Asymptotically Optimal Transmission Policies for Large-Scale Low-Power Wireless Sensor Networks[*]

Ioannis Ch. Paschalidis[†],   Wei Lai[‡],   David Starobinski[§],

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 and 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. Furthermore, we establish that the utility maximization problem we consider can, in principle, be solved in polynomial time. Numerical results show that our policy is near-optimal, even at high power levels, and far superior to the best known heuristics at low 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.

Index Terms— Wireless sensor networks, transmission scheduling, routing, mathematical programming/optimization.

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]. The sensors have limited on-board capabilities, communicate wirelessly and are usually powered by limited and non-replenishable energy resources. Hence, energy preservation is critical. In this setting, 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, 3]. A plethora of applications fit this paradigm, including process control, machine condition monitoring, indoor location detection, and homeland defense [4–6].

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 [4]. 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 to 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 optimizing network operations while guaranteeing a minimum lifetime for the network is another significant problem [7].

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 [7], to optimize the use of scarce resources, future sensor network architectures must address all these challenges (multi-access control – MAC –, routing, QoS, and power control) using an integrated approach.

In this paper, we address the joint routing, power control and scheduling optimization problem and propose a new efficient computational approach to solve it in a centralized fashion. Specifically, we consider the regime of low transmission powers (e.g., on the order of mWatts) with a moderate node density (e.g., on the order of 1 node per 10 square meters), in which many sensor networks operate. In this regime, the transmission rate of sensors scales linearly with the transmission power. Based on this assumption, we present a new methodology to derive the optimal transmission policy for the sensors. Our approach employs a dual cutting-plane approach 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 devise an algorithm whose running time grows polynomially with the size of the input. The polynomial complexity of the subproblem implies, under certain technical conditions, the polynomial solvability of the master problem. Namely, the optimal transmission policy can, in principle, be computed in polynomial time, which is of independent interest.

Our methodology dramatically improves the size of problems that can be solved. A similar optimization framework to check whether a transmission rate is achievable has been provided in [8], wherein the proposed approach to characterize the capacity region amounts to an enumeration of all the possible transmission strategies. Unfortunately, this enumer-

† Corresponding author. Center for Information & Systems Eng., and Dept. of Manufacturing 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., and Dept. of Manufacturing Eng., Boston University, e-mail: `wei@bu.edu`.

§ Center for Information & Systems Eng., and ECE Dept., Boston University, e-mail: `staro@bu.edu`.

ation 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. Instead, we are able to solve instances with 50 or so nodes in less than a minute. Furthermore, although our policy is provably optimal only in the regime of low power levels, we show that it can easily be adapted to the regime where the linear approximation is not in effect. To demonstrate its wider applicability, we present numerical results that show that our policy performs close to optimal even at high transmission powers.

The optimization problem we formulate is a *utility maximization* problem that can accommodate lifetime constraints, fairness constraints, and potential interdependencies among sensor objectives. Therefore, throughput maximization [11] and maximization of separable utility functions [13, 14] are special cases of the more general problem we consider.

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 networks), 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 approach can speed-up the convergence rate of the algorithm by close to two orders of magnitude.

Finally, we compare the performance of our policy with some of the best known joint scheduling-routing heuristics proposed in the literature [13]. We consider throughput maximization in randomly generated instances of low-power sensor networks having 40 nodes. Our comparisons show that our policy increases total throughput by 50%, on average, where the gain exceeds 100% in some cases.

The rest of the paper is organized as follows. In Sec. II we present the system model and formulate the utility maximization problem. Sec. III discusses the (undesirable) implications of not enforcing fairness constraints. Sec. IV presents our decomposition algorithm and establishes its convergence. Sec. V demonstrates how to solve the subproblem in polynomial time. In Sec. VI we show that the master problem can be solved in polynomial time. In Sec. VII we discuss how to obtain a policy when the linear approximation of rates is not in effect. Optimization over power limits is considered in Sec. VIII. In Sec. IX we show how to trade-off achieved utility vs. network lifetime. We deal with node failures in Sec. X. Numerical results are in Sec. XI and conclusions in Sec. XII.

Notational Conventions: All vectors are assumed to be column vectors and are denoted by lower case boldface letters; 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 $\mathbf{x}$, $\mathbf{0}$ the vector of all zeroes, and $\mathbf{e}$ the vector of all ones. $||\mathbf{x}||_\infty$ and $||\mathbf{x}||_1$ denote the L_∞ and L_1 norms, respectively. We use upper case boldface letters to denote matrices. We use script letters to define sets and denote by $\mathrm{Conv}(\mathscr{A})$ 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 sensors, the network uses M gateways which receive information from sensors and relay it to some other processing infrastructure. Gateways can only receive information and, without loss of generality, they are allowed to receive data from multiple nodes simultaneously. Henceforth, we will refer to all $M + N$ sensor and gateway nodes alike as *nodes of the WSNET*. Nodes $1, \ldots, N$ correspond to sensors 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 $i, j = 1, \ldots, N + M$, $k = 1, \ldots, K$. We will refer to such a transmission as the (i, j, k) transmission. Note that multihop schemes are also covered in the description. 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 $i, j = 1, \ldots, M + N$, $k = 1, \ldots, K$, 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.

Consider an (i, j, k) transmission. The *Signal to Interference and Noise Ratio (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.,

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

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 $r_{ijk} \leq 0$. Therefore, r_{ijk} and r_{jik} have opposite signs, but the same absolute value. 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 mWatts and the node density is moderate. In this case 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]. In particular, taking the Taylor series expansion of (3) around $\mathbf{p} = \mathbf{0}$ and maintaining up to first order terms we obtain

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

In matrix notation we have $\mathbf{r} = \mathbf{H}\mathbf{p}$, where the matrix $\mathbf{H}$ is appropriately defined. We use this linear approximation to devise the structure of the optimal 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 establishes some properties of $\mathscr{P}$, $\mathscr{R}$, and their convex hulls; the proof is given in Appendix I, which due to space limitations is not included and is available at [15].

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 we choose 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 final 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 rate vector $\mathbf{r}$. We assume that $F(\mathbf{r})$ is continuous, concave, and bounded in Conv($\mathscr{R}$). 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, $F(\mathbf{r})$ need 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 sensors, e.g., when 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 instance we can use *linear* fairness constraints to require the transmission rate of each node to exceed a certain minimum value in order to prevent starvation, or maintain a certain relationship between the throughputs of a number of nodes. Moreover, if the "needs" of the network can be formulated as a set of linear constraints, then they can also be treated as fairness constraints in our model. We elaborate further on this point in Sec. III. Let $\mathscr{S}$ be the set of rates satisfying fairness constraints and flow conservation, i.e.,

$$\mathscr{S} \triangleq \left\{ \mathbf{r} \mid \mathbf{A}\mathbf{r} \leq \mathbf{b},\ \sum_{j=1}^{N+M} r_{ijk} = 0, \forall i \neq s(k), d(k), \forall k \right\}$$

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

We can formulate the utility optimization problem as

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

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 Conv($\mathscr{R}$). Any $\mathbf{r} \in$ 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

$$
\begin{aligned}
\min \quad & -F(\mathbf{r}) \\
\text{s.t.} \quad & \mathbf{r} - \sum_{n=1}^{L} \alpha_n \mathbf{r}^n = 0, \\
& \sum_{n=1}^{L} \alpha_n = 1, \\
& \sum_{j=1}^{N+M} r_{ijk} = 0, \ \forall i \neq s(k), d(k), \forall k, \\
& \mathbf{A}\mathbf{r} \leq \mathbf{b}, \\
& \alpha_n \geq 0, \ n = 1, \dots, L,
\end{aligned}
\tag{5}
$$

where $\mathbf{r}$ and $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_L)$ are the decision variables. Note that $\mathbf{r}^1, \dots, \mathbf{r}^L$ are also points of $\mathscr{R}$, thus, there exist corresponding valid transmission schemes (i.e., points in $\mathscr{P}$) $\mathbf{p}^1, \dots, \mathbf{p}^L$ with $\mathbf{r}^n = \mathbf{H}\mathbf{p}^n$ for all $n = 1, \dots, L$. 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 for which very efficient algorithms exist.

Of course, $\mathrm{Conv}(\mathscr{R})$ can have a humongous number of extreme points and this is the key challenge in solving (5). A simpler version of (5) with no fairness constraints was used in [8] to check whether a transmission rate is achievable and it was proposed to be solved by simply enumerating all extreme points. As indicated in [8], and clearly illustrated in Sec. XI, this approach can only handle very small networks. As we will see in Sec. IV, there are more efficient ways to solve (5). The decomposition algorithm we propose does not need to know $\mathbf{r}^1, \dots, \mathbf{r}^L$ (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. According to our model in Sec. II the gateway can receive data from multiple nodes simultaneously. The objective is to maximize total throughput. This problem can be casted in the general framework of Sec. II. More specifically, $M = 1$ and there are N sensor nodes each of which transmits traffic intended for the gateway. In this example, we assume that each sensor only generates one class of traffic. Therefore, traffic classes differ from each other only in their respective sources. Thus, there are N traffic classes and we let class i be associated with sensor node i for $i = 1, \dots, N$. Let us adopt the notation of Sec. 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. (5)):

$$
\begin{aligned}
\max \quad & \sum_{i=1}^{N} \sum_{j=1}^{N+1} r_{iji} \\
\text{s.t.} \quad & \mathbf{r} \in \mathrm{Conv}(\mathscr{R}), \\
& \sum_{j=1}^{N+1} r_{ijk} = 0, \ \forall i \neq k, N+1, \forall k.
\end{aligned}
\tag{6}
$$

Theorem III.1 *Optimality for problem (6) 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

$$
\begin{aligned}
\max \quad & \sum_{i=1}^{N} \sum_{j=1}^{N+1} r_{iji} \\
\text{s.t.} \quad & \mathbf{r} \in \mathrm{Conv}(\mathscr{R}).
\end{aligned}
\tag{7}
$$

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 $\mathrm{Conv}(\mathscr{R})$. The optimal solution $\mathbf{r}^*$ is also in $\mathscr{R}$, thus, no time division is needed to achieve optimality. Since there is no time-sharing and $\mathbf{r}^*$ has to be valid, the optimal strategy for problem (7) is for every node to send directly to the gateway. For such an $\mathbf{r}^*$ conservation constraints are satisfied and $\mathbf{r}^*$ solves (6). ∎

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 (5). 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.

We will be referring to (5) as the *master problem*. Let $(\boldsymbol{\lambda}, \mu, \boldsymbol{\sigma}, \boldsymbol{\nu})$ be the dual vector, then the dual function is

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

where

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

Let

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

and note that

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

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

and $\mathscr{D}_1$ is independent of $\mathbf{r}^1, \ldots, \mathbf{r}^L$. Then the dual of (5) is

$$\begin{aligned} \max \quad & G_1(\boldsymbol{\lambda}, \boldsymbol{\nu}, \boldsymbol{\sigma}) - \mu - \boldsymbol{\sigma}'\mathbf{b} \\ \text{s.t.} \quad & (\boldsymbol{\lambda}, \boldsymbol{\nu}, \boldsymbol{\sigma}) \in \mathscr{D}_1, \\ & \mu - \boldsymbol{\lambda}'\mathbf{r}^n \geq 0, \quad n = 1, \ldots, L, \\ & \boldsymbol{\sigma} \geq \mathbf{0}. \end{aligned} \tag{8}$$

Since (5) is a convex problem there is no duality gap.

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

$$\begin{aligned} \min \quad & -F(\mathbf{r}) \\ \text{s.t.} \quad & \mathbf{r} - \sum_{n=1}^m \alpha_n \mathbf{r}^n = 0, \\ & \sum_{n=1}^m \alpha_n = 1, \\ & \sum_{j=1}^{N+M} r_{ijk} = 0, \ \forall i \neq s(k), d(k), \forall k, \\ & \mathbf{A}\mathbf{r} \leq \mathbf{b}, \\ & \alpha_n \geq 0, \ n = 1, \ldots, m, \end{aligned} \tag{9}$$

which we call the *restricted master problem* at the mth iteration. Suppose we solve this problem to optimality. The dual of (9) is identical to (8) 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 $\mu^{(m)} - \boldsymbol{\lambda}^{(m)'}\mathbf{r}^n \geq 0$ for all $n = 1, \ldots, L$ then we have a primal-dual pair for (5) and we are done. Otherwise, we need to generate an extreme point, say $\mathbf{r}^{m+1}$, of $\text{Conv}(\mathscr{R})$ 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 an extreme point that violates dual feasibility.

A. The subproblem

At the mth iteration we seek an extreme point $\mathbf{r}^{m+1}$ of $\text{Conv}(\mathscr{R})$ satisfying $\mu^{(m)} - \boldsymbol{\lambda}^{(m)'}\mathbf{r}^{m+1} < 0$. As we argued earlier, the extreme points of $\text{Conv}(\mathscr{R})$ 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

$$\begin{aligned} \max \quad & \boldsymbol{\lambda}'\mathbf{r} \\ \text{s.t.} \quad & \mathbf{r} = \mathbf{H}\mathbf{p}, \\ & \mathbf{p} \in \mathscr{P}, \end{aligned} \tag{10}$$

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

We now establish some properties of (10). $\boldsymbol{\lambda} \in \mathbb{R}^{(M+N)^2 K}$ is the dual vector corresponding to the first constraint of (5). Denote by λ_{ijk} the element of $\boldsymbol{\lambda}$ corresponding to r_{ijk} and let $\pi_{ijk} = \lambda_{ijk} - \lambda_{jik}$. Then

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

hence the subproblem is equivalent to

$$\begin{aligned} \max \quad & \sum_{k=1}^K \sum_{i=1}^{N+M} \sum_{j=1}^{N+M} \frac{\pi_{ijk}G_{ij}}{\eta} p_{ijk} \\ \text{s.t.} \quad & \mathbf{p} \in \mathscr{P}. \end{aligned} \tag{11}$$

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

Proposition IV.1 *Problem (11) is equivalent to the ILP:*

$$\begin{aligned} \max \quad & \sum_{(i,j,k)|\psi_{ijk}>0} \psi_{ijk} s_{ijk} \\ \text{s.t.} \quad & \sum_{j=1}^{N+M} \sum_{k=1}^K s_{ijk} + \sum_{j=1}^{N+M} \sum_{k=1}^K s_{jik} \leq 1, \ \forall i \leq N, \\ & 0 \leq s_{ijk} \leq I_{ijk}, \\ & s_{ijk} \in \{0, 1\}, \end{aligned} \tag{12}$$

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

Proof: Note that there always exists an optimal solution $\mathbf{p}^*$ to the problem (11) 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 (11) is equivalent to:

$$\begin{aligned} \max \quad & \sum_{(i,j,k)|\psi_{ijk}>0} \psi_{ijk} s_{ijk} \\ \text{s.t.} \quad & s_{ijk} + s_{uiv} \leq 1, \ \forall i, j, k, u, v, \\ & s_{ijk} + s_{iuv} \leq 1, \ \forall (j,k) \neq (u,v), \\ & s_{ijk} + s_{ujv} \leq 1, \ \forall (i,k) \neq (u,v), \ j \leq n, \\ & 0 \leq s_{ijk} \leq I_{ijk}, \\ & s_{ijk} \in \{0, 1\}. \end{aligned} \tag{13}$$

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

$$p_{ijk}^* = \begin{cases} \bar{p}_{ijk}, & \text{if } s_{ijk} = 1, \\ 0, & \text{otherwise,} \end{cases} \tag{14}$$

is an optimal solution of (11). Writing (13) in a more compact way we obtain (12). ■

We summarize the discussion on the subproblem as follows: to compute an optimal solution $\mathbf{r}^*$ of (10) we first solve (12) to obtain an optimal solution $\mathbf{s}^*$, then compute $\mathbf{p}^*$ as in (14), 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 (5) is feasible; we will discuss at the end of this Section how this assumption can be relaxed.

Theorem IV.2 *Assume that (5) is feasible. Then the decomposition algorithm of Fig. 1 terminates with an optimal solution of (5) 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

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 (9) 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 (10) with cost vector $\boldsymbol{\lambda}^{(m)}$ as outlined in Sec. 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 (5). Otherwise, set $m := m+1$ and go to step 2a.

Fig. 1. The decomposition algorithm.

all $\mathbf{r} \in \mathscr{R}$. Since all extreme points of $\text{Conv}(\mathscr{R})$ 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 $\text{Conv}(\mathscr{R})$. 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 (5) 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 $\text{Conv}(\mathscr{P})$. Let $\mathbf{p}^1, \ldots, \mathbf{p}^m$ be 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 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*

$$
\begin{aligned}
\min \quad & -\sum_{i=1}^R y_i \\
\text{s.t.} \quad & \mathbf{r} - \sum_{n=1}^L \alpha_n \mathbf{r}^n = 0, \\
& \sum_{n=1}^L \alpha_n = 1, \\
& \sum_{j=1}^{N+M} r_{ijk} = 0, \ \forall i \neq s(k), d(k), \forall k, \\
& \mathbf{Ar} + \mathbf{y} = \mathbf{b}, \\
& \alpha_n \geq 0, \ n = 1, \ldots, L,
\end{aligned}
\tag{15}
$$

where we introduce the 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 (15) is almost identical to (8) 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 (15) satisfies $\mathbf{y} \geq \mathbf{0}$ then we are done as we have a feasible solution of (5) to initialize our algorithm. Otherwise, (5) 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 Sec. IV-A, solving the subproblem amounts to solving an ILP. General ILPs are hard to solve; they are NP-complete. In this section, we establish that (12) 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

$$
w_{ij} = \begin{cases}
\max\limits_{k=1,\ldots,K} \max\{\psi_{ijk}, \psi_{jik}\}, & \forall i, j \in \mathscr{A}, \\
\max\limits_{k=1,\ldots,K} \max\{\psi_{i,N+l,k}, 0\}, & \forall i \in \mathscr{A}, j \in \mathscr{B}_l, \\
\max\limits_{k=1,\ldots,K} \max\{\psi_{j,N+l,k}, 0\}, & \forall i \in \mathscr{B}_l, j \in \mathscr{A}, \\
0, & \text{otherwise.}
\end{cases}
\tag{16}
$$

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 (16).

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

$$
\begin{aligned}
max \quad & \sum_{(i,j) \in \mathscr{E}} w_{ij} x_{ij} \\
s.t. \quad & \sum_{j \mid (i,j) \in \mathscr{E}} x_{ij} \leq 1, & \forall i \\
& x_{ij} = x_{ji}, & \forall i, j, \\
& x_{ii} = 0, & \forall i, \\
& x_{ij} \in \{0, 1\}, & \forall i, j.
\end{aligned}
\tag{17}
$$

Then, an optimal solution $\mathbf{s}^*$ to the subproblem (12) satisfies

$$
s_{ijk}^* = \begin{cases}
1_{\mathscr{K}}(i,j,k) x_{ij}^*, & \text{if } 1 \leq i, j \leq N, \\
1_{\mathscr{K}}(i,j,k) \sum\limits_{v \in \mathscr{B}_{j-N}} x_{iv}^*, & \text{if } 1 \leq i \leq N, \ j \geq N+1, \\
0, & \text{otherwise.}
\end{cases}
$$

Proof: We first show $\mathbf{s}^*$ is a feasible solution to the subproblem (12). Note that if $\psi_{ijk} \leq 0$ then $I_{ijk} = 0$ and $1_{\mathscr{K}}(i,j,k) = 0$, thus $s^*_{ijk} = 0$. Also, by construction, s^*_{ijk} can only take values in $\{0,1\}$ for any i,j,k. Finally, for any sensor node i ($1 \leq i \leq N$), we have

$$\begin{aligned}
&\textstyle\sum_{j=1}^{N+M}\sum_{k=1}^{K} s^*_{ijk} + \sum_{j=1}^{N+M}\sum_{k=1}^{K} s^*_{jik}\\
&= \sum_{j=1}^{N}\sum_{k=1}^{K} 1_{\mathscr{K}}(i,j,k)x^*_{ij} + \sum_{j=1}^{N}\sum_{k=1}^{K} 1_{\mathscr{K}}(j,i,k)x^*_{ji}\\
&\quad + \sum_{j=N+1}^{N+M}\sum_{k=1}^{K} 1_{\mathscr{K}}(i,j,k)\sum_{v\in\mathscr{B}_{j-N}} x^*_{iv}\\
&= \sum_{j=1}^{N} x^*_{ij}\sum_{k=1}^{K}[1_{\mathscr{K}}(i,j,k) + 1_{\mathscr{K}}(j,i,k)]\\
&\quad + \sum_{j=N+1}^{N+M}\sum_{v\in\mathscr{B}_{j-N}} x^*_{iv}\sum_{k=1}^{K} 1_{\mathscr{K}}(i,j,k)\\
&\leq \sum_{j=1}^{N} x^*_{ij} + \sum_{j=N+1}^{N+M}\sum_{v\in\mathscr{B}_{j-N}} x^*_{iv}\\
&= \sum_{v|(i,v)\in\mathscr{E}} x^*_{iv} \leq 1,
\end{aligned}$$

where the first equality above follows from the definition of $\mathbf{s}^*$. The third equality follows from the fact that $x^*_{ij} = x^*_{ji}$. For the first inequality, note that if $(i,j,k) \in \mathscr{K}$, then no other triplet with the same (i,j) is in $\mathscr{K}$. Furthermore, if $(i,j,k) \in \mathscr{K}$ then $\psi_{ijk} > 0$ which implies $\psi_{jik} < 0$, hence $(j,i,k) \notin \mathscr{K}$.

Next we relate the objective values of problems (12) and (17), respectively. We have

$$\begin{aligned}
&\textstyle\sum_{(i,j)\in\mathscr{E}} w_{ij}x^*_{ij}\\
&= \sum_{i\in\mathscr{A}}\sum_{j\in\mathscr{A}} w_{ij}x^*_{ij} + 2\sum_{l=1}^{M}\sum_{i\in\mathscr{A}}\sum_{j\in\mathscr{B}_l} w_{ij}x^*_{ij}\\
&= \sum_{i\in\mathscr{A}}\sum_{j\in\mathscr{A}} \max_{k=1,\ldots,K}\max\{\psi_{ijk},\psi_{jik}\}x^*_{ij}\\
&\quad + 2\sum_{l=1}^{M}\sum_{i\in\mathscr{A}}\sum_{j\in\mathscr{B}_l}\max_{k=1,\ldots,K}\max\{\psi_{i,l+N,k},0\}x^*_{ij}\\
&= \sum_{i=1}^{N}\sum_{j=1}^{N}[\sum_{k=1}^{K} 1_{\mathscr{K}}(i,j,k)\psi_{ijk} + \sum_{k=1}^{K} 1_{\mathscr{K}}(j,i,k)\psi_{jik}]x^*_{ij}\\
&\quad + 2\sum_{l=1}^{M}\sum_{i=1}^{N}\sum_{j\in\mathscr{B}_l}\sum_{k=1}^{K} 1_{\mathscr{K}}(i,l+N,k)\psi_{i,l+N,k}x^*_{ij}\\
&= 2\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{K} \psi_{ijk}1_{\mathscr{K}}(i,j,k)x^*_{ij}\\
&\quad + 2\sum_{i=1}^{N}\sum_{l=1}^{M}\sum_{k=1}^{K} \psi_{i,l+N,k}1_{\mathscr{K}}(i,l+N,k)\sum_{j\in\mathscr{B}_l} x^*_{ij}\\
&= 2\sum_{i=1}^{N+M}\sum_{j=1}^{N+M}\sum_{k=1,\psi_{ijk}>0}^{K} \psi_{ijk}s^*_{ijk}\\
&= 2\sum_{(i,j,k)|\psi_{ijk}>0} \psi_{ijk}s^*_{ijk}
\end{aligned}$$

where the first equality above follows from (16) and the fact that $x^*_{ij} = x^*_{ji}$. The last two equalities follow from the definition of $\mathbf{s}^*$.

In the remainder of this proof, we proceed to prove that $\mathbf{s}^*$ is an optimal solution to problem (12) by contradiction. Suppose $\mathbf{s}^*$ is not optimal, then there exists a feasible solution $\hat{\mathbf{s}}$ to problem (12) with strictly higher objective value. From $\hat{\mathbf{s}}$ construct a solution $\hat{\mathbf{x}}$ such that

$$\hat{x}_{ij} = \begin{cases}
\sum_{k=1}^{K}(\hat{s}_{ijk} + \hat{s}_{jik}), & \text{if } 1 \leq i,j \leq N,\ i \neq j,\\
\sum_{k=1}^{K} \hat{s}_{i,l+N,k}, & \text{if } 1 \leq i \leq N,\ j = Nl+i,\\
\sum_{k=1}^{K} \hat{s}_{j,l+N,k}, & \text{if } 1 \leq j \leq N,\ i = Nl+j,\\
0, & \text{otherwise.}
\end{cases}$$

Note that by the construction of $\hat{\mathbf{x}}$, we have $\hat{x}_{ij} = \hat{x}_{ji}$ and $\hat{x}_{ii} = 0, \forall i,j$. To show that $\hat{\mathbf{x}}$ is a feasible solution to problem (17), we distinguish between two cases.

Case 1: $1 \leq i \leq N$. In this case, as $\hat{\mathbf{s}}$ is a feasible solution to problem (12), we have

$$\begin{aligned}
&\textstyle\sum_{j|(i,j)\in\mathscr{E}} \hat{x}_{ij}\\
&= \sum_{j=1}^{N}\sum_{k=1}^{K}(\hat{s}_{ijk} + \hat{s}_{jik}) + \sum_{l=1}^{M}\sum_{k=1}^{K} \hat{s}_{i,l+N,k}\\
&= \sum_{j=1}^{N+M}\sum_{k=1}^{K} \hat{s}_{ijk} + \sum_{j=1}^{N+M}\sum_{k=1}^{K} \hat{s}_{jik} \leq 1.
\end{aligned}$$

where the last equality follows from the fact that $\hat{s}_{jik} = 0$ for any $j = N+1,\ldots,N+M$.

Case 2: $Nl+1 \leq i \leq Nl+N$ for some l such that $1 \leq l \leq M$. In this case, we have

$$\textstyle\sum_{j|(i,j)\in\mathscr{E}} \hat{x}_{ij} = \hat{x}_{i,i-Nl} = \sum_{k=1}^{K} \hat{s}_{i-Nl,l+N,k} \leq 1.$$

We conclude that $\hat{\mathbf{x}}$ is a feasible solution to problem (17).

As the last step of the proof, note that

$$\begin{aligned}
&\textstyle\sum_{(i,j)\in\mathscr{E}} w_{ij}\hat{x}_{ij}\\
&= \sum_{i\in\mathscr{A}}\sum_{j\in\mathscr{A}} w_{ij}\hat{x}_{ij} + \sum_{l=1}^{M}\sum_{i\in\mathscr{A}}\sum_{j\in\mathscr{B}_l} w_{ij}\hat{x}_{ij}\\
&\quad + \sum_{l=1}^{M}\sum_{i\in\mathscr{B}_l}\sum_{j\in\mathscr{A}} w_{ij}\hat{x}_{ij}\\
&= \sum_{i\in\mathscr{A}}\sum_{j\in\mathscr{A}} w_{ij}\hat{x}_{ij} + 2\sum_{l=1}^{M}\sum_{i\in\mathscr{A}} w_{i,Nl+i}\hat{x}_{i,Nl+i}\\
&= \sum_{i=1}^{N}\sum_{j=1}^{N}\left(\max_{t=1,\ldots,K}\max\{\psi_{ijt},\psi_{jit}\}\right)\sum_{k=1}^{K}(\hat{s}_{ijk} + \hat{s}_{jik})\\
&\quad + 2\sum_{l=1}^{M}\sum_{i=1}^{N}\left(\max_{t=1,\ldots,K}\max\{\psi_{i,l+N,t},0\}\right)(\sum_{k=1}^{K}\hat{s}_{i,l+N,k})\\
&\geq 2\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1,\psi_{ijk}>0}^{K} \psi_{ijk}\hat{s}_{ijk}\\
&\quad + 2\sum_{l=1}^{M}\sum_{i=1}^{N}\sum_{k=1,\psi_{i,l+n,k}>0}^{K} \psi_{i,l+N,k}\hat{s}_{i,l+N,k}\\
&= 2\sum_{(i,j,k)|\psi_{ijk}>0} \psi_{ijk}\hat{s}_{ijk}
\end{aligned}$$

With this we can see that

$$\begin{aligned}
\textstyle\sum_{(i,j)\in\mathscr{E}} w_{ij}\hat{x}_{ij} &\geq 2\sum_{(i,j,k)|\psi_{ijk}>0} \psi_{ijk}\hat{s}_{ijk}\\
&> 2\sum_{(i,j,k)|\psi_{ijk}>0} \psi_{ijk}s^*_{ijk} = \sum_{(i,j)\in\mathscr{E}} w_{ij}x^*_{ij},
\end{aligned}$$

which contradicts the assumption that $\mathbf{x}^*$ is optimal for (17), therefore, $\mathbf{s}^*$ is optimal for the subproblem. ∎

Remark : It should be noted that (17) 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 (17) can be solved in $O(|\mathscr{V}|^3)$ amount of time [16], 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. POLYNOMIAL SOLVABILITY OF THE MASTER PROBLEM

The complexity of the subproblem suggests that, in principle, the master problem (5) can also be solved in polynomial time under mild technical conditions. First, we consider the following problems concerning any given polytope $\mathscr{L}$.

1) *Separation problem* of $\mathscr{L}$ (**SP**($\mathscr{L}$)): *Given an arbitrary vector $\mathbf{r}$ of suitable dimension, decide whether $\mathbf{r} \in \mathscr{L}$, and if not, find a hyperplane that separates $\mathbf{r}$ from $\mathscr{L}$.*

2) *Optimization problem* of $\mathscr{L}$ ($\mathbf{OP}(\mathscr{L})$): *Given an arbitrary vector $\mathbf{c}$ of suitable dimension, find a vector $\mathbf{r} \in \mathscr{L}$ that maximizes $\mathbf{c}'\mathbf{r}$ over $\mathscr{L}$ or assert that $\mathscr{L}$ is empty.*

$\mathbf{SP}(\mathscr{L})$ and $\mathbf{OP}(\mathscr{L})$ are polynomially equivalent, that is, if one is polynomially solvable, so is the other (cf. [17, Theorem 6.4.9]). This equivalence holds when $\mathscr{L}$ is *well-described*, namely, the *facet-complexity* (or equivalently, the *vertex-complexity*) is bounded (cf. [17, Lemma 6.2.4]). Here, we consider the boundedness of the vertex-complexity which, roughly, means that all vertices of the polyhedron can be written as a binary number –the encoding length– with a number of bits bounded by a constant ([17, Definition 6.2.2]). For example, it can be shown that $\text{Conv}(\mathscr{R})$ is a well-described polyhedron, assuming $\bar{p}_i$, G_{ij} and η are all rational numbers, for any i, j. Specifically, note that the encoding length for $\frac{\bar{p}_i G_{ij}}{\eta}$ is bounded, for any i, j. Letting U be an upper bound on this encoding length it follows that the vertex-complexity of $\text{Conv}(\mathscr{R})$ is at most $(M + N)^2 K U$. As $\mathbf{OP}(\text{Conv}(\mathscr{R}))$ is polynomially solvable, the following lemma holds.

Lemma VI.1 $\mathbf{SP}(Conv(\mathscr{R}))$ *is polynomially solvable.*

Now let us consider the master problem of the form

$$\begin{aligned} \max \quad & F(\mathbf{r}) \\ \text{s.t.} \quad & \mathbf{r} \in \text{Conv}(\mathscr{R}), \qquad (18) \\ & \mathbf{r} \in \mathscr{S}, \end{aligned}$$

and note that $\boldsymbol{\alpha}$ (cf. (5)) does not appear in the formulation. Furthermore, the constraints concerning $\text{Conv}(\mathscr{R})$ are not explicitly given. For simplicity let us write the constraint $\mathbf{r} \in \mathscr{S}$ as $\mathbf{Dr} \leq \mathbf{d}$, for an appropriate matrix $\mathbf{D}$ and a vector $\mathbf{d}$. Next, we define the *infeasibility measure* of problem (18) as $B(\mathbf{r}) = \max\{B_1(\mathbf{r}), B_2(\mathbf{r})\}$, where

$$B_1(\mathbf{r}) = \min\{||\mathbf{t}||_\infty \mid \mathbf{r} - \mathbf{t} \in \text{Conv}(\mathscr{R})\},$$
$$B_2(\mathbf{r}) = \min\{v \geq 0 \mid \mathbf{Dr} - v\mathbf{e} \leq \mathbf{d}\},$$

and $\mathbf{e}$ is the vector of all 1's. It can be seen that $B_1(\mathbf{r})$ is the minimum L_∞ distance of $\mathbf{r}$ from the set $\text{Conv}(\mathscr{R})$. The next lemma establishes that $B(\mathbf{r})$ is a *valid* infeasibility measure; the proof is in Appendix II available at [15].

Lemma VI.2 $B(\mathbf{r})$ *is a valid infeasibility measure, i.e., (i) $B(\mathbf{r}) \geq 0$, and $B(\mathbf{r}) = 0$ if $\mathbf{r} \in Conv(\mathscr{R}) \cap \mathscr{S}$; (ii) $B(\mathbf{r})$ is convex in $\mathbf{r}$.*

To proceed we will need to establish a technical property for the infeasibility measure $B(\cdot)$. Let us again consider $B_1(\mathbf{r})$ and $B_2(\mathbf{r})$ separately, and define the ϵ-level sets $\mathscr{L}_\epsilon^{B_1} = \{\mathbf{r} \mid B_1(\mathbf{r}) \leq \epsilon\}$, $\mathscr{L}_\epsilon^{B_2} = \{\mathbf{r} \mid B_2(\mathbf{r}) \leq \epsilon\}$, and $\mathscr{L}_\epsilon^B = \{\mathbf{r} \mid B(\mathbf{r}) \leq \epsilon\}$ where ϵ is an arbitrary rational number. Clearly, we have $\mathscr{L}_\epsilon^B = \mathscr{L}_\epsilon^{B_1} \cap \mathscr{L}_\epsilon^{B_2}$. The following lemma establishes some useful results about $\mathscr{L}_\epsilon^{B_1}$.

Lemma VI.3 *(i) $\mathscr{L}_\epsilon^{B_1}$ is a polytope. (ii) $\mathbf{OP}(\mathscr{L}_\epsilon^{B_1})$ is polynomially solvable. (iii) $\mathbf{SP}(\mathscr{L}_\epsilon^{B_1})$ is polynomially solvable.*

Proof: *(i)* Note that

$$\mathscr{L}_\epsilon^{B_1} = \left\{\mathbf{r} \mid \exists\, \mathbf{r}^0 \in \text{Conv}(\mathscr{R}),\ ||\mathbf{r} - \mathbf{r}^0||_\infty \leq \epsilon\right\},$$

where $\text{Conv}(\mathscr{R})$ and $\{\mathbf{r} \mid ||\mathbf{r}||_\infty \leq \epsilon\}$ are polytopes, therefore $\mathscr{L}_\epsilon^{B_1}$ is also a polytope.

(ii) For arbitrary vector $\mathbf{c}$, consider the optimization problem $\max\{\mathbf{c}'\mathbf{r} \mid \mathbf{r} \in \mathscr{L}_\epsilon^{B_1}\}$, and note that it is equivalent to separately consider the following two optimization problems

$$\begin{array}{ll} \max & \mathbf{c}'\mathbf{r}^0 \\ \text{s.t.} & \mathbf{r}^0 \in \text{Conv}(\mathscr{R}), \end{array} \qquad \begin{array}{ll} \max & \mathbf{c}'\mathbf{t}^0 \\ \text{s.t.} & ||\mathbf{t}^0||_\infty \leq \epsilon, \end{array}$$

because of the separable structure of the original problem. Note that both problems above can be solved in polynomial time, and the polynomial complexity of $\mathbf{OP}(\mathscr{L}_\epsilon^{B_1})$ follows.

(iii) Let the encoding length of $\pm\epsilon$ be bounded by U_0, then encoding length of each extreme point of $\mathscr{L}_\epsilon^{B_1}$ is upper bounded by $(M + N)^2 K(U + U_0)$ where U is an upper bound on the encoding length of $\frac{\bar{p}_i G_{ij}}{\eta}$. Therefore, $\mathscr{L}_\epsilon^{B_1}$ is well-described and property *(ii)* implies that $\mathbf{SP}(\mathscr{L}_\epsilon^{B_1})$ is polynomially solvable. ∎

The following Lemma establishes a useful property of $\mathscr{L}_\epsilon^B$.

Lemma VI.4 *For any given $\mathbf{r}$ and $\epsilon > 0$ we can determine whether $B(\mathbf{r}) \leq \epsilon$, or, if not, separate $\mathbf{r}$ from the set $\mathscr{L}_\epsilon^B$. Moreover, this can be done in running time that is polynomial in $(M + N)^2 K U$ and $\log(1/\epsilon)$.*

Proof: Fix $\mathbf{r}$ and $\epsilon > 0$. Computing $B_2(\mathbf{r})$ amounts to solving a linear programming problem which can be done in polynomial time. Due to Lemma VI.3 we can determine in polynomial time whether $\mathbf{r} \in \mathscr{L}_\epsilon^{B_1}$. We distinguish four cases:
Case 1: $\mathbf{r} \in \mathscr{L}_\epsilon^{B_1}$ and $B_2(\mathbf{r}) \leq \epsilon$. In this case we have $B(\mathbf{r}) = \max\{B_1(\mathbf{r}), B_2(\mathbf{r})\} \leq \epsilon$.
Case 2: $\mathbf{r} \in \mathscr{L}_\epsilon^{B_1}$ and $B_2(\mathbf{r}) > \epsilon$. Finding a separating hyperplane that separates $\mathbf{r}$ from $\mathscr{L}_\epsilon^{B_2}$ can also be done in polynomial time since checking whether $\mathbf{r} \in \mathscr{L}_\epsilon^{B_2}$ amounts to checking feasibility of a well-described linear programming problem with a finite number of constraints. In particular, it suffices to find i such that $(\mathbf{Dr})_i - d_i \geq (\mathbf{Dr})_j - d_j, \forall j \neq i$. The separating hyperplane is defined by the i-th row of $\mathbf{D}$.
Case 3: $\mathbf{r} \notin \mathscr{L}_\epsilon^{B_1}$ and $B_2(\mathbf{r}) \leq \epsilon$. By Lemma VI.3, $\mathbf{SP}(\mathscr{L}_\epsilon^{B_1})$ can be solved polynomially.
Case 4: $\mathbf{r} \notin \mathscr{L}_\epsilon^{B_1}$ and $B_2(\mathbf{r}) > \epsilon$. In this case we can follow the procedure in either *Case 2* or *Case 3*. ∎

Now we are ready to prove the polynomial complexity of problem (18) under the following assumption on $F(\cdot)$.

Assumption A
For every instance of (18) and given an $\mathbf{r} \in \mathbb{R}^{(N+M)^2 K}$: (i) we can compute $F(\mathbf{r})$ and a subgradient of $F(\cdot)$ evaluated at $\mathbf{r}$ in time that is polynomial in the size of the instance of (18); and (ii) $F(\mathbf{r})$ grows polynomially in $\mathbf{r}$, that is, $|F(\mathbf{r})|$ is upper bounded by a polynomial function of $|\mathbf{r}|_1$ with degree equal to a power of the size of the instance of (18).

Theorem VI.5 *Let Assumption A prevail. Then the convex optimization problem (18) is polynomially solvable.*

Proof: To prove the polynomial-time solvability, it suffices to check polynomial computability, polynomial growth and polynomial boundedness of the feasible set (cf. [18, Theorem 5.3.1]).

(1) Polynomial computability: $F(\mathbf{r})$ satisfies Assumption A(i) and $B(\mathbf{r})$ satisfies the statement of Lemma VI.4.

(2) Polynomial growth: $F(\mathbf{r})$ is of polynomial growth by Assumption A(ii). To justify the polynomial growth of $B(\mathbf{r})$, we consider $B_1(\mathbf{r})$ and $B_2(\mathbf{r})$ separately. First, note that $B_1(\mathbf{r}) \leq ||\mathbf{r}||_\infty \leq ||\mathbf{r}||_1$ because $\mathbf{0} \in \mathrm{Conv}(\mathscr{R})$. To upper bound $B_2(\mathbf{r})$ we have

$$\begin{aligned} B_2(\mathbf{r}) \quad &= \max(0, \max_i((\mathbf{D}\mathbf{r})_i - d_i)) \\ &\leq ||\mathbf{D}\mathbf{r} - \mathbf{d}||_1 \leq ||\mathbf{D}||_1 \cdot ||\mathbf{r}||_1 + ||\mathbf{d}||_1. \end{aligned}$$

Therefore $B(\mathbf{r})$ is also of polynomially growth.

(3) Polynomial boundedness of the feasible set: It suffices to establish that the feasible set of every instance of problem (18) is bounded and is contained in a ball centered at the origin with radius that is bounded by a polynomial in the size of the instance of (18). This is evident for $\mathrm{Conv}(\mathscr{R})$. ∎

Remark : Assumption A on $F(\mathbf{r})$ is not as restrictive as it seems to be. Assumption A(i) is satisfied by most utility functions of interest. Assumption A(ii) appears more restrictive as several functions violate it. However, it can be easily satisfied by making certain transformations. For example, if $F(\mathbf{r}) = -\sum_i e^{r_i}$, then we can equivalently consider the function $-\log(\sum_i e^{r_i})$ which is of polynomial growth.

Equipped with Theorem VI.5, we know that there exists an algorithm that finds an optimal solution $\mathbf{r}^*$ to problem (18) in polynomial time. As the last step to solving the master problem (5), we also have to compute $\boldsymbol{\alpha}^*$ in polynomial time. This problem amounts to finding a polynomial number of extreme points of $\mathrm{Conv}(\mathscr{R})$ such that $\mathbf{r}^*$ can be expressed as a convex combination of those. By Caratheodory's theorem, $(M + N)^2 K + 1$ extreme points suffice. Despite the fact that the form of $\mathrm{Conv}(\mathscr{R})$ is not explicitly given, we know that $\mathbf{OP}(\mathrm{Conv}(\mathscr{R}))$ can be solved in polynomial time, and that the polytope $\mathrm{Conv}(\mathscr{R})$ is well-described. Therefore, [17, Theorem 6.5.11] guarantees the existence of a polynomial time algorithm achieving this goal. We summarize the major result of this section in the following theorem.

Theorem VI.6 *Let Assumption A prevail. Then the convex optimization problem (5) is polynomially solvable.*

Theorem VI.6 is an interesting theoretical result. It asserts that even though $\mathrm{Conv}(\mathscr{R})$ has an exponential number of extreme points [8], and computing its explicit form is expensive (i.e., it takes exponential time), we can still solve (5) in polynomial time. To that end, however, one would have to use the ellipsoid method, which is *typically* slow and impractical even for moderate-sized problems. Instead we expect the decomposition method we have presented to be more efficient in practice.

VII. A 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). With the exact rate function $\mathrm{Conv}(\mathscr{R})$ becomes difficult to characterize and makes (5) intractable. Earlier attempts in the literature used either approximation techniques, e.g., discretization, or restricted

routing strategies [8, 9]. Here, we use linearization to obtain the structure of the policy and then remove the linearization to devise a policy under the exact expression (3).

More specifically, we first solve (5) using the linear approximation in (4) as outlined in Sec. 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 earlier sections, we know that for each transmission vector $\mathbf{r}^n$ ($n = 1, \ldots, D$), $r_{ijk}^n > 0$ 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

$$\tilde{r}_{ijk}^n = W \log(1 + \tilde{\gamma}_{ijk}^n) - W \log(1 + \tilde{\gamma}_{jik}^n),$$

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

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

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

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

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 $\mathbf{p} \to \mathbf{0}$. Numerical examples in Sec. XI illustrate that the policy remains close to optimal even when $\mathbf{p}$ is far away from $\mathbf{0}$.

VIII. 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 Sec. II as parametrized by the vector of power limits, denoted by $\bar{\mathbf{p}} = (\bar{p}_1, \ldots, \bar{p}_N)$. Consider

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

where we write $\mathscr{R}(\bar{\mathbf{p}})$ to explicitly denote the fact that the set of rate vectors depends on $\bar{\mathbf{p}}$. $\bar{F}(\bar{\mathbf{p}})$ denotes the optimal

value. The first, and rather intuitive, property we show is monotonicity; the proof is almost immediate.

Lemma VIII.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 $\bar{F}(\bar{\mathbf{p}}^1) \geq \bar{F}(\bar{\mathbf{p}}^2)$.*

Next we show that the optimal utility is concave in $\bar{\mathbf{p}}$ if we scale the power limits uniformly.

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

Proof: First observe that due to Lemma II.1 for each extreme point $\mathbf{r}$ of $\mathrm{Conv}(\mathscr{R})$ there exists an extreme point $\mathbf{p}$ of $\mathrm{Conv}(\mathscr{P})$ such that $\mathbf{r} = \mathbf{H}\mathbf{p}$. Moreover, extreme points of $\mathrm{Conv}(\mathscr{P})$ are in the set $\bar{\mathscr{P}} = \{\mathbf{p} \in \mathscr{P} \mid p_{ijk} \in \{0, \bar{p}_{ijk}\}\}$. It follows that each extreme point $\mathbf{r}^n$ of $\mathrm{Conv}(\mathscr{R})$ can be written as $\mathbf{r}^n = \mathbf{Q}_n\bar{\mathbf{p}}$ for some appropriate matrix $\mathbf{Q}_n$, for $n = 1, \ldots, L$. Consequently, the problem in (20) can be equivalently formulated as

$$
\begin{aligned}
\bar{F}(\bar{\mathbf{p}}) = \max \quad & F(\mathbf{r}) \\
\text{s.t.} \quad & \mathbf{r} = \sum_n \alpha_n \mathbf{Q}_n \bar{\mathbf{p}}, \\
& \mathbf{r} \in \mathscr{S}, \\
& \sum_n \alpha_n = 1, \ \alpha_n \geq 0, \ n = 1, \ldots, L.
\end{aligned} \tag{21}
$$

Let $\bar{\mathbf{p}}^1, \bar{\mathbf{p}}^2 \in \mathscr{W}$, or equivalently, $\bar{\mathbf{p}}^1 = \phi_1\bar{\mathbf{p}}^0$ and $\bar{\mathbf{p}}^2 = \phi_2\bar{\mathbf{p}}^0$ for some $\phi_1, \phi_2 > 0$. Since $F(\mathbf{r})$ is concave, for any $\theta \in [0, 1]$

$$
\begin{aligned}
& \theta\bar{F}(\bar{\mathbf{p}}^1) + (1-\theta)\bar{F}(\bar{\mathbf{p}}^2) \\
& \leq F(\textstyle\sum_k \mathbf{Q}_n(\alpha_n\theta\phi_1 + \beta_n(1-\theta)\phi_2)\bar{\mathbf{p}}_0),
\end{aligned} \tag{22}
$$

where $(\mathbf{r}^*(\bar{\mathbf{p}}^1), \boldsymbol{\alpha})$ and $(\mathbf{r}^*(\bar{\mathbf{p}}^2), \boldsymbol{\beta})$ are optimal solutions of (21) corresponding to $\bar{\mathbf{p}}^1$ and $\bar{\mathbf{p}}^2$, respectively. Note that $\theta\phi_1 + (1-\theta)\phi_2 > 0$. Let

$$
\tau_n = \frac{\alpha_n\theta\phi_1 + \beta_n(1-\theta)\phi_2}{\theta\phi_1 + (1-\theta)\phi_2},
$$
$$
\mathbf{r}_\tau = \sum_n \mathbf{Q}_n \tau^n(\theta\phi_1 + (1-\theta)\phi_2)\bar{\mathbf{p}}_0,
$$

where $\tau_n \geq 0$. Then, since $\mathscr{S}$ is a polyhedron,

$$
\sum_n \tau_n = \frac{\sum_n \alpha_n\theta\phi_1 + \sum_n \beta_n(1-\theta)\phi_2}{\theta\phi_1 + (1-\theta)\phi_2} = 1,
$$
$$
\mathbf{r}_\tau = \theta\mathbf{r}^*(\bar{\mathbf{p}}^1) + (1-\theta)\mathbf{r}^*(\bar{\mathbf{p}}^2) \in \mathscr{S}.
$$

These imply that $(\mathbf{r}_\tau, \boldsymbol{\tau})$, where $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_L)$, is a feasible solution to problem (21) when the power limit vector equals $(\theta\phi_1 + (1-\theta)\phi_2)\bar{\mathbf{p}}_0$. Therefore, using (22), we conclude

$$
\bar{F}(\theta\bar{\mathbf{p}}^1 + (1-\theta)\bar{\mathbf{p}}^2) \geq F(\mathbf{r}_\tau) \geq \theta\bar{F}(\bar{\mathbf{p}}^1) + (1-\theta)\bar{F}(\bar{\mathbf{p}}^2).
$$

∎

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

$$
\begin{aligned}
\min \quad & \phi \\
\text{s.t.} \quad & \bar{F}(\phi\bar{\mathbf{p}}_0) \geq F_{\min},
\end{aligned} \tag{23}
$$

where $\bar{F}(\phi\bar{\mathbf{p}}_0)$ is defined in (20). Thm. VIII.2 asserts that the above is a convex optimization problem, thus, a global minimum, say ϕ^*, can be obtained using standard gradient-based algorithms. 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.

IX. The Lifetime of Sensor Networks

In this section we consider the implications of power optimization on 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.[1]

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; hence, the change of T does not affect the time horizon considered in Sec. II. Let $\mathbf{p}$ be the long-term 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, $\mathbf{C}\mathbf{p} \leq \boldsymbol{\chi}/T$, where $\mathbf{C} \in \mathbb{R}^{N \times (N+M)^2 K}$ 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:

$$
\begin{aligned}
\hat{F}(T) = \max \quad & F(\mathbf{H}\mathbf{p}) \\
\text{s.t.} \quad & \mathbf{p} = \sum_{n=1}^{J} \alpha_n \mathbf{p}^n, \\
& \mathbf{H}\mathbf{p} \in \mathscr{S}, \\
& \mathbf{C}\mathbf{p} \leq \boldsymbol{\chi}/T, \\
& \sum_{n=1}^{J} \alpha_n = 1, \ \alpha_n \geq 0, \ n = 1, \ldots, J.
\end{aligned} \tag{24}
$$

where $\mathbf{p}^1, \ldots, \mathbf{p}^J$ are the extreme points of $\mathrm{Conv}(\mathscr{P})$. Notice that in problem (24), 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 (24) is a convex programming problem. (Note that the objective $F(\mathbf{H}\mathbf{p})$ is concave in $\mathbf{H}\mathbf{p}$ and therefore concave in $\mathbf{p}$, and the constraints are linear in $\mathbf{p}$ and $\boldsymbol{\alpha}$.) Thus, we have a problem similar to (5) and a complete analog of the algorithm in Fig. 1 can be used to solve large-scale instances of (24). The details are in Appendix III available at [15].

The results of Sec. VI can also be extended to include problem (24). Namely, if $F(\mathbf{H}\mathbf{p})$ satisfies Assumption A then problem (24) is also solvable in polynomial time. The proof is almost identical with Theorem VI.6 and is omitted for brevity.

---

[1]This assumption can actually be relaxed if we impose some additional technical conditions.

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 (24) 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 (24) is a concave maximization problem, a standard argument shows that $\hat{F}(T)$ is concave in $1/T$. We have the following theorem.

Theorem IX.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. IX.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.

Finally, we note that the network lifetime may have different definitions depending on the application we are concerned with. For instance, lifetime can also be defined as the time that the energy level of each node drops below a certain threshold, or some critical routers run out of energy. Our model can be extended to handle those cases and all the results still hold.

X. Dealing with node failures

Next we discuss how to accommodate node failures in our decomposition framework. Suppose we solve problem (5) 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. 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

$$
\begin{aligned}
\max \quad & F(\tilde{\mathbf{r}}) && (26)\\
\text{s.t.} \quad & \tilde{\mathbf{r}} - \sum_{n=1}^{D} \tilde{\alpha}_n \tilde{\mathbf{r}}^n = 0, \\
& \sum_{j=1}^{N+M} \tilde{r}_{ijk} = 0, \ \forall i \neq s(k), d(k), \forall k, \\
& \mathbf{A}\tilde{\mathbf{r}} \leq \mathbf{b}, \\
& \sum_{n=1}^{D} \tilde{\alpha}_n = 1, \ \tilde{\alpha}_n \geq 0, \ n = 1, \ldots, D,
\end{aligned}
$$

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 (10). 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.

XI. 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 WS-NET with sensor nodes uniformly distributed in the box $[-10m, 10m] \times [-10m, 10m]$. The network has a single gateway at the origin which can receive data from multiple nodes simultaneously. We use the same identical parameters as in [8]. In particular, $G_{ij} = K S_{ij}(d_0/d_{ij})^\alpha$, 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 Sec. VII, 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 (5) by enumerating all feasible transmission rate vectors in $\text{Conv}(\mathscr{R})$. 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

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

maximize total throughput (reported in bps) and the fairness constraints have the form $\rho_{i+1} \leq 2\rho_i$, $i = 1, \ldots, N - 1$, 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
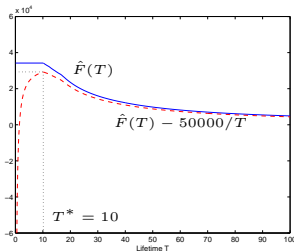
Fig. 2.   Utility as a function of network lifetime.

column reports the throughput achieved by the single-hop strategy, i.e., when each node sends directly to the gateway. [2]

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 operate at lower power 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 problems with 50 nodes in less than 1 minute. Third, 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%.

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

Node failure: Next we consider the re-optimization technique of Sec. X. We first calculate the optimal utility and rate vectors for a WSNET with N sensor nodes and then let node N fail. 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 Sec. IX, our framework allows us to trade-off utility vs. the lifetime of the WSNET. In Fig 2 we plot the system throughput of the 30-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 $\zeta = 50000$. 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 much higher power levels. Again we compare the approach of Sec. VII with the enumeration approach. The results are reported in table II, where $\bar{p}$ is the maximum available power (Watts) for every

---

[2]All the programs were run on a computer running Redhat Linux 8.0 with an Intel Xeon CPU of 3.06 GHz, and 3.6 Gbytes main memory.

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 HIGH POWER LEVELS

$\bar{p}$	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 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 shows, 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

Example 3: In the last example, we compare our method with the heuristics proposed in [13]. Specially, in [13], the authors proposed a similar framework in the context of ultra-wideband (UWB) networks. They found that the optimal power allocation should be such that each node either transmits with the maximum power or remains silent. Further, to find a good joint routing and scheduling scheme for an arbitrary network, they proposed a number of routing (MELR, MER, DIR, etc.) and scheduling (Exclusion Region, Total Exclusion, All-at-Once, etc.) heuristics. Then, they considered all the possible combinations of routing and scheduling heuristics. Among all the utilities achieved by any of these combinations, we choose the maximum as the one to compare with.
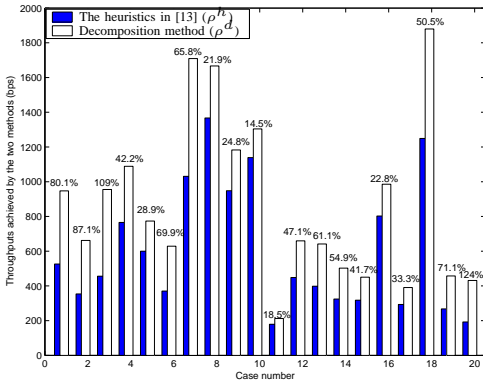
Fig. 3. Comparison between the utilities achieved by the decomposition algorithm and the heuristics in [13].

The rate function in [13] is linear in SINR. In the context of our work, the linear relationship translates to

$$r_{ijk} = C(\gamma_{ijk} - \gamma_{jik}), \quad \forall i,j,k, \qquad (27)$$

where C is a constant. This is different from the logarithmic rate function (3) used in this work. However, our approximation scheme can be easily modified such that the rate function (27) can fit in our framework. In particular, we first use the linear approximation (4) to derive the structure of the transmission schemes, then modify the transmission vectors according to (27) assuming maximum transmission power, following a method analogous to the method described in Sec. VII. Finally, we optimize over the modified transmission vectors in order to find a near-optimal transmission policy.

To numerically compare the results of the two methods, we consider the throughput maximization problem of a network with 40 nodes uniformly distributed in the box $[0m, 100m] \times [0m, 100m]$. In addition, there is a single gateway located at $(50m, 50m)$ and all the information is intended for this gateway. The fairness constraints are $\rho_{i+1} \le 2\rho_i$, $i = 1, \ldots, N-1$, where ρ_i denotes the throughput of node i.[3] We let $C = 10^5$. The maximum power of each node is 10^{-3} Watts and the background noise is 10^{-5} Watts. The channel gain between node i and j is given by $G_{ij} = d_{ij}^{-4}$, where d_{ij} is the distance between node i and j. To proceed, we randomly generated 20 networks and for each network we compare the throughputs obtained by the heuristics in [13] and the decomposition algorithm proposed in this paper. The results are shown in Fig. 3. For each case, we denote by ρ^h the throughput achieved by the heuristics in [13], and by ρ^d the throughput achieved by the decomposition method. The number on the top of each column is $\frac{\rho^d - \rho^h}{\rho^h} \times 100\%$, namely, the percentage by which the decomposition method outperforms the heuristics in [13].

In all the cases shown in Fig. 3, the throughput calculated by the decomposition algorithm is higher than the heuristics

in [13]. The average gain in throughput is 50%, but in some cases it is as much as 100% or more.

We should point out that the numerical results presented in this example are obtained in the region of small powers, which is the practical and interesting operating region for sensor networks. We do not claim that our method is unconditionally superior to the method in [13]. The heuristics in [13] can outperform our method when we significantly deviate from the small power region or when the node density increases.

To assess the effect of the node density, we adopted the same setting as above, fixed the number of nodes to 20, and varied the coverage area. That is, the nodes are uniformly distributed in the box $[0, x] \times [0, x]$ and the gateway is at $(x/2, x/2)$ with x varying in $[0, 100]$m. For each case, we randomly generate 10 networks and compare the average throughput obtained by the heuristics in [13] with our approach. Our tests show that when the network density is around 1 sensor per 10 m^2, the performance of the heuristics in [13] is competitive to that of our algorithm. However, when the density decreases our algorithm quickly outperforms the heuristic in [13]. For example, when the density is 1 sensor per 20 m^2, our algorithm outperforms by about 20% on average.

XII. CONCLUSIONS

We considered the problem of scheduling transmissions in WSNETs to maximize the total system utility subject to fairness constraints. In our setting, utility is a function of long-term average transmission rates. We proposed a decomposition algorithm and established its convergence. 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 efficiently solving 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 moderately dense WSNETs or denser WSNETs operating in noisy environments. Still, the subproblem is an integer linear programming problem. Nevertheless, we exploit its structure and show that it is polynomially solvable. 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. We note that the policy we obtain can be implemented in a causal fashion, provided that the network is initialized with a sufficient (and finite) number of packets at every node. This initial condition does not affect the long-term average transmission rates.

A result of independent interest is that when this linear approximation is in effect the problem of computing the optimal policy can, in principle, be solved in polynomial time. Yet, this does not lead into a practical algorithm. We expect the decomposition method to be more efficient in practice.

[3]The heuristics in [13] are intended for proportional fair rate allocation, but with some slight modifications their method can also be applied to the throughput maximization with fairness constraints.

PSfrag replacements

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 point on a lifetime vs. utility curve.

The numerical results we presented suggest that our approach is accurate even for power levels higher than typically encountered. 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]). Moreover, our approach substantially outperforms some of the best known joint scheduling and routing heuristics [13] in the regime of low power transmissions.

We end by noting that our results can also be used to provide a benchmark on achieved utility to which one can compare alternative transmission scheduling and routing policies.

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-efficient communication protocols for wireless microsensor networks," in *Proc. 33rd Hawaiian International Conference on Systems Science (HICSS-33)*, Maui, HI, USA, 2000.

[3] R. Krishnan and D. Starobinski, "Efficient clustering algorithms for self-organizing wireless sensor networks," *J. Ad-Hoc Networks*, 2004.

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

[5] S. Ray, D. Starobinski, A. Trachtenberg, and R. Ungrangsi, "Robust location detection with sensor networks," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 1016–1025, 2004.

[6] 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*, Paradise Island, The Bahamas, December 2004, pp. 1478–1483.

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

[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, 2003.

[9] 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.

[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 routing in UWB networks," *IEEE J. Select. Areas Commun.*, vol. 22, no. 7, pp. 1252 – 1270, 2004.

[14] X. Lin and N. B. Shroff, "Joint rate control and scheduling in multihop wireless networks," Preprint, February 2004.

[15] I. Paschalidis, W. Lai, and D. Starobinski, "Asymptotically optimal transmission policies for large-scale low-power wireless sensor networks," On-line version, http://ionia.bu.edu/, 2005.

[16] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. Prentice Hall, 1982.

[17] M. Grotschel, L. Lovasz, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*. Springer-Verlag, 1988.

[18] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization*. MPS-SIAM Series on Optimization, 2001.

PLACE PHOTO HERE

Ioannis Ch. Paschalidis (M'96) received the Diploma in ECE from the National Technical University of Athens, Athens, Greece, in 1991, and the S.M. and Ph.D. degrees in EECS from the Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts, in 1993 and 1996, respectively. He joined Boston University in 1996 where he is an Associate Professor of Manufacturing Engineering and the Co-Director of the Center for Information and Systems Engineering (CISE). He is also the Academic Director of the Sensor Network Consortium (SNC) – an industry consortium consisting of 14 companies. 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, optimization, networking, operations research, and computational biology. The main application areas he is targeting include communication and sensor networks, manufacturing systems, and supply chains.

He has received an NSF CAREER award (2000), the second prize in the 1997 George E. Nicholson paper competition by INFORMS, and was an invited participant at the 2002 Frontiers of Engineering Symposium, organized by the National Academy of Engineering. He is an associate editor of Operations Research Letters, Automatica, and the IEEE Transactions on Automatic Control.

PLACE PHOTO HERE

Wei Lai (S'03) received his B.E. in Automatic Control Engineering and the M.E. in Systems Engineering both from Huazhong University of Science and Technology, China, in 1999 and 2001, respectively. He is currently working towards his Ph.D. degree at Boston University. His research interests include the performance analysis and optimization of wireless and sensor networks.

PLACE PHOTO HERE

David Starobinski received his Ph.D. in Electrical Engineering (1999) from the Technion-Israel Institute of Technology. In 1999-2000, he was a visiting post-doctoral researcher in the EECS department at UC Berkeley. Since September 2000, he has been an assistant Professor at Boston University.

Dr. Starobinski received a CAREER award from the U.S. National Science Foundation and an Early Career Principal Investigator (ECPI) award from the U.S. Department of Energy. His research interests are in the modeling and performance evaluation of high-speed, wireless, and sensor networks.

APPENDIX I
PROOF OF LEMMA II.1

(i) Consider the following set

$$\bar{\mathscr{P}} = \{\mathbf{p} \mid p_{ijk} \in \{0, \bar{p}_{ijk}\}, \mathbf{p} \in \mathscr{P}\}$$

and note that every feasible $\mathbf{p} \in \mathscr{P}$ can be expressed as a convex combination of the points in $\bar{\mathscr{P}}$. Thus $\text{Conv}(\mathscr{P})$ is finitely generated by the points in $\bar{\mathscr{P}}$ and is therefore a polyhedron. Clearly, it is also bounded, hence a polytope (i.e., a bounded polyhedron). Since $\mathscr{R}$ is a linear mapping of $\mathscr{P}$, the same argument establishes that $\text{Conv}(\mathscr{R})$ is a polytope.

(ii) Next, let $\mathbf{r} \in \text{Conv}(\mathscr{R})$ and suppose that the extreme points of $\text{Conv}(\mathscr{R})$ are $\mathbf{r}^1, \ldots, \mathbf{r}^L$. These extreme points are also elements of $\mathscr{R}$, so they can be written as $\mathbf{r}^n = \mathbf{H}\mathbf{p}^n$ for some $\mathbf{p}^n \in \mathscr{P}$ and $n = 1, \ldots, L$. Thus, for $\alpha_n \geq 0$ satisfying $\sum_{n=1}^{L} \alpha_n = 1$,

$$\mathbf{r} = \sum_n \alpha_n \mathbf{r}^n = \sum_n \alpha_n \mathbf{H}\mathbf{p}^n = \mathbf{H}\mathbf{p},$$

where $\mathbf{p} = \sum_n \alpha_n \mathbf{p}^n \in \text{Conv}(\mathscr{P})$. Now, let $\mathbf{p} \in \text{Conv}(\mathscr{P})$ and let $\mathbf{p}^1, \ldots, \mathbf{p}^J$ be the extreme points of $\text{Conv}(\mathscr{P})$. For $\beta_i \geq 0$ satisfying $\sum_i \beta_i = 1$, $\mathbf{p} = \sum_i \beta_i \mathbf{p}^i$. These extreme points are also elements of $\mathscr{P}$. Set $\mathbf{r}^i = \mathbf{H}\mathbf{p}^i$; we have $\mathbf{r}^i \in \mathscr{R}$. It follows

$$\mathbf{H}\mathbf{p} = \sum_i \beta_i \mathbf{H}\mathbf{p}^i = \sum_i \beta_i \mathbf{r}^i \in \text{Conv}(\mathscr{R}).$$

(iii) Let $\mathbf{r}$ be an extreme point of $\text{Conv}(\mathscr{R})$. Then there exists a cost vector $\mathbf{c} \in \mathbb{R}^{(N+M)^2 K}$ such that $\mathbf{r}$ is the unique minimizer of $\mathbf{c}'\mathbf{x}$ over $\mathbf{x} \in \text{Conv}(\mathscr{R})$. Using part (ii), this latter problem is

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} \\ \text{s.t.} \quad & \mathbf{x} = \mathbf{H}\mathbf{p}, \ \mathbf{p} \in \text{Conv}(\mathscr{P}). \end{aligned} \tag{28}$$

Let $(\mathbf{r}, \mathbf{p}^*)$ be an optimal solution. Then $\mathbf{p}^*$ is an optimal solution to

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{H}\mathbf{p} \\ \text{s.t.} \quad & \mathbf{p} \in \text{Conv}(\mathscr{P}). \end{aligned} \tag{29}$$

(If not, there exists an optimal solution $\tilde{\mathbf{p}}$ of (29) and $\mathbf{H}\tilde{\mathbf{p}} \neq \mathbf{r}$ is optimal for (28), which contradicts the fact that $\mathbf{r}$ uniquely solves (28).) Now, there also exists some extreme point $\mathbf{p}^i$ of $\text{Conv}(\mathscr{P})$ which is optimal for (29). In particular, $\mathbf{c}'\mathbf{H}\mathbf{p}^i = \mathbf{c}'\mathbf{H}\mathbf{p}^* = \mathbf{c}'\mathbf{r}$ and $\mathbf{H}\mathbf{p}^i$ also solves (28). Since $\mathbf{r}$ is a unique solution of (28) we have $\mathbf{r} = \mathbf{H}\mathbf{p}^i$.

APPENDIX II
PROOF OF LEMMA VI.2

(i) As $B_1(\mathbf{r})$ and $B_2(\mathbf{r})$ both satisfy condition *(i)*, $B(\mathbf{r})$ also satisfies this condition.

(ii) Let $\mathbf{r}^1$, $\mathbf{r}^2$, and $\mathbf{t}^1$, $\mathbf{t}^2$ such that $\mathbf{r}^i - \mathbf{t}^i \in \text{Conv}(\mathscr{R})$ and $B_1(\mathbf{r}^i) = ||\mathbf{t}^i||_\infty$, for $i = 1, 2$. Then for any $\beta \in [0, 1]$, we have

$$\beta\mathbf{r}^1 + (1 - \beta)\mathbf{r}^2 - \beta\mathbf{t}^1 - (1 - \beta)\mathbf{t}^2 \in \text{Conv}(\mathscr{R}),$$

and

$$\begin{aligned} B_1(\beta\mathbf{r}^1 + (1 - \beta)\mathbf{r}^2) \quad &\leq ||\beta\mathbf{t}^1 + (1 - \beta)\mathbf{t}^2||_\infty \\ &\leq \beta||\mathbf{t}^1||_\infty + (1 - \beta)||\mathbf{t}^2||_\infty \\ &= \beta B_1(\mathbf{r}^1) + (1 - \beta)B_1(\mathbf{r}^2), \end{aligned}$$

where the first inequality is due to the definition of $B_1(\mathbf{r})$ and the second inequality follows from the convexity of L_∞ norm.

Next, note that $B_2(\mathbf{r}) = \max(0, \max_i((\mathbf{D}\mathbf{r})_i - d_i))$, where $(\mathbf{D}\mathbf{r})_i$ denotes the ith element of $\mathbf{D}\mathbf{r}$. It follows that $B_2(\mathbf{r})$ is convex in $\mathbf{r}$. As the pointwise maximum of two convex functions $B_1(\mathbf{r})$ and $B_2(\mathbf{r})$, $B(\mathbf{r})$ is also convex.

APPENDIX III
THE DECOMPOSITION ALGORITHM FOR UTILITY MAXIMIZATION SUBJECT TO A LIFETIME CONSTRAINT

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

$$\begin{aligned} \min \quad & -F(\mathbf{H}\mathbf{p}) \\ \text{s.t.} \quad & \mathbf{p} - \sum_{n=1}^{m} \alpha_n \mathbf{p}^n = 0, \\ & \sum_{n=1}^{m} \alpha_n = 1, \\ & \sum_{j=1}^{N+M} \frac{p_{ijk}G_{ij} - p_{jik}G_{ji}}{\eta} = 0, \ \forall i \neq s(k), d(k), \forall k, \\ & \mathbf{A}\mathbf{H}\mathbf{p} \leq \mathbf{b}, \\ & \mathbf{C}\mathbf{p} \leq \boldsymbol{\chi}/T, \\ & \alpha_n \geq 0, \ n = 1, \ldots, m, \end{aligned} \tag{30}$$

and the corresponding subproblem is

$$\begin{aligned} \max \quad & \boldsymbol{\lambda}'\mathbf{p} \\ \text{s.t.} \quad & \mathbf{p} \in \mathscr{P}, \end{aligned} \tag{31}$$

with cost vector $\boldsymbol{\lambda} = \boldsymbol{\lambda}^{(m)}$, where as before we have $(\mathbf{p}^{(m)}, \boldsymbol{\alpha}^{(m)}; \boldsymbol{\lambda}^{(m)}, \boldsymbol{\mu}^{(m)}, \boldsymbol{\nu}^{(m)}, \boldsymbol{\sigma}^{(m)}, \boldsymbol{\xi}^{(m)})$ as the optimal primal-dual pair for the restricted master problem (30) 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 (24) in a finite number of iterations. The argument is almost identical and is therefore omitted for brevity. Furthermore, problem (31) is still equivalent to the maximum weighted matching problem constructed in a similar way as in Sec. V, and is solvable in polynomial time. In particular, to solve problem (31), 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

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

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

satisfying the conditions

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

and

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

and let (i, j, k) be an element of $\mathcal{K}$. Given the graph $\mathcal{G}$ constructed above, we can obtain the optimal solution of (31) in the same way as in Sec. V.