Asymptotically Optimal Transmission Policies for Low-Power Wireless Sensor Networks

Ioannis Ch. Paschalidis*, Wei Lai†, David Starobinski‡

*Center for Information and Systems Engineering and
Department of Manufacturing Engineering
Boston University
e-mail: yannisp@bu.edu, url: http://ionia.bu.edu/
†Center for Information and Systems Engineering and
Department of Manufacturing Engineering
Boston University
e-mail: wei@bu.edu
‡Center for Information and Systems Engineering and
Department of Electrical and Computer Engineering
Boston University
e-mail: staro@bu.edu

***Abstract*— We consider wireless sensor networks with multiple gateways and multiple classes of traffic carrying data generated by different sensory inputs. The objective is to devise joint routing, power control and transmission scheduling policies in order to gather data in the most efficient manner while respecting the needs of different sensing tasks (fairness). We formulate the problem as maximizing the utility of transmissions subject to explicit fairness constraints. We propose an efficient decomposition algorithm drawing upon large-scale decomposition ideas in mathematical programming. We show that our algorithm terminates in a finite number of iterations and produces a policy that is asymptotically optimal at low transmission power levels. Moreover, numerical results establish that this policy is near-optimal even at high power levels. We also demonstrate how to adapt our algorithm to accommodate energy constraints and node failures. The approach we introduce can efficiently determine near-optimal transmission policies for dramatically larger problem instances than an alternative enumeration approach.**

I. Introduction

Wireless sensor networks consist of a potentially large number of typically small devices – the sensor nodes or sensors – used to monitor some physical process or system [1]. Wireless sensors have limited computational capabilities, communicate wirelessly, and often operate in noisy and potentially adverse environments. Furthermore, as sensors are usually powered by limited and non-replenishable energy resources, energy preservation is also regarded as one of the keys to unlock the full potential of sensor networks. As a result, efficient resource allocation and aggressive optimization of network operations is not merely a desirable luxury but rather an indispensable necessity.

In this paper, we view the sensor network as a network that collects data to relay them to some other processing or communication infrastructure. To that end, it utilizes a host of *gateways* whose role is information collection (and fusion) from the sensor nodes [2–4]. A plethora of applications fit this paradigm, including process control, industrial automation, condition monitoring in manufacturing systems, indoor location detection, environmental monitoring, military, and homeland defense [5–7].

Sensor network architectures create several new and interesting challenges. For instance, traditional carrier sensing and random access strategies, as used in the IEEE 802.11 protocol, are often seen as inefficient and energy wasteful in sensor network applications [5]. Furthermore, it is unclear what is the transmission power at which sensors should communicate and to which node they should forward their data. For instance, is it preferable to communicate directly with the gateways, possibly using high transmission power, or via other sensors at a lower power? Another problem is how to differentiate between data generated by different sensing tasks so that more sensitive data is given higher priority. Finally, the question of how to optimize network operations while guaranteeing a minimum life-time for the network is another problem of significant theoretical and practical importance [4].

Several works have addressed subsets of these problems, showing, for instance, the crucial role played by multi-hop routing and power control [8–14]. However, as mentioned in [4], to optimize the use of scarce resources, future sensor network architectures must address all these challenges (MAC, routing, QoS, and power control) using an integrated approach. In an earlier attempt to address this joint optimization problem, the work in [8] develops a computational approach for deriving an optimal policy for sensor transmissions. This approach amounts to an enumeration of all the possible transmission strategies. Unfortunately, this enumeration requires a formidable amount of computations and is applicable in practice only to networks of at most 5 or 6 nodes, which is admittedly very small.

In this paper, we address the joint optimization problem of [8] and propose a new and much more efficient computational approach to solve it. Specifically, we consider the regime of low transmission powers (i.e., 100 mW or less), in which most sensor networks operate. In this regime, the transmission rate of sensors scales linearly with the transmission power. Based on this assumption, and building on some preliminary work in [15], we present a new methodology to derive the optimal transmission policy for the sensors. Our approach employs a column-generation method that consists of a *master problem* and a *subproblem*. Apart from establishing the convergence of the proposed algorithm, a key contribution is the efficient solution of the subproblem; we establish that it can be solved in polynomial-time. Our methodology dramatically improves the size of problems that can be solved. Compared to very small instances solvable by enumeration [8], we are able to solve instances with 50 or so nodes in less than a minute. Furthermore, although our derived policy is provably optimal only in the regime of very low power levels, in this paper we show that it can easily be adapted to the regime where the linear approximation is not in effect. In particular, we present numerical results that show that our policy is nearly-optimal even at very high transmission powers.

The optimization problem, as formulated in this paper, is a *utility maximization* problem that can accommodate lifetime constraints, fairness constraints, and potential interdependencies among sensor objectives. Therefore, throughput maximization [8, 11] and maximization of separable utility functions [14] are special cases of the more general problem considered in our paper.

As part of the derivation of our solution, we show that the optimal policy involves time-sharing among several transmission schemes. We also show that in the case of a node failure (which is a likely event in sensor network environments), we do not have to re-compute the transmission schemes from scratch. Instead, we introduce an optimization technique which reuses (after appropriate modifications) the previously obtained transmission schemes an an input to the algorithm, thus allowing the algorithm to converge much faster. Numerical results show that this optimization technique can speed-up the convergence rate of the algorithm by close to two orders of magnitude.

The rest of the paper is organized as follows. In Section II we present the system model and formulate the utility maximization problem. Section III discusses the (undesirable) implications of not enforcing fairness constraints. Section IV presents our decomposition algorithm and establishes its convergence. In Section V we discuss how to solve the subproblem efficiently (in polynomial-time). In Section VI we use the policy structure obtained from the decomposition algorithm and discuss how to obtain a policy when the linear approximation of rates is not in effect. Optimization over power limits is considered in Section VII. In Section VIII we discuss how to trade-off achieved utility vs. desirable network lifetime. We deal with node failures in Section IX. Some illustrative numerical results are presented in Section X. Concluding remarks are in Section XI.

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 $\mathbf{x} = (x_1, \ldots, x_R)$ 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 Conv$(\mathscr{A})$ the convex hull of a set $\mathscr{A}$, and by $|\mathscr{A}|$ its cardinality. We denote by $1_{\mathscr{A}}(\mathbf{x})$ the indicator function of $\mathbf{x} \in \mathscr{A}$. When $\mathscr{A}$ 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 *Wireless Sensor NETwork (WSNET)* with N sensor nodes each of which can receive, transmit and relay information with a single port/antenna that it carries. We assume that sensor nodes do not multicast information, so each transmission 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. In addition to the sensor nodes, the network uses M gateways which receive information from the sensors and relay it to some other networking or processing infrastructure. In our model, gateways can only receive information. However, they are allowed to receive data from multiple nodes simultaneously (our model does not assume that gateways must be able to receive multiple transmissions simultaneously; it just allows to model this capability). Henceforth, we will refer to all $M+N$ sensor and gateway nodes alike as *nodes of the WSNET*. Nodes $1, \ldots, N$ will correspond to sensor nodes and nodes $N+1, \ldots, N+M$ to gateways.

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 gateway) nodes. As a result, the WSNET carries multiple types of traffic, differing in information content and utility associated with their successful transmission. We use the term *traffic class* to refer to types of traffic with a particular origin and destination. Let K be the total number of traffic classes. We denote by $s(k)$ and $d(k)$ the source and destination of class k, for $k = 1, \ldots, K$.

We model the background noise in the WSNET as a single source of additive, white and Gaussian noise, with power spectral density η and bandwidth W. Let p_{ijk} denote the power used by node i to transmit class k traffic to node j, for $i = 1, \ldots, N+M$, $j = 1, \ldots, N+M$, $k = 1, \ldots, K$. We will refer to such a transmission as the (i, j, k) transmission. Let G_{ij} be the channel gain between nodes i and j when i is transmitting. When node i transmits class k traffic the received power at node j is $p_{ijk}G_{ij}$. Sensor nodes have limited power resources; we let $\bar{p}_i$ denote the maximum power available at node i for $i = 1, \ldots, N$. Thus, for any $i, j = 1, \ldots, M+N$,

$k = 1, \ldots, K$, it follows that p_{ijk} is upper bounded by

$$\bar{p}_{ijk} \triangleq \begin{cases} 0, & \text{if } i = N+1, \ldots, N+M, \\ & \text{or } i = d(k), \text{ or } i = j, \\ \bar{p}_i, & \text{otherwise}, \end{cases} \quad (1)$$

where $\bar{p}_{ijk}$ denotes the maximum power available for the (i, j, k) transmission. The first branch of (1) is an immediate consequence of the assumptions we made.

Consider an (i, j, k) transmission. The SINR, γ_{ijk}, is

$$\gamma_{ijk} = \frac{p_{ijk} G_{ij}}{\eta W + \sum_{v=1}^{K} \sum_{l=1, l \neq i}^{N+M} \sum_{u=1}^{N+M} p_{luv} G_{lj}}. \quad (2)$$

We use the Shannon capacity to determine the maximum rate for an (i, j, k) transmission and assume that the sending node i transmits with the maximum possible rate. Let r_{ijk} denote the net flow rate for an (i, j, k) or a (j, i, k) transmission, i.e.,

$$r_{ijk} = W \log_2(1 + \gamma_{ijk}) - W \log_2(1 + \gamma_{jik}). \quad (3)$$

When an (i, j, k) transmission is in progress, and under the transmission restrictions adopted, it follows that $\gamma_{ijk} \geq 0$, $\gamma_{uiv} = 0$ for all u, v, and $r_{ijk} \geq 0$. Otherwise, when a (j, i, k) transmission is in progress, $\gamma_{jik} \geq 0$, $\gamma_{ujv} = 0$ for all u, v, and $r_{ijk} \leq 0$. Clearly, $r_{ijk} = -r_{jik}$. We write $\mathbf{r}$ for the $(N+M)^2 K$-dimensional vector of r_{ijk}'s and denote by r_{ijk} its component that corresponds to the net flow rate for an (i, j, k) or a (j, i, k) transmission. Similarly, we write $\mathbf{p}$ for the $(N+M)^2 K$-dimensional vector of powers and denote by p_{ijk} its component corresponding to the (i, j, k) transmission.

In this work, we first concentrate on sensor networks in which power levels are on the order of mWatt or even lower and the transmission rates in (3) can be well approximated by a linear function of transmitting powers. A linear approximation is also used in the literature, as long as nodes do not transmit at high rates [10, 12, 13]. In particular, taking the Taylor series expansion of (3) around $\mathbf{p} = \mathbf{0}$ and maintaining up to first order terms we obtain

$$r_{ijk} = \frac{p_{ijk} G_{ij}}{\eta \ln 2} - \frac{p_{jik} G_{ji}}{\eta \ln 2}, \quad \forall i, j, k. \quad (4)$$

In matrix notation we have $\mathbf{r} = \mathbf{H}\mathbf{p}$, where the matrix $\mathbf{H}$ is appropriately defined. As outlined in the Introduction, we use this linear approximation in devising the structure of the optimal transmission policy. Later on in the paper, we abandon the linear approximation and derive policies using the exact form of transmission rates (cf. (3)).

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

$$\begin{aligned} & p_{ijk} p_{uiv} = 0, && \forall i, j, k, u, v, \\ & p_{ijk} p_{iuv} = 0, && \forall (j, k) \neq (u, v), \\ & p_{ijk} p_{ujv} = 0, && \forall (i, k) \neq (u, v), \ j \leq N, \\ & 0 \leq p_{ijk} \leq \bar{p}_{ijk}, && \forall i, j, k. \end{aligned}$$

These conditions respectively state that at any point in time *(i)* nodes cannot transmit and receive simultaneously, *(ii)* can only transmit traffic of a single class to a single node and, *(iii)* except for gateways, nodes can receive only a single traffic class from a single node. We denote by $\mathscr{P}$ the set of all $\mathbf{p} \in \mathbb{R}^{(N+M)^2 K}$ that satisfy the conditions above. We call *valid* a transmission scheme with $\mathbf{p} \in \mathscr{P}$. Clearly $\mathscr{P}$ is bounded. We also denote $\mathscr{R} = \{\mathbf{r} \mid \mathbf{r} = \mathbf{H}\mathbf{p}, \ \mathbf{p} \in \mathscr{P}\}$.

The next lemma, whose proof is omitted due to space limitations, establishes some useful properties of $\mathscr{P}$, $\mathscr{R}$, and their convex hulls.

Lemma II.1 *(i) Conv($\mathscr{P}$) and Conv($\mathscr{R}$) are polytopes (i.e., bounded polyhedra). (ii) Conv($\mathscr{R}$) $= \{\mathbf{r} \mid \mathbf{r} = \mathbf{H}\mathbf{p}, \ \mathbf{p} \in$ Conv($\mathscr{P}$)$\}$. (iii) For any extreme point $\mathbf{r} \in$ Conv($\mathscr{R}$), there exists an extreme point $\mathbf{p} \in$ Conv($\mathscr{P}$) such that $\mathbf{r} = \mathbf{H}\mathbf{p}$.*

Suppose next that there are totally L valid transmission schemes. To every valid transmission scheme n corresponds a rate vector in $\mathscr{R}$, say $\mathbf{r}^n$. Let us consider the information flow in the network in a potentially large but finite time interval. Normalize the length of this interval to 1. At different times, the network may employ different transmission schemes, e.g., in order to implement multi-hop routing. Suppose that during this time interval, the network uses the L selected schemes only and spends a fraction of time α_n transmitting according to scheme $n = 1, \ldots, L$. Then the total amount of information delivered during this interval is characterized by $\mathbf{r} = \sum_{n=1}^{L} \alpha_n \mathbf{r}^n$. This is also the long-term average transmission rate vector.

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

$$\sum_{j=1}^{N+M} r_{ijk} = 0, \qquad \forall i \neq s(k), d(k), \ \forall k,$$

that is class k traffic flow into i equals class k traffic outflow from node i.

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 Conv($\mathscr{R}$). 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 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 $\mathbf{A} \in \mathbb{R}^{R \times (N+M)^2 K}$ and $\mathbf{b} \in \mathbb{R}^{(N+M)^2 K}$ are given. For example, these constraints can impose equality among all transmission rates. Let $\mathscr{S}$ be the set of rates that satisfy fairness constraints and flow conservation, i.e.,

$$\mathscr{S} \triangleq \Big\{ \mathbf{r} \mid \mathbf{A}\mathbf{r} \leq \mathbf{b}, \sum_{j=1}^{N+M} r_{ijk} = 0,$$

$$\forall i \neq s(k), d(k), \forall k \Big\} \quad (5)$$

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

We can formulate the utility optimization problem as

$$\begin{array}{ll} \max & F(\mathbf{r}) \\ \text{s.t.} & \mathbf{r} \in \text{Conv}(\mathscr{R}) \cap \mathscr{S}. \end{array}$$

An important observation is that we seek to maximize utility over the convex hull of $\mathscr{R}$ rather than $\mathscr{R}$ itself (as for example in earlier work, e.g., [9]). This is bound to yield higher system utility and as we have seen the WSNET operates by time-sharing among different transmission schemes.

Let $\mathbf{r}^1, \ldots, \mathbf{r}^L$ denote the extreme points of $\text{Conv}(\mathscr{R})$. Any $\mathbf{r} \in \text{Conv}(\mathscr{R})$ can be expressed as a convex combination of those. Incorporating the definition of $\mathscr{S}$ and writing it as a minimization problem, the problem above is equivalent to

$$\begin{array}{ll} \min & -F(\mathbf{r}) \\ \text{s.t.} & \mathbf{r} - \sum_{n=1}^{L} \alpha_n \mathbf{r}^n = 0, \\ & \sum_{n=1}^{L} \alpha_n = 1, \\ & \sum_{j=1}^{N+M} r_{ijk} = 0, \ \forall i \neq s(k), d(k), \forall k, \\ & \mathbf{A}\mathbf{r} \leq \mathbf{b}, \\ & \alpha_n \geq 0, \ n = 1, \ldots, L. \end{array} \quad (6)$$

Note that $\mathbf{r}^1, \ldots, \mathbf{r}^L$ are also points of $\mathscr{R}$, thus, there exist corresponding valid transmission schemes (i.e., points in $\mathscr{P}$) $\mathbf{p}^1, \ldots, \mathbf{p}^L$ with $\mathbf{r}^n = \mathbf{H}\mathbf{p}^n$ for all $n = 1, \ldots, L$. The problem above maximizes a concave function over a polyhedron. It can be solved using, for example, the *conditional gradient method* [16]. If $F(\mathbf{r})$ is linear, then it is a linear programming problem for which very efficient algorithms exist.

Of course, $\text{Conv}(\mathscr{R})$ can have a humongous number of extreme points and this is the key challenge in solving (6). A simpler version of (6), maximizing throughput and with no fairness constraints, was considered in [8] and proposed to be solved by simply enumerating all extreme points and including them in the formulation (6). As indicated in [8] and clearly illustrated in Section X, this approach can quickly run out of steam (that is, memory) in very small networks. As we will see in Section IV, there are more efficient ways to solve (6). The decomposition algorithm we propose does not need to know $\mathbf{r}^1, \ldots, \mathbf{r}^L$ (or equivalently, the corresponding transmission schemes) in advance. It generates them as needed and identifies the ones that should be used to achieve optimality.

III. The Importance of Being Fair

Before we proceed with our agenda we demonstrate why it is important to explicitly include fairness constraints in the proposed framework. To this end, we consider a special case.

Consider a WSNET with a single gateway where all information transmitted by the sensor nodes is intended for this gateway. The objective is to maximize total throughput. This problem can be casted in the general framework of Section II. More specifically, $M = 1$ and there are N sensor nodes each of which transmits traffic intended for the gateway. Thus, there are N traffic classes and we let class i be associated with sensor node i for $i = 1, \ldots, N$. Let us adopt the notation of Section II and suppose no fairness constraints are enforced. The net flow out of node i equals $\sum_{j=1}^{N+1} r_{iji}$, thus, the total throughput is given by $\sum_{i=1}^{N} \sum_{j=1}^{N+1} r_{iji}$. The throughput maximization problem becomes (cf. (6)):

$$\begin{array}{ll} \max & \sum_{i=1}^{N} \sum_{j=1}^{N+1} r_{iji} \\ \text{s.t.} & \mathbf{r} \in \text{Conv}(\mathscr{R}), \\ & \sum_{j=1}^{N+1} r_{ijk} = 0, \ \forall i \neq k, N+1, \forall k. \end{array} \quad (7)$$

Theorem III.1 *Optimality for problem (7) can be achieved without time division. Furthermore, it is optimal for every node to transmit directly to the gateway.*

Proof: Let us first relax the flow conservation constraints and consider the following problem

$$\begin{array}{ll} \max & \sum_{i=1}^{N} \sum_{j=1}^{N+1} r_{iji} \\ \text{s.t.} & \mathbf{r} \in \text{Conv}(\mathscr{R}). \end{array} \quad (8)$$

The objective function is linear and the feasible set is a polytope, hence, there always exists an optimal solution $\mathbf{r}^*$ which is an extreme point of $\text{Conv}(\mathscr{R})$. $\mathbf{r}^*$ is also in $\mathscr{R}$, thus, no time division is needed to achieve optimality. Since there is no time-sharing and the transmission scheme corresponding to $\mathbf{r}^*$ has to be valid, it follows that the optimal strategy for problem (8) is for every node to send directly to the gateway. For such an $\mathbf{r}^*$ conservation constraints are satisfied and $\mathbf{r}^*$ solves (7). ■

Theorem III.1 states that the throughput is maximized when all nodes transmit directly to the gateway at the maximum rate allowed by the Shannon capacity. This implies that nodes close to the gateway (i.e., with high channel gains) have a significant advantage over nodes that happen to be further away. This is a rather unfair operation of the WSNET and is due to the wireless medium rather than nodes' actual needs. In WSNETs collecting data, for example, it can introduce a "geographic" bias into the data collection process. One way to mitigate it is to explicitly introduce fairness constraints into the problem formulation. The resulting strategy could use multi-hop routing (i.e., where nodes far away use other nodes as relays to reach the gateway) to achieve a more balanced operation.

IV. A Decomposition Method

In this section we propose a decomposition method for solving (6). For linear utilities the method is a *column generation* method for solving large-scale linear programming problems [17]. 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 (6), to which we will be referring as the *master problem*. Let $(\boldsymbol{\lambda}, \mu, \boldsymbol{\sigma}, \boldsymbol{\nu})$ be the dual vectors, then the dual function $G(\boldsymbol{\lambda}, \mu, \boldsymbol{\nu}, \boldsymbol{\sigma})$ is given by

$$G(\boldsymbol{\lambda}, \mu, \boldsymbol{\nu}, \boldsymbol{\sigma}) = \inf_{\boldsymbol{\alpha} \geq \mathbf{0}, \mathbf{r}} \Big\{ -F(\mathbf{r}) + \boldsymbol{\lambda}'(\mathbf{r} - \textstyle\sum_n \alpha_n \mathbf{r}^n) + \mu(\sum_n \alpha_n - 1) + \boldsymbol{\sigma}'(\mathbf{A}\mathbf{r} - \mathbf{b})$$

$$+ \sum_k \sum_{i\neq s(k),d(k)} \nu_{ik} \sum_j r_{ijk} \Big\}$$
$$= G_1(\boldsymbol{\lambda},\boldsymbol{\nu},\boldsymbol{\sigma}) + G_2(\boldsymbol{\lambda},\mu) - \mu - \boldsymbol{\sigma}'\mathbf{b},$$

where

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

Let

$$\mathscr{D}_1 = \{(\boldsymbol{\lambda},\boldsymbol{\nu},\boldsymbol{\sigma}) \mid G_1(\boldsymbol{\lambda},\boldsymbol{\nu},\boldsymbol{\sigma}) > -\infty\}$$
$$\mathscr{D}_2 = \{(\boldsymbol{\lambda},\mu) \mid G_2(\boldsymbol{\lambda},\mu) > -\infty\}$$

and note that

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

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

and $\mathscr{D}_1$ is independent of $\mathbf{r}^1,\ldots,\mathbf{r}^L$. Then the dual of the master problem (6) is

$$\begin{array}{ll} \max & G_1(\boldsymbol{\lambda},\boldsymbol{\nu},\boldsymbol{\sigma}) - \mu - \boldsymbol{\sigma}'\mathbf{b} \\ \text{s.t.} & (\boldsymbol{\lambda},\boldsymbol{\nu},\boldsymbol{\sigma}) \in \mathscr{D}_1, \\ & \mu - \boldsymbol{\lambda}'\mathbf{r}^n \geq 0, \quad n = 1,\ldots,L, \\ & \boldsymbol{\sigma} \geq \mathbf{0}. \end{array} \tag{9}$$

Since the master problem is a convex optimization problem there is no duality gap [16].

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

$$\begin{array}{ll} \min & -F(\mathbf{r}) \\ \text{s.t.} & \mathbf{r} - \sum_{n=1}^m \alpha_n \mathbf{r}^n = 0, \\ & \sum_{n=1}^m \alpha_n = 1, \\ & \sum_{j=1}^{N+M} r_{ijk} = 0,\ \forall i \neq s(k), d(k), \forall k, \\ & \mathbf{A}\mathbf{r} \leq \mathbf{b}, \\ & \alpha_n \geq 0,\ n = 1,\ldots,m, \end{array} \tag{10}$$

which we call the *restricted master problem* at the mth iteration. Suppose we solve this problem to optimality. The dual of this problem is identical to (9) with the exception that only constraints $\mu - \boldsymbol{\lambda}'\mathbf{r}^n \geq 0$, for $n = 1,\ldots,m$, appear. We refer to this latter problem as the *restricted dual problem* at the mth 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 $(\boldsymbol{\lambda}^{(m)}, \boldsymbol{\nu}^{(m)}, \boldsymbol{\sigma}^{(m)}) \in \mathscr{D}_1$, $\boldsymbol{\sigma}^{(m)} \geq \mathbf{0}$, and $\mu^{(m)} - \boldsymbol{\lambda}^{(m)'}\mathbf{r}^n \geq 0$, for all $n = 1,\ldots,m$. If it happens that $\mu^{(m)} - \boldsymbol{\lambda}^{(m)'}\mathbf{r}^n \geq 0$ for all $n = 1,\ldots,L$ then we have a primal-dual pair for (6) and we are done. Otherwise, we need to generate an extreme point, say $\mathbf{r}^{m+1}$, of Conv($\mathscr{R}$) 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 at every step an extreme point that violates dual feasibility.

A. The subproblem

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

$$\begin{array}{ll} \max & \boldsymbol{\lambda}'\mathbf{r} \\ \text{s.t.} & \mathbf{r} = \mathbf{H}\mathbf{p}, \\ & \mathbf{p} \in \mathscr{P}, \end{array} \tag{11}$$

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

Before we proceed showing that the proposed decomposition algorithm converges we establish some properties of (11). $\boldsymbol{\lambda} \in \mathbb{R}^{(M+N)^2K}$ is the dual vector corresponding to the first constraint of (6). Denote by λ_{ijk} the element of $\boldsymbol{\lambda}$ corresponding to r_{ijk} and let $\pi_{ijk} = \lambda_{ijk} - \lambda_{jik}$. Then

$$\begin{aligned} \boldsymbol{\lambda}'\mathbf{H}\mathbf{p} &= \sum_{k=1}^K \sum_{i=1}^{N+M} \sum_{j=1}^{N+M} \lambda_{ijk} \frac{p_{ijk}G_{ij} - p_{jik}G_{ji}}{\eta \ln 2} \\ &= \sum_{k=1}^K \sum_{i=1}^{N+M} \sum_{j=1}^{N+M} \frac{\pi_{ijk}G_{ij}}{\eta \ln 2} p_{ijk}, \end{aligned}$$

hence the subproblem is equivalent to

$$\begin{array}{ll} \max & \sum_{k=1}^K \sum_{i=1}^{N+M} \sum_{j=1}^{N+M} \frac{\pi_{ijk}G_{ij}}{\eta \ln 2} p_{ijk} \\ \text{s.t.} & \mathbf{p} \in \mathscr{P}. \end{array} \tag{12}$$

Next we reduce it to an *integer linear programming problem (ILP)*.

Proposition IV.1 *Problem (12) is equivalent to the ILP:*

$$\begin{array}{ll} \textit{max} & \sum_{(i,j,k)|\psi_{ijk}>0} \psi_{ijk} s_{ijk} \\ \textit{s.t.} & \sum_{j=1}^{N+M} \sum_{k=1}^{K} s_{ijk} + \sum_{j=1}^{N+M} \sum_{k=1}^{K} s_{jik} \leq 1,\ \forall i \leq N, \\ & 0 \leq s_{ijk} \leq I_{ijk}, \\ & s_{ijk} \in \{0,1\}, \end{array} \tag{13}$$

where $\psi_{ijk} = \frac{\pi_{ijk}\bar{p}_{ijk}G_{ij}}{\eta \ln 2}$ *and* $I_{ijk} = 1(\psi_{ijk} > 0)$.

Proof: Note that there always exists an optimal solution $\mathbf{p}^*$ to the problem (12) satisfying the conditions

$$\begin{aligned} p^*_{ijk} &\in \{0, \bar{p}_{ijk}\}, && \text{if } \psi_{ijk} > 0, \\ p^*_{ijk} &= 0, && \text{otherwise.} \end{aligned}$$

Letting

$$s_{ijk} = \begin{cases} 1, & \text{if } p_{ijk} = \bar{p}_{ijk} > 0, \\ 0, & \text{otherwise,} \end{cases}$$

we obtain that the problem (12) is equivalent to:

$$\begin{array}{ll} \max & \sum_{(i,j,k)|\psi_{ijk}>0} \psi_{ijk} s_{ijk} \\ \text{s.t.} & s_{ijk} + s_{uiv} \leq 1,\ \forall i,j,k,u,v, \\ & s_{ijk} + s_{iuv} \leq 1,\ \forall (j,k) \neq (u,v), \\ & s_{ijk} + s_{ujv} \leq 1,\ \forall (i,k) \neq (u,v),\ j \leq n, \\ & 0 \leq s_{ijk} \leq I_{ijk}, \\ & s_{ijk} \in \{0,1\}. \end{array} \tag{14}$$

In particular, $\mathbf{s}^*$ is an optimal solution of the above if and only if $\mathbf{p}^*$ satisfying

$$p^*_{ijk} = \begin{cases} \bar{p}_{ijk}, & \text{if } s_{ijk} = 1, \\ 0, & \text{otherwise}, \end{cases} \tag{15}$$

is an optimal solution of (12). Writing (14) in a more compact way we obtain (13). ■

We summarize the discussion on the subproblem as follows: to compute an optimal solution $\mathbf{r}^*$ of (11) we first solve (13) to obtain an optimal solution $\mathbf{s}^*$, then compute $\mathbf{p}^*$ as in (15), and finally compute $\mathbf{r}^* = \mathbf{H}\mathbf{p}^*$. It is evident from the proof of Proposition IV.1 that $\mathbf{s}^*$ prescribes how to operate the network under the transmission scheme $\mathbf{p}^*$: (i, j, k) transmissions occur only if $s_{ijk} = 1$ and if so at maximum power.

B. The decomposition algorithm

We now have all the ingredients to present the decomposition algorithm and show its convergence. The algorithm is in Fig. 1 and the next theorem establishes its convergence. In the sequel, we assume that (6) is feasible; we will discuss at the end of this Section how this assumption can be relaxed.

---

1) **Initialization**: Let $\mathbf{r}^1 \in \text{Conv}(\mathscr{R}) \cap \mathscr{S}$ and set $m = 1$.
2) m**-th iteration**:
   a) Solve the restricted master problem (10) with $\mathbf{r}^1, \ldots, \mathbf{r}^m$ 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 (11) with cost vector $\boldsymbol{\lambda}^{(m)}$ as outlined in Section IV-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 (6). Otherwise, set $m := m + 1$ and go to step 2a.

Fig. 1. The decomposition algorithm.

Theorem IV.2 *Assume that (6) is feasible. Then the decomposition algorithm of Fig. 1 terminates with an optimal solution of (6) in a finite number of iterations.*

Proof: Recall that at the m-th iteration the subproblem minimizes $\mu^{(m)} - \boldsymbol{\lambda}^{(m)'}\mathbf{r}$ over $\mathbf{r} \in \mathscr{R}$. 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 $\mathbf{r} \in \mathscr{R}$. Since all extreme points of Conv($\mathscr{R}$) are in $\mathscr{R}$, the latter condition implies that $\mu^{(m)} - \boldsymbol{\lambda}^{(m)'}\mathbf{r}^n \geq 0$ for all extreme points $\mathbf{r}^1, \ldots, \mathbf{r}^L$ of Conv($\mathscr{R}$). 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 (6) and the algorithm terminates.

Next note that due to Proposition IV.1 and the resulting structure of the subproblem solutions, at each iteration we generate a transmission scheme in $\{\mathbf{p} \in \mathscr{P} \mid p_{ijk} \in \{0, \bar{p}_{ijk}\}\}$ which contains all extreme points of Conv($\mathscr{P}$). Let $\mathbf{p}^1, \ldots, \mathbf{p}^m$ the transmission schemes generated up to the m-th iteration and suppose the algorithm does not terminate at the m-th iteration. The next transmission scheme to be generated, $\mathbf{p}^{m+1}$, 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

$$\mu^{(m)} - \boldsymbol{\lambda}^{(m)'}\mathbf{H}\mathbf{p}^n \geq 0, \quad n = 1, \ldots, m,$$
$$\mu^{(m)} - \boldsymbol{\lambda}^{(m)'}\mathbf{H}\mathbf{p}^{m+1} < 0.$$

Thus, at each iteration we generate a new point of the finite set $\{\mathbf{p} \in \mathscr{P} \mid p_{ijk} \in \{0, \bar{p}_{ijk}\}\}$. Hence, the algorithm terminates 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 $\mathbf{r}^1 \in \text{Conv}(\mathscr{R}) \cap \mathscr{S}$. 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 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*

$$\begin{array}{ll} \min & -\sum_{i=1}^{R} y_i \\ \text{s.t.} & \mathbf{r} - \sum_{n=1}^{L} \alpha_n \mathbf{r}^n = 0, \\ & \sum_{n=1}^{L} \alpha_n = 1, \\ & \sum_{j=1}^{N+M} r_{ijk} = 0, \ \forall i \neq s(k), d(k), \forall k, \\ & \mathbf{A}\mathbf{r} + \mathbf{y} = \mathbf{b}, \\ & \alpha_n \geq 0, \ n = 1, \ldots, L, \end{array} \tag{16}$$

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 (16) is almost identical to (9) 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 (16) satisfies $\mathbf{y} \geq \mathbf{0}$ then we are done as we have a feasible solution of (6) to initialize the algorithm in Fig. 1 (this could involve time-sharing between several transmission schemes). Otherwise, (6) is infeasible.

V. Solving the Subproblem

The efficiency of the algorithm of Fig. 1 critically depends on how efficiently we can solve the subproblem. As outlined in Section IV-A, solving the subproblem amounts to solving an ILP. 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. Fortunately, our subproblem has enough of special structure that makes it polynomially solvable. In this section, we establish that (13) is equivalent to a *maximum weighted matching* problem, which is polynomially solvable.

Let us define the following sets: $\mathscr{A} = \{1, \ldots, N\}$ and $\mathscr{B}_l = \{Nl + 1, \ldots, Nl + N\}$ for $l = 1, \ldots, M$. Each element of $\mathscr{A}$ corresponds to a sensor node of the WSNET and set $\mathscr{B}_l$ corresponds to the gateway l of the WSNET. Let $\mathscr{V} = \mathscr{A} \cup (\cup_{l=1}^{M} \mathscr{B}_l)$ and consider the undirected graph $\mathscr{G} = (\mathscr{V}, \mathscr{E})$, where $\mathscr{E}$ is the complete set of edges between nodes in $\mathscr{V}$. With each edge $(i,j) \in \mathscr{E}$ we associate a weight w_{ij} such that

$$w_{ij} = \begin{cases} \max_{k=1,\ldots,K} \max\{\psi_{ijk}, \psi_{jik}\}, & \forall i, j \in \mathscr{A}, \\ \max_{k=1,\ldots,K} \max\{\psi_{i,N+l,k}, 0\}, & \forall i \in \mathscr{A}, j \in \mathscr{B}_l, \\ \max_{k=1,\ldots,K} \max\{\psi_{j,N+l,k}, 0\}, & \forall i \in \mathscr{B}_l, j \in \mathscr{A}, \\ 0, & \text{otherwise}. \end{cases} \tag{17}$$

Note that $w_{ij} = w_{ji} \geq 0, \forall i, j$. Also for any i, u, v, if $u, v \in \mathscr{B}_l$ for some l, then $w_{iu} = w_{iv}$, that is, the weight of the link between i and any node in $\mathscr{B}_l$ is the same. Let us also construct a set $\mathscr{K}$ as follows: for each $1 \leq i \leq N, 1 \leq j \leq N + M$, we select only one, if any, k satisfying the conditions

$$k = \begin{cases} \text{argmax}_{t=1,\ldots,K} \max\{\psi_{ijt}, \psi_{jit}\}, & \text{if } j \leq N, \\ \text{argmax}_{t=1,\ldots,K} \max\{\psi_{ijt}, 0\}, & \text{otherwise}, \end{cases}$$

and $\psi_{ijk} > 0$, and let (i, j, k) be an element of $\mathscr{K}$.

The next theorem establishes that solving the subproblem amounts to solving a maximum weighted matching for graph $\mathscr{G}$ where edge weights are given in (17). We omit the proof due to space limitations.

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

$$\begin{array}{rll} \textit{max} & \sum_{(i,j)\in\mathscr{E}} w_{ij} x_{ij} & \\ \textit{s.t.} & \sum_{j|(i,j)\in\mathscr{E}} x_{ij} \leq 1, & \forall i \\ & x_{ij} = x_{ji}, \quad \forall i, j, & \\ & x_{ii} = 0, \quad \forall i, & \\ & x_{ij} \in \{0, 1\}, \quad \forall i, j. & \end{array} \tag{18}$$

Then, an optimal solution $\mathbf{s}^*$ to the subproblem (13) satisfies

$$s^*_{ijk} = \begin{cases} 1_{\mathscr{K}}(i,j,k) x^*_{ij}, & \textit{if } 1 \leq i, j \leq N, \\ 1_{\mathscr{K}}(i,j,k) \sum_{v \in \mathscr{B}_{j-N}} x^*_{iv}, & \textit{if } 1 \leq i \leq N, \\ & \textit{and } j \geq N+1, \\ 0, & \textit{otherwise.} \end{cases}$$

Remark : It should be noted that (18) is always feasible ($\mathbf{x} = \mathbf{0}$ is a feasible solution), thus, it is always possible to obtain an optimal solution of the subproblem as specified above.

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 (18) can be solved in $O(|\mathscr{V}|^3)$ amount of time [18], that is, polynomial in the size of the input. In our case, $|\mathscr{V}| = (M + 1)N$ and it takes $O(KN(N + M))$ additional time to calculate the weights and obtain $\mathbf{s}^*$ from $\mathbf{x}^*$, thus, subproblem's complexity is $O(KN(N + M) + (M + 1)^3 N^3)$.

VI. A TRANSMISSION POLICY WITHOUT THE LINEAR APPROXIMATION

In this section we outline how to remove the linear approximation of transmission rates (cf. (4)) and obtain a policy under the exact expression of (3). We should note that with the exact rate function, Conv($\mathscr{R}$) becomes extremely difficult to characterize, and makes problem (6) intractable. Earlier attempts in the literature used either approximation techniques, e.g., discretization, or restricted routing strategies [8, 9]. In this work, we use linearization to obtain the structure of the policy and then remove the linearization to devise a policy under the exact expression (3) for transmission rates.

More specifically, we first solve (6) using the linear approximation in (4) as outlined in Section IV and obtain a set of transmission schemes under which the network will operate; let $\mathbf{r}^1, \ldots, \mathbf{r}^D$ be the corresponding rates. Based on the discussion in the previous sections, we know that for each transmission vector $\mathbf{r}^n$ $(n = 1, \ldots, D)$, $r^n_{ijk} > 0$ implies that node i transmits class k traffic to node j. Now letting each node use the maximum available power if it transmits, the modified transmission vector $\tilde{\mathbf{r}}^n$ corresponding to $\mathbf{r}^n$ is given by

$$\tilde{r}^n_{ijk} = W \log_2(1 + \tilde{\gamma}^n_{ijk}) - W \log_2(1 + \tilde{\gamma}^n_{jik}),$$

where for any (i, j, k) and $n = 1, \ldots, D$

$$\tilde{\gamma}^n_{ijk} = \frac{1(r^n_{ijk} > 0)\bar{p}_{ijk} G_{ij}}{\eta W + \sum_{v=1}^{K} \sum_{l=1, l\neq i}^{N+M} \sum_{u=1}^{N+M} 1(r^n_{luv} > 0)\bar{p}_{luv} G_{lj}}.$$

Next we use the modified transmission vectors and solve the following utility maximization problem

$$\begin{array}{rl} \max & F(\tilde{\mathbf{r}}) \\ \text{s.t.} & \tilde{\mathbf{r}} - \sum_{n=1}^{D} \tilde{\alpha}_n \tilde{\mathbf{r}}^n = 0, \\ & \sum_{n=1}^{D} \tilde{\alpha}_n = 1, \\ & \sum_{j=1}^{N+M} \tilde{r}_{ijk} = 0, \ \forall i \neq s(k), d(k), \forall k, \\ & \mathbf{A}\tilde{\mathbf{r}} \leq \mathbf{b}, \\ & \tilde{\alpha}_n \geq 0, \ n = 1, \ldots, D, \end{array} \tag{19}$$

with decision variables $\tilde{\mathbf{r}}$ and $\tilde{\boldsymbol{\alpha}}$. The optimal solution provides a transmission policy time-sharing among the schemes with rates $\tilde{\mathbf{r}}^1, \ldots, \tilde{\mathbf{r}}^D$. Note that we solve problem (19) without further iterations of the decomposition method to add more transmission schemes. That is, we adopt and fix the transmission schemes obtained under the linear approximation. It can be seen that if $\mathbf{b} \geq \mathbf{0}$, then problem (19) is feasible if we add $\mathbf{0}$ to the allowable transmission schemes $\tilde{\mathbf{r}}^1, \ldots, \tilde{\mathbf{r}}^D$. (This is always possible if we set $\bar{p}_{ijk} = 0$ for all (i, j, k).) However, feasibility in general is not guaranteed.

The line of development so far implies that the policy obtained from (19) is asymptotically optimal as $\mathbf{p}$ approaches $\mathbf{0}$. As we will illustrate later on with numerical examples the devised policy remains close to the optimal even when $\mathbf{p}$ is far away from $\mathbf{0}$. To summarize, once we have a set of transmission schemes obtained under our linear approximation we can easily produce a policy for the original system under the exact expressions for transmission rates. In the sequel, we turn back to discuss some other interesting properties related to problem (6).

VII. Optimization over Power Limits

So far we have assumed that the power limits $\bar{p}_i$ of all sensor nodes $i = 1, \ldots, N$ are fixed. As we will see higher power limits lead to higher utility, but, of course, higher energy consumption. As energy preservation is critical in WSNETs, it becomes of interest to optimize the power limits used by the sensor nodes to achieve a certain utility target. In this section we discuss how this can be accomplished.

Let us view the utility maximization problem formulated in Section II as parametrized by the vector of power limits, denoted by $\bar{\mathbf{p}} = (\bar{p}_1, \ldots, \bar{p}_N)$. Consider

$$\begin{aligned} \bar{F}(\bar{\mathbf{p}}) \stackrel{\triangle}{=} \max \quad & F(\mathbf{r}) \\ \text{s.t.} \quad & \mathbf{r} \in \text{Conv}(\mathscr{R}(\bar{\mathbf{p}})) \cap \mathscr{S}, \end{aligned} \tag{20}$$

where we write $\mathscr{R}(\bar{\mathbf{p}})$ to explicitly denote the fact that the set of transmission rate vectors depends on $\bar{\mathbf{p}}$. We denote by $F(\bar{\mathbf{p}})$ the optimal value of the above. The first, and rather intuitive, property we show is monotonicity.

Theorem VII.1 *(Monotonicity) Suppose $\bar{\mathbf{p}}^1$ and $\bar{\mathbf{p}}^2$ are two vectors of power limits. If $\bar{\mathbf{p}}^1 \geq \bar{\mathbf{p}}^2$, then $\bar{F}(\bar{\mathbf{p}}^1) \geq \bar{F}(\bar{\mathbf{p}}^2)$.*

Proof: Suppose we have two WSNETs A and B with no difference other than the power limit vectors. Let $\bar{\mathbf{p}}^1$ and $\bar{\mathbf{p}}^2$ the power limit vectors in WSNETs A and B, respectively. As $\bar{\mathbf{p}}^1 \geq \bar{\mathbf{p}}^2$, A can implement the optimal transmission strategy for B. The optimal strategy for A can be no worse. ■

Next we show that the optimal utility is concave in $\bar{\mathbf{p}}$ if we scale the power limits uniformly. The proof is omitted due to the space limitations.

Theorem VII.2 *(Concavity) Suppose the power limit vector $\bar{\mathbf{p}}$ belongs to the set $\mathscr{W} = \{\bar{\mathbf{p}} \mid \bar{\mathbf{p}} = \phi\bar{\mathbf{p}}^0, \phi > 0\}$ where $\bar{\mathbf{p}}^0 > \mathbf{0}$ is a constant vector. Then $\bar{F}(\bar{\mathbf{p}})$ is concave in $\bar{\mathbf{p}}$ over $\mathscr{W}$.*

The above theorem is critical in trading-off energy consumption with achieved utility. Suppose we are interested in minimizing energy consumption subject to achieving a utility level equal to some given value, say $F_{\min}$. Assuming that power limits are scaled uniformly for the whole network by a factor ϕ, we can formulate the problem as

$$\begin{aligned} \min \quad & \phi \\ \text{s.t.} \quad & \bar{F}(\phi\bar{\mathbf{p}}_0) \geq F_{\min}, \end{aligned} \tag{21}$$

where $\bar{F}(\phi\bar{\mathbf{p}}_0)$ is defined in (20). Theorem VII.2 asserts that the above is a convex optimization problem, thus, a global minimum, say ϕ^*, can be obtained using standard gradient-based algorithms [16]. One complicating factor is that closed form expressions for $\bar{F}(\phi\bar{\mathbf{p}}_0)$ and its derivative are not available. The decomposition algorithm of Fig. 1 can evaluate $\bar{F}(\phi\bar{\mathbf{p}}_0)$ and its derivative can be obtained using finite differences.

VIII. The Lifetime of Sensor Networks

In this section we consider what are the implications of power optimization to the lifetime of the network. Let us assume that the energy expended by the sensors to receive and decode information is negligible compared to the energy expended while transmitting.

We define the lifetime T of a WSNET as the length of time during which no node runs out of energy resources. As we have seen, the transmission policies we consider time-share among several transmission schemes. Assume that T is in a much longer time-scale than the time-scale in which the policy switches among the various transmission schemes. Let $\mathbf{p}$ be the vector of average powers consumed during a long time-interval, i.e., $\mathbf{p}$ is the time average of the power vectors corresponding to all transmission schemes employed by the transmission policy. For each node $i = 1, \ldots, N$ set $\mathbf{c}_i$ such that $\mathbf{c}_i'\mathbf{p} = \sum_{j=1}^{M+N} \sum_{k=1}^{K} p_{ijk}$. Then $T \leq \frac{\chi_i}{\mathbf{c}_i'\mathbf{p}}$, for all $i = 1, \ldots, N$, where χ_i is the available energy at sensor node i. In matrix notation, we write $\mathbf{Cp} \leq \boldsymbol{\chi}/T$, where $\mathbf{C}$ is an $N \times (N+M)^2 K$ matrix whose ith row equals $\mathbf{c}_i'$ and $\boldsymbol{\chi} = (\chi_1, \ldots, \chi_N)$.

To capture the trade-off between system utility and the lifetime of the WSNET, we propose the following utility maximization problem with parameter T:

$$\begin{aligned} \hat{F}(T) = \max \quad & F(\mathbf{Hp}) \\ \text{s.t.} \quad & \mathbf{p} = \textstyle\sum_{n=1}^{J} \alpha_n \mathbf{p}^n, \\ & \textstyle\sum_{n=1}^{J} \alpha_n = 1, \\ & \mathbf{Hp} \in \mathscr{S}, \\ & \mathbf{Cp} \leq \boldsymbol{\chi}/T, \\ & \alpha_n \geq 0, \quad n = 1, \ldots, J. \end{aligned} \tag{22}$$

where $\mathbf{p}^1, \ldots, \mathbf{p}^J$ are the extreme points of Conv($\mathscr{P}$). Notice that in problem (22), we still seek to maximize the system utility with time division, under fairness and flow conservation constraints. The difference here is that we add a hard constraint on the lifetime of the WSNET.

Problem (22) is a convex programming problem. (Note that the objective function $F(\mathbf{Hp})$ is concave in $\mathbf{Hp}$ and therefore concave in $\mathbf{p}$, and all the constraints are linear in $\mathbf{p}$ and $\boldsymbol{\alpha}$.) Consequently, we have a problem very similar to problem (6) and a complete analog of the decomposition method in Fig. 1 can be used to solve large-scale instances of (22).

In particular, suppose we have an extreme point $\mathbf{p}^1$ of Conv($\mathscr{P}$), which belongs to $\{\mathbf{p} \mid \mathbf{Hp} \in \mathscr{S}, \mathbf{Cp} \leq \boldsymbol{\chi}/T\}$ and let $m \in \{1, \ldots, J\}$, then the restricted master problem at mth iteration is

$$\begin{aligned} \min \quad & -F(\mathbf{Hp}) \\ \text{s.t.} \quad & \mathbf{p} - \textstyle\sum_{n=1}^{m} \alpha_n \mathbf{p}^n = 0, \\ & \textstyle\sum_{n=1}^{m} \alpha_n = 1, \\ & \sum_{j=1}^{N+M} \frac{p_{ijk}G_{ij} - p_{jik}G_{ji}}{\eta \ln 2} = 0, \ \forall i \neq s(k), d(k), \forall k, \\ & \mathbf{AHp} \leq \mathbf{b}, \\ & \mathbf{Cp} \leq \boldsymbol{\chi}/T, \\ & \alpha_n \geq 0, \ n = 1, \ldots, m, \end{aligned} \tag{23}$$

and the corresponding subproblem is

$$\begin{array}{ll} \max & \boldsymbol{\lambda}'\mathbf{p} \\ \text{s.t.} & \mathbf{p} \in \mathscr{P}, \end{array} \tag{24}$$

with cost vector $\boldsymbol{\lambda} = \boldsymbol{\lambda}^{(m)}$, where as before we have $(\mathbf{p}^{(m)}, \boldsymbol{\alpha}^{(m)}; \boldsymbol{\lambda}^{(m)}, \mu^{(m)}, \boldsymbol{\nu}^{(m)}, \boldsymbol{\sigma}^{(m)}, \boldsymbol{\xi}^{(m)})$ as the optimal primal-dual pair for the restricted master problem (23) with $\boldsymbol{\xi}^{(m)}$ being the optimal dual variable corresponding to the lifetime constraint.

Following the same recipe as described in Fig. 1, we can obtain an optimal solution to problem (22) in a finite number of iterations. The argument is almost identical and is therefore omitted for brevity. Furthermore, problem (24) is still equivalent to the maximum weighted matching problem constructed in a similar way as in Section V, and is solvable in polynomial time. In particular, to solve problem (24), we construct the same undirected graph $\mathscr{G} = (\mathscr{V}, \mathscr{E})$. Let $\delta_{ijk} = \lambda_{ijk}\bar{p}_{ijk}$ $\forall i, j, k$, and the weight for each edge $(i, j) \in \mathscr{E}$ is given by

$$w_{ij} = \begin{cases} \max\limits_{k=1,\ldots,K} \max\{\delta_{ijk}, \delta_{jik}, 0\}, & \forall i, j \in \mathscr{A}, \\ \max\limits_{k=1,\ldots,K} \max\{\delta_{i,N+l,k}, 0\}, & \forall i \in \mathscr{A}, j \in \mathscr{B}_l, \\ \max\limits_{k=1,\ldots,K} \max\{\delta_{j,N+l,k}, 0\}, & \forall i \in \mathscr{B}_l, j \in \mathscr{A}, \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, we construct the set $\mathscr{K}$ as follows. For each $1 \le i \le N$, $1 \le j \le N+M$, we select only one, if any, k satisfying the conditions

$$k = \begin{cases} \text{argmax}_{t=1,\ldots,K} \max\{\delta_{ijt}, \delta_{jit}, 0\}, & \text{if } j \le N, \\ \text{argmax}_{t=1,\ldots,K} \max\{\delta_{ijt}, 0\}, & \text{otherwise,} \end{cases}$$

and

$$\begin{cases} \delta_{ijk} = \max\{\delta_{ijk}, \delta_{jik}, 0\} > 0, & \text{if } j \le N, \\ \delta_{ijk} > 0, & \text{otherwise,} \end{cases}$$

and let (i, j, k) be an element of $\mathscr{K}$. Given the graph $\mathscr{G}$ constructed above, we can obtain the optimal solution of (24) in the same way as in Section V.

It is interesting to examine how the lifetime parameter T affects the system utility. The first observation is monotonicity. Namely, if T decreases, the feasible set of (22) gets larger, thus, the optimal system utility can be no smaller. The maximum possible system utility is corresponding to the case $T = 0$. Furthermore, as problem (22) is a concave maximization problem, a standard convexity argument shows that $\hat{F}(T)$ is concave in $1/T$. We summarize the above observations in the following theorem.

Theorem VIII.1 *$\hat{F}(T)$ is monotonically nonincreasing in T, upper bounded by $\hat{F}(0)$, and concave in $1/T$.*

Though $\hat{F}(T)$ is in general non-convex in T, Thm. VIII.1 suggests how to trade-off utility vs. the lifetime T. The first observation is that our algorithm can be used to efficiently obtain a transmission policy for any desirable T. Moreover, it allows us to solve an optimization problem of the form

$$\max_T (\hat{F}(T) - \zeta/T), \tag{25}$$

for some scalar ζ. This can be interpreted as maximizing utility while paying a cost for short lifetime. Problem (25) is concave in $1/T$ and can be solved very efficiently using line search techniques (see [16]).

IX. Dealing with Node Failures

Next we discuss how to accommodate node failures in our decomposition framework. As we will see, our approach enables us to re-optimize and adjust accordingly the transmission policy in response to node failures.

Suppose we solve problem (6) for a WSNET and obtain the transmission vectors $\mathbf{r}^1, \ldots, \mathbf{r}^D$. If we detect that node l has failed, we do not have to solve the utility maximization problem from scratch. Instead, we make use of the following re-optimization technique: reuse the obtained transmission vectors and modify them to obtain a set of valid transmission vectors for the modified WSNET. In particular, the modified transmission vector $\tilde{\mathbf{r}}^n$ corresponding to $\mathbf{r}^n$ is given by

$$\tilde{r}^n_{ijk} = \begin{cases} 0, & \text{if } i = l, \text{ or } j = l, \text{ or } s(k) = l, \text{or } d(k) = l, \\ r^n_{ijk}, & \text{otherwise.} \end{cases}$$

Then we consider the following problem

$$\begin{array}{ll} \max & F(\tilde{\mathbf{r}}) \\ \text{s.t.} & \tilde{\mathbf{r}} - \sum_{n=1}^{D} \tilde{\alpha}_n \tilde{\mathbf{r}}^n = 0, \\ & \sum_{n=1}^{D} \tilde{\alpha}_n = 1, \\ & \sum_{j=1}^{N+M} \tilde{r}_{ijk} = 0, \ \forall i \ne s(k), d(k), \forall k, \\ & \mathbf{A}\tilde{\mathbf{r}} \le \mathbf{b}, \\ & \tilde{\alpha}_n \ge 0, \ n = 1, \ldots, D, \end{array} \tag{26}$$

and view the above problem as the restricted master problem. Let $\bar{p}_{ljk} = 0$ $\forall j, k$, and the subproblem has the same form as problem (11). Starting from (26) and transmission schemes with rates $\tilde{\mathbf{r}}^1, \ldots, \tilde{\mathbf{r}}^D$ iterate using the algorithm of Fig. 1 to derive an optimal transmission policy for the modified WSNET (where node l is removed).

Note that the coefficient matrix in problem (26) is sparse, as all the rows related to node l are forced to $\mathbf{0}$; this can be exploited to reduce the size of the problem and the resulting running time of the algorithm. In several cases, this re-optimization procedure results in much shorter running time than solving the original problem from scratch; we provide numerical results in the next section. We close this section by noting that multiple node failures can be similarly handled.

X. Numerical Results

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

Example 1: The first example we consider is a WSNET with sensor nodes uniformly distributed in the box $[-10m, 10m] \times [-10m, 10m]$. The network has a single gateway at the origin. We use the same identical parameters as in [8]. In particular,

$G_{ij} = KS_{ij}(d_0/d_{ij})^\alpha$, where $K = 10^{-6}$, $d_0 = 10$, d_{ij} is the distance between nodes i and j, $\alpha = 4$, $S_{ij} = S_{ji}$ are independent and identically generated from a lognormal distribution with a mean of 0dB and variance 8dB, and $\bar{p}_i = 0.1$ Watts for all nodes i. The noise is characterized by $\eta = 10^{-10}$ and $W = 10^6$.

Comparison with enumeration: We obtain a transmission policy using the approach outlined in Section VI, namely, we make the linear approximation to obtain the structure of the policy and use this structure to devise a policy under the exact transmission rate expressions of (3). We compare the policy we obtain in this fashion with what we call the *enumeration* approach proposed in [8]. This latter approach does not make the linear approximation we made in (4); it instead uses directly the exact expression for transmission rates given in (3). It solves (6) by enumerating all feasible transmission rate vectors in Conv($\mathscr{R}$). To that end, it discretizes the possible values $\mathbf{p} \in \mathscr{P}$ can take, generates all possible transmission schemes, and from those it derives the corresponding rate vectors $\mathbf{r}$. Table I contains the results. In all cases, the objective is to maximize total throughput (reported in bps) and the fairness constraints have the form $\rho_{i+1} \leq 2\rho_i$, $i = 1, \dots, N$, where ρ_i denotes the throughput of node i. The 1st column of Table I lists the number of nodes in the network. The 2nd and 3rd columns list the throughput achieved by the enumeration approach and the corresponding CPU time in seconds. The 4th and 5th columns list the throughput achieved by our algorithm and the corresponding CPU time in seconds. Finally, the last column reports the throughput achieved by the single-hop strategy, i.e., when each node sends directly to the gateway.

TABLE I

COMPUTATIONAL EFFICIENCY COMPARISON (SINGLE GATEWAY).

N	Enumeration	Time	Decomposition	Time	Single-hop
2	14.44	0.02	14.44	0.01	14.4
3	122.28	0.02	122.28	0.01	122.2
4	689.16	0.13	689.16	0.02	167.6
5	7962.63	63.4	7960.87	0.02	582.3
6	out of memory	-	6339.97	0.03	191.9

A couple of remarks are in order. First, comparing columns 2 and 4 of Table I suggests that even at power levels of 0.1 Watts our approach is very accurate. Typical sensor networks will operate at lower power levels which is bound to improve accuracy. Second, the inherent combinatorial explosion of possible transmission schemes limits the use of the enumeration method to very small instances (in the 6-node case we run out of memory). In comparison, computational requirements in our method scale rather nicely. Without particular effort at optimizing the code we can currently solve problem instances with 50 nodes in less than 1 minute. Third, it is interesting to note that time-sharing (multi-hop strategy) can dramatically improve performance over the naive single-hop strategy. For the cases reported in Table I the improvement is on the order of 3000%.

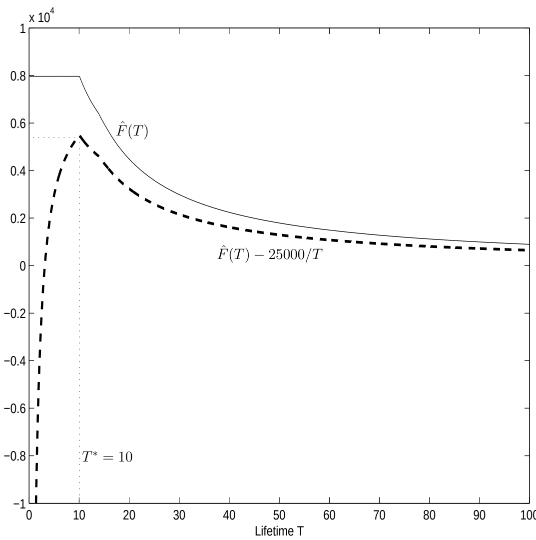

Fig. 2. Utility as a function of network lifetime.

Optimization over power limits: To demonstrate the effects of power optimization we considered the 5-node case in Table I. Setting $F_{\text{min}} = 6500$, the approach of Section VII yields $\phi = 0.81$. That is, sensors can scale down their power by ϕ and this is sufficient to achieve a throughput equal to F_{min}.

Node failure: The next part considers the re-optimization technique described in Sec. IX. We first calculate the optimal utility and rate vectors for a WSNET with N sensor nodes and then let node N fail. As outlined in Sec. IX we modify the available rate vectors and re-optimize to compute the optimal system utility for the network of $N-1$ sensor nodes. For a test network with $N = 35$, the CPU time for re-optimization was 0.57 seconds, while it takes 24.4 seconds to solve the modified problem from scratch. The optimal values are of course the same.

Utility vs. life-time: As we discussed in Section VIII, our framework allows us to trade-off utility vs. the lifetime of the WSNET. In Fig 2 we plot the system throughput of the 5-node network when T varies from 0 to 100, and the energy for each node is 1 unit. The curve $\hat{F}(T)$ shows the monotonicity with respect to T. Obviously $\hat{F}(T)$ is not a convex function and is upper bounded by $\hat{F}(0)$. We also depict (dashed line) $\hat{F}(T) - \zeta/T$ for $\zeta = 25000$. Solving problem (25) yields an optimal lifetime of $T^* = 10$.

High-power levels: The last part of this example explores the accuracy of our approach in WSNETs with power levels far away from 0 (i.e., the linear approximation regime). The setup is the same except that we now consider much larger power levels. Again we compare the approach outlined in Section VI with the enumeration approach. The results are reported in table II, where $\bar{p}$ is the maximum available power (Watt) for every node. Note that for the cases reported the SINR is typically on the order of 40dB ($\bar{p} = 5000$ case) and it can be much greater for some cases. The results verify that our approach is fairly accurate even at these unrealistically high power levels.

TABLE II

ACCURACY OF OUR APPROACH WITH POWER LEVELS FAR AWAY FROM THE LOW POWER REGIME.

$\bar{p}$	Our Approach	Enumeration	Gap
1000	4.486×10^6	4.486×10^6	0%
5000	1.216×10^7	1.249×10^7	3%
8000	1.481×10^7	1.611×10^7	8%

Example 2: Our next example explores the benefits of multi-hop in a larger WSNET. The objective is total throughput maximization and the fairness constraints mandate equal throughput for all nodes. The test network consists of a gateway located at $(0, 0)$ and two clusters containing equal number of nodes: one cluster contains nodes uniformly distributed in the box $[10m, 20m] \times [10m, 20m]$ and the other cluster consists of uniformly distributed nodes in the box $[25m, 35m] \times [25m, 35m]$. Let us denote by $\mathscr{C}_A$ and $\mathscr{C}_B$, respectively, these two clusters. All the other settings are identical to our first example, except that we use the expected value of S_{ij} to calculate the channel gains throughout. Table III compares the throughput (in bps) of our algorithm with two alternative policies: a single-hop and and a 2-hop policy. According to the latter one, nodes in $\mathscr{C}_B$ transmit to nodes in $\mathscr{C}_A$ for a 50% fraction of time and the remaining 50% fraction of time nodes in $\mathscr{C}_A$ transmit directly to the gateway. Note that due to the special (and deliberate) structure of the WSNET, this 2-hop policy would be quite effective. Indeed, as Table III illustrates, the 2-hop policy performs quite well. Still, our policy can improve throughput by up to 37.4% (30-node case). The performance of the single-hop policy is understandably dismal.

TABLE III

COMPARISON OF DIFFERENT POLICIES IN LARGER WSNETS.

N	Decomposition	Single-hop	2-hop
20	672.99	26.59	625.07
26	851.20	27.09	633.41
30	949.77	31.26	691.12

XI. CONCLUSION

We considered the problem of scheduling transmissions in WSNETs to maximize the total system utility subject to fairness constraints. We proposed a decomposition algorithm and established its convergence in a finite number of iterations. The resulting policy involves time-sharing over a number of feasible transmission schemes. Time-sharing convexifies the achievable region for transmission rate vectors and thus, achieves higher utility than any individual scheme. To the best of our knowledge, there is no alternative in the existing literature other than enumerating all feasible transmission schemes (the enumeration approach) for solving this problem in the general setting we consider.

The efficiency of our decomposition algorithm rests on our ability to efficiently solve a subproblem that identifies "promising" transmission schemes. To that end, we adopt a linear approximation of achievable rates which is asymptotically exact in the regime of low power levels. This regime is appropriate for WSNETs with rather dispersed nodes or operating in noisy environments. Still, the subproblem is an integer linear programming problem. Nevertheless, we show that it is polynomially solvable due to its structure. The linear approximation yields the structure of the transmission policy which we use to derive a policy under the exact (Shannon) expressions for transmission rates.

Our framework allows us to optimize sensor power levels to achieve a given utility target. This can translate into significant energy savings with a certain quality of service guarantee on the system utility. In our setting, we can solve the utility optimization problem subject to a hard constraint on the network lifetime. Alternatively, we are able to efficiently find a desirable operating time on a lifetime vs. utility curve.

The numerical results we presented suggest that our approach is very accurate even for power levels that are much higher than typical WSNET applications. They also convincingly demonstrate that our approach can handle sizable instances of the problem. For example, we are able to solve problems with 50 or so nodes in less than a minute. This is a dramatic improvement over what is computationally feasible with an enumeration approach (e.g., as in [8]).

ACKNOWLEDGMENTS

Research partially supported by the NSF under CAREER awards ANI-9983221 and ANI-0132802 and grants DMI-0330171, ANI-0240333, ECS-0426453, CNS-0435312, and DMI-0300359, and by the ARO under the ODDR&E MURI2001 Program Grant DAAD19-01-1-0465 to the Center for Networked Communicating Control Systems.

REFERENCES

[1] D. Estrin, D. Culler, K. Pister, and G. Sukhatme, "Connecting the physical world with pervasive networks," *IEEE Pervasive Computing*, vol. 1, no. 1, pp. 59–69, 2002.

[2] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-effi cient communication protocols for wireless microsensor networks," in *Proceedings of the 33rd Hawaiian International Conference on Systems Science (HICSS-33)*, Maui, HI, USA, 2000.

[3] R. Krishnan and D. Starobinski, "Effi cient clustering algorithms for self-organizing wireless sensor networks," *Journal of AD-Hoc Networks*, 2004.

[4] E. Gehrke and S. Madden, "Query processing in sensor networks," *IEEE Pervasive Computing*, vol. 3, no. 1, pp. 46–55, 2004.

[5] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, 2002.

[6] S. Ray, D. Starobinski, A. Trachtenberg, and R. Ungrangsi, "Robust location detection with sensor networks," *IEEE Journal on Selected Areas in Communications*, 2004, special issue on the fundamental performance limits of wireless sensor networks, in press.

[7] I. C. Paschalidis and S. Ray, "Optimizing location detection services in wireless sensor networks," in *Proceedings of the 43rd IEEE Conference on Decision and Control*, 2004, to appear.

[8] S. Toumpis and A. J. Goldsmith, "Capacity region for wireless ad hoc networks," *IEEE Transaction on Wireless Communications*, vol. 2, no. 4, pp. 736–748, September 2001.

[9] M. Johansson, L. Xiao, and S. Boyd, "Simultaneous routing and resource allocation in CDMA wireless data networks," *Proceedings of IEEE International Conference on Communications*, vol. 1, pp. 51–55, May 2003.

[10] P. Marbach and R. Berry, "Downlink resource allocation and pricing for wireless networks," *INFOCOM 2002*, June 23-27 2002.
[11] M. Kodialam and T. Nandagopal, "Characterizing achievable rates in multi-hop wireless networks: The joint routing and scheduling problem," *MobiCom'03*, pp. 42–55, September 14-19 2003.
[12] M. Bhardwaj, A. Chandrakasan, and T. Garnett, "Upper bounds on the lifetime of sensor networks," *Proc. of IEEE International COnf. on Communications*, pp. 785–790, 2001.
[13] B. Radunovic and J. Y. L. Boudec, "Optimal power control, scheduling and uwb networks," to appear, December 2004.
[14] X. Lin and N. B. Shroff, "Joint rate control and scheduling in multihop wireless networks," Preprint, February 2004.
[15] I. C. Paschalidis, W. Lai, and D. Starobinski, "On maximizing the utility of uplink transmissions in sensor networks under explicit fairness constraints," in *Proceedings of the 43rd IEEE Conference on Decision and Control*, 2004, to appear.
[16] D. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA: Athena Scientific, 1999.
[17] D. Bertsimas and J. Tsitsiklis, *Introduction to Linear Optimization*. Belmont, MA: Athena Scientific, 1997.
[18] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. Prentice Hall, 1982.