OPTIMIZATION OF ENERGY CONSUMPTION IN WIRELESS SENSOR NETWORKS

*Angela Feistel**, *Marcin Wiczanowski**, *Sławomir Stańczak**†

* Heinrich-Hertz Chair, EECS,
University of Technology Berlin,
Einsteinufer 25, 10587 Berlin, Germany
angela.feistel@tu-berlin.de
marcin.wiczanowski@tu-berlin.de

† Fraunhofer German-Sino Lab
for Mobile Communications,
Einsteinufer 37,
10587 Berlin, Germany
stanczak@hhi.fhg.de

ABSTRACT

In wireless networks, a widely studied approach is to minimize transmit powers subject to some quality-of-service constraints. Minimizing transmit powers however is not equivalent to minimizing the energy consumption. In wireless sensor networks that operate in a low-rate regime, the discrepancy between power consumption and the overall energy consumption is even more evident since the energy expenditure for signal processing cannot be neglected. The paper addresses the problem of minimizing the overall energy consumption in wireless networks, including the energy expenditure for hardware. The objective is to give some valuable insights into the problem. In particular, we point out basic properties of the optimal power allocation and discuss properties of the general energy minimization problem. Due to the complexity, we focus on the problem of finding a power-time tradeoff by determining the energy-optimal number of parallel data streams per link for a certain SIR requirement.

1. INTRODUCTION

A crucial and important notion of a certain class of wireless networks is the energy consumption of network nodes. Indeed, the minimization of the energy consumption aims at saving energy costs and maximizing the lifetime duration of the network nodes. Especially in the context of sensor networks, energy consumption is of main interest. Due to their properties wireless sensor networks are attractive for a wide range of applications including surveillance, environmental monitoring, home automation, applications in industrial automation and control, and logistics.

However, some applications in wireless sensor networks not only require a long lifetime duration, yet also need to support certain QoS (Quality of Service) requirements. In most cases the requirements are in terms of bit-error-rate and/ or data-rate and can be uniquely mapped onto an SIR requirement. There exists a power control theory framework applicable to wireless networks and based on the SIR [1, 2, 3, 4, 5]; requiring a certain SIR target at each communicating node-pair, the sum transmit power of the wireless network has to be minimized. There also exist distributed algorithmic solutions for this problem, which are suitable for a self-organizing network [1].

In general, minimizing the total transmit power is not equivalent to minimizing the total energy consumption. The total energy consumption should be understood as transmission energy consumption together with hardware (or circuit) energy consumption. This notion of total energy depends on transmission time and sleeping time of the network. Thus, when energetic optimality in the above sense is desired, both transmission time and transmit power should be optimized jointly.

More precisely, a certain transmit power level is necessary to satisfy certain QoS requirements. If, e.g., the data-rate increases, the required transmit power level increases as well. However, at the same time the transmission time decreases, so that the change in energy spent for transmission mirrors the resulting shift in the trade-off between transmission time and transmission power. Consequently, to minimize the energy consumption we have to find the optimal power-time tradeoff subject to the given SIR requirements.

The paper is organized as follows. First we summarize some known results of the power minimization problem and point out basic properties of the optimal power allocation. Then we discuss the general problem of energy minimization and its relation to the power minimization problem. Due to the fact, that the general problem of energy minimization is difficult to solve we focus on a simplified version that provides a power-time tradeoff. Finally, we propose an algorithm that determines the power-time tradeoff by finding the energy-optimal number of data streams per link for a certain SIR requirement and give some numerical results.

2. SYSTEM MODEL

2.1. Network model

In this paper we consider a multiple antenna wireless network with K transmitter-receiver pairs, where each transmitter and receiver is equipped with N antennas. Let $\mathbf{H}^{(k,l)}$ de-

This work was partly supported by the German Ministry for Education and Research (BMBF) under Grant 01BU0680.

note the channel between the transmitter of the l-th link and the receiver of the k-th link. Unless otherwise stated, assume that $\mathrm{rank}(\mathbf{H}^{(k,l)}) = L$, with $1 \le L \le N$. Therefore, on each link, $L \ge 1$ data streams can be transmitted in parallel. Let $\mathbf{U}^{(k)} = (\mathbf{u}^{(k,j)})_{1\le j\le L}$ and $\mathbf{V}^{(k)} = (\mathbf{v}^{(k,j)})_{1\le j\le L}$ denote link-specific matrices whose columns are transmit and receive beamforming vectors of the L data streams of the k-th link, respectively. Throughout the paper, the beamforming vectors are arbitrary but fixed. Without loss of generality, we can assume that $\|\mathbf{u}^{(k,j)}\|_2 = 1$ and $\|\mathbf{v}^{(k,j)}\|_2 = 1$, $1 \le k \le K$, $1 \le j \le L$. We group the transmit powers assigned to the data streams of link k in the vector $\mathbf{p}^{(k)} = (p_1^{(k)}, \dots, p_L^{(k)})$, $1 \le k \le K$. Due to the power constraints on each transmitting node, we always have $\|\mathbf{p}^{(k)}\|_1 \le P, 1 \le k \le K$, for some given constant P. We further define $\mathbf{P}^{(k)} = diag(\mathbf{p}^{(k)}), \quad 1 \le k \le K$, and group the spatial link power allocations and link beamformers in the sets $\mathcal{P} = \{\mathbf{P}^{(k)}\}_{1\le k\le K}$, $\mathcal{U} = \{\mathbf{U}^{(k)}\}_{1\le k\le K}$, $\mathcal{V} = \{\mathbf{V}^{(k)}\}_{1\le k\le K}$. Thus, provided that all links are simultaneously active in the same frequency band, the SIR of the l-th data stream of the k-th link can be written as

$$\mathrm{SIR}_{k,j} := \frac{p_j^{(k)} D_j^{(k)}}{\sum_{l=1}^{K}\sum_{i=1}^{L} p_i^{(l)} G_{j,i}^{(k,l)} + \sigma^2},$$

where $G_{j,j}^{(k,k)} := 0, 1 \le k \le K, 1 \le j \le L$ and

$$D_j^{(k)} = |\mathbf{v}^{(k,j)H}\mathbf{H}^{(k,k)}\mathbf{u}^{(k,j)}|^2,$$
$$G_{j,i}^{(k,l)} = |\mathbf{v}^{(k,j)H}\mathbf{H}^{(k,l)}\mathbf{u}^{(l,i)}|^2$$

and $\sigma^2 > 0$ is the variance of the spatially uncorrelated white Gaussian noise on each link. Note that in general, the SIR defined above depends on transmit powers and beamformers of all links as well as on the receive beamformer of the desired link. We call $D_j^{(k)}$ and $G_{j,i}^{(k,l)}$ the desired and interference channel gains, respectively. Let

$$\mathbf{D}^{(k)} = \mathrm{diag}\big(D_1^{(k)}, \dots, D_L^{(k)}\big)$$

be the $L \times L$ diagonal matrix of the desired channel gains of link k and

$$\mathbf{G}^{(k,l)} = \big(G_{j,i}^{(k,l)}\big)_{1\le i,j\le L}$$

the $L\times L$ matrix of the interference gains between transmitter l and receiver k. Note that by assumption , $\mathrm{trace}(\mathbf{G}^{(k,k)}) = 0$ for each $1 \le k \le K$.

2.2. SIR metrics

In our optimization approach, we put certain requirements on a suitably chosen SIR metric. More precisely, we require that

$$\theta(\mathrm{SIR}_{k,1}, \dots, \mathrm{SIR}_{k,L}) \ge \gamma_k, 1 \le k \le K \qquad (1)$$

where $\theta : \mathbb{R}_+^L \to \mathbb{R}_+$ is a SIR metric that assigns a single nonnegative value to the several SIRs of the data streams. Given a specific SIR metric, this value can be interpreted as a "virtual" SIR that should reflect the quality-of-service perceived by a link with multiple data streams. In this paper, we distinguish two types of the SIR metric. The first one is the minimum metric defined to be

$$\theta(\mathrm{SIR}_{k,1}, \dots, \mathrm{SIR}_{k,L}) = \min_{1\le i\le L} \mathrm{SIR}_{k,i}, 1 \le k \le K\,. \qquad (2)$$

By (1), this metric is suitable when data streams of each link k are treated independently and each data stream needs to satisfy the common SIR requirement γ_k. Note that if the data rate is a strictly increasing function of the SIR, a parallel transmission of L data streams under the assumption of (1) and (2) increases the data rate of each link by the factor of L, when compared with the one data stream strategy. However, the requirement that each data stream must satisfy a common SIR requirement is likely to be infeasible in multiple antenna systems with independent data streams (see also sections 5 and 6). Even if the SIR requirement is feasible, the necessary transmit powers for achieving it is usually relatively high. This motivates the second metric of interest in this paper, the so-called average metric defined to be ($1 \le k \le K$)

$$\theta(\mathrm{SIR}_{k,1}, \dots, \mathrm{SIR}_{k,L}) = \exp\Big(\frac{1}{L}\sum_{i=1}^{L} \log(1 + \mathrm{SIR}_{k,i})\Big) - 1\,. \qquad (3)$$

In order to motivate this metric, assume for a moment that there is only one data stream on each link and the SIR of link k (the virtual SIR) is denoted by SIR_k. If a linear receiver is then followed by single-user decoders, one for each link, the data rate achieved on link k under an independent Gaussian input distribution is equal to $\log(1 + \mathrm{SIR}_k)$ nats per channel use (with some possible constants that are neglected here), where $\log(x), x > 0$, denotes the natural logarithm. Consequently, if there are L data streams on link k and each data stream is equipped with a single-user decoder, the achievable data rate is equal to $\sum_l \log(1 + \mathrm{SIR}_{k,l})$. Now, as in the case of the minimum metric, we require that a parallel transmission of L data streams provides L times increase in the data rate so that we can write

$$L\log(1 + \mathrm{SIR}_k) = \sum_l \log(1 + \mathrm{SIR}_{k,l}),\ 1 \le k \le K\,.$$

From this, it follows that SIR_k in the case of $L = 1$ which is necessary to support the same link rate as in the case of L parallel data streams is given by

$$\mathrm{SIR}_k = \theta(\mathrm{SIR}_{k,1}, \dots, \mathrm{SIR}_{k,L}), 1 \le k \le K$$

where θ is the average metric defined by (3).

The average metric is suitable when the L streams of a single link are treated separately and are allowed to operate at different SIR levels in order to support different modulation levels, and hence different data rates. The problem of choosing SIR levels for different data streams is addressed in Section 5.3.

Note that the simplified version of the average metric with $\log(1+\mathrm{SIR}) \approx \mathrm{SIR}$ represents an aggregated SIR measure of the link k: $\theta(\mathrm{SIR}_{k,1}, \ldots, \mathrm{SIR}_{k,L}) = 1/L \sum_{i=1}^{K} \mathrm{SIR}_{k,i}$. However, this metric is not considered here.

2.3. Standard, hardware and energy model

The network operation time is partitioned into frames of duration T_F, with each frame consisting of the transmit time and sleep time, in which the transmitter and receiver hardware are in the sleep (stand-by) mode. In this paper, it is assumed that the symbol rate (or the symbol duration) on each parallel data stream is fixed to comply with some practical constraints. In other words, the symbol rates of the L data streams are the same and equal to some predetermined basic symbol rate.[1] Therefore, given a set of beamformers and transmit powers, the symbol rate per link can be increased only through the increase of the number of parallel data streams L. As an immediate consequence, the transmit time of link k is equal to T_A/L, where T_A is the time which is necessary to transmit a data packet of fixed length using the corresponding single data stream strategy with a L times lower data rate. Consequently, for the sleep time T_S we have $T_S = T_F - T_A/L^2$.

With the expressions for transmit time and sleep time, the energy-consumption $E = E_A + E_S$ can be determined, where E_A and E_S are the energy consumption in transmit mode and sleep mode, respectively. The transceiver hardware of each node is assumed to consist of a microcontroller and N transceiver chips, each supporting one antenna element. Let P_A and P_S denote the hardware-related power consumption in active and sleep mode. Then we have

$$\begin{aligned} E = E_A + E_S &= \sum_{k=1}^{K} T_A \frac{\mathrm{trace}(\mathbf{P}^{(k)}) + P_A}{L} + K T_S P_S \\ &= \sum_{k=1}^{K} T_A \frac{\mathrm{trace}(\mathbf{P}^{(k)}) + P_A - P_S}{L} + K T_F P_S, \end{aligned} \tag{4}$$

where

$$\begin{aligned} P_A &= P_A^C(N) + P_A^M \\ P_S &= P_S^C(N) + P_S^M. \end{aligned} \tag{5}$$

Here P_A^C, P_S^C denote the power consumption of the transceiver chips(including both nodes of the link) in active and sleep mode, P_A^M, P_S^M denote the power consumption of the microcontroller (including both nodes of the link) in active and sleep mode.

[1] However, note that the transmitted symbols can carry different numbers of information bits.

[2] In the following we neglect the time to transmit control data or pilot symbols for channel estimation.

3. TRANSMIT POWER MINIMIZATION

Multiple antenna systems are capable of providing high data rate transmissions in a fading environment without the need of increasing the signal bandwidth. This suggests the question whether they can be used to reduce the energy consumption in wireless networks. Before addressing this problem, we must first understand the problem of minimizing transmit powers subject to some SIR requirements. In fact, note that for a fixed transmit time T_A, a fixed number of antennas N and fixed beamformers $\mathbf{U}^{(k)}, \mathbf{V}^{(k)}$ at each transmitter-receiver pair, the problem of energy minimization subject to the SIR requirements is equivalent to minimizing transmit powers subject to the same SIR requirements. In this section, we briefly summarize some known results on the power minimization problem and point out basic properties of the optimal power allocation.

Let $\gamma_l^{(k)}$ be the SIR requirement for the lth data stream of link k. We define

$$\mathbf{\Gamma} = \mathrm{diag}\Big(\gamma_1^{(1)}, \ldots, \gamma_L^{(1)}, \ldots, \gamma_1^{(K)}, \ldots, \gamma_L^{(K)})\Big)$$

and, without loss of generality, assume that $\mathbf{\Gamma}$ is positive definite. Now we say that $\mathbf{\Gamma}$ is feasible if there exists a power vector $\mathbf{p} > 0$ (called a valid power vector) such that $\mathrm{SIR}_{k,l}(\mathbf{p}) \geq \gamma_l^{(k)}$ for all k and l. It is well-known [2, 4, 5] that $\mathbf{\Gamma}$ is feasible if and only if

$$\rho(\mathbf{\Gamma}\mathbf{D}^{-1}\mathbf{G}) < 1 \tag{6}$$

where $\rho(\mathbf{\Gamma}\mathbf{D}^{-1}\mathbf{G})$ is the spectral radius of $\mathbf{\Gamma}\mathbf{D}^{-1}\mathbf{G}$, $\mathbf{D} = \mathrm{diag}(\mathbf{D}^{(1)}, \ldots, \mathbf{D}^{(K)})$, and $\mathbf{G} = (\mathbf{G}^{(k,l)})$ denotes a matrix whose (k,l)th entry is the $L \times L$ matrix $\mathbf{G}^{(k,l)}$ as defined in section 2.1. Hence, $\mathbf{\Gamma}$ is feasible if and only if $\mathbf{\Gamma} \in \mathrm{F}$ where

$$\mathrm{F} := \{\mathbf{\Gamma} : \rho(\mathbf{\Gamma}\mathbf{D}^{-1}\mathbf{G}) < 1\} \tag{7}$$

Now, if condition (6) is fulfilled, there is a unique valid power vector $\mathbf{p}^*(\mathbf{\Gamma}) = (p_1^*(\mathbf{\Gamma}), \ldots, p_K^*(\mathbf{\Gamma})) > 0$ such that $\mathbf{p}^*(\mathbf{\Gamma}) \leq \mathbf{p}$ for all other valid power vectors $\mathbf{p}$ [5]. In words, $\mathbf{p}^*(\mathbf{\Gamma})$ is the unique component-wise minimal valid power vector. Furthermore, if (6) holds, then

$$\begin{aligned} \mathbf{p}^*(\mathbf{\Gamma}) &= \sigma^2 \Big(\mathbf{\Gamma}^{-1}\mathbf{D} - \mathbf{G}\Big)^{-1} \mathbf{1} \\ &= \sigma^2 \Big[\sum_{l=0}^{\infty} (\mathbf{\Gamma}\mathbf{D}^{-1}\mathbf{G})^l\Big] \mathbf{\Gamma}\mathbf{D}^{-1}\mathbf{1} \\ &= \sigma^2 \mathbf{\Gamma}\mathbf{D}^{-1}\mathbf{1} + \sigma^2 \Big[\sum_{l=1}^{\infty} (\mathbf{\Gamma}\mathbf{D}^{-1}\mathbf{G})^l\Big] \mathbf{\Gamma}\mathbf{D}^{-1}\mathbf{1}, \end{aligned} \tag{8}$$

where we used the Neumann series in the second step [6]. We point out that $\mathbf{p}^*(\mathbf{\Gamma})$ is unique and positive for any nonnegative (not necessarily irreducible) matrix $\mathbf{D}^{-1}\mathbf{G}$, provided that $\mathbf{\Gamma}$ is positive definite [5]. Obviously, $\mathbf{\Gamma}$ is positive definite if and only if all data streams have a (strictly) positive SIR requirement. By (8), we see that for sufficiently small SIR requirements and interference gains in $\mathbf{G}$, one can approximate $\mathbf{p}^*(\mathbf{\Gamma})$ by a finite series

$$\mathbf{p}^*(\mathbf{\Gamma}) = \sigma^2\mathbf{\Gamma}\mathbf{D}^{-1}\mathbf{1} + \sigma^2\Big[\sum_{l=1}^{t}(\mathbf{\Gamma}\mathbf{D}^{-1}\mathbf{G})^l\Big]\mathbf{\Gamma}\mathbf{D}^{-1}\mathbf{1}$$

for some $t \geq 1$.

Now since $\mathbf{p}^*(\mathbf{\Gamma})$ is the component-wise minimal, the total power consumption is equal to

$$P_{\min}(\mathbf{\Gamma}) := \|\mathbf{p}^*(\mathbf{\Gamma})\|_1 = \sigma^2\mathbf{1}^T\Big(\mathbf{\Gamma}^{-1}\mathbf{D}-\mathbf{G}\Big)^{-1}\mathbf{1} > 0\,. \quad (9)$$

Of course,

$$P_{\min}(\mathbf{\Gamma}) \to \begin{cases} 0 & \mathbf{\Gamma}\to\mathbf{0}, \mathbf{\Gamma}\in C(\mathbf{0}) \\ +\infty & \mathbf{\Gamma}\to\hat{\mathbf{\Gamma}}, \mathbf{\Gamma}\in C(\hat{\mathbf{\Gamma}}), \hat{\mathbf{\Gamma}}\in\partial\mathrm{F}, \end{cases} \quad (10)$$

where $C(\mathbf{\Gamma}) \in \mathrm{F}$ is any path (curve) in F that ends at the point $\mathbf{\Gamma}$ and $\partial\mathrm{F} = \{\mathbf{\Gamma} : \rho(\mathbf{\Gamma}\mathbf{D}^{-1}\mathbf{G}) = 1\}$ is the boundary of the feasible SIR region F defined by (7).

From practical point of view, it is worth pointing out that the power vector (8) can be determined in a distributed manner using the iterative algorithm of [1, 7]. In what follows, $\mathbf{p}^*(\mathbf{\Gamma})$ is referred to as the optimal power vector.

3.1. Some properties of the optimal power vector

In the following, we show some properties of the optimal power vector $\mathbf{p}^*(\gamma\mathbf{I})$ under the assumption that all SIR requirements are equal, that is, we have $\mathbf{\Gamma} = \gamma\mathbf{I}$. Without loss of generality, we further assume that $\sigma^2 = 1$.

First, let us consider the first derivation of

$$\begin{aligned} p_k^*(\gamma\mathbf{I}) &= \mathbf{e}_k^T\Big(\frac{1}{\gamma}\mathbf{I}-\mathbf{D}^{-1}\mathbf{G}\Big)^{-1}\mathbf{D}^{-1}\mathbf{1} \\ &= \mathbf{e}_k^T\Big(\frac{1}{\gamma}\mathbf{D}-\mathbf{G}\Big)^{-1}\mathbf{1} \end{aligned} \quad (11)$$

where $\mathbf{e}_k = (0,\ldots,0,1,0,\ldots,0)$ denotes the unity vector with 1 at the k-th position. We use the following relation: for each $x \in D \subseteq \mathbb{R}$ (D is an open set) and an arbitrary invertible matrix $\mathbf{A}(x)$, whose entries are twice continuously differentiable functions of $x \in D$, there holds $\mathbf{A}(x)\mathbf{A}^{-1}(x) = \mathbf{I}$, for all $x \in D \subseteq \mathbb{R}$, and hence we have

$$\frac{d\mathbf{A}^{-1}}{dx}(x) = -\mathbf{A}^{-1}(x)\frac{d\mathbf{A}(x)}{dx}\mathbf{A}^{-1}(x),\ x\in D\,.$$

Now we define $\mathbf{A}(\gamma) = (1/\gamma\mathbf{D}-\mathbf{G})$ with $\gamma \in D = (0, 1/\rho(\mathbf{D}^{-1}\mathbf{G}))$. Note, that $D = (0,+\infty) = \mathbb{R}_{++}$, if $\rho(\mathbf{D}^{-1}\mathbf{G}) = 0$. Now since $(d\mathbf{A}/d\gamma)(\gamma) = -1/\gamma^2\mathbf{D}$, we have

$$\begin{aligned} \frac{dp_k^*}{d\gamma}(\gamma\mathbf{I}) &= \mathbf{e}_k^T(\mathbf{D}-\gamma\mathbf{G})^{-1}\mathbf{D}(\mathbf{D}-\gamma\mathbf{G})^{-1}\mathbf{1} \\ &= \mathbf{e}_k^T(\mathbf{I}-\gamma\mathbf{D}^{-1}\mathbf{G})^{-2}\mathbf{D}^{-1}\mathbf{1} > 0\,. \end{aligned}$$

Due to the fact, that $\mathbf{D}$ is a positive definite matrix, the first derivation is positive for each k. Consequently $p_k^*(\gamma\mathbf{I})$ is a strictly monotonically increasing function. Further, we have

$$P'_{\min}(\gamma\mathbf{I}) = \frac{dP_{\min}}{d\gamma}(\gamma\mathbf{I}) = \mathbf{1}^T(\mathbf{I}-\gamma\mathbf{D}^{-1}\mathbf{G})^{-2}\mathbf{D}^{-1}\mathbf{1} > 0\,.$$

Both derivations show how fast $p_k(\gamma\mathbf{I})$ and $P_{\min}(\gamma\mathbf{I})$ change depending on $\gamma \in (0, 1/\rho(\mathbf{D}^{-1}\mathbf{G}))$. If for example $\gamma > 0$ is sufficiently small, then

$$P'_{\min}(\gamma\mathbf{I}) \approx \mathbf{1}^T(\mathbf{I}+\gamma\mathbf{D}^{-1}\mathbf{G})^2\mathbf{D}^{-1}\mathbf{1},\ \mathbf{G}\neq\mathbf{0}\,.$$

In particular, since $\lim_{\gamma\to 0} P'_{\min}(\gamma\mathbf{I}) = \mathbf{1}^T\mathbf{D}^{-1}\mathbf{1}$, the slope of $P_{\min}$ for sufficiently small values of $\gamma > 0$ only depends on the diagonal elements of the matrix $\mathbf{D}^{-1}$, which are determined by the quality of the desired channels.

Considering the second derivative

$$\begin{aligned} &\frac{d^2\mathbf{A}^{-1}}{dx^2}(x) = \\ &\mathbf{A}^{-1}(x)\Big(2\frac{d\mathbf{A}(x)}{dx}\mathbf{A}^{-1}(x)\frac{d\mathbf{A}(x)}{dx} - \frac{d^2\mathbf{A}(x)}{dx^2}\Big)\mathbf{A}^{-1}(x) \end{aligned}$$

and the fact that $(d^2\mathbf{A}/d\gamma^2)(\gamma) = 2/\gamma^3\mathbf{D}, \gamma \in (0, 1/\rho(\mathbf{D}^{-1}\mathbf{G}))$, one obtains after some elementary calculations

$$\begin{aligned} \frac{d^2p_k}{d\gamma^2}(\gamma\mathbf{I}) &= 2\mathbf{e}_k^T(\mathbf{I}-\gamma\mathbf{D}^{-1}\mathbf{G})^{-3}\mathbf{D}^{-1}\mathbf{G}\mathbf{D}^{-1}\mathbf{1} \\ \frac{d^2P_{\min}}{d\gamma^2}(\gamma\mathbf{I}) &= 2\mathbf{1}^T(\mathbf{I}-\gamma\mathbf{D}^{-1}\mathbf{G})^{-3}\mathbf{D}^{-1}\mathbf{G}\mathbf{D}^{-1}\mathbf{1}\,. \end{aligned}$$

The second derivations are non-negative for all $\gamma \in D = (0, 1/\rho(\mathbf{D}^{-1}\mathbf{G}))$ because $(\mathbf{I}-\gamma\mathbf{D}^{-1}\mathbf{G})^{-3}$ is a non-negative matrix. This implies that both $p_k^*(\gamma\mathbf{I})$ and $P_{\min}(\gamma\mathbf{I})$ are convex functions on D. Clearly, the second derivative of $P_{\min}$ is strictly positive unless $\mathbf{G} \geq 0$ is a zero matrix. Hence $P_{\min}(\gamma\mathbf{I})$ is a strictly convex function (except in the orthogonal case). Now, if $\mathbf{G} \geq 0$ is irreducible, then $p_k^*(\gamma\mathbf{I})$ is strictly convex for all k. This follows from the fact that $\mathbf{D}^{-1}\mathbf{G}\mathbf{D}^{-1}\mathbf{1} \neq \mathbf{0}$ is a non-negative vector and $(\mathbf{I}-\gamma\mathbf{D}^{-1}\mathbf{G})^{-1}$ is a positive matrix for all $\gamma \in (0, 1/\rho(\mathbf{D}^{-1}\mathbf{G}))$ and any irreducible matrix $\mathbf{G}$. Multiplying a non-negative vector (except the zero vector) with a positive matrix results in a positive vector. The positivity of $(\mathbf{I}-\gamma\mathbf{D}^{-1}\mathbf{G})^{-1}$ follows from the Neumann series and the fact, that a quadratic non-negative matrix $\mathbf{A}$ is irreducible if and only if for all pairs (i,j), there exists a number $n \in \mathbb{N}$, such that $(\mathbf{A}^n)_{i,j} > 0$ [8, 5].

3.2. Discussion

The problem of minimizing transmit powers subject to SIR requirements is well understood. The optimal power vector can be explicitly computed when the channel matrices $\mathbf{D}$ and $\mathbf{G}$ are known. If they are not known, each link can iteratively compute its transmit power, provided that there is some coarse synchronization between links and a low-rate feedback channel for each transmitter-receiver pair [1].

By Section 3.1, it follows that for sufficiently small SIR requirements, the total transmit power $P_{\min}(\gamma\mathbf{I})$ can be assumed to linearly depend on the SIR requirement γ, with the slope determined by the diagonal entries of $\mathbf{D}$. Since the second derivative is positive (except for the case of mutually orthogonal links), the growth rate of the total transmit power increases as γ increases. Furthermore, by (10), it diverges to infinity as γ approaches $1/\rho(\mathbf{D}^{-1}\mathbf{G})$.

As a consequence, if the operating point is in the low SIR regime, then transmit and receive beamformers should be matched to the desired channel in the sense that $\mathbf{D}^{(k)} = \text{diag}(\lambda_1^{(k)}, \ldots, \lambda_L^{(k)})$ where $\lambda_1^{(k)}, \ldots, \lambda_L^{(k)} > 0$ are the largest *positive* eigenvalues of the matrix $\mathbf{H}^{(k,k)}\mathbf{H}^{(k,k)H}$. This is not true for relatively large SIR requirements γ, in which case it may be better to choose the beamformers so as to avoid the interference. Note that for a sufficiently large SIR requirement, the total transmit power is (to a large extent) determined by the value of the spectral radius $\rho(\mathbf{D}^{-1}\mathbf{G})$, which is strongly influenced by the matrix of interference gains $\mathbf{G}$. In fact, by [9], we know that for sufficiently large $\gamma \in (0, 1/\rho(\mathbf{D}^{-1}\mathbf{G}))$,

$$P_{\min}(\gamma\mathbf{I}) \approx \frac{\gamma\,\alpha}{1 - \gamma\rho(\mathbf{D}^{-1}\mathbf{G})}$$

where $\alpha > 0$ is a constant that depends on the matrices $\mathbf{D}$ and $\mathbf{G}$. Finally, note that supporting a common SIR requirement on L independent parallel data streams can be very insufficient in terms of the power consumption since some of the L positive eigenvalues of the matrix $\mathbf{H}^{(k,k)}\mathbf{H}^{(k,k)H}$ may be relatively small.

4. THE PROBLEM OF ENERGY MINIMIZATION

The previous section was devoted to the problem of minimizing transmit powers for given fixed system parameters $(N, L, \mathcal{U}, \mathcal{V})$. In particular, this implies a fixed transmit power $\mathcal{P}$ and makes the power vector in (8) be also optimal in the sense of minimizing the overall energy consumption. However, as mentioned in the introduction, it may be possible to achieve further energy savings by jointly optimizing transmit powers and transmit time for a given set of transmit and receive beamformers. In our setting (see Section 2 and in particular Section 2.3), this means to minimize the overall (total) energy consumption with respect to $(\mathcal{P}, L)$ where $\mathcal{P}$ satisfies the power constraints and $L \in \{1, \ldots, N\}$. In this section, we also discuss a more general (and significantly more complicated) problem, in which links can choose different numbers of parallel data streams $L_1, \ldots, L_K \in \{1, \ldots, N\}$. Note that in this case, $\mathcal{P}$ is a set of time varying transmit power matrices such that the instantaneous transmit powers depend on the overall number of active data streams. This approach is not of interest in practice due to its complexity so that our focus later in the paper will be on the case of a common number of data streams L per link.

Multiple antenna techniques may require significantly more hardware. Indeed, each antenna element is connected with its own transceiver chip, and hence nodes equipped with multiple antennas need more energy. Moreover, additional channel estimation and the need for more control data will further increase the energy consumption. On the other hand, multiple antenna techniques may improve the robustness of the network or may support higher SIR requirements. This is of interest in networks, where the QoS-requirements are more stringent and the requirements on outage probability are quite low. Another possibility is to use multiple antenna techniques to increase the modulation level of the data stream or to transmit multiple data streams. Then the transmit time may be decreased and energy may be saved. Indeed, diverse multiple antenna techniques can be used. More precisely, one can think of classical beamforming, space time codes, the use of diversity in form of antenna selection or the transmission of multiple data streams, where each of these techniques has its advantages and disadvantages. Besides one can think of optimizing transmit power, transmit beamformers and receive beamformers jointly [10, 11, 12].

Nevertheless, optimizing over all variables $(N, L, \mathcal{U}, \mathcal{P}, \mathcal{V})$ – that means number of antennas, number of data streams, transmission strategy and transmit power – is a problem, that is difficult to solve. So we intend to investigate a simplified version of this problem, where $(N, \mathcal{U}, \mathcal{V})$ are fixed. Remember, the energy consumption of the hardware has a significant influence on the total energy consumption. Clearly, the hardware energy consumption depends on the number of antennas, where each antenna element requires its own transceiver chip, the transmission time and the sleeping time of the network. Considering the second factor due to fixed N, the transmission time may be shortened, if the links transmit at a higher rate. A higher rate can be achieved by increasing the modulation level, that implicitly requires to adapt the SIR requirements, or by transmitting several data streams. Hence, as can be easily seen, the transmit power is a function of the transmission time. More precisely, if the data rate increases, the required transmit power level increases as well. However, at the same time the transmission time decreases. Now the question arises, whether the change of energy spent for transmission mirrors the change of energy spent for hardware. Therefore, in the following we minimize the energy consumption finding the optimal power-time tradeoff subject to given SIR requirements.

As assumed in section 2.3, the data rate per link can be increased only through the increase of the number of parallel data streams L. This leads to a transmission time of T_A/L. Now, assuming that $N, \mathcal{U}, \mathcal{V}$ are fixed, we consider the restricted problem $E(\mathcal{P}, L)$ that minimizes the energy consumption subject to certain requirements on link SIR metrics. The optimization problem, that provides a power-time tradeoff, can be written as follows

$$\min_{L,\mathcal{P}} E(L,\mathcal{P}) \quad s.t.$$
$$\theta(\mathrm{SIR}_{k,1}(L,\mathcal{P}),\ldots,\mathrm{SIR}_{k,L}(L,\mathcal{P})) \geq \gamma_k, \quad 1 \leq k \leq K. \tag{12}$$

Note, that the upper problem can be thought of as more general, if the links choose different numbers of parallel data streams $L_1,\ldots,L_K$. Under the assumption that the same data amount is transmitted during a frame it follows that the links transmit during the transmit times $T_A/L_1,\ldots,T_A/L_K$ that may be different long. As an immediate consequence, the optimal transmit power vector changes over time. We have maximal $\max(L_1,\ldots,L_K) \leq N$ different power vectors, that need to satisfy the SIR requirements. Note, that this problem is a combinatorial one. Clearly, if the maximum number of parallel data streams is N, there exist N^K possible combinations of active parallel data streams $(L_1,\ldots,L_K)$. Assuming, that the number of antennas and the number of active links in the neighbourhood is limited, the problem can be solved by comparing all possible combinations. Otherwise a simple heuristic is needed. However, due to its complexity this approach is not of interest in practice. Nevertheless, it provides a lower bound.

5. ADAPTIVE ENERGY MINIMIZATION FOR FIXED BEAMFORMERS

In this section we first characterize problem (12) and then derive an algorithm to solving it in order to find the energy-optimal number of data streams per link L^*, which uniquely determines the energy-optimal power allocation $\mathcal{P}^*$, for a certain SIR requirement. Note, that the beamformers $\mathbf{U}^{(k)}$,$\mathbf{V}^{(k)}$, $\forall_k$ and the number of antennas N are fixed.

5.1. Problem characterization

By (12), the problem is to find an optimal tradeoff between transmit powers trace$(\mathbf{P}^{(k)})$ (which should be as low as possible) and the number of parallel data streams L (which should be as large as possible in order to switch off the hardware components). Note that these objectives are opposed since a larger number of parallel data streams implies higher transmit powers that are necessary to satisfy the SIR requirements. To achieve the optimal trade-off, we discuss the problem in an incremental way. Define

$$t_\Sigma(L_0) = \min_{\mathcal{P}} \frac{1}{L_0}\sum_{k=1}^{K}(\mathrm{trace}(\mathbf{P}^{(k)}) + P_A - P_S) \quad s.t. \tag{13}$$
$$\theta(\mathrm{SIR}_{k,1},\ldots,\mathrm{SIR}_{k,L}) \geq \gamma_k, \quad 1 \leq k \leq K,$$

where $t_\Sigma(L_0)$ denotes the minimum sum power to fulfill the requirements for a certain number L_0 of parallel data streams.

We start with $L = L_0$. If L increases to $L = L + 1$, an energetic advantage can be recognized, if $t_\Sigma(L) - t_\Sigma(L-1) < 0$. Then an additional data stream decreases the energy consumption. As an immediate consequence, we simple have to solve the following problem

$$L^* = \max L \quad s.t. \quad t_\Sigma(L) - t_\Sigma(L-1) < 0. \tag{14}$$

Hence, the energy optimal strategy can be achieved by maximal L^* calculations of energy optimal strategies.

5.2. Algorithm

The following algorithm is derived from section 5.1 and can be applied using an arbitrary SIR metric. It presents a simple way to solve (12) and determines the energy-optimal number of data streams per link L^* and the energy-optimal power allocation for a certain SIR requirement. The algorithm works as follows (for an arbitrary $\mathbf{\Gamma}$):

i Initialization $L = 0$, $L^* = 0$

ii $L = L + 1$

iii if $E(\mathbf{\Gamma}, L-1) < E(\mathbf{\Gamma}, L)$ $\qquad L^* = L - 1$
else (if $L < N$ goto (ii) else $L^* = L$)

5.3. Average metric

Given the SIR requirement γ_k for each link k, in this section we propose a heuristic approach to specify the SIR requirements of each data stream $\gamma_l^{(k)}$ for the average metric and discuss whether the heuristic is convenient.

If the beamformers have to be fixed, a common and simple choice is to adapt them to the channel. The singular value decomposition of the channel of link k provides the transmit beamformers and receive beamformers and the singular values of the data streams: $\mathbf{H}^{(k,k)} = \mathbf{V}^{(k)}\mathbf{S}^{(k)}\mathbf{U}^{(k)}$, where the diagonal elements of $\mathbf{S}^{(k)}$ represent the ordered singular values $(s_1^{(k)},\ldots,s_N^{(k)}) \geq 0$ of link k. If no interference occurs, the SIR of link k and data stream l is simply $\mathrm{SIR}_{k,l} = p_l^{(k)} {s_l^{(k)}}^2/\sigma^2$. According to this relation it seems reasonable to determine the SIR requirements for each data stream such that equation (3) is fulfilled and we have

$$\frac{\gamma_1^{(k)}}{{s_1^{(k)}}^2} = \ldots = \frac{\gamma_L^{(k)}}{{s_L^{(k)}}^2}, \forall_k.$$

Clearly, in practice there is interference on each link and hence this choice may be not optimal.

6. NUMERICAL RESULTS

This section present some numerical results. We consider a simple network with $N = 4$ antennas per node and $K =$

15 transmitter-receiver pairs. The simulation parameters are chosen as follows: $T_F = 3$s, $T_A = 4$ms, $P_A^C = 0.035$W, $P_A^M = 0.003$W, $P_S = 0.00005$W. The channel is modeled as a channel with a line-of-sight path and several non-line-of-sight paths, where the Rice-factor is chosen to be 4. The average distance of the transmitter-receiver pairs is $d = 5$m, the path-loss exponent $n = 3.5$. Figure 1 depicts the en-

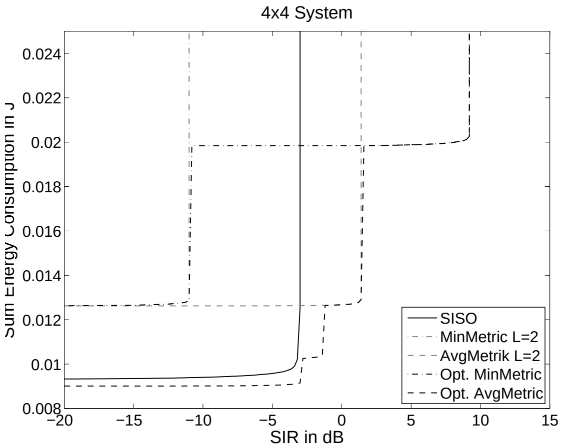

Fig. 1. Energy consumption over SIR requirements, $N = 4$

ergy consumption over SIR requirements for the following 5 transmission strategies. We compare the simple SISO-case with the proposed algorithm that determines adaptively the energy-optimal number of data streams for both average and minimum metric. Further, we depict the energy consumption of the average and minimum metric for a fixed number $L = 2$ of data streams. In both cases (adaptive and fixed number of data streams) the beamformers are adapted to the channel.

The simulation results show for both minimum and average metric, that with increasing SIR requirements γ_k, $1 \leq k \leq K$ the energy-optimal number of data streams L^* decreases. The stepwise changes of the curve follow from the discrete transitions from L to $L-1$ as energy optimal number of data streams. Indeed, if the number of data streams is fixed (see for an example the grey lines with $L = 2$), only a part of the performance curve can be achieved. As expected it can be seen, that the maximum SIR that can be supported by the adaptive strategy, is achieved by concentrating the power on one data stream (with the maximum singular value of the channel).

Now, comparing both minimum and average metric, the impact of the SIR metric on the performance can be evaluated. The transmission strategy that is based on the minimum metric is energy suboptimal. This is due to the fact, that the performance is determined by the active data stream with the worst singular value of the channel. Indeed, the minimum metric can be interpreted as a max-min-fair concept, that suffers by its inefficiency. This is true if the singular values of the channel differ as is in general the case.

The comparison between the adaptive strategy and SISO shows that both metrics can outperform SISO in the high SIR region. In the low SIR region the minimum metric suffers from its fairness, whereas the average metric performs better than SISO. The performance gains highly depend on the relation between the energy consumption of the hardware and the energy consumption for transmitting data. Consequently, hardware properties and the transmission distance between the nodes play a decisive role. A larger distance between nodes improves the gains of multiple antenna concepts as the influence of hardware energy consumption reduces.

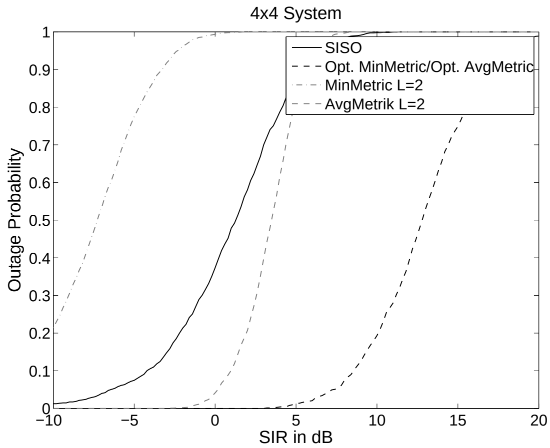

Fig. 2. Outage probability over SIR requirements averaged over 1000 channel realizations, $N = 4$

In addition, figure 2 depicts the outage probability over the SIR requirements for the 5 transmission strategies. The outage probability is defined as follows: $P_{out}(\boldsymbol{\gamma}) = P(\rho(\boldsymbol{\Gamma}\mathbf{D}^{-1}\mathbf{G}) \geq 1)$. Regardless of the energy consumption, it can be seen that the adaptive approach supports high SIR requirements and that the minimum metric for a fixed number of active data streams $L = 2$ again suffers from its fairness.

7. CONCLUSIONS

In wireless networks, a widely studied approach is to minimize transmit powers subject to some QoS constraints. However, minimizing the transmit powers is not equivalent to minimizing the energy consumption. In this paper we addressed the problem of minimizing the overall energy consumption in wireless networks including the energy consumption for hardware. Therefore, we first pointed out some basic properties of the optimal power allocation. In order to give some insights, we then discussed the general energy minimization problem that depends on system parameters as the number of antennas N, on the transmission strategy represented by beamformers and number of parallel data streams and on transmit powers. Due to the fact that the general problem is quite complex we focused on a restricted problem. Considering the relation between transmission time and transmit power, we optimized both jointly to find an energy-optimal power-time tradeoff. More precisely, we proposed an algorithm that determines the energy-optimal number of data streams per link for a certain SIR requirement.

To gain further insights into the energy minimization problem, it has to be considered for assumptions that may be more general or give another perspective on the problem. Further, note that the notion of energy minimization is not restricted to sensor networks. Thus, in future work the optimization problem may also include other aims and constraints.

8. REFERENCES

[1] G.J. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," *IEEE Trans. on Vehicul. Technol.*, vol. 42, no. 4, pp. 641–646, November 1993.

[2] J. Zander, "Distributed cochannel interference control in cellular radio systems," *IEEE Trans. Veh. Technol.*, vol. 41, pp. 305–311, August 1992.

[3] S.A. Grandhi, J. Zander, and R. Yates, "Constrained power control," *Wireless Personal Communications, Kluwer*, vol. 2, no. 3, pp. 257–270, 1995.

[4] N. Bambos, S.C. Chen, and G.J. Pottie, "Channel access algorithms with active link protection for wireless communication networks with power control," *IEEE/ACM Trans. Networking*, vol. 8, no. 5, pp. 583–597, October 2000.

[5] S. Stanczak, M. Wiczanowski, and H. Boche, *Theory and Algorithms for Resource Allocation in Wireless Networks*, Lecture Notes in Computer Science (LNCS). Springer, Berlin, 2006.

[6] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia, 2000.

[7] R.D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Select. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, September 1995.

[8] E. Seneta, *Non-Negative Matrices and Markov Chains*, Springer, Berlin, 1981.

[9] H. Boche and S. Stanczak, "Convexity of some feasible QoS regions and asymptotic behavior of the minimum total power in CDMA systems," *IEEE Trans. Commun.*, vol. 52, no. 12, pp. 2190–2197, December 2004.

[10] F. Rashid-Farrokhi, J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE JSAC*, vol. 16, no. 8, pp. 1437–1450, Oct 1998.

[11] M. Bengtsson and B. Ottersten, "Optimal downlink beamforming using semidefinite optimization," in *Proc. Allerton Conference*, Sep 1999.

[12] M. Schubert and H. Boche, *QoS-Based Resource Allocation and Transceiver Optimization*, Now the essence of knowledge. Verdu, Princeton, foundations and trends edition, 2006.