Energy Efficient Collaborative Beamforming in Wireless Sensor Networks

Benjamín Béjar Haro, *Student Member, IEEE*, Santiago Zazo, *Member, IEEE*, and Daniel P. Palomar, *Fellow, IEEE*

***Abstract*—Energy efficiency is a major design issue in the context of Wireless Sensor Networks (WSN). If the acquired data is to be sent to a far-away base station, collaborative beamforming performed by the sensors may help to distribute the communication load among the nodes and to reduce fast battery depletion. However, collaborative beamforming techniques are far from optimality and in many cases we might be wasting more power than required. We consider the issue of energy efficiency in beamforming applications. Using a convex optimization framework, we propose the design of a virtual beamformer that maximizes the network lifetime while satisfying a pre-specified Quality of Service (QoS) requirement. We derive both centralized and distributed algorithms for the solution of the problem using convex optimization and consensus algorithms. In order to account for other sources of battery depletion different from that of communications beamforming, we consider an additional random energy term in the consumption model. The formulation then switches to a probabilistic design that generalizes the deterministic case. Conditions under which the general problem is convex are also provided.**

***Index Terms*— Wireless sensor networks, energy efficiency, optimization, distributed processing, antenna arrays.**

I. Introduction

THE use of multi-antenna systems is well motivated by the increasing demand of reliable, high data-rate communications. Beamforming techniques adjust the antenna weights in order to mitigate fading channel or interference effects, thus enhancing the quality of the signal of interest. In the context of a Wireless Sensor Network (WSN), it may happen that the area of interest to be sensed is located in a remote region with difficult access. To overcome the problem of retrieving the gathered data, nodes can cooperate to form a virtual beamformer in order to send the acquired data to a far-away base station, where further processing and analysis could be done. At the same time, a certain Quality of Service (QoS) measure must be imposed at the receiver side (i.e., base station) that allows reliable signal decoding.

One possible solution to this end is the concept of collaborative beamforming [1], where nodes synchronize their phases to add constructively at the base station. The statistical properties of the average radiation pattern have been analyzed for the case of uniformly distributed nodes over a disc of a certain radius [1]. It is demonstrated that as the number of nodes increases, the average directivity of the virtual array approaches its maximum. In the same direction, the works in [2] and [3] show that better properties in terms of sidelobe level can be obtained if nodes are deployed following a Gaussian distribution over the disc. Obviously, this comes at the expense of some degradation in terms of directivity. An alternative selection mechanism is proposed in [4], where it is shown that if nodes are chosen within a disc of an appropriate radius the beamwidth can be reduced and, at the same time, connectivity of the network is easily preserved and energy consumption may be reduced.

Although the average properties of the radiation pattern are insightful they only hold asymptotically when the number of nodes is very large. There are several issues regarding the collaborative beamforming strategy that should be pointed out. The first one is that channel effects are usually ignored and the only source of signal attenuation considered is due to propagation losses. This is an important modeling limitation since the effect of the wireless channel can change the radiation pattern drastically. Another point is that in many situations we may be wasting more power than necessary (far from optimality) or even we might be violating some spatial radiation power constraints such as interference level caused to coexisting systems. In order to meet some QoS at the receiver, it would be more energy-efficient to optimize the individual antenna weights so as to maximize the network's lifetime (i.e., the time that the network is going to be operative), using the more mature beamforming technology for centralized scenarios, otherwise we may cause rapid energy depletion at the nodes, shortening their time of activity.

In the last few years, the application of convex optimization techniques to beamforming problems has been proven very successful, see [5], [6] and references therein. The use of convex optimization can help to produce optimal or close-to-optimal solutions in many beamforming problems. When designing a beamformer, two common approaches are found in the literature that either try to maximize the SNR subject to (individual or total) power constraints or that consider energy minimization while ensuring a specified QoS. In the context of WSN, energy efficiency is a major design issue that should be looked at carefully. It is desirable for such networks to be autonomous and capable of working for long periods of time without battery

Manuscript received April 10, 2012; revised February 19, 2013 and July 27, 2013; accepted October 09, 2013. Date of publication October 30, 2013; date of current version December 27, 2013. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Eduard Axel Jorswieck. This work was supported in part by the Spanish Ministry of Science and Innovation under Grant TEC2010-21217-C02-02-CR4HFDVL; the program CONSOLIDER-INGENIO 2010 under the grant CSD2008-00010 COMONSENS; the Consejo Social de la UPM; and by the Hong Kong RGC 617810 research grant.

B. Béjar Haro is with the LCAV, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Swizerland (e-mail: benjamin.bejarharo@epfl.ch).

S. Zazo is with the GAPS, Universidad Politécnica de Madrid (UPM), Madrid, Spain (e-mail: santiago@gaps.ssr.upm.es).

D. P. Palomar is with the ECE Department, Hong Kong University of Science and Technology (HKUST), Clear Water Bay, Hong Kong SAR (e-mail: palomar@ust.hk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2013.2288080

replacement. There has been little attention to this issue in the context of beamforming applications. In [7] they consider energy-efficiency when collaborative beamforming is used. However, the work in [7] is oriented towards routing optimization instead of energy efficient beamforming. A routing scheduling policy has been recently proposed in [8] that helps to extend the network's lifetime. Again, in [7], [8], collaborative beamforming is considered with no weight optimization; therefore, the network may be using more energy than strictly required. It becomes clear that the development of distributed optimization techniques that take into account energy efficiency are of paramount importance in the context of sensor networks. Several distributed beamforming approaches have been proposed for the network relay-beamforming scenario [9]–[14]. However, these are based on centralized optimization and reduced feedback between the relay network and the transmitter/receiver pair. In our context of WSNs, such approaches may have limited applicability since they require constant feedback between sensors and base station (e.g., battery level and channel state information). It is therefore preferable to devise distributed algorithms that do not require any communication with the base station. Another differentiating aspect to the relay-beamforming scenario is that, in the context of sensor networks, nodes are battery-powered devices with finite energy resources.

In this paper, we consider the distributed beamforming problem with QoS constraints where the metric to be optimized is the network's lifetime (e.g., the time that the network can guarantee the specified QoS requirement). This work represents a major extension of [15] where only an idealized sensing scenario was considered. We derive closed-form expressions for the optimal beamformer and provide iterative algorithms for its numerical computation. Using only local information about battery status and channel conditions, we use consensus algorithms [16] to propose fully distributed solutions to the problem that only require local communication among nodes.

In the last part of the paper we consider the case where the energy consumption at the nodes is not deterministic. It has been shown that the energy consumption in a wireless sensor network can be modeled as random quantity [17] that depends on several parameters related to data processing and sensing characteristics, node to node communication, transmission rate, duty cycle, MAC layer protocol, etc. In order to account for other sources of battery depletion different from far-away transmission to the base station we consider an additional random energy consumption term in our formulation. The problem then switches to a probabilistic design that generalizes the original problem. Conditions under which the general problem is convex are also provided. In some specific scenarios, the more general problem is amenable for its solution in a distributed fashion. However, this is not the case in general and further work needs to be done in that direction.

The paper is organized as follows: We first provide some general definitions and describe the problem in Section II. In Section III we formulate the energy-efficient beamforming problem and derive simple expressions for its computation. Section IV provides both centralized and distributed algorithms for the computation of the optimal beamvector. Section V extends the results for the case of random energy consumption. Numerical simulations are provided in Section VI and conclusions are drawn in Section VII.

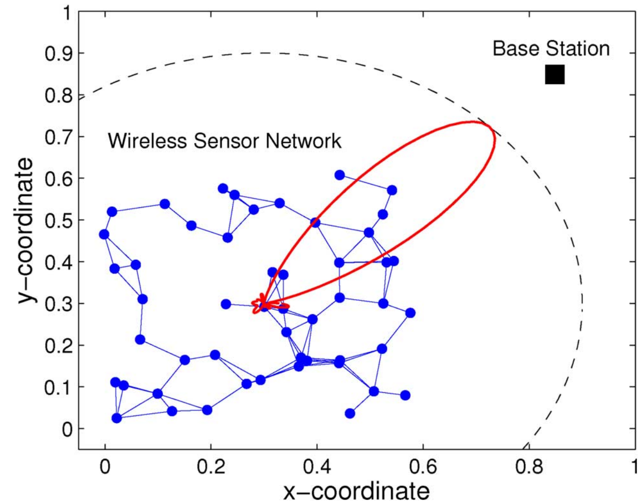

Fig. 1. Beamforming scenario between the nodes and the far-away base station.

Notation: Vector-valued quantities are denoted using bold lower-case letters. The optimal value of a variable $\mathbf{x}$ in an optimization problem is denoted by $\mathbf{x}^\star$. The symbols $\mathbb{C}$, $\mathbb{R}$, and $\mathbb{R}_+$ denote the set of complex, real, and non-negative real numbers, respectively. For a scalar a, its complex-conjugate is denoted by a^*. For vectors, $\mathbf{x}^\mathsf{T}$ denotes transposition while $\mathbf{x}^\mathsf{H} = (\mathbf{x}^\mathsf{T})^*$ denotes complex-conjugate transposition.

II. System Model

Consider a WSN composed of M nodes scattered over a certain area. Nodes are battery-powered elements equipped with a single-antenna whose purpose is to sense and retrieve information from the environment. The information sensed is to be sent to a far-away base station where the data is further processed and analyzed. In order to reach the base station, nodes need to cooperatively form a virtual beamformer for transmitting the acquired data to the base station, see Fig. 1. At the base station, a minimum QoS requirement must be fulfilled that allows reliable decoding of the received signal. We aim to maximize the time that the network remains operative without human intervention (i.e., battery replacement). The problem is then to design a virtual beamformer that meets the required QoS at the base station while maximizing the network's lifetime. It is further assumed that all elements (nodes and base station) lie on the same plane.

Consider a scenario where all nodes of the network have access to a (possibly) noisy version of the signal to be transmitted (e.g., they may measure independently the same quantity or they might have obtained it through a joint estimation process). We assume that the network has two distinct modes of operation, namely sensing (data harvesting) and data reporting. Once enough measurements have been acquired, it is necessary to report those to the base station through a noisy wireless communication channel. Let $s_m[k]$ denote the discrete-time signal available at the mth node that needs to be transmitted at time k

$$s_m[k] = s[k] + \epsilon_m[k], \tag{1}$$

where $s[k]$ is the true signal to be transmitted and $\epsilon_m[k]$ corresponds to the observation noise at node m, which is assumed to be independent and identically distributed (i.i.d.) with zero-mean and variance σ_ϵ^2.

Let $w_m^*[k]$ denote the complex antenna weight of the mth node at the kth sampling instant. Each node multiplies its measured signal by its corresponding weight in order to send it to the base station. The received signal $y[k]$ at the base station can be expressed as

$$\begin{aligned} y[k] &= \sum_{m=1}^{M} h_m w_m^*[k] s_m[k] + n[k] \\ &= \sum_{m=1}^{M} h_m w_m^*[k] s[k] + \sum_{m=1}^{M} h_m w_m^*[k] \epsilon_m[k] + n[k], \end{aligned} \quad (2)$$

where $h_m \neq 0$ is the channel coefficient (assumed flat during the transmission) between the mth node and the base station, and $n[k]$ represents the discrete-time measurement noise process. The noise samples $n[k]$ are assumed to be i.i.d. distributed random variables, also independent of $\epsilon_m[k]$, with zero mean and variance σ_n^2.

The (instantaneous) received Signal to Noise Ratio (SNR) at destination is then given by

$$\begin{aligned} \Gamma[k] &= \frac{\left| \sum_{m=1}^{M} h_m w_m^*[k] \right|^2 P_s}{\sum_{m=1}^{M} |h_m w_m^*[k]|^2 \sigma_\epsilon^2 + \sigma_n^2} \\ &= \frac{\left| \mathbf{w}^{\mathsf{H}}[k]\mathbf{h} \right|^2 \rho_0}{\mathbf{w}^{\mathsf{H}}[k] \mathrm{diag}(\mathbf{h} \odot \mathbf{h}^*) \mathbf{w}[k] \sigma_r^2 + 1}, \end{aligned} \quad (3)$$

where $\mathbf{h} = [h_1, \ldots, h_m]^{\mathsf{T}}$, $\mathbf{w}^{\mathsf{H}}[k] = [w_1^*[k], \ldots, w_m^*[k]]$, and where the expectation is taken over the noise $n[k]$ and the symbols $s[k]$, with $P_s = \mathrm{E}[|s[k]|^2]$, $\mathrm{E}[s[k]] = 0$. The symbol $\odot$ denotes element-wise product, $\rho_0 = P_s/\sigma_n^2$ and $\sigma_r^2 = \sigma_\epsilon^2/\sigma_n^2$ is the ratio between the observation noise power and the communication channel noise power.

We will use the instantaneous received SNR in (3) as our QoS measure, i.e., we want to design our beamformer in order to ensure that the received SNR is above some threshold. Besides, we also seek to maximize the network's lifetime so that it can be operative for the longest period of time. Several measures of network's lifetime have been proposed in the literature (see [18] and references therein) attending to different criteria such as percentage of alive nodes, coverage area, or connectivity, among others. In our problem, a natural measure of the network's lifetime is the time that the network can satisfy the QoS constraint. We will show later that, in some cases, such lifetime criterion is equivalent to maximizing the time that takes for the first node to deplete its battery.

It is important to mention that, since our discretized transmission model is an approximation of a continuous-time transmission, we will consider the network's lifetime as a real-valued variable in the optimization process.

III. Energy-Efficient Beamforming

The energy consumption at a node varies according to its mode of operation (e.g., idle, sleep, transmitting, receiving, sensing, etc) [19]. When it comes to data transmission, a common model for the energy consumption is to have a cost associated with circuitry consumption plus a term that is proportional to the power put into the antenna [19]–[22]. For low-range, node to node communications, the necessary power is dependent on the distance between the nodes (propagation losses). However, for far-away communications, the required power will be determined by the propagation channel. Furthermore, it is well known that when nodes are transmitting the dominant term in energy consumption is the one that corresponds to data transmission [19].

Since we are concerned about energy consumption related with the process of communicating with the base station, let us start considering a simple model where battery depletion is only due to far-away communication with the base station. Therefore, the amount of energy consumed during the kth sampling period at node m would be proportional to the power put into the antenna, that is

$$e_m[k] = T\,|w_m[k]|^2, \quad (4)$$

where T is some proportionality constant (e.g., the sampling period). Later in Section V we will consider the more general case where additional energy is dedicated to other tasks different from far-away transmissions.

In order to allow for a reliable signal decoding at the base station we want to satisfy a minimum QoS constraint. Such constraint can be expressed in terms of the instantaneous SNR of (3) as

$$\Gamma[k] = \frac{\left| \mathbf{w}^{\mathsf{H}}[k]\mathbf{h} \right|^2 \rho_0}{\mathbf{w}^{\mathsf{H}}[k] \mathrm{diag}(\mathbf{h} \odot \mathbf{h}^*) \mathbf{w}[k] \sigma_r^2 + 1} \geq \rho, \quad (5)$$

where ρ is the target SNR at destination.

As stated in the previous section, a natural measure of the network's lifetime is the longest time (or maximum time) that we can guarantee the QoS constraint (5) to be fullfilled. We then consider the following definition of network lifetime:

Definition 1 (Deterministic Network Lifetime): The lifetime of the network is the longest time (or maximum time) that the QoS constraint (5) can be satisfied.

Our goal is then to find the sequence of beamforming vectors $\{\mathbf{w}[k]\}$ that maximizes the network lifetime given in Definition 1. Since nodes are battery-equipped elements with limited power resources, we also constrain the maximum transmission power at node m to be less than or equal to $p_m, m = 1, \ldots, M$.

Let denote $K^\star$ as the maximum time that the QoS constraint (5) can be satisfied, then the problem of finding the optimal beamvectors $\{\mathbf{w}[1], \ldots, \mathbf{w}[K^\star]\}$ can be expressed as

$$\begin{aligned} \text{find} \quad & \{\mathbf{w}[1], \ldots, \mathbf{w}[K^\star]\} \in \mathbb{C}^M \\ \text{subject to} \quad & \frac{\left| \mathbf{w}^{\mathsf{H}}[k]\mathbf{h} \right|^2}{\mathbf{w}^{\mathsf{H}}[k] \mathrm{diag}(\mathbf{h} \odot \mathbf{h}^*) \mathbf{w}[k] \sigma_r^2 + 1} \geq \frac{\rho}{\rho_0}, \ \forall k \\ & |w_m[k]|^2 \leq p_m, \ \text{for all } m, k \\ & T \sum_{k=1}^{K^\star} |w_m[k]|^2 \leq E_m, \ m = 1, \ldots, M \end{aligned} \quad (6)$$

where the last constraint ensures that no node can waste more energy than its actual battery level E_m.

Note that in order to constructively add at the receiver side, the phase of the beamformer must be matched to that of the channel. Therefore, we can get rid of the phase and reduce problem (6) to a power allocation problem. Let $\bar{w}_m[k] = |w_m[k]|$ and $\bar{h}_m = |h_m|$, denote the magnitude of the mth beamweight and the mth channel coefficient, respectively. We can fix the phase of the beamvectors $\{\mathbf{w}[1], \ldots, \mathbf{w}[K^\star]\}$

to match that of the channel and replace problem (6) by the following real-valued feasibility power allocation problem

$$\begin{aligned} \text{find} \quad & \{\bar{\mathbf{w}}[1], \ldots, \bar{\mathbf{w}}[K^\star]\} \in \mathbb{R}^M \\ \text{subject to} \quad & \frac{\left(\bar{\mathbf{w}}^\mathsf{T}[k]\bar{\mathbf{h}}\right)^2}{\bar{\mathbf{w}}^\mathsf{T}[k]\mathrm{diag}(\bar{\mathbf{h}} \odot \bar{\mathbf{h}})\bar{\mathbf{w}}[k]\sigma_r^2 + 1} \geq \frac{\rho}{\rho_0}, \ \forall k \\ & \bar{w}_m^2[k] \leq p_m, \ \text{for all } m, k \\ & T \sum_{k=1}^{K^\star} \bar{w}_m^2[k] \leq E_m, \ m = 1, \ldots, M \end{aligned} \tag{7}$$

where $\bar{\mathbf{w}}[k] = [\bar{w}_1[k], \ldots, \bar{w}_M[k]]^\mathsf{T}$ and $\bar{\mathbf{h}} = [\bar{h}_1, \ldots, \bar{h}_M]^\mathsf{T}$. Note that problem (7) is not convex due to the SNR constraint. However, the set of SNR constraints can be equivalently expressed as a set of Second Order Cone (SOC) constraints. Therefore, we can reformulate (7) as the following (equivalent) convex feasibility power allocation problem

$$\begin{aligned} \text{find} \quad & \{\bar{\mathbf{w}}[1], \ldots, \bar{\mathbf{w}}[K^\star]\} \in \mathbb{R}^M \\ \text{subject to} \quad & [\bar{\mathbf{h}}^T \ 0]\begin{bmatrix}\bar{\mathbf{w}}[k] \\ 1\end{bmatrix} \geq \left\| \mathbf{D}\begin{bmatrix}\bar{\mathbf{w}}[k] \\ 1\end{bmatrix}\right\|, \ \forall k \\ & \bar{w}_m^2[k] \leq p_m, \ \text{for all } m, k \\ & T \sum_{k=1}^{K^\star} \bar{w}_m^2[k] \leq E_m, \ m = 1, \ldots, M \end{aligned} \tag{8}$$

where the matrix $\mathbf{D}$ is given by

$$\mathbf{D} = \sqrt{\frac{\rho}{\rho_0}} \begin{bmatrix} \sigma_r \mathrm{diag}(\bar{\mathbf{h}}) & 0 \\ 0 & 1 \end{bmatrix}. \tag{9}$$

Since the problem at hand is convex we can now show that a constant beamformer (independent of time) suffices to optimally solve problem (8).

Lemma 1 (Constant Beamformer): Assume that the feasibility problem (8) is feasible. Then, there exist an optimal solution $\{\bar{\mathbf{w}}^\star[1], \ldots, \bar{\mathbf{w}}^\star[K^\star]\}$ to (8) such that $\bar{\mathbf{w}}^\star[i] = \bar{\mathbf{w}}^\star$ for all $i = 1, \ldots, K^\star$.

Proof: We prove it using a convexity argument. Assume $\{\bar{\mathbf{w}}^\star[1], \ldots, \bar{\mathbf{w}}^\star[K^\star]\}$ is a solution to the feasibility problem (8). Then, any permutation of the optimal sequence $\{\bar{\mathbf{w}}^\star[1], \ldots, \bar{\mathbf{w}}^\star[K^\star]\}$ is also optimal. Since the problem is convex, any convex combination of optimal solutions is also optimal. Take, for example, the mean over all permutations, e.g., $\bar{\mathbf{w}}^\star = \frac{1}{K^\star}\sum_{k=1}^{K^\star} \bar{\mathbf{w}}[k]$. Therefore, we have found a constant beamvector that optimally solves the problem. ■

Using the result in Lemma 1, feasibility problem (7) simplifies to

$$\begin{aligned} \text{find} \quad & \bar{\mathbf{w}} \\ \text{subject to} \quad & \frac{\left(\bar{\mathbf{w}}^\mathsf{T}\bar{\mathbf{h}}\right)^2}{\bar{\mathbf{w}}^\mathsf{T}\mathrm{diag}(\bar{\mathbf{h}} \odot \bar{\mathbf{h}})\bar{\mathbf{w}}\sigma_r^2 + 1} \geq \frac{\rho}{\rho_0} \\ & \bar{w}_m^2 \leq p_m, \ m = 1, \ldots, M \\ & TK^\star \bar{w}_m^2 \leq E_m, \ m = 1, \ldots, M \end{aligned} \tag{10}$$

and thus, the lifetime maximization problem can be expressed as

$$\begin{aligned} \underset{\bar{\mathbf{w}}, K}{\text{maximize}} \quad & K \\ \text{subject to} \quad & \frac{\left(\bar{\mathbf{w}}^\mathsf{T}\bar{\mathbf{h}}\right)^2}{\bar{\mathbf{w}}^\mathsf{T}\mathrm{diag}(\bar{\mathbf{h}} \odot \bar{\mathbf{h}})\bar{\mathbf{w}}\sigma_r^2 + 1} \geq \frac{\rho}{\rho_0} \\ & \bar{w}_m^2 \leq p_m, \ m = 1, \ldots, M \\ & KT\bar{w}_m^2 \leq E_m, \ m = 1, \ldots, M \end{aligned} \tag{11}$$

Alternatively, we can reformulate the optimal power allocation problem (11) as an equivalent minimization problem by the change of variables $t = 1/K$. This way, the last set of constraints become a set of convex SOC constraints, too. The equivalent lifetime maximization problem reads

$$\begin{aligned} \underset{\bar{\mathbf{w}}, t}{\text{minimize}} \quad & t \\ \text{subject to} \quad & \frac{\left(\bar{\mathbf{w}}^\mathsf{T}\bar{\mathbf{h}}\right)^2}{\bar{\mathbf{w}}^\mathsf{T}\mathrm{diag}(\bar{\mathbf{h}} \odot \bar{\mathbf{h}})\bar{\mathbf{w}}\sigma_r^2 + 1} \geq \frac{\rho}{\rho_0} \\ & \bar{w}_m^2 \leq p_m, \ m = 1, \ldots, M \\ & T\bar{w}_m^2 \leq tE_m, \ m = 1, \ldots, M \end{aligned} \tag{12}$$

Intuitively, one would expect that in order to maximize the lifetime of the network, the SNR constraint must hold with equality. Otherwise, we would be wasting more power than necessary. This intuition is formalized in Lemma 2.

Lemma 2 (SNR With Equality): Assume that the problem (12) is solvable. Then, at the optimum, the SNR constraint must hold with equality.

Proof: We prove it by contradiction. Assume $\bar{\mathbf{w}}^\star$ is an optimal solution to problem (12) such that $\frac{\bar{\mathbf{w}}^\mathsf{T}\bar{\mathbf{h}}\bar{\mathbf{h}}^\mathsf{T}\bar{\mathbf{w}}}{\bar{\mathbf{w}}^\mathsf{T}\mathrm{diag}(\bar{\mathbf{h}}\odot\bar{\mathbf{h}})\bar{\mathbf{w}}\sigma_r^2+1} > \frac{\rho}{\rho_0}$. Let $\tilde{\mathbf{w}} = \sqrt{\varepsilon}\bar{\mathbf{w}}$, with $\sqrt{\varepsilon} \in (0, 1)$. Since the function $f(x) = \frac{xa}{xb+c}$ is monotone increasing for positive a, b and c, then there exist a positive number $\varepsilon < 1$ such that $\frac{\varepsilon\bar{\mathbf{w}}^\mathsf{T}\bar{\mathbf{h}}\bar{\mathbf{h}}^\mathsf{T}\bar{\mathbf{w}}}{\varepsilon\bar{\mathbf{w}}^\mathsf{T}\mathrm{diag}(\bar{\mathbf{h}}\odot\bar{\mathbf{h}})\bar{\mathbf{w}}\sigma_r^2+1} = \frac{\rho}{\rho_0}$. Therefore, we have found a vector $\tilde{\mathbf{w}}$ that is feasible and allows for a reduction in the objective function. We conclude then that $\bar{\mathbf{w}}^\star$ cannot be optimal and that the only possibility is that the SNR constraint holds with equality. ■

As we mentioned earlier, it turns out that in our particular setting, the QoS lifetime maximization criterion given in Definition 1 coincides with maximizing the time that takes for the first node to deplete its battery (i.e., 1st node depletion criterion).

Lemma 3 (Lifetime and First Node): The maximum time that the network can satisfy the QoS constraint (5) is equal to the maximum time that takes for the first node to deplete its battery.

Proof: In order to prove the claim, we need to show that problem (12) is equivalent to maximizing the time for the first node to deplete its battery. For that purpose, let's define the forecast longevity of node m (i.e., expected node lifetime) as

$$l_m = \frac{E_m}{\bar{w}_m^2 T}, \quad m = 1, \ldots, M. \tag{13}$$

Maximizing the time for the first node to deplete its battery can be then expressed as the maximization of the minimum of l_m. Alternatively, it could be expressed as the minimization of the maximum of $1/l_m$. If we follow the latter approach we can express the 1st node lifetime maximization problem as the following minmax optimization problem:

$$\begin{aligned} \underset{\bar{\mathbf{w}}}{\text{minimize}} \quad & \max\left(\frac{\bar{w}_1^2 T}{E_1}, \ldots, \frac{\bar{w}_M^2 T}{E_M}\right) \\ \text{subject to} \quad & \frac{\left(\bar{\mathbf{w}}^\mathsf{T}\bar{\mathbf{h}}\right)^2}{\bar{\mathbf{w}}^\mathsf{T}\mathrm{diag}(\bar{\mathbf{h}} \odot \bar{\mathbf{h}})\bar{\mathbf{w}}\sigma_r^2 + 1} \geq \frac{\rho}{\rho_0} \\ & \bar{w}_m^2 \leq p_m, \ m = 1, \ldots, M. \end{aligned} \tag{14}$$

It can be easily verified that problem (12) is the epigraph form of problem (14). ■

It is clear that we can express problem (12) in convex form by expressing the SNR constraints as SOC constraints as in (8).

Therefore, it can be solved efficiently using any general purpose optimization software. However, problem (12) allows the derivation of simple expressions for the computation of the optimal solution. These expressions, in turn, can be used to devise very efficient algorithms for the computation of the optimal solution.

Proposition 1 (Optimal Power Allocation): Suppose problem (12) is solvable, then the optimal power allocation is given by

$$\bar{w}_m^\star = \min\left(t^\star\sqrt{\frac{E_m}{T}}, \sqrt{p_m}\right), \tag{15}$$

with

$$t^\star = \frac{\sqrt{(\rho/\rho_0 + d)(a^2 - c) + b^2 c} - ab}{a^2 - c}, \tag{16}$$

where

$$a = \sum_{m\in\mathcal{M}} h_m \sqrt{\frac{E_m}{T}} \quad b = \sum_{m\notin\mathcal{M}} h_m\sqrt{p_m}$$
$$c = \sigma_r^2 \frac{\rho}{\rho_0} \sum_{m\in\mathcal{M}} h_m^2 \frac{E_m}{T} \quad d = \sigma_r^2 \frac{\rho}{\rho_0} \sum_{m\notin\mathcal{M}} h_m^2 p_m \tag{17}$$

and $\mathcal{M} = \{m \mid \bar{w}_m^2 < p_m\}$ is the set of nodes not transmitting at their maximum power.

Proof: In order to prove the result let us rewrite problem (12) in convex form as

$$\begin{aligned} \underset{\bar{\mathbf{w}},t}{\text{minimize}} \quad & t \\ \text{subject to} \quad & \bar{\mathbf{w}}^\mathsf{T}\bar{\mathbf{h}} \geq \sqrt{\frac{\rho}{\rho_0}\left(1 + \sigma_r^2 \bar{\mathbf{w}}^\mathsf{T}\text{diag}(\bar{\mathbf{h}}\odot\bar{\mathbf{h}})\bar{\mathbf{w}}\right)} \\ & \bar{w}_m \leq \sqrt{p_m},\ m = 1,\ldots,M \\ & \bar{w}_m \leq \sqrt{tE_m/T},\ m = 1,\ldots,M. \end{aligned} \tag{18}$$

In order to get rid of the $\sqrt{t}$ term we will consider the following problem

$$\begin{aligned} \underset{\bar{\mathbf{w}},t}{\text{minimize}} \quad & t \\ \text{subject to} \quad & \bar{\mathbf{w}}^\mathsf{T}\bar{\mathbf{h}} \geq \sqrt{\frac{\rho}{\rho_0}\left(1 + \sigma_r^2 \bar{\mathbf{w}}^\mathsf{T}\text{diag}(\bar{\mathbf{h}}\odot\bar{\mathbf{h}})\bar{\mathbf{w}}\right)} \\ & \bar{w}_m \leq \sqrt{p_m},\ m = 1,\ldots,M \\ & \bar{w}_m \leq t\sqrt{E_m/T},\ m = 1,\ldots,M. \end{aligned} \tag{19}$$

which can be easily shown to yield to the same optimal power allocation as (18). The Lagrangian of (19) is given by

$$\begin{aligned} &\mathcal{L}\left(t, \{\bar{w}_m\}; \lambda, \{\mu_m\}, \{\sigma_m\}\right) \\ &= t + \lambda\left(\sqrt{\frac{\rho}{\rho_0}\left(1 + \sigma_r^2 \sum_{m=1}^M \bar{h}_m^2 \bar{w}_m^2\right)} - \sum_{m=1}^M \bar{h}_m \bar{w}_m\right) \\ &\quad + \sum_{m=1}^M \mu_m(\bar{w}_m - t\sqrt{E_m/T}) \\ &\quad + \sum_{m=1}^M \sigma_m(\bar{w}_m - \sqrt{p_m}) \end{aligned} \tag{20}$$

where $\lambda, \mu_m, \sigma_m \in \mathbb{R}_+$ are the Lagrange multipliers associated with the constraints of the problem. The Karush-Kuhn-Tucker (KKT) optimality conditions are then given by

Feasibility:

$$\bar{\mathbf{w}}^\mathsf{T}\bar{\mathbf{h}} \geq \sqrt{\frac{\rho}{\rho_0}\left(1 + \sigma_r^2 \sum_m \bar{h}_m^2 \bar{w}_m^2\right)}, \qquad \lambda \geq 0$$
$$\bar{w}_m - t\sqrt{E_m/T} \leq 0, \quad \mu_m \geq 0$$
$$\bar{w}_m - \sqrt{p_m} \leq 0, \quad \sigma_m \geq 0$$

Complementary Slackness:

$$\lambda\left(\sqrt{\frac{\rho}{\rho_0}\left(1 + \sigma_r^2 \sum_m \bar{h}_m^2 \bar{w}_m^2\right)} - \sum_{m=1}^M \bar{h}_m \bar{w}_m\right) = 0 \tag{21}$$

$$\mu_m(\bar{w}_m - t\sqrt{E_m/T}) = 0 \tag{22}$$

$$\sigma_m(\bar{w}_m - \sqrt{p_m}) = 0 \tag{23}$$

Zero gradient of the Lagrangian: By differentiating (20) with respect to t and $\bar{w}_m$ and equating to zero we get

$$\sum_m \mu_m \sqrt{E_m/T} = 1 \tag{24}$$

$$\lambda\left(\frac{\sigma_r^2\sqrt{\rho/\rho_0}\,\bar{h}_m^2 \bar{w}_m}{\sqrt{1 + \sigma_r^2 \sum_m \bar{h}_m^2 \bar{w}_m^2}} - \bar{h}_m\right) + \mu_m + \sigma_m = 0 \tag{25}$$

Note that (20) is differentiable everywhere in its domain since $\mathbf{D}[\bar{\mathbf{w}}^\mathsf{T}\ 1]^\mathsf{T} \geq 1 > 0$. Recall that by Lemma 2 the SNR constraint must hold with equality, that is

$$\sqrt{1 + \sigma_r^2 \sum_m \bar{h}_m^2 \bar{w}_m^2} = \frac{\sum_m \bar{h}_m \bar{w}_m}{\sqrt{\rho/\rho_0}}. \tag{26}$$

Therefore, we can rewrite (25) as

$$\lambda(\alpha u_m - 1)\bar{h}_m + \mu_m + \sigma_m = 0, \tag{27}$$

where $\alpha = \sigma_r^2 \frac{\rho}{\rho_0}$ and $u_m = \frac{\bar{h}_m \bar{w}_m}{\sum_m \bar{h}_m \bar{w}_m}$. Solving for λ in (27) leads to

$$\lambda = -\frac{\mu_m + \sigma_m}{(\alpha u_m - 1)\bar{h}_m} \geq 0, \tag{28}$$

where the last inequality follows from the feasibility conditions of λ. Consider now the complementary slackness conditions of the individual transmission power constraints (22) and (23). Assume that $\bar{w}_m < \sqrt{p_m}$ for some $m \in \{1,\ldots,M\}$, which implies by (23) that $\sigma_m = 0$. Consider now the two possibilities $\mu_m = 0$ and $\mu_m > 0$. If $\mu_m = 0$, then we have by (27) that $\lambda = 0$. From (28) it would follow that, $\sigma_m = 0$ and $\mu_m = 0$ for all m, which is not a valid solution since (24) wouldn't be satisfied. Therefore, it must be the case that $\mu_m > 0$, and thus, $\bar{w}_m = t\sqrt{E_m/T}$. The other possibility is that $\bar{w}_m^2 = p_m$, i.e., the node transmits at its maximum power. Using the fact that, at the optimum, the SNR constraint must hold with equality we can now write

$$\sqrt{\frac{\rho}{\rho_0}\left(1+t^2\sigma_r^2\sum_{m\in\mathcal{M}}\bar{h}_m^2\frac{E_m}{T}+\sigma_r^2\sum_{m\notin\mathcal{M}}\bar{h}_m^2 p_m\right)} = t\sum_{m\in\mathcal{M}}\bar{h}_m\sqrt{\frac{E_m}{T}}+\sum_{m\notin\mathcal{M}}\bar{h}_m\sqrt{p_m}, \quad (29)$$

where $\mathcal{M}=\{m\,|\,\bar{w}_m^2<p_m\}$ is the set of nodes not transmitting at their maximum power.

In order to simplify notation, let us denote $a = \sum_{m\in\mathcal{M}} h_m\sqrt{\frac{E_m}{T}}$, $b = \sum_{m\notin\mathcal{M}} h_m\sqrt{p_m}$, $c=\sigma_r^2\frac{\rho}{\rho_0}\sum_{m\in\mathcal{M}} h_m^2\frac{E_m}{T}$ and $d=\sigma_r^2\frac{\rho}{\rho_0}\sum_{m\notin\mathcal{M}} h_m^2 p_m$. With these definitions in mind, and by taking the square at both sides of (29) we get

$$\frac{\rho}{\rho_0}+t^2c+d=(at+b)^2. \quad (30)$$

Solving for the above quadratic equation yields that its only positive solution $t^\star$ is given by

$$t^\star=\frac{\sqrt{(\rho/\rho_0+d)(a^2-c)+b^2c}-ab}{a^2-c}, \quad (31)$$

which completes the proof. ■

Corollary 2: Under the same assumptions as in Proposition 1, the optimal beamformer is then given by

$$\mathbf{w}^\star=\left[\bar{w}_1^\star\frac{h_1}{|h_1|},\ldots,\bar{w}_M^\star\frac{h_M}{|h_M|}\right]^{\mathsf{T}}. \quad (32)$$

where $\bar{w}_m^\star$ is given by (15), (16) and (17).

Proof: The result follows directly from Proposition 1 and from the fact that the phase of the beamformer must be matched to that of the channel. ■

There is one caveat with Proposition 1 and it is the fact that it does not tell you how to find the optimal set $\mathcal{M}$. In Section IV we will come back to that issue and will present an iterative procedure that allows to determine the set $\mathcal{M}$ and converges to the optimal solution stated in Proposition 1.

It is important to note that under the assumptions of Proposition 1 and taking into account the expressions (15) and (16) it is easy to see that, at the optimum, all nodes must be active (i.e., $\bar{w}_m>0$ for all m). Further, since all nodes transmitting below its maximum allowed transmission power have the same ratio $E_m/|w_m^\star|^2$, $m\in\mathcal{M}$ (i.e., share the same value of $t^\star$), then they will deplete their batteries at the same time. These results have important implications since the most energy efficient policy (when it comes to far-away transmissions) requires the collaboration of all nodes in the network.

Ideal Sensing: One special case of problem (12) is that when all nodes have perfect knowledge about the signal to be transmitted to the base station. This situation may be considered as a special case of high SNR scenario with respect to the sensing channel or it may correspond to a situation where the nodes have previously agreed on the information to be sent to the base station, e.g., by *flooding* or by an aggregation mechanism as in [23]. We refer to this situation as *ideal sensing* since all nodes have common knowledge about the information to be transmitted. Such assumption has been commonly employed in collaborative beamforming problems [1]–[4].

Assuming a noiseless sensing model is equivalent to say that $\sigma_\epsilon^2=0$. Hence, problem (12) simplifies to:

$$\begin{aligned}&\underset{\bar{\mathbf{w}},t}{\text{minimize}} && t\\ &\text{subject to} && \bar{\mathbf{w}}^{\mathsf{T}}\bar{\mathbf{h}}\ge\sqrt{\rho/\rho_0}\\ &&& \bar{w}_m^2\le p_m && m=1,\ldots,M\\ &&& \bar{w}_m^2\le tE_m/T && m=1,\ldots,M.\end{aligned} \quad (33)$$

which is nothing but problem (12) particularized for $\sigma_r^2=0$. From Proposition 1 it is easy to see that the optimal power allocation of problem (33) can be computed as $\bar{w}_m^\star=\min(t^\star\sqrt{E_m/T},\sqrt{p_m})$ with

$$t^\star=\frac{\sqrt{\rho/\rho_0}-\sum_{m\notin\mathcal{M}}\bar{h}_m\sqrt{p_m}}{\sum_{m\in\mathcal{M}}\bar{h}_m\sqrt{E_m/T}} \quad (34)$$

where $\mathcal{M}$ is, as previously defined, the set of nodes not transmitting at maximum power.

IV. Algorithms

So far we have analyzed the lifetime maximization problem (12) providing simple expressions for the computation of the optimal beamweights. In this section we develop algorithms for the computation of the solution both in a centralized and in a distributed way. First, we consider an iterative (centralized) method based on Proposition 1. We then consider a distributed version of the iterative algorithm that allows the computation of the optimal beamforming weights by means of local node to node communication only (using a consensus algorithm).

A. Centralized Iterative Algorithm

From (16), (17) we can devise an iterative (centralized) algorithm that allows the computation of the optimal beamformer of problem (12) in a number of iterations that is at most the number of nodes in the network. We use the superscript $(\cdot)^{(k)}$ to denote the kth iteration of the algorithm. The algorithm works as follows: At the first stage of the algorithm, all nodes are initialized to belong to the set $\mathcal{M}^{(0)}$ that is, they are all assumed to be transmitting below their maximum power. Then we enter a loop where we compute the power allocation $[\bar{w}_1^{(0)},\ldots,\bar{w}_M^{(0)}]$ using (15) and (16), and update the set $\mathcal{M}^{(k)}$ accordingly. We keep iterating until there are no changes in the set $\mathcal{M}^{(k)}$ with respect to the previous iteration. Once we have converged to the optimal power allocation, the optimal beamformer is obtained by setting the phase of each node to match that of its respective channel coefficient. The process is summarized in Algorithm 1.

Algorithm 1: Iterative Centralized Algorithm

1: $k\to 0$, $\mathcal{M}^{(0)}\to\{1,\ldots,M\}$
2: **repeat**
3: $k\to k+1$
4: Compute $t^{(k)}$ using (16) and (17) with $\mathcal{M}=\mathcal{M}^{(k-1)}$
5: $\bar{w}_m^{(k)}\to t^{(k)}\sqrt{\frac{E_m}{T}}$, for all $m=1,\ldots,M$
6: $\mathcal{M}^{(k)}\to\{m\mid\bar{w}_m^{(k)}<\sqrt{p_m}\}$
7: **until** $\mathcal{M}^{(k)}=\mathcal{M}^{(k-1)}$
8: $w_m\to\min(\bar{w}_m^{(k)},\sqrt{p_m})\frac{h_m}{|h_m|}$, $m=1,\ldots,M$

Proposition 3: Assume that problem (12) is feasible. Then, Algorithm 1 converges to the optimal solution of the problem in at most M iterations.

Proof: Let denote $t^\star$ and $\mathcal{M}^\star$ the optimal values for the solution of problem (12) according to Proposition 1. At the first iteration ($k=1$) of the algorithm $t^{(1)}$ is computed assuming that none of the nodes is transmitting at maximum power, which means $\mathcal{M}^{(0)} = \{1,\ldots,M\}$. From $t^{(1)}$ we can compute the optimal power allocation as per (15) and the corresponding set $\mathcal{M}^{(1)}$. If $\mathcal{M}^{(1)} = \mathcal{M}^{(0)}$ then by Proposition 1 it must be the case that $t^\star = t^{(1)}$. In that case the algorithm stops and we have found the optimal power allocation. The other possibility is that $\mathcal{M}^{(1)} \subset \mathcal{M}^{(0)}$ which implies that $t^\star \neq t^{(1)}$. The fact that we are assuming that none of the nodes is transmitting at maximum power means that $t^{(1)}$ would be the solution to problem (12) but where the individual power constraints have been increased so that none of them is attained. Since this is equivalent to say that we are enlarging the search space it must hold that $t^{(1)} \leq t^\star$. We can then conclude that $t^\star > t^{(1)}$ which also implies that $\mathcal{M}^\star \subseteq \mathcal{M}^{(1)}$. Following a similar argument at every iteration it can be shown that the algorithm produces sequences of the form

$$t^{(1)} < t^{(2)} < \ldots < t^{(k)} < t^{k+1} = t^\star \tag{35}$$

with corresponding sets

$$\mathcal{M}^{(0)} \subset \mathcal{M}^{(1)} \subset \mathcal{M}^{(2)} \subset \ldots \subset \mathcal{M}^{(k)} \subseteq \mathcal{M}^{(k+1)} = \mathcal{M}^\star. \tag{36}$$

In order to keep the iterations going, it is necessary that at least one of the elements in the current set $\mathcal{M}^{(k)}$ is removed for the next iteration. It is clear then that the maximum number of iterations that the algorithm can perform is M. In particular, M iterations would only happen in the case where all nodes transmit at maximum power ($\mathcal{M} = \emptyset$). Since, by assumption, the problem is feasible, there exist $k \leq M$ such that $t^\star = t^{(k)}$ otherwise it would imply that $t^\star > t^{(M)}$. ■

B. Consensus-Based Distributed Implementation

The computation of the optimal solution to problem (12) via Algorithm 1 requires a central entity with full knowledge about channels, battery levels, and power constraints. This implies that the information should be transferred to a central node that performs the computation of the optimal beamforming weights as per (15). This approach is not practical in the context of WSNs. Instead, it would be preferable to arrive at the same solution through a distributed procedure where each node uses only local information. We now present an approach for arriving at the same centralized solution of (12) in a distributed way. The approach is based on direct observation of the solution via the iterative Algorithm 1 and how to distribute it by means of consensus [24].

It is clear from Proposition 1 that in order to arrive to the optimal solution in a distributed way, it will suffice to compute the optimal value $t^\star$ in a distributed fashion. By direct inspection of Algorithm 1 it is easy to realize that a distributed counterpart based on consensus is possible provided that the number of nodes in the network is known. The idea is very simple and is based on the observation that $t^{(k)}$ in step 4 of Algorithm 1 is a function of four terms, each of which can be computed by means of consensus. To that end, consider four variables per node α_m, β_m, γ_m and η_m, each one contributing to a, b, c and d of (16), respectively. Initially, we assume that all nodes are transmitting below their maximum power (the same as in Algorithm 1) so that $\alpha_m^{(0)} = \bar{h}_m\sqrt{E_m/T}$ and $\gamma_m^{(0)} = \sigma_r^2(\rho/\rho_0)\bar{h}_m^2 E_m/T$ while $\beta_m^{(0)}$ and $\eta_m^{(0)}$ are set to zero. If we perform an average consensus over these initial quantities, they converge to $\frac{1}{M}\sum_m \alpha_m^{(0)}$ and $\frac{1}{M}\sum_m \gamma_m^{(0)}$ at each node [24]. By multiplying by M these two average quantities we recover a and c of (17). Using these two quantities we can arrive at the same initial value of $t^{(0)}$ as in the centralized Algorithm 1. After that, each node computes its power allocation $\bar{w}_m^{(0)}$ as in (15). If a node is required to transmit at its maximum power then, it sets $\alpha_m^{(1)} = 0$, $\beta_m^{(1)} = \bar{h}_m\sqrt{p_m}$, $\gamma_m^{(1)} = 0$ and $\eta_m^{(1)} = \sigma_r^2(\rho/\rho_0)\bar{h}_m^2 p_m$ for the next iteration. After the nodes have reached a consensus, they scale their local variables by multiplying them by M in order to recover a, b, c and d of (17). The process is then repeated until convergence (e.g., the change in the optimal t does not change significantly). The complete algorithm is shown in Algorithm 2. It is easy to see that both Algorithms 1 and 2 yield to the same solution.

Algorithm 2: Consensus-based Iterative Algorithm

$k \to 0$
$\alpha_m^{(0)} \to \bar{h}_m\sqrt{E_m/T}$, $\beta_m^{(0)} \to 0$ for all $m = 1,\ldots,M$
$\gamma_m^{(0)} \to \sigma_r^2(\rho/\rho_0)\bar{h}_m^2 E_m/T$, $\eta_m^{(0)} \to 0$ for all $m = 1,\ldots,M$
repeat
  $k \to k+1$
  begin consensus
    $\alpha_c^{(k)} \to \frac{1}{M}\sum_{m=1}^M \alpha_m^{(k-1)}$, $\beta_c^{(k)} \to \frac{1}{M}\sum_{m=1}^M \beta_m^{(k-1)}$
    $\gamma_c^{(k)} \to \frac{1}{M}\sum_{m=1}^M \gamma_m^{(k-1)}$, $\eta_c^{(k)} \to \frac{1}{M}\sum_{m=1}^M \eta_m^{(k-1)}$
  end consensus
  Locally, at every node $m = 1,\ldots,M$ compute:
  $[a,b,c,d] \to M * [\alpha_c^{(k)}, \beta_c^{(k)}, \gamma_c^{(k)}, \eta_c^{(k)}]$
  $t_m^{(k)} \to (\sqrt{(\rho/\rho_0 + d)(a^2-c) + b^2 c} - ab)/(a^2-c)$
  $\bar{w}_m^{(k)} \to \min(t_m^{(k)}\sqrt{E_m/T}, \sqrt{p_m})$
  if $\bar{w}_m^{(k)} == \sqrt{p_m}$ **then**
    $\alpha_m \to 0$, $\beta_m^{(k)} \to \bar{h}_m\sqrt{p_m}$
    $\gamma_m \to 0$, $\eta_m^{(k)} \to \sigma_r^2(\rho/\rho_0)\bar{h}_m^2 p_m$
  end if
until $t^{(k)} - t^{(k-1)} \leq \varepsilon$
$w_m \to \bar{w}_m^{(k)} \frac{h_m}{|h_m|}$ for all $m = 1,\ldots,M$

V. Beamforming With Additional Random Energy Consumption

So far we have considered that energy consumption at the nodes is only due to the communication process with the far-away base station. However, there are other tasks that cause battery depletion such as sensing, data processing, or node to node communication, among others. In this section we analyze the beamforming problem also considering the energy demands of processes other than communications beamforming. We do it by adding an extra random energy term in the consumption model that allows us to abstract the communications part (e.g., beamforming) from the rest of the of the tasks. We follow a fairly general approach and do not assume any specific distribution of the random energy term since it may vary depending on the application, the hardware used, the protocol characteristics or the connectivity of the network [17]. We first consider that the additional energy consumption is random among nodes but fixed over time (e.g., following the general model of fix cost associated with circuitry consumption plus a term corresponding to the power dissipated through the antenna [19]–[22]) and then we look at the more general case where energy consumption is a stochastic process on every node [17]. Due to the introduction of randomness into the problem, we have to switch from a deterministic to a probabilistic design. We will provide conditions under which the problem is convex and can be efficiently solved. Additionally, we will also consider distributed implementations for a restricted class of problems using dual decomposition methods and consensus.

A. Random Energy Consumption Among the Nodes

Assume that each node has a fixed energy consumption due to non-beamforming communications but that this energy consumption is random among the nodes. We denote the additional energy consumed per sampling period at node m as ξ_m. The total energy consumed in one sampling period at node m is

$$\tilde{e}_m(\bar{w}_m) = \xi_m + T\bar{w}_m^2. \tag{37}$$

After K time samples, the battery level at node m will be reduced by an amount of $K\tilde{e}_m(w_m)$ units of energy. Our goal is then to design a beamformer that maximizes the time that the network will be operative with a certain probability.

Definition 2 (Probabilistic Network Lifetime): The δ-lifetime of the network $(0 \leq \delta \leq 1)$ is the longest time (or maximum time) after which the network will have all its nodes active with probability no less than δ.

Let denote δ as the probability that the network is operating after K sampling periods and, assuming independence of the energy consumption at each node, we can formulate the problem as

$$\begin{aligned} \underset{\mathbf{w},K}{\text{maximize}} \quad & K \\ \text{subject to} \quad & [\bar{\mathbf{h}}^T \ 0]\begin{bmatrix}\bar{\mathbf{w}}\\ 1\end{bmatrix} \geq \left\|\mathbf{D}\begin{bmatrix}\bar{\mathbf{w}}\\ 1\end{bmatrix}\right\| \\ & \bar{w}_m^2 \leq p_m, \text{ for all } m = 1,\ldots,M \\ & \prod_{m=1}^{M} \Pr\left[E_m - K\tilde{e}_m(\bar{w}_m) \geq 0\right] \geq \delta. \end{aligned} \tag{38}$$

We also assume a constant beamvector over time in problem (38). It can be shown that if the underlying pdf's of ξ_m are log-concave, then a constant beamvector maximizes the lifetime of the network.

Note that the above problem is a generalization of the original problem (11) and that both problems coincide if $\xi_m = 0$ (i.e., no randomness). Consider now the change of variables $t = 1/K$ and take the $\log(\cdot)$ in the last constraint so that we can express problem (38) as the following equivalent minimization problem

$$\begin{aligned} \underset{\bar{\mathbf{w}},t}{\text{minimize}} \quad & t \\ \text{subject to} \quad & [\bar{\mathbf{h}}^T \ 0]\begin{bmatrix}\bar{\mathbf{w}}\\ 1\end{bmatrix} \geq \left\|\mathbf{D}\begin{bmatrix}\bar{\mathbf{w}}\\ 1\end{bmatrix}\right\| \\ & \bar{w}_m^2 \leq p_m, \text{ for all } m = 1,\ldots,M \\ & \sum_m \log \Pr\left[E_m - \frac{\xi_m + T\bar{w}_m^2}{t} \geq 0\right] \geq \log\delta. \end{aligned} \tag{39}$$

The new problem (39) can be shown to be convex under certain conditions on the random variables ξ_m. Let $p_{\xi_m}(\xi_m)$ denote the probability density function (pdf) of ξ_m.

Lemma 4: If the family of probability density functions $p_{\xi_m}(\xi_m)$, $m = 1,\ldots,M$ are log-concave, then problem (39) is convex.

Proof: In order to show convexity of (39) it suffices to show that the last constraint in (39) is, indeed, convex. For that purpose recall from [25] that the Cumulative Distribution Function (CDF) of a random variable with log-concave pdf is also log-concave. We also have that the composition of a non-decreasing concave function with a concave function is also concave [25]. Keeping in mind these results, note that

$$\begin{aligned} \Pr\left[E_m - \tilde{e}_m(w_m)/t \geq 0\right] &= \Pr\left[\xi_m \leq tE_m - T\bar{w}_m^2\right] \\ &= F_{\xi_m}\left(tE_m - T\bar{w}_m^2\right) \end{aligned} \tag{40}$$

where $F_{\xi_m}(\cdot)$ is the CDF of ξ_m. We can now rewrite the last inequality in (39) as

$$\sum_{m=1}^{M} \log\left(F_{\xi_m}\left(tE_m - T\bar{w}_m^2\right)\right) \geq \log\delta \tag{41}$$

Since, by assumption, $p_{\xi_m}(\xi_m)$ is log-concave, so it is $F_{\xi_m}(\xi_m)$. Note that the argument of $F_{\xi_m}(\cdot)$ is concave in t (actually, affine) and $\bar{w}_m$ and note also that $\log F_{\xi_m}(\cdot)$ is concave and non-decreasing since it is a CDF. Then, applying the composition rule of concave functions we conclude that $\log(F_{\xi_m}(tE_m - T\bar{w}_m^2))$ is concave. The addition of concave functions is also concave and therefore the above constraint defines a convex set (i.e., it is a convex constraint). ■

The result in Lemma 4 is very useful since many common distributions (i.e., Gaussian, exponential, uniform) have log-concave pdf's, see Fig. 2.

Example—Uniform Case: We provide here an illustrative example where we consider that the additional random energy consumption due to non-beamforming task ξ_m follows a uniform distribution on some interval. More precisely, assume that $\xi_m \sim \mathcal{U}(e_{\min}, e_{\max})$; then, it is easy to check that

$$\begin{aligned} F_{\xi_m}(\xi_m) &= \int_{-\infty}^{\xi_m} p_{\xi_m}(x)dx \\ &= \begin{cases} 0 & \xi_m < e_{\min} \\ \frac{\xi_m - e_{\min}}{e_{\max} - e_{\min}} & e_{\min} \leq \xi_m \leq e_{\max} \\ 1 & \xi_m > e_{\max} \end{cases} \end{aligned} \tag{42}$$

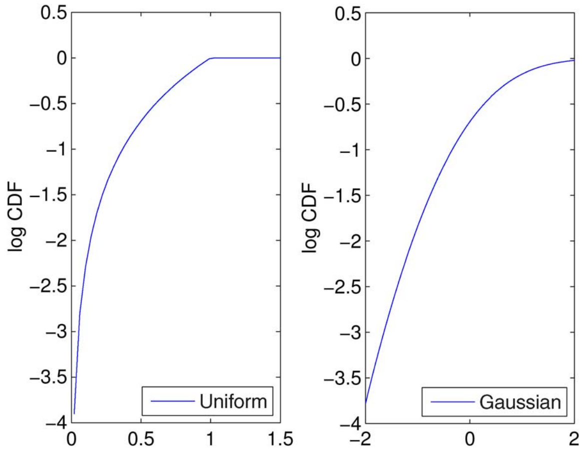

Fig. 2. Examples of log-concave CDFs, Uniform (left) and Gaussian (right).

So we can write the last constraint in (39) as

$$\sum_{m=1}^{M} \log\left(\min\left(\frac{tE_m - T\bar{w}_m^2 - e_{\min}}{e_{\max} - e_{\min}}, 1\right)\right) \geq \log\delta \quad (43)$$

together with the additional set of constraints

$$tE_m - T_s\bar{w}_m^2 \geq e_{\min}\ m = 1, \ldots, M. \quad (44)$$

Observe that, as expected from Lemma 4, the expression in (43) defines a convex set. In particular, the left-hand-side (LHS) of (43) is the sum of concave functions (note that the log of a concave function is also concave).

Example—Gaussian Case: Assume that ξ_m follows a Gaussian distribution of mean μ_m and variance σ_m^2, that is $\xi_m \sim \mathcal{N}(\mu_m, \sigma_m^2)$. Let the CDF of a normalized Gaussian be defined as

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2} dt. \quad (45)$$

Then, the last constraint in (39) can be written as

$$\sum_{m=1}^{M} \log\Phi\left(\frac{tE_m - T\bar{w}_m^2}{\sigma_m}\right) \geq \log\delta, \quad (46)$$

whose LHS is concave as it is the sum of concave functions (note that $\Phi(x)$ is log-concave [25]).

B. Stochastic Energy Consumption

If the network has no fixed way of operation, the amount of energy consumed on each node for non-beamforming tasks may vary over time. In this section we consider the more general case where ξ_m is a stochastic process [17] (i.e., ξ_m is a random variable at each time instant). Denote $\xi_m[k]$ as the random energy consumption at time instant k then, the battery level at node m after K transmissions will be equal to

$$\tilde{E}_m[K] = E_m - \sum_{k=1}^{K} \xi_m[k] - T\sum_{k=1}^{K} \bar{w}_m^2[k] \quad (47)$$

Again, we aim to maximize the network lifetime using Definition 2 (i.e., the longest time that guarantees that all nodes will be alive with a certain probability). As a result we end up with a family of power allocation problems indexed by K:

$$\begin{aligned} &\text{find} && \{\bar{\mathbf{w}}[1], \ldots, \bar{\mathbf{w}}[K]\} \in \mathbb{R}^M \\ &\text{subject to} && [\bar{\mathbf{h}}^T\ 0]\begin{bmatrix}\bar{\mathbf{w}}[k] \\ 1\end{bmatrix} \geq \left\|\mathbf{D}\begin{bmatrix}\bar{\mathbf{w}}[k] \\ 1\end{bmatrix}\right\|, \forall k \\ &&& \bar{w}_m^2[k] \leq p_m, \forall m, k \\ &&& \sum_{m=1}^{M} \log\Pr\left[E_m - \sum_{k=1}^{K}\xi_m[k] \right. \\ &&& \left. \quad -T\sum_{k=1}^{K}\bar{w}_m^2[k] \geq 0\right] \geq \log\delta, \forall k. \end{aligned} \quad (48)$$

The feasibility problem (48) need not to be convex in general, but under some relaxed conditions it is.

Lemma 5: If the family of probability density functions $p_{\xi_m}(\xi_m)$, $m = 1, \ldots, M$ are log-concave, then the feasibility problem (48) is convex.

Proof: The first constraint set of constraints are SOC convex constraints, while the second set of constraints define a set of balls and therefore, are convex, too. It only remains to show that the last constraint defines a convex set. Let $\eta_m = \sum_{k=1}^{K}\xi_m[k]$ and denote $p_{\eta_m}(\eta_m)$ and $F_{\eta_m}(\eta_m)$ its probability density function and cumulative distribution function, respectively. Note that $p_{\eta_m}(\eta_m)$ is the convolution of a set of log-concave functions. Since log-concavity is preserved under convolution [25], $p_{\eta_m}(\eta_m)$ is log-concave and so it is $F_{\eta_m}(\eta_m)$. Now rewrite the last constraint in (48) as

$$\sum_{m=1}^{M} \log F_{\eta_m}\left(E_m - T\sum_{k=1}^{K}\bar{w}_m^2[k]\right) \geq \log\delta. \quad (49)$$

Since $\log F_{\eta_m}(\cdot)$ is concave and non-decreasing (since it is a CDF) and $E_m - T\sum_{k=1}^{K}\bar{w}_m^2[k]$ is concave in $\bar{w}_m[k]$, then applying the composition rule of concave functions we conclude that the functions $\log F_{\eta_m}(E_m - T\sum_{k=1}^{K}\bar{w}_m^2[k])$ are concave and, hence, the last constraint defines a convex set. ■

Note that, the pdf's $p_{\xi_m}(\cdot)$ could vary from time to time as long as they are log-concave.

We have seen that if the involved density functions are log-concave the problem is convex. However, problem (48) can be shown to be convex asymptotically even when log-concavity does not hold.

Lemma 6: Assume, without loss of generality, that $\xi_m[k]$ are i.i.d. random variables with $\mathrm{E}[\xi_m[k]] = 0$ and $\mathrm{E}[\xi_m^2[k]] = 1$. As the number of slots K tends to infinity, the feasibility problem (48) is asymptotically convex regardless of the distribution of ξ_m.

Proof: In order to show convexity of (48) for K sufficiently large, it only suffices to show that the third constraint is asymptotically convex as K increases. Rewrite the last constraint in (48) as

$$\sum_{m=1}^{M} \log\Pr\left[E_m - \sqrt{K}\frac{1}{\sqrt{K}}\sum_{k=1}^{K}\xi_m[k] - T\sum_{k=1}^{K}\bar{w}_m^2[k] \geq 0\right] \geq \log\delta. \quad (50)$$

Let $\phi_m = \frac{1}{\sqrt{K}} \sum_{k=1}^{K} \xi_m[k]$. As the number of slots K increases and, by virtue of the Central Limit Theorem, ϕ_m can be well approximated by a Gaussian distribution ($\phi_m \sim \mathcal{N}(0,1)$). We can now approximate (50) by

$$\sum_{m=1}^{M} \log \Phi \left(\frac{1}{\sqrt{K}} \left(E_m - T \sum_{k=1}^{K} \bar{w}_m^2[k] \right) \right) \geq \log \delta. \quad (51)$$

Since, for fixed K, the function $\frac{1}{\sqrt{K}}(E_m - T\sum_{k=1}^{K} \bar{w}_m^2[k])$ is concave in $\bar{w}_m[k]$, $k = 1, \ldots, K$, and the Gaussian distribution $\Phi(x)$ is log-concave [25] and non-decreasing, we have by the composition rule of concave functions [25] that (51) defines a convex set. Therefore, the feasibility problem (48) is convex. ■

If problem (48) is convex we can further show that a constant power allocation (independent of the slot index k) can be built to solve for (48).

Lemma 7: Assume that the last constraint in (48) is convex then, there exist an optimal solution $\{\bar{\mathbf{w}}^\star[1], \ldots, \bar{\mathbf{w}}^\star[K]\}$ to (48) such that $\bar{\mathbf{w}}^\star[i] = \bar{\mathbf{w}}^\star$ for all $i = 1, \ldots, K$. In particular, as K goes to infinity and for $\xi_m[k]$ i.i.d. such a solution always exist.

Proof: If the last constraint in (48) is convex then the problem is convex. Then, the same argument as in Lemma 1 applies. ■

Proposition 4: Assume, without loss of generality, that $\xi_m[k]$ are i.i.d. with zero mean and unit variance. For K sufficiently large, feasibility problem (48) can be well-approximated by the following problem:

$$\begin{aligned} &\text{find} && \bar{\mathbf{w}} \\ &\text{subject to} && [\bar{\mathbf{h}}^T \; 0] \begin{bmatrix} \bar{\mathbf{w}} \\ 1 \end{bmatrix} \geq \left\| \mathbf{D} \begin{bmatrix} \bar{\mathbf{w}} \\ 1 \end{bmatrix} \right\| \\ &&& \bar{w}_m^2 \leq p_m, \; m = 1, \ldots, M \\ &&& \sum_m \log \Phi \left(\frac{E_m - KT\bar{w}_m^2}{\sqrt{K}} \right) \geq \log \delta. \end{aligned} \quad (52)$$

Proof: The result follows directly from Lemmas 6 and 7.■

Note that feasibility problem (52) is convex. Then, we can approximately find the beamformer that maximizes the network lifetime by first finding the optimal power allocation and then setting the phase to match that of the channel. The optimal power allocation can be obtained by solving a sequence of feasibility problems by performing a bisection search over K. A lower bound on the optimal value of K can be easily obtained as

$$K_{\min} = \frac{1}{T} \min \left(\frac{E_1}{p_1}, \ldots, \frac{E_M}{p_M} \right). \quad (53)$$

An upper bound can be obtained as follows: If we assume that there is no additional energy consumption (i.e., $\xi_m = 0$ for all $m = 1, \ldots, M$) then, finding the optimal beamformer is equivalent to solving the original problem (11). Then, the optimal value of (11) will give an upper bound on the value of K. We can now run a bisection search over K between these two values in order to find the optimal beamformer.

C. Distributed Computation

Contrary to the case with no random energy consumption, problems (39) and (52) are not easily solved in a distributed fashion due to the SOC SNR coupling constraints. However, under the *ideal sensing* scenario we would like to argue that a distributed implementation is possible. Observe first that under the assumption of *ideal sensing*, the SOC SNR constraint becomes a linear constraint consisting of the addition of quantities computable at each node. Note also that the last constraint in both problems (39) and (52) can also be formed by the addition of node-dependent quantities. Using similar ideas as in [26]–[29] and using primal-dual decomposition methods it is possible to formulate the problem in such a way that each node has to solve a small optimization problem while the master problem for the dual updates can be computed by means of consensus. In particular, consider problem (39) under the *ideal sensing* assumpmtion. It is clear that the problem can be written as

$$\begin{aligned} &\underset{\bar{\mathbf{w}}, t \geq 0}{\text{minimize}} && t \\ &\text{subject to} && \bar{\mathbf{w}}^\mathsf{T} \bar{\mathbf{h}} \geq \sqrt{\rho/\rho_0} \\ &&& \bar{w}_m \leq \sqrt{p_m}, \; m = 1, \ldots, M \\ &&& \sum_{m=1}^{M} \log F_{\xi_m} \left(tE_m - T\bar{w}_m^2 \right) \geq \log \delta. \end{aligned} \quad (54)$$

Consider now the replication of variable t at every node keeping a global constraint that makes all of them equal. If we do so, then the optimal power allocation can be computed as the solution of the following problem

$$\begin{aligned} &\underset{\bar{\mathbf{w}}, t_m \geq 0, t}{\text{minimize}} && \sum_m^M t_m \\ &\text{subject to} && \sum_m^M \bar{w}_m \bar{h}_m \geq \sqrt{\rho/\rho_0} \\ &&& \sum_{m=1}^{M} \log F_{\xi_m} \left(t_m E_m - T\bar{w}_m^2 \right) \geq \log \delta \\ &&& \bar{w}_m \leq \sqrt{p_m}, \; m = 1, \ldots, M \\ &&& t_m = t, \; m = 1, \ldots, M. \end{aligned} \quad (55)$$

As shown in [26], [28] it is not necessary for all the nodes to agree on a common variable since it suffices for every node to agree with its neighbors (provided that the graph is strongly connected). Therefore, we can write the power allocation problem as

$$\begin{aligned} &\underset{\bar{\mathbf{w}}, t_m \geq 0}{\text{minimize}} && \sum_m^M t_m \\ &\text{subject to} && \sum_m^M \bar{w}_m \bar{h}_m \geq \sqrt{\rho/\rho_0} \\ &&& \sum_{m=1}^{M} \log F_{\xi_m} \left(t_m E_m - T\bar{w}_m^2 \right) \geq \log \delta \\ &&& \bar{w}_m \leq \sqrt{p_m}, \; m = 1, \ldots, M \\ &&& t_m = t_j, \; j \in \mathcal{N}_m, \; m = 1, \ldots, M, \end{aligned} \quad (56)$$

where $\mathcal{N}_m$ is the set of neighbors of node m. Consider now the equivalent problem where a quadratic penalty term on the values of t_m is also included in the objective so that we end up with the following problem

$$\begin{aligned} \underset{\bar{\mathbf{w}}, t_m \geq 0}{\text{minimize}} \quad & \sum_m^M t_m + c \sum_m \sum_{j \in \mathcal{N}_m} (t_m - t_j)^2 \\ \text{subject to} \quad & \sum_m^M \bar{w}_m \bar{h}_m \geq \sqrt{\rho/\rho_0} \\ & \sum_{m=1}^M \log F_{\xi_m} \left(t_m E_m - T \bar{w}_m^2 \right) \geq \log \delta \\ & \bar{w}_m \leq \sqrt{p_m}, \; m = 1, \ldots, M \\ & t_m = t_j, \; j \in \mathcal{N}_m, \; m = 1, \ldots, M, \end{aligned} \quad (57)$$

where c is a parameter. Note that for any feasible point, the quadratic term in the objective function of problem (57) equals zero and does not affect the optimal value of the problem. The only purpose of the quadratic term is to make the cost function strictly convex allowing the computation of its solution in a distributed way [26], [28]. The Lagrangian of problem (57) can be written as

$$\begin{aligned} \mathcal{L}(\bar{\mathbf{w}}, \{t_m\}) = & \sum_m^M t_m^2 + c \sum_m \sum_{j \in \mathcal{N}_m} (t_m - t_j)^2 \\ & + \lambda \left(\sqrt{\frac{\rho}{\rho_0}} - \sum_m \bar{w}_m \bar{h}_m \right) + \sum_m \sigma_m (\bar{w}_m - \sqrt{p_m}) \\ & + \mu \left(\log \delta - \sum_{m=1}^M \log F_{\xi_m} \left(t_m E_m - T \bar{w}_m^2 \right) \right) \\ & + \sum_m \sum_{j \in \mathcal{N}_m} \beta_{mj} (t_m - t_j) \end{aligned} \quad (58)$$

It is clear from the Lagrangian that problem (56) is amenable to be solved in a distributed fashion using dual decomposition techniques together with consensus. Using the ideas in [26], [28] it can be easily shown that the above problem can be solved by the following updates:

i) Each node computes $\bar{t}_m^{(k+1)}$ and $\bar{w}_m^{(k+1)}$ as the solution to

$$\begin{aligned} \underset{\bar{w}_m \leq \sqrt{p_m}, t_m \geq 0}{\text{minimize}} \quad & t_m + \tilde{\beta}_m t_m - \lambda \bar{w}_m \bar{h}_m \\ & - \mu \log F_{\xi_m} \left(t_m E_m - T \bar{w}_m^2 \right) \\ & + c \sum_{j \in \mathcal{N}_m} \left(t_m - \frac{1}{2} \left(t_m^{(k)} + t_j^{(k)} \right) \right)^2, \end{aligned} \quad (59)$$

where $\tilde{\beta}_m$ are equivalent Lagrange multipliers that replace β_{mj} as described in [28].

ii) Neighboring nodes exchange their values $t_m^{(k+1)}$ and update their multipliers

$$\tilde{\beta}_m = \tilde{\beta}_m + c \sum_{j \in \mathcal{N}_m} \left(t_m^{(k+1)} - t_j^{(k+1)} \right). \quad (60)$$

iii) Nodes agree on the value of the coupling constraints (e.g., by consensus and assuming that M is known)

$$\sum_m \bar{w}_m^{(k+1)} \bar{h}_m$$

$$\sum_m \log F_{\xi_m} \left(t_m^{(k+1)} E_m - T \left(\bar{w}_m^{(k+1)} \right)^2 \right)$$

iv) Global multiplier update at every node

$$\lambda = \lambda + \epsilon_\lambda \left(\sqrt{\rho/\rho_0} - \sum_m \bar{w}_m^{(k+1)} \bar{h}_m \right) \quad (61)$$

$$\mu = \mu + \epsilon_\mu \left(\log \delta - \sum_m \log F_{\xi_m} \left(t_m^{(k+1)} E_m - T \left(\bar{w}_m^{(k+1)} \right)^2 \right) \right) \quad (62)$$

where ϵ_λ and ϵ_μ are the step-sizes for the updates of the multipliers λ and μ, respectively. At the beginning of the algorithm all multipliers are initially set to zero.

VI. Numerical Simulations

In this section we provide some numerical results in order to illustrate the proposed solution and algorithms. We will start considering a centralized scenario, providing a comparison of the proposed method versus collaborative beamforming with uniform power allocation among the nodes. We will then show how the distributed consensus-based method of Algorithm 2 provides the same solution as its centralized counterpart, i.e., Algorithm 1. The last set of experiments will consider the case with random energy consumption. We will explore the additional energy consumption term affects the optimal beamformer as compared to the deterministic case.

1) Centralized and Distributed Algorithms: In order to illustrate the benefits of the proposed approach, we have generated a random network of $M = 20$ nodes over the unit square. The channel coefficients have been generated at random following a circularly symmetric complex Gaussian distribution of zero mean and unit variance (i.e., Rayleigh fading). The proportionality constant T has been set to one for all experiments.

In our first experiment, we want to quantify the degradation in terms of SNR when the observation noise is neglected. For that purpose, we run a simulation where we have varied the observation noise power (equivalently the observation to measurement noise power ratio σ_r^2). The battery levels of the nodes have been generated at random from a uniform distribution $\mathcal{U}(0.5, 1)$ while the maximum transmission power of the nodes has been set to $p_m = 1$ for all m. The measurement noise power is set to $\sigma_n^2 = 0.1, P_s = 1$ and the target SNR has been set to $\rho = 20\,\text{dB}$. Under this setup, we have computed the optimal beamformers using the iterative Algorithm 1 for different values of σ_r^2. We have also computed the optimal beamformers as is there were no observation noise (*ideal sensing*). In Fig. 3 we have depicted the achieved SNR at the base station for the two considered cases. As it can be observed, there is a fast degradation in terms of SNR when neglecting the observation noise while the optimal beamformer always achieves the target SNR.

In Fig. 4 we have illustrated the battery depletion over time when using the optimal power allocation of Proposition 1. In the top plot of Fig. 4 we have depicted the optimal power allocation. Note that, in this particular example, there are a few nodes transmitting at maximum power (dashed lines). In the bottom plot of Fig. 4 we have illustrated the corresponding battery level over time of each of the nodes. It is easy to see that, at the optimum, all nodes that are transmitting below their maximum power deplete their batteries at the same time. Those nodes with favorable channel conditions could be operative for a longer period of time (thus they can transmit at maximum power). This figure

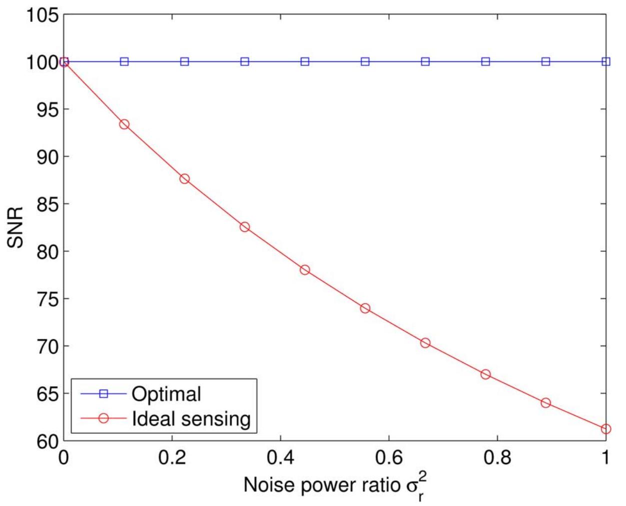

Fig. 3. SNR degradation when neglecting the observation noise for a network with 20 nodes and target SNR $\rho = 20$ dB. The measurement noise power is set to $\sigma_n^2 = 0.1$ and $\rho_0 = 10$.

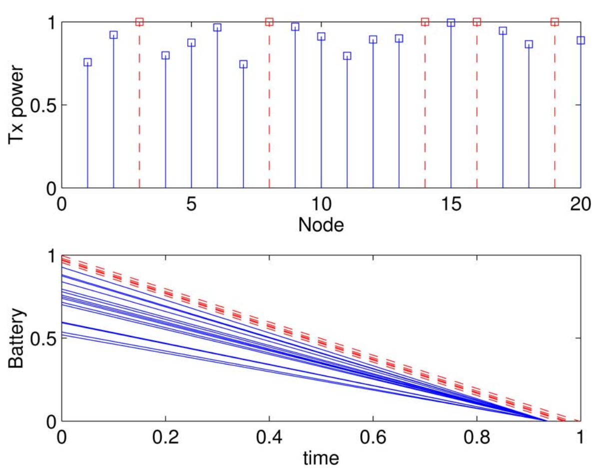

Fig. 4. Optimal power allocation (top) and corresponding battery depletion for a network of 20 nodes.

also serves to illustrate that, in our particular scenario the first node depletion criterion for lifetime maximization and the one of Definition 1 are equivalent.

In order to quantify the benefits of the proposed methodology we have performed several simulations to compare against more conventional collaborative beamforming strategies. We compare our approach to a more efficient version of the CB strategy that we call "Greedy CB". The way it works is as follows: At the beginning all nodes participate in the transmission and they scale their power (the same for all nodes) so that the SNR constraint is achieved with equality. Every time a node depletes its battery, the transmission power is recalculated again. This procedure is repeated until the SNR constraint cannot be met. We have run 1000 realizations of the experiment with the following parameters: $\rho_0 = \rho = 20\,\text{dB}, \sigma_r^2 = \sigma_n^2 = 0.01$ and $E_m \sim \mathcal{U}(0.5, 1)$, $p_m = 1$ for $m = 1, \ldots, 20$. We have plot the empirical CDFs of the greedy CB approach for the time that takes for the first node to deplete its battery as well as the time that the QoS SNR constraint can be satisfied. In Fig. 5 we can compare the CDFs of the "Greedy CB" versus the optimized solution "Optimized CB". It can be appreciated that the optimized beamforming weights outperform the greedy approach. We can also observe that there is a gap between the times resulting from the first node

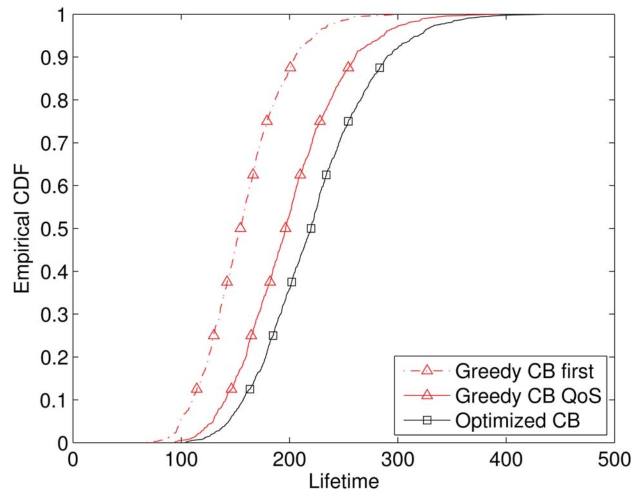

Fig. 5. Empirical CDFs of the greedy CB versus the optimized CB strategy.

TABLE I
AVERAGE AND STANDARD DEVIATION OF THE NETWORK'S LIFETIME FOR DIFFERENT APPROACHES

	Feng et. al. [8]		Greedy CB		Optimized CB
	1st	QoS	1st	QoS	
Avg.	5.77	6.14	156.84	199.43	222.41
Std.	0.52	0.47	37.10	47.49	52.72

depletion criterion of "Greedy CB" versus the QoS criterion of Definition 1. On contrast, in "Optimized CB" there is no such gap since both times coincide. We have also performed a comparison with the selection mechanism proposed in [8]. Since, we assume that nodes can adjust their phases to add constructively at the receiver, the selection mechanism is based on the remaining battery level of the nodes and their channel coefficient. In [8] only battery level is considered since all nodes are assumed to experience the same attenuation. However, in our case different nodes see different channels. Therefore, we select the nodes according to their battery levels and their channels. This approach gives better results than considering the batteries only. In Table I we have displayed the mean and standard deviation values of the network's lifetime based on 1000 realizations for the considered approaches. As it can be seen from Table I, the selection mechanism of [8] does not give good results and performs far from the other considered approaches. This result is not surprising since, as stated before, at the optimum all nodes must be active. Therefore, considering only a small subset of nodes does not help to improve the network's lifetime. Also note that the SNR increases as the square of the sum of individual powers, which means better efficiency as more nodes participate.

In order to illustrate the convergence of the proposed distributed algorithm, in Fig. 6 we have depicted the evolution of the consensus-based iterative algorithm (Algorithm 2) as a function of the number of consensus rounds. In the top plot of Fig. 6 we observe the evolution of the SNR at the base station as the number of iterations increases. In the bottom plot, we have displayed the difference between the optimal beamformer weights and the weights at each iteration. As it can be observed, the distributed consensus-based algorithm reaches the same solution as in the centralized case.

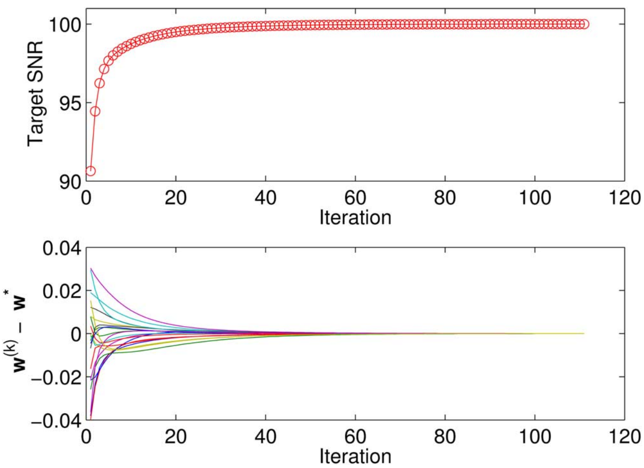

Fig. 6. Convergence of the distributed algorithm to the centralized solution and to the target QoS SNR.

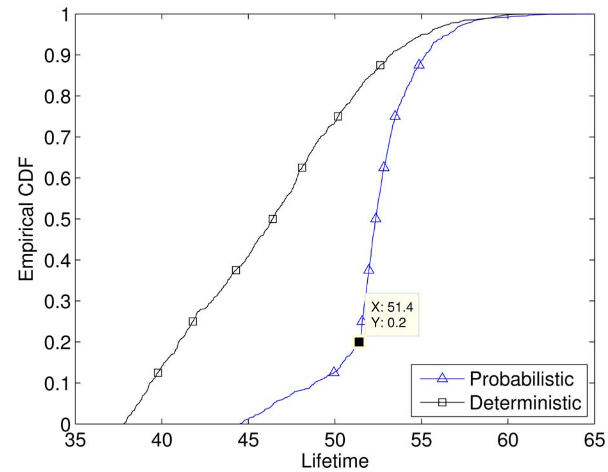

Fig. 8. Empirical CDF of the network lifetime using the static random model where there is an extra energy consumption term that is fix over time but random among the nodes.

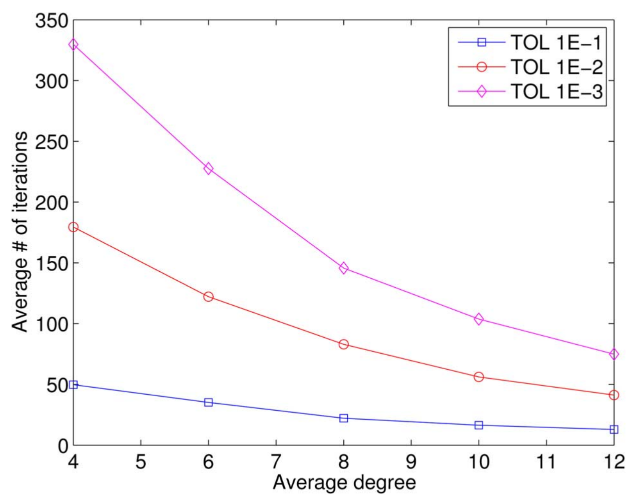

Fig. 7. Average number of iterations required for the convergence of the consensus-based iterative algorithm to the centralized solution. The average number of iterations is represented as a function of the average degree for a network of 20 nodes, and for different thresholds to decide upon convergence.

Continuing with the analysis of Algorithm 2 we have randomly generated different networks of $M = 20$ nodes with different connectivities. We varied the average degree from 4 to 12 and generated 1000 realizations for each case. For each realization, the number of iterations required to converge has been evaluated. The information is mixed in the consensus part using a weighting matrix equal to $\mathbf{W} = \mathbf{I} - \alpha\mathbf{L}$, where $\mathbf{I}$ is the identity matrix and $\mathbf{L}$ is the Laplacian of the graph. We have set $\alpha = 0.9/\sigma_{\max}(\mathbf{L})$, where $\sigma_{\max}$ is the maximum singular value of the Laplacian matrix. Although not optimal, such a simple choice of the weighting matrix can be shown to converge [30]. We decide that the algorithm has converged when the standard deviation of the relative error between the nodes and the centralized solution is smaller than some threshold. In Fig. 7 we have plotted the results for different values of the threshold (TOL curves). As expected, higher precision (smaller TOL values) require a higher number of iterations. Similarly, increasing the average degree of the network reduces the averaging time since nodes can reach a larger number of neighbors at every iteration. We would like to point out that the results presented in Fig. 7 can be substantially improved by appropriately choosing the weights (the $\mathbf{W}$ matrix) for mixing the information in the consensus part as described in [30].

2) Random Energy Consumption Model: We now proceed to show some simulations when considering that there is an additional random energy consumption term due to tasks other than beamforming for communications such as sensing, local (node to node) communication, and data processing. Consider first the case where nodes have an additional random (but fixed over the optimization period) energy consumption due to the aforementioned additional tasks. We provide here an illustrative example of such situation by assuming that the distribution of ξ_m is uniform on some interval. For the simulations we have set a target probability of $\delta = 0.8$ that is, we want to find the maximum time that our network is going to be alive at least in $\delta \times 100\%$ of the realizations. We set the target SNR to $\rho = 20$ dB and perform a simulation over 1000 realizations. The random energy consumption is drawn from a uniform distribution $\mathcal{U}(0, 0.01)$ units of energy. In Fig. 8 we can see the CDF of the optimized network lifetime when we take into account the presence of an additional energy consumption (Probabilistic) compared to the case where no additional consumption is considered (Deterministic). We obtain through the solution of the optimization problem that the maximum time that the network will be alive $\delta \times 100\%$ of the time is 51.48. As it can be seen in the figure the results are in agreement with the empirical CDF.

A similar simulation has been performed but, in this case, we consider that the energy consumption varies from slot to slot (stochastic model). The energy consumption has been modeled as a stochastic process with $\xi_m[k] \sim \mathcal{U}(0, 7 \times 10^{-3})$. We have computed the empirical CDF's using three different approaches: deterministic (no random consumption considered), static (random but fixed over the optimization period with uniform distribution) and using the Gaussian approximation (52). As expected, taking into account the random energy term provides an improvement in the network lifetime as compared to the deterministic case. Also observe that the Gaussian approximation outperforms the assumption that the additional random energy term follows the wrong distribution.

Finally, we have included in Fig. 10 an illustration on how the outlined procedure in Section V-C converges to the centralized solution under the assumption of ideal sensing. We would

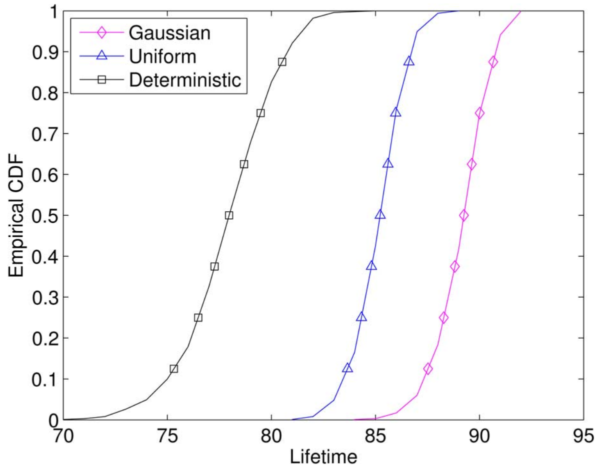

Fig. 9. Empirical CDF of the network's lifetime using the stochastic random model where there is an extra random energy consumption over time.

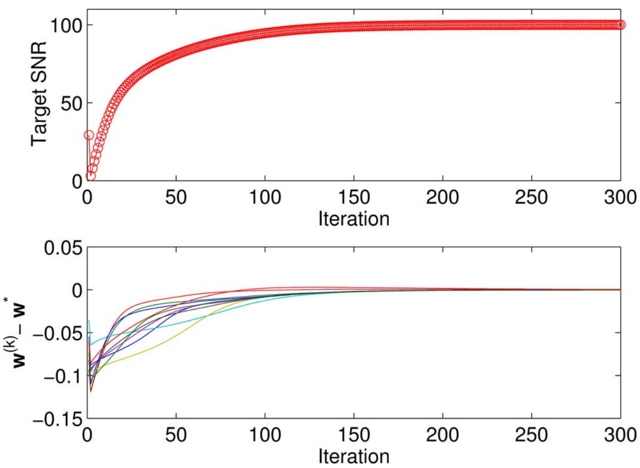

Fig. 10. Convergence of the distributed implementation to the centralized solution for the case with ideal sensing and random energy consumption.

like to point out that the convergence speed of the algorithm will depend on several parameters such as the choice of the penalty parameter c, the step-sizes of the global Lagrange multiplier updates ε_λ and ε_μ, the weighs used for the consensus part, the degree of the nodes, etc. Depending on the network's characteristics the optimal choices of these parameters will change and a significant speed-up in convergence time can be achieved by an appropriate choice of them.

VII. Conclusions

We have presented a new approach to energy efficient beamforming in sensor networks using convex optimization tools and consensus algorithms. The proposed strategy takes into account the remaining battery level at each node in order to optimize for the network lifetime while guaranteeing a specified QoS requirement. We have provided analytical expressions for the computation of the optimal solution and have devised both centralized and distributed (consensus-based) algorithms for its computation. We have validated by means of simulations that the proposed scheme outperforms existing collaborative beamforming strategies and that the distributed consensus-based algorithm converges to the same centralized solution of the problem.

In order to account for other possible causes of energy depletion like sensing, local communications, or data processing, we have added an additional random energy term into the energy consumption model. We have shown that the problem is convex under certain general assumptions on the distributions of the additional energy waste and therefore, it can still be solved efficiently using general purpose convex optimization software. Furthermore, we have provided asymptotic conditions under which the problem is convex regardless the distribution of the additional energy consumption due to non-beamforming tasks. We have shown that accounting for the additional energy waste results in a significant improvement of the network's lifetime.

Limitations and Future Directions: One of the assumptions of our model is that the information to be transmitted is common to all nodes in the network. This means that the nodes have acquired the signal either through a joint estimation process (e.g., joint localization of a source) or that they have spread the information to be transmitted. Therefore, our framework is more suitable for applications where data transmission can be done offline after enough data have been collected. In our formulation, we have assumed that the channel conditions don't change during the transmission. Whenever channel conditions change, it would be necessary to update the optimal beamformer. Since our approach uses consensus, it requires some iterations to converge to the optimal solution, therefore our approach will be applicable depending on on how fast the channel conditions change and on the particular size and topology of the network (i.e., how fast information is fused).

In order to address these issues, as future directions we could study the effect on varying channel conditions into the problem and try to simultaneously track the channel and update the beamformer weights. Alternatively, one could use a robust formulation where only the statistics of the channel are considered. Assuming that all nodes share the same information might become unrealistic when the network is very large. In such cases, it might be more convenient to cluster the nodes and form several beamformers rather than a single one.

Finally we have observed that when we include randomness into the problem, finding a distributed strategy to solve it becomes more challenging. For some scenarios, we were still able to provide a distributed algorithm for solving the problem. In the more general case further work needs to be done into that direction, too.

Acknowledgment

The authors would like to thank the editor Prof. Eduard Jorswieck and the anonymous reviewers for their useful comments and constructive criticism during the review process. Their valuable feedback has certainly contributed to improve the quality of the present work.

References

[1] H. Ochiai, P. Mitran, H. V. Poor, and V. Tarokh, "Collaborative beamforming for distributed wireless *ad hoc* sensor networks," *IEEE Trans. Signal Process.*, vol. 53, pp. 4110–4124, 2005.

[2] M. F. A. Ahmed and S. A. Vorobyov, "Performance characteristics of collaborative beamforming for wireless sensor networks with Gaussian distributed sensor nodes," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP 2008)*, 2008, pp. 3249–3252.

[3] M. F. A. Ahmed and S. A. Vorobyov, "Collaborative beamforming for wireless sensor networks with gaussian distributed sensor nodes," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 638–643, 2009.

[4] K. Zarifi, S. Affes, and A. Ghrayeb, "Distributed beamforming for wireless sensor networks with random node location," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2009, pp. 2261–2264.

[5] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 62–75, 2010.

[6] D. P. Palomar and Y. C. Eldar, *Convex Optimization in Signal Processing and Communications*. Cambridge, UK: Cambridge Univ. Press, 2010.

[7] Z. Han and H. V. Poor, "Lifetime improvement in wireless sensor networks via collaborative beamforming and cooperative transmission," *IET Microw., Antennas Propag.*, vol. 1, no. 6, pp. 1103–1110, 2007.

[8] J. Feng, C.-W. Chang, S. Sayilir, Y.-H. Lu, B. Jung, D. Peroulis, and Y. C. Hu, "Energy-efficient transmission for beamforming in wireless sensor networks," in *Proc. 7th Ann. IEEE Commun. Soc. Conf. Sens. Mesh and Ad Hoc Commun. Netw. (SECON)*, 2010, pp. 1–9.

[9] V. Havary-Nassab, S. Shahbazpanahi, A. Grami, and Z.-Q. Luo, "Distributed beamforming for relay networks based on second-order statistics of the channel state information," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4306–4316, 2008.

[10] G. Zheng, K.-K. Wong, A. Paulraj, and B. Ottersten, "Robust collaborative-relay beamforming," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3130–3143, 2009.

[11] G. Zheng, K.-K. Wong, A. Paulraj, and B. Ottersten, "Collaborative-relay beamforming with perfect CSI: Optimum and distributed implementation," *IEEE Signal Process. Lett.*, vol. 16, no. 4, pp. 257–260, 2009.

[12] L. Chen, K. Wong, H. Chen, J. Liu, and G. Zheng, "Optimizing transmitter-receiver collaborative-relay beamforming with perfect CSI," *IEEE Commun. Lett.*, vol. 15, no. 3, pp. 314–316, 2011.

[13] Z. Yi and I.-M. Kim, "Joint optimization of relay-precoders and decoders with partial channel side information in cooperative networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 447–458, 2007.

[14] H. Chen, A. B. Gershman, and S. Shahbazpanahi, "Distributed peer-to-peer beamforming for multiuser relay networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP 2009)*, 2009, pp. 2265–2268.

[15] B. Bejar, S. Zazo, and D. P. Palomar, "Lifetime maximization for beamforming applications in wireless sensor networks," in *Proc. Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2012, pp. 2849–2852.

[16] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.

[17] Y. Wang, M. C. Vuran, and S. Goddard, "Stochastic analysis of energy consumption in wireless sensor networks," in *Proc. 7th Ann. IEEE Commun. Soc. Conf. Sens. Mesh and Ad Hoc Commun. Netw. (SECON)*, 2010, pp. 1–9.

[18] I. Dietrich and F. Dressler, "On the lifetime of wireless sensor networks," *ACM Trans. Sen. Netw.* vol. 5, pp. 5:1–5:39, Feb. 2009 [Online]. Available: http://doi.acm.org/10.1145/1464420.1464425

[19] J. Polastre, J. Hill, and D. Culler, "Versatile low power media access for wireless sensor networks," in *Proc. 2nd Int. Conf. Embedded Netw. Sens. Syst.*, 2004, ser. SenSys'04, pp. 95–107 [Online]. Available: http://doi.acm.org/10.1145/1031495.1031508

[20] Q. Wang, M. Hempstead, and W. Yang, "A realistic power consumption model for wireless sensor network devices," in *Proc. 3rd Annu. IEEE Commun. Soc. Sens. Ad Hoc Commun. Netw. (SECON'06)* , Sep. 2006, vol. 1, pp. 286–295.

[21] Q. Wang and W. Yang, "Energy consumption model for power management in wireless sensor networks," in *Proc. 4th Annu. IEEE Commun. Soc. Conf. Sens., Mesh and Ad Hoc Commun. Netw. (SECON'07)*, Jun. 2007, pp. 142–151.

[22] H. Zhou, D. Luo, Y. Gao, and D. Zuo, "Modeling of node energy consumption for wireless sensor networks," *Wireless Sens. Netw.*, vol. 3, pp. 18–23, 2011.

[23] M. Z. Siam, M. Krunz, and O. Younis, "Energy-efficient clustering/routing for cooperative MIMO operation in sensor networks," in *Proc. INFOCOM*, 2009, pp. 621–629.

[24] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.

[25] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.

[26] M. G. Rabbat, R. D. Nowak, and J. A. Bucklew, "Generalized consensus computation in networked systems with erasure links," in *Proc. IEEE 6th Workshop Signal Process. Adv. Wireless Commun.*, 2005, pp. 1088–1092.

[27] B. Johansson, C. M. Carretti, and M. Johansson, "On distributed optimization using peer-to-peer communications in wireless sensor networks," in *Proc. 5th Ann. IEEE Commun. Soc. Conf. Sens., Mesh and Ad Hoc Commun. Netw. (SECON'08)*, 2008, pp. 497–505.

[28] J. Li, E. Elhamifar, I.-J. Wang, and R. Vidal, "Consensus with robustness to outliers via distributed optimization," in *Proc. 49th IEEE Conf. Decision Contr. (CDC)*, 2010, pp. 2111–2117.

[29] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[30] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Control Lett.*, vol. 53, pp. 65–78, 2004.

Benjamín Béjar Haro (S'12) received the Electrical Engineering degree from both the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, and the Technische Universität Darmstadt (TUD), Darmstadt, Germany, in 2006 under the framework of the Double Degree Exchange program. He received the Ph.D. degree in Electrical Engineering from the Universidad Politécnica de Madrid (UPM), Madrid, Spain, in 2012.

From September 2011 to September 2013, he was a member of the Vision, Dynamics and Learning Lab at the Johns Hopkins University (JHU), Baltimore, MD, as part of the MSE program in Biomedical Engineering. He currently serves as a Postdoctoral Fellow at the Audiovisual Communications Laboratory (LCAV), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Swizerland. He has held research appointments as a visiting Ph.D. student at the Università degli Studi di Udine, Udine, Italy, in 2009 and, at the Hong Kong University of Science and Technology (HKUST), Hong Kong, China, in 2011. His current research interests include the development and application of signal processing, convex optimization and machine learning techniques to cooperative/distributed communication systems and surgical robotics.

Dr. Haro e was awarded with the 2012 Best Paper Award in Medical Robotics and Computer Assisted Intervention Systems from the Medical Image Computing and Computer Assisted Intervention Society (MICCAI).

Santiago Zazo (M'12) received the Telecom Engineer and Dr. Eng. degrees by the Universidad Politecnica de Madrid (UPM), Madrid, Spain, in 1990 and 1995, respectively.

From 1991 to 1994, he was with the University of Valladolid and from 1995 to 1997 with the University Alfonso X El Sabio, Madrid. In 1998, he joined UPM as Associate Professor in Signal Theory and Communications. His main research activities are in the field of Signal Processing with applications to Audio, Communications and Radar. More recently, he has been mostly focused on MIMO communications and Wireless Sensor Networks from both physical layer and networking points of view. Since 1990, he has been author/coauthor of more than 20 journal papers and more than 100 conference papers. He has also led many private and national Spanish projects and has participated in six European projects, leading three of them.

Daniel P. Palomar (S'99–M'03–SM'08–F'12) received the Electrical Engineering and Ph.D. degrees from the Technical University of Catalonia (UPC), Barcelona, Spain, in 1998 and 2003, respectively.

He is an Associate Professor with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology (HKUST), Hong Kong, which he joined in 2006. Since 2013, he has been a Fellow of the Institute for Advance Study (IAS) at HKUST. He had previously held several research appointments, namely, at King's College London (KCL), London, UK; Technical University of Catalonia (UPC), Barcelona; Stanford University, Stanford, CA; Telecommunications Technological Center of Catalonia (CTTC), Barcelona; Royal Institute of Technology (KTH), Stockholm, Sweden; University of Rome "La Sapienza", Rome, Italy; and Princeton University, Princeton, NJ. His current research interests include applications of convex optimization theory, game theory, and variational inequality theory to financial systems and communication systems.

Dr. Palomar is recipient of a 2004/06 Fulbright Research Fellowship, the 2004 Young Author Best Paper Award by the IEEE Signal Processing Society, the 2002/03 best Ph.D. prize in Information Technologies and Communications by the Technical University of Catalonia (UPC), the 2002/2003 Rosina Ribalta first prize for the Best Doctoral Thesis in Information Technologies and Communications by the Epson Foundation, and the 2004 prize for the best Doctoral Thesis in Advanced Mobile Communications by the Vodafone Foundation and COIT. He serves as an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY, and has been an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, a Guest Editor of the IEEE SIGNAL PROCESSING MAGAZINE 2010 Special Issue on "Convex Optimization for Signal Processing," the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS 2008 Special Issue on "Game Theory in Communication Systems," and the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS 2007 Special Issue on "Optimization of MIMO Transceivers for Realistic Communication Networks."