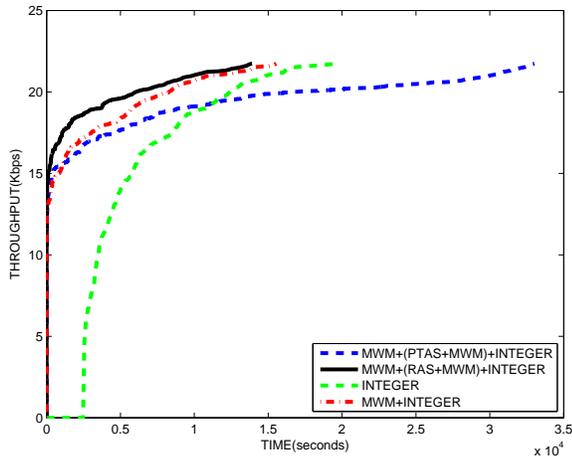


Fig. 8. Comparison between various cuts in the case of two available channels.

TABLE II
RUNNING TIME (IN SECONDS) COMPARISON BETWEEN VARIOUS CUTS
WITH A SINGLE AND TWO CHANNELS.

Type of Cuts	1 Channel	2 Channels
INTEGER	688	11572
MWM	488	6980
MWM and PTAS+MWM	8663	12925
MWM and RAS+MWM	1428	4792

cuts may actually increase the running time. On the other hand, for larger instances (two frequency channels) RAS+MWM cuts become much more helpful, reducing the running time of the algorithm with only MWM-based cuts by 31.35%. Third, with a proper setting of the maximum number of trials, RAS+MWM cuts enable a much more efficient use of the decomposition algorithm than PTAS+MWM cuts, and the running times are reduced by 83.52% and 62.92% for one and two frequency channels, respectively.

As the last remark in this section, we note another evidence illustrating that MWM and RAS+MWM supply quality cuts for the decomposition algorithms. In particular, before achieving 90% of the optimal throughput in the instance of two frequency channels with the algorithm with MWM and RAS+MWM cuts, the decomposition algorithm performs 1317 iterations, out of which there are 784 times (or 59.53%) where the MWM algorithm provides a valid transmission vector to the restricted master problem, RAS+MWM 482 times (36.6%), and we only solve the ILP (13) 14 times, i.e., 3.87% of the total number of iterations.

VI. CONCLUSIONS

We considered the problem of efficient transport in WSNs with multiple frequency channels and developed an approach to jointly optimize transmission scheduling, power control, channel selection, and routing. The objective is to maximize a utility function of average transmission rates (e.g., weighted throughput) subject to fairness constraints. We proposed a decomposition algorithm and established its convergence. The resulting policy involves time-sharing among

a number of feasible transmission schemes. Time-sharing convexifies the achievable region and achieves higher utility than any individual scheme.

Because we optimize over the convex hull of achievable transmission rate vectors, the problem formulation grows exponentially with the size of the network. Yet, the decomposition method identifies “promising” transmission schemes as needed, similarly as in column generation methods for large-scale linear programs. These transmission schemes are generated by a dual separation subproblem and the efficiency of our algorithm largely depends on how efficiently we can solve this subproblem. To that end, we exploited the subproblem structure and developed a toolset based on a maximum weighted matching relaxation and polynomial time approximation schemes. The numerical results we presented convincingly demonstrate the benefits of multiple channels and show that our approach can efficiently solve large instances of WSN problems.

We should note that the approach we presented is offline and centralized – admittedly an implementation drawback for large WSNs. However, the structure of the decomposition algorithm allows us to distribute some of the work with online computation. Specifically, master problem iterations can be performed at a central gateway. Note that given the set of transmission schemes $\mathbf{r}^1, \dots, \mathbf{r}^m$ used at the m -th iteration no other information is needed, especially no connectivity or interference information that may only be known by individual nodes. Once a master problem iteration is performed the values of the dual variables needed for solving the subproblem can be sent over to the respective nodes. If we give up on reaching an exact optimal solution, feasible solutions to the subproblem (generating dual cuts for the master problem) can be obtained in a distributed manner and the generated transmission schemes passed back to the gateway which solves the master problem. To solve the subproblem, one can employ distributed approximation algorithms for the maximum weighted matching relaxation [29], [30]. Furthermore, the PTAS schemes lend themselves to a distributed implementation as the various $O(\Theta^2)$ rectangular subregions are localized. The exact implementation details of such a distributed approach we outlined and whether it is practical for large WSNs is left to future investigations.

VII. APPENDICES

A. Proof of Theorem III.1

First we show $(\mathbf{s}^*, \mathbf{p}^*)$ is feasible for problem (14). To see this, note that by construction (\mathbf{s}, \mathbf{p}) satisfies the constraints (6) and (7). Moreover, by setting $s_{ijkc}^* = 0$ for any (i, j, k, c) such that $\pi_{ijkc} = 0$ we have \mathbf{s}^* satisfying constraint (5). Also, by writing the constraints (2) – (4) in a more compact form,

$$\sum_{(j,k,c)} s_{ijkc} + \sum_{(j,k,c)} s_{jikc} \leq 1, \forall i, \quad (22)$$

we can see that $(\mathbf{s}^*, \mathbf{p}^*)$ satisfies constraint (22), that is,

$$\begin{aligned} & \sum_{(j,k,c)} s_{ijkc}^* + \sum_{(j,k,c)} s_{jikc}^* \\ &= \sum_{(j,k,c)} 1_{\mathcal{X}}(i, j, k, c) x_{ijc}^* + \sum_{(j,k,c)} 1_{\mathcal{X}}(j, i, k, c) x_{jic}^* \\ &\leq \sum_{(j,c)} x_{ijc}^* + \sum_{(j,c)} x_{jic}^* \leq 1. \end{aligned}$$

Similarly, let us write constraint (8) as

$$s_{ijkc} + \sum_{u=1}^N s_{uvwc} \leq 1, \quad (23)$$

such that $\forall c, \forall (j, k) \neq (v, w)$ and $\forall v \in \mathcal{Q}_l(i, p_{ijkc}, c)$, and note that by definition

$$\mathcal{Q}_l(i, p_{ijkc}^*, c) = \begin{cases} \mathcal{Q}_l(i, \hat{p}_{ijkc}, c) & \text{if } s_{ijkc}^* = 1, \\ \emptyset & \text{otherwise,} \end{cases}$$

$\forall i, j, k, c$. Then it is evident that constraint (23) holds when $s_{ijkc}^* = 0, \forall i, j, k, c$. When $s_{ijkc}^* = 1 \forall i, j, k, c$, we have

$$\sum_{u=1}^N s_{uvwc}^* = \sum_{u=1}^N \mathbf{1}_{\mathcal{X}}(u, v, w, c) x_{uv}^* \leq \sum_{u=1}^N x_{uv}^* = 0,$$

$\forall (j, k) \neq (v, w), v \in \mathcal{Q}_l(i, p_{ijkc}, c)$, where the last equality is due to the fact $x_{ijc}^* = 1$. As such, we conclude that $(\mathbf{s}^*, \mathbf{p}^*)$ is feasible for problem (14).

Next we proceed to show $(\mathbf{s}^*, \mathbf{p}^*)$ is optimal for problem (14). Suppose it is not. Then there exists an optimal solution $(\tilde{\mathbf{s}}, \tilde{\mathbf{p}})$ to problem (14) with a strictly better objective value. Without losing feasibility or optimality, let us set $\tilde{p}_{ijkc} = 0$ if $\tilde{s}_{ijkc} = 0, \forall i, j, k, c$. Next, pick any (i, j, k, c) transmission present in the optimal solution, i.e., such that $\tilde{s}_{ijkc} = 1$. Clearly, $\tilde{p}_{ijkc} \geq \hat{p}_{ijkc}$ because otherwise j would not be reachable by i and the transmission can not take place (cf. (5)). Lower the power of the transmission to \hat{p}_{ijkc} and note that j is still reachable and $\mathcal{Q}_l(i, \hat{p}_{ijkc}, c) \subseteq \mathcal{Q}_l(i, \tilde{p}_{ijkc}, c)$ due to Assumption 1. As a result, by changing \tilde{p}_{ijkc} to \hat{p}_{ijkc} we can still have $\tilde{s}_{ijkc} = 1$ and maintain feasibility of all constraints (2)–(8) since the transmissions that can “collide” with (i, j, k, c) get reduced (cf. (8)). Furthermore, the value of the objective function in (14) remains unchanged since it only depends on \mathbf{s} . We can use the same power reduction strategy for any transmission present in the optimal solution $(\tilde{\mathbf{s}}, \tilde{\mathbf{p}})$, thus constructing a new solution $(\check{\mathbf{s}}, \check{\mathbf{p}})$ which remains optimal for problem (14).

Now let us construct a solution $\check{\mathbf{x}}$ to problem (15) such that $\check{x}_{ijc} = \sum_{k=1}^K \check{s}_{ijkc}, \forall i, j, c$. The constructed solution $\check{\mathbf{x}}$ is feasible for problem (15). To see this, note that by construction $\check{x}_{ijc} \in \{0, 1\}, \forall i, j, c$. Further, since $(\check{\mathbf{s}}, \check{\mathbf{p}})$ is feasible for problem (14) we can see

$$\sum_{(j,c)} \check{x}_{ijc} + \sum_{(j,c)} \check{x}_{jic} = \sum_{(j,k,c)} \check{s}_{ijkc} + \sum_{(j,k,c)} \check{s}_{ijkc} \leq 1.$$

Finally, for any i, j, c , if (a) $\check{x}_{ijc} = 1$ then we have $\check{s}_{ijkc} = 1$ for some k and thus $\check{s}_{uvwc} = 0, \forall (i, j, k) \neq (u, v, w), v \in \mathcal{Q}_l(i, \check{p}_{ijkc}, c)$ due to constraint (8). As a result $\check{x}_{uv} = 0, \forall u, \forall v \neq j, v \in \mathcal{Q}_l(i, \check{p}_{ijkc}, c)$, and thus $\check{x}_{ijc} + \sum_{u=1}^N \check{x}_{uv} \leq 1$; if (b) $\check{x}_{ijc} = 0$, then by construction we can see $\check{s}_{ijkc} = 0$ for all k , and consequently $\check{p}_{ijkc} = 0$. Since $\mathcal{Q}_l(i, \check{p}_{ijkc}, c) = \emptyset$ we also have $\check{x}_{ijc} + \sum_{u=1}^N \check{x}_{uv} \leq 1, \forall v \neq j, v \in \mathcal{Q}_l(i, \check{p}_{ijkc}, c)$. To this end we have proved that $\check{\mathbf{x}}$ is feasible for problem (15).

Last, recall that we suppose $(\tilde{\mathbf{s}}, \tilde{\mathbf{p}})$ is a better solution than $(\mathbf{s}^*, \mathbf{p}^*)$ to problem (14). Then

$$\begin{aligned} \sum_{(i,j,c)} \omega_{ijc} \check{x}_{ijc} &\geq \sum_{(i,j,k,c)} \pi_{ijkc} \check{s}_{ijkc} > \sum_{(i,j,k,c)} \pi_{ijkc} s_{ijkc}^* \\ &= \sum_{(i,j,k,c)} \pi_{ijkc} \mathbf{1}_{\mathcal{X}}(i, j, k, c) x_{ijc}^* = \sum_{(i,j,c)} \omega_{ijc} x_{ijc}^*, \end{aligned}$$

where the first and last equalities are due to the definition of ω_{ijc} . This contradicts the fact that \mathbf{x}^* is optimal for problem (14) and the conclusion follows.

B. Proof of Theorem III.2

Recall that at the m -th iteration the subproblem minimizes $\mu^{(m)} - \boldsymbol{\lambda}^{(m)'} \mathbf{r}$ over $\mathbf{r} \in \mathcal{R}$. Thus, if $\mu^{(m)} - \boldsymbol{\lambda}^{(m)'} \mathbf{r}^{m+1} \geq 0$ it follows that $\mu^{(m)} - \boldsymbol{\lambda}^{(m)'} \mathbf{r} \geq 0$ for all $\mathbf{r} \in \mathcal{R}$. Since all extreme points of $\text{Conv}(\mathcal{R})$ are in \mathcal{R} , the latter condition implies that $\mu^{(m)} - \boldsymbol{\lambda}^{(m)'} \mathbf{r}^n \geq 0$ for all extreme points $\mathbf{r}^1, \dots, \mathbf{r}^L$ of $\text{Conv}(\mathcal{R})$. Therefore, $(\mathbf{r}^{(m)}, \boldsymbol{\alpha}^{(m)}; \boldsymbol{\lambda}^{(m)}, \mu^{(m)}, \boldsymbol{\nu}^{(m)}, \boldsymbol{\sigma}^{(m)})$ is an optimal primal-dual pair for (10) and the algorithm terminates.

Next note that due to Thm. III.1 and the resulting structure of the subproblem solutions, at each iteration we generate a valid transmission rate vector $\mathbf{r} \in \mathcal{R}$. Let $\mathbf{r}^1, \dots, \mathbf{r}^m$ be the transmission rates generated up to the m -th iteration and suppose the algorithm does not terminate at the m -th iteration. The next transmission rate to be generated, \mathbf{r}^{m+1} , is different from the ones generated earlier since they are separated by a hyperplane. In particular, since $\mu^{(m)}, \boldsymbol{\lambda}^{(m)}$ are feasible for the restricted dual problem at the m -th iteration we have

$$\begin{aligned} \mu^{(m)} - \boldsymbol{\lambda}^{(m)'} \mathbf{r}^n &\geq 0, \quad n = 1, \dots, m, \\ \mu^{(m)} - \boldsymbol{\lambda}^{(m)'} \mathbf{r}^{m+1} &< 0. \end{aligned}$$

Thus, at each iteration we generate a new point of the finite set \mathcal{R} . Hence, the algorithm terminates in a finite number of iterations.

C. Proof of Theorem IV.1

Let us first show \mathbf{x}^* is a feasible solution to problem (18). To see this, note that x_{ijc}^* by construction takes values in $\{0, 1\}$ for all i, j, c . Further,

$$\begin{aligned} &\sum_{j=1}^N \sum_{c=1}^C x_{ijc}^* + \sum_{j=1}^N \sum_{c=1}^C x_{jic}^* \\ &= \sum_{j=1}^N \sum_{c=1}^C \mathbf{1}_{\mathcal{C}}(i, j, c) y_{ij}^* + \sum_{j=1}^N \sum_{c=1}^C \mathbf{1}_{\mathcal{C}}(j, i, c) y_{ji}^* \\ &= \sum_{j=1}^N y_{ij}^* \sum_{c=1}^C [\mathbf{1}_{\mathcal{C}}(i, j, c) + \mathbf{1}_{\mathcal{C}}(j, i, c)] \\ &= \sum_{j=1}^N y_{ij}^* \leq 1, \end{aligned}$$

where the second equality is due to $y_{ij}^* = y_{ji}^*$ and the third equality follows from the definition of \mathcal{C} .

Next we relate the objective values of problem (18) and (20). We have

$$\begin{aligned} &\sum_{(i,j) \in \mathcal{E}} \hat{\omega}_{ij} y_{ij}^* \\ &= \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{V}} y_{ij}^* [\max_{1 \leq c \leq C} \max\{\omega_{ijc}, \omega_{jic}\}] \\ &= \sum_{(i,j,c)} y_{ij}^* [\mathbf{1}_{\mathcal{C}}(i, j, c) \omega_{ijc} + \mathbf{1}_{\mathcal{C}}(j, i, c) \omega_{jic}] \\ &= 2 \sum_{(i,j,c)} \omega_{ijc} x_{ijc}^*, \end{aligned}$$

where the second equality is due to the definition of \mathcal{C} .

Now suppose \mathbf{x}^* is not optimal for problem (18) and there exists a feasible solution $\hat{\mathbf{x}}$ with a strictly better objective value. Let us construct a solution $\hat{\mathbf{y}}$ to problem (20) as

$$\hat{y}_{ij} = \sum_{c=1}^C \hat{x}_{ijc} + \sum_{c=1}^C \hat{x}_{jic}, \quad \forall 1 \leq i, j \leq N.$$

It is evident that by construction \hat{y}_{ij} is 0 or 1, and $\hat{y}_{ij} = \hat{y}_{ji}$ for any i, j . Furthermore, for any $i = 1, \dots, N$ we have

$$\sum_{j|(i,j) \in \mathcal{E}} \hat{y}_{ij} = \sum_{j=1}^N \sum_{c=1}^C (\hat{x}_{ijc} + \hat{x}_{jic}) \leq 1,$$

and thus $\hat{\mathbf{y}}$ is feasible for problem (18).

To this end, note that

$$\begin{aligned} & \sum_{(i,j) \in \mathcal{E}} \hat{\omega}_{ij} \hat{y}_{ij} \\ &= \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{V}} \left[\max_{1 \leq c \leq C} \max\{\omega_{ijc}, \omega_{jic}\} \right] \left[\sum_{c=1}^C (\hat{x}_{ijc} + \hat{x}_{jic}) \right] \\ &\geq 2 \sum_{(i,j,c)} \omega_{ijc} x_{ijc}. \end{aligned}$$

Therefore, we have

$$\begin{aligned} \sum_{(i,j) \in \mathcal{E}} \hat{\omega}_{ij} \hat{y}_{ij} &\geq 2 \sum_{(i,j,c)} \omega_{ijc} \hat{x}_{ijc} \\ &> 2 \sum_{(i,j,c)} \omega_{ijc} x_{ijc}^* = \sum_{(i,j) \in \mathcal{E}} \hat{\omega}_{ij} y_{ij}^*, r \end{aligned}$$

which contradicts the assumption that \mathbf{y}^* is optimal for problem (18) and the conclusion follows.

D. Proof of Theorem IV.2

The proof is similar to [28] with some differences due to our problem settings, that is, the structure of the interference constraints and the presence of multiple frequency channels. Denote by $\tilde{\mathcal{G}} = (\tilde{\mathcal{V}}, \tilde{\mathcal{E}})$ the graph induced by the network under consideration, that is, $\tilde{\mathcal{V}} = \{1, \dots, N\}$ and $\tilde{\mathcal{E}} = \{(i, j) \mid j \in \mathcal{Q}_t(i, \hat{p}_{ijc}, c) \text{ for some } c \in \{1, \dots, C\}, \forall i, j \in \tilde{\mathcal{V}}\}$.

We denote by $\mathcal{G}_\theta = (\mathcal{V}_\theta, \mathcal{E}_\theta)$ the subgraph created by removing from $\tilde{\mathcal{G}}$ all the nodes and links in each strip \mathcal{B}_m such that $m \bmod \Theta + 1 = \theta$ (cf. Sec. IV-C), for $\theta = 0, \dots, \Theta$. Define $\hat{\mathcal{G}}_\theta = (\hat{\mathcal{V}}_\theta, \hat{\mathcal{E}}_\theta) \triangleq \tilde{\mathcal{G}} \setminus \mathcal{G}_\theta$ and denote by $\mathcal{G}_{(\theta,t)} = (\mathcal{V}_{(\theta,t)}, \mathcal{E}_{(\theta,t)})$ the subgraph of $\tilde{\mathcal{G}}$ with all the nodes inside the subregion $\mathcal{B}_{(\theta,t)}$, for $t = 1, \dots, T(\theta)$. Similarly, though with slight abuse of notation, we denote by $\mathcal{G}_{(\theta',t)}^{\theta'}$ the subgraph induced by removing from subregion $\mathcal{B}_{(\theta',t)}$ each strip $\mathcal{B}_{(\theta',t)}^m$ such that $m \bmod \Theta + 1 = \theta'$, for $\theta' = 0, \dots, \Theta$.

Consider the following ILP problem defined on an arbitrary graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}) \subseteq \tilde{\mathcal{G}}$:

$$\begin{aligned} \max \quad & \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{V}} \sum_{c=1}^C \omega_{ijc} x_{ijc} \\ \text{s.t.} \quad & \sum_{j \in \mathcal{V}} \sum_{c=1}^C x_{ijc} + \sum_{j \in \mathcal{V}} \sum_{c=1}^C x_{jic} \leq 1, \quad \forall i \in \mathcal{V}, \\ & x_{ijc} + \sum_{u \in \mathcal{V}} x_{uvc} \leq 1, \quad \forall c, \text{ such that} \\ & \quad \forall v \neq j, v \in \mathcal{Q}_t(i, \hat{p}_{ijc}, c), v \in \mathcal{V}, \\ & x_{ijc} \in \{0, 1\}, \quad \forall i, j \in \mathcal{V}, \forall c = 1, \dots, C, \end{aligned} \quad (24)$$

where ω_{ijc} is defined in Sec. IV-C, $\forall i, j, c$, and let $\tilde{\mathbf{x}}(\mathcal{G})$ and $\tilde{f}(\mathcal{G})$ be the optimal solution and the optimal value of problem (24) associated with graph \mathcal{G} , respectively. Note that problem (24) is identical to problem (15) if $\mathcal{G} = \tilde{\mathcal{G}}$, and thus $\tilde{f}(\tilde{\mathcal{G}}) = \Gamma^*$. For convenience we denote by $\tilde{\mathcal{E}}(\mathbf{x})$ the set of links induced by any \mathbf{x} feasible for problem (24), namely,

$$\tilde{\mathcal{E}}(\mathbf{x}) = \{(i, j) \mid x_{ijc} = 1 \text{ for some } c = 1, \dots, C, \forall i, j \in \mathcal{V}\}.$$

Moreover, for any set $\mathcal{W} \subseteq \mathcal{E}$ let us define $w(\mathcal{W}) \triangleq \tilde{f}(\tilde{\mathcal{G}}(\mathcal{W}))$ where $\tilde{\mathcal{G}}(\mathcal{W})$ is the graph spanned by \mathcal{W} .

To proceed we will need the following lemma.

Lemma VII.1 $\max_{\theta} \tilde{f}(\mathcal{G}_\theta) \geq \frac{\Theta}{\Theta+1} \Gamma^*$.

Proof: First note that by construction $\hat{\mathcal{E}}_{\theta_1} \cap \hat{\mathcal{E}}_{\theta_2} = \emptyset$ if $\theta_1 \neq \theta_2, \forall \theta_1, \theta_2 = 0, \dots, \Theta$. Moreover, $\cup_{\theta=0}^{\Theta} \hat{\mathcal{E}}_{\theta} \subseteq \tilde{\mathcal{E}}$.

Let $\mathcal{W}_\theta = \hat{\mathcal{E}}_{\theta} \cap \tilde{\mathcal{E}}(\tilde{\mathbf{x}}(\tilde{\mathcal{G}})), \forall \theta = 0, \dots, \Theta$. Note that the links in \mathcal{W}_θ are among the ones that are induced by the optimal solution $\tilde{\mathbf{x}}(\tilde{\mathcal{G}})$ but reside in the region that is later removed. Then it can be seen that

$$\sum_{\theta} w(\mathcal{W}_\theta) = w(\cup_{\theta} \mathcal{W}_\theta) \leq w(\tilde{\mathcal{E}}(\tilde{\mathbf{x}}(\tilde{\mathcal{G}}))) = \tilde{f}(\tilde{\mathcal{G}}) = \Gamma^*,$$

since we consider disjoint parts of the optimal solution. As a result we can see that $\min_{\theta} w(\mathcal{W}_\theta) \leq \frac{\Gamma^*}{\Theta+1}$.

Let $\theta' = \operatorname{argmin}_{\theta} w(\mathcal{W}_\theta)$. Then

$$\begin{aligned} \max_{\theta} \tilde{f}(\mathcal{G}_\theta) &\geq \tilde{f}(\mathcal{G}_{\theta'}) \\ &\geq w(\tilde{\mathcal{E}}(\tilde{\mathbf{x}}(\tilde{\mathcal{G}}))) - w(\mathcal{W}_{\theta'}) \\ &\geq \Gamma^* - \frac{\Gamma^*}{\Theta+1} = \frac{\Theta}{\Theta+1} \Gamma^*, \end{aligned}$$

where the second inequality is due to the definition of \tilde{f} . ■

Now we have all the ingredients to prove Theorem IV.2.

Let us first suppose $\theta'_0 = \operatorname{argmax}_{\theta} \tilde{f}(\mathcal{G}_\theta)$ and $\theta'_t = \operatorname{argmax}_{\theta} \tilde{f}(\mathcal{G}_{(\theta',t)}^{\theta'})$, $\forall t = 1, \dots, T(\theta'_0)$. Apply Lemma VII.1 to graph $\mathcal{G}_{(\theta'_0,t)}$ and we have

$$\tilde{f}(\mathcal{G}_{(\theta'_0,t)}^{\theta'_0}) \geq \frac{\Theta}{\Theta+1} \tilde{f}(\mathcal{G}_{(\theta'_0,t)}^{\theta'_0}), \quad \forall t = 1, \dots, T(\theta'_0).$$

Note that $\sum_{t=1}^{T(\theta'_0)} \tilde{f}(\mathcal{G}_{(\theta'_0,t)}^{\theta'_0}) = \tilde{f}(\mathcal{G}_{\theta'_0})$. Therefore

$$\begin{aligned} \Gamma(\theta_0^*; \theta_1^*, \dots, \theta_{T(\theta_0^*)}^*) &\geq \sum_{t=1}^{T(\theta'_0)} \tilde{f}(\mathcal{G}_{(\theta'_0,t)}^{\theta'_0}) \\ &\geq \sum_{t=1}^{T(\theta'_0)} \frac{\Theta}{\Theta+1} \tilde{f}(\mathcal{G}_{(\theta'_0,t)}^{\theta'_0}) \\ &= \frac{\Theta}{\Theta+1} \tilde{f}(\mathcal{G}_{\theta'_0}) \geq \left(\frac{\Theta}{\Theta+1}\right)^2 \Gamma^*, \end{aligned}$$

where the last inequality follows from another application of Lemma VII.1.

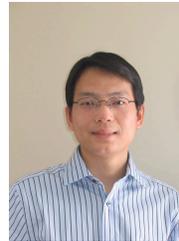
REFERENCES

- [1] R. Madan, S. Cui, S. Lall, and A. Goldsmith, "Cross-layer design for lifetime maximization in interference-limited wireless sensor networks," *IEEE Trans. on Wireless Communications*, vol. 5, no. 11, 2006.
- [2] I. C. Paschalidis, W. Lai, and D. Starobinski, "Asymptotically optimal transmission policies for large-scale low-power wireless sensor networks," *IEEE/ACM Trans. Networking*, vol. 15, no. 1, pp. 105–118, 2007.
- [3] S. Ray, W. Lai, and I. C. Paschalidis, "Statistical location detection with sensor networks," *Joint special issue IEEE/ACM Trans. Networking and IEEE Trans. Information Theory*, vol. 52, no. 6, pp. 2670–2683, 2006.
- [4] I. C. Paschalidis and D. Guo, "Robust and distributed stochastic localization in sensor networks: Theory and experimental results," *ACM Trans. Sensor Networks*, 2008, in print.
- [5] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, "Wireless sensor networks for habitat monitoring," *Proc. 1st ACM Intl. Workshop on Wireless Sensor Networks and Applications (WSNA02)*, pp. 88–97, 2002.
- [6] T. Teixeira, D. Lymberopoulos, E. Culurciello, Y. Aloimonos, and A. Savvides, "A lightweight camera sensor network operating on symbolic information," in *Proceedings of First Workshop on Distributed Smart Cameras, held in conjunction with ACM SenSys*, Boulder, Colorado, October 2006.
- [7] U. C. Kozat, I. Koutsopoulos, and L. Tassiulas, "A framework for cross-layer design of energy-efficient communication with qos provisioning in multi-hop wireless networks," *IEEE Infocom*, 2004.
- [8] X. Lin, N. Shroff, and R. Srikant, "A Tutorial on Cross-Layer Optimization in Wireless Networks," *IEEE J. Select. Areas Commun.*, vol. 24, no. 8, pp. 1452–1463, 2006.

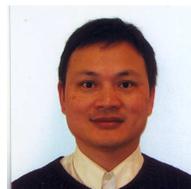
- [9] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Trans. Automat. Contr.*, vol. 37, no. 12, pp. 1936–1948, 1992.
- [10] E. Modiano, D. Shah, and G. Zussman, "Maximizing throughput in wireless networks via gossiping," in *Proceedings of the joint international conference on Measurement and modeling of computer systems (ACM SIGMETRICS/IFIP Performance)*, June 2006, pp. 27–38.
- [11] L. Tassiulas, "Linear complexity algorithms for maximum throughput in radio networks and input queued switches," in *Proceedings of the INFOCOM Conference*, 1998, pp. 533–539.
- [12] X. Wu and R. Srikant, "Regulated Maximal Matching: A Distributed Scheduling Algorithm for Multi-Hop Wireless Networks With Node-Exclusive Spectrum Sharing," in *Proceedings of the 44th IEEE Conference on Decision and Control (CDC-ECC'05)*, Seville, Spain, December 2005, pp. 5342–5347.
- [13] S. Sanghavi, L. Bui, and R. Srikant, "Distributed link scheduling with constant overhead," in *Proceedings of the ACM SIGMETRICS*, 2007, pp. 313–324.
- [14] X. Lin and S. Rasool, "Constant-time distributed scheduling policies for ad hoc wireless networks," in *Proceedings of the IEEE CDC conference*, San Diego, California, 2006.
- [15] S. Sarkar and K. Kar, "Achieving 2/3 throughput approximation with sequential maximal scheduling under primary interference constraints," in *Proceedings of 44th Annual Allerton Conference on Communication, Control and Computing*, Monticello, Illinois, 2006, pp. 27–29.
- [16] S. Sarkar and S. Ray, "Arbitrary throughput versus complexity tradeoffs in wireless networks using graph partitioning," *IEEE Trans. Automat. Contr.*, 2008, to appear, available at <http://www.seas.upenn.edu/~swati/\verb>.
- [17] L. Xiao, M. Johansson, H. Hindi, S. Boyd, and A. Goldsmith, *Joint Optimization of Wireless Communication and Networked Control Systems*, ser. Switching and Learning, Springer Lecture Notes in Computer Science 3355. Springer, 2005, pp. 248–272.
- [18] T. Chahed and B. Tuffin, "Network control and optimization," *Springer Lecture Notes in Computer Science 4465*, 2007.
- [19] M. Johansson, L. Xiao, and S. Boyd, "Simultaneous routing and resource allocation in CDMA wireless data networks," *Proc. IEEE International Conference on Communications*, vol. 1, pp. 51–55, May 2003.
- [20] I. C. Paschalidis, W. Lai, and X. Song, "Optimal transmission scheduling policies for wireless sensor networks with frequency diversity," in *Proceedings of the 46th IEEE Conference on Decision and Control*, New Orleans, Louisiana, December 2007, pp. 512–517.
- [21] —, "A decomposition method for transmission scheduling in multi-channel wireless sensor networks," in *Proceedings of the IEEE INFOCOM Conference*, Phoenix, Arizona, April 2008.
- [22] M. Neely, E. Modiano, and C. Li, "Fairness and optimal stochastic control for heterogeneous networks," *IEEE Infocom*, 2005.
- [23] A. Stolyar, "Maximizing queueing network utility subject to stability: Greedy primal-dual algorithm," *Queueing Systems: Theory and Applications*, vol. 50, pp. 401–457, August 2005.
- [24] A. Eryilmaz and R. Srikant, "Joint congestion control, routing and MAC for stability and fairness in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 1514–1524, August 2006.
- [25] L. Chen, S. H. Low, M. Chiang, and J. C. Doyle, "Jointly optimal congestion control, routing and scheduling for wireless ad hoc networks," *IEEE Infocom*, 2006.
- [26] G. Sharma, R. Mazumdar, and N. Shroff, "On the complexity of scheduling in wireless networks," in *Proceedings of the MobiCom Conference*, Los Angeles, California, September 2006, pp. 227–238.
- [27] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. Prentice Hall, 1982.
- [28] H. Hunt, M. Marathe, V. Radhakrishnan, S. Ravi, D. J. Rosenkrantz, and R. Stearns, "NC-approximation schemes for NP- and PSPACE-hard problems for geometric graphs," *Journal of Algorithms*, vol. 26, no. 2, pp. 238–274, 2004.
- [29] Z. Lotker, B. Patt-Shamir, and A. Rosen, "Distributed approximate matching," in *Proceedings of the 26th annual ACM symposium on Principles of distributed computing*, 2007, pp. 167–174.
- [30] Y. Cheng, M. Neely, and K. Chugg, "Iterative Message Passing Algorithm for Bipartite Maximum Weighted Matching," in *Information Theory, 2006 IEEE International Symposium on*, 2006, pp. 1934–1938.



Ioannis Ch. Paschalidis (M'96, SM'06) is an Associate Professor at Boston University with appointments in the Department of Electrical and Computer Engineering and the Division of Systems Engineering. He is a Co-Director of the Center for Information and Systems Engineering (CISE) and the Academic Director of the Sensor Network Consortium. He completed his graduate education at the Massachusetts Institute of Technology (MIT) receiving an M.S. (1993) and a Ph.D. (1996), both in Electrical Engineering and Computer Science. In September 1996 he joined Boston University where he has been ever since. He has held visiting appointments with MIT, and the Columbia University Business School. His current research interests lie in the fields of systems and control, networking, applied probability, optimization, operations research and computational biology.



Wei Lai received the B.E. degree in Automatic Control and the M.E. degree in Systems Engineering, both from Huazhong University of Science and Technology, China, in 1999 and 2001, respectively, and the Ph.D. degree in Systems Engineering from Boston University in 2007.



Xiangdong Song received the M.S. in geography in 2003 from Boston University. He is currently working towards the Ph.D degree in Systems Engineering at Boston University. His research interests include optimization and wireless sensor networks.