Sensor Selection and Rate Distribution Based Beamforming in Wireless Acoustic Sensor Networks

Jie Zhang, Richard Heusdens, and Richard C. Hendriks
Circuits and Systems (CAS) Group, Delft University of Technology, the Netherlands
E-mail: {j.zhang-7; r.heusdens; r.c.hendriks}@tudelft.nl

***Abstract*—Power usage is an important aspect of wireless acoustic sensor networks (WASNs) and reducing the amount of information that is to be transmitted is one effective way to save it. In previous contributions, we presented sensor selection as well as rate distribution methods to reduce the power usage of beamforming algorithms in WASNs. Taking only transmission power into account, it was shown that rate distribution is a generalization of sensor selection and that rate distribution is more efficient than sensor selection with respect to the power usage versus performance trade-off. However, this excludes the energy consumption that it takes to keep the WASN nodes activated. In this paper, we present a more detailed comparison between sensor selection and rate-allocation by taking also into account the power to keep sensors activated for centralized WASNs. The framework is formulated by minimizing the total power usage, while lower bounding the noise reduction performance. Numerical results show that whether rate distribution is more efficient than sensor selection depends on the actual power that is used to keep sensors activated.**

***Index Terms*—Wireless acoustic sensor networks, beamforming, sensor selection, rate distribution, energy consumption.**

I. Introduction

Power usage is a vital factor for the applications in wireless acoustic sensor networks (WASNs), e.g., beamforming based noise reduction, since usually the sensors are battery-driven and have limited power budget [1]. It is desirable that the tasks at hand consume as little energy as possible, such that the network lifetime can be maximized. In this work, we focus on the centralized WASNs, which exploit a fusion center (FC) to gather and process the data. The sensor nodes are connected with the FC via wireless links and transmit their recordings to the FC at a certain communication rate. Obviously, the actual rate influences the performance that can be achieved; lowering the rate will increase the amount of quantization noise.

Sensor selection [2]–[6] and *rate allocation* [7]–[11] are two often-used strategies to reduce the power usage for WASNs. Sensor selection chooses a subset of sensors by optimizing a certain performance criterion, while constraining the cardinality of the selected subset, or the other way around. As the sensor measurements in a WASN are usually highly correlated, there is often no need to gather all data from the complete network to achieve a certain prescribed performance. Hence, with sensor selection the most informative subset of sensors that should be involved can be determined, leading to reduced data transmission. On the other hand, since the wireless transmission cost is directly related to the bit-rate, another strategy is to distribute the transmission bits to sensors such that a certain prescribed performance is obtained. The difference between these two techniques lies in the hard binary decision to select a node versus the soft decision by assigning a certain amount of bits for transmission. In other words, the sensors selected by the sensor selection algorithms use either zero or maximum rate to communicate with the FC, while the rate allocation methods allow the sensors to use any possible rate between zero and the maximum value.

In [10], we proposed a rate-distributed linearly constrained minimum variance (LCMV) beamformer based noise reduction algorithm for energy-aware WASNs. The problem was formulated by minimizing the total transmission cost between the sensors and the FC under a constraint on the desired noise reduction performance. The optimal rate distribution over the network was shown to lead to a reduced power consumption compared to sensor selection. Moreover, by representing rate distribution as a Boolean optimization problem, it was shown that sensor selection is a special case of rate allocation. Although it followed from [10] that rate-allocation is always more efficient than sensor selection in terms of the power usage versus performance trade-off, this strongly depends on how power usage is defined. In [10], only the power consumption for transmission was taken into account, while the sensor self-cost, i.e., the power needed to keep sensors activated (even if they transmit at a very low rate), was neglected. Taking the power to keep nodes activated into account is expected to change the outcome of the comparison between rate allocation and sensor selection and will be investigated in this paper. Particularly, we generalize the framework in [10] by taking the sensor self-cost into account. The proposed framework is formulated by optimizing the trade-off between the total power usage and the noise reduction capability. Using an LCMV beamformer, the problem is formulated as a semi-definite program for both sensor selection and rate allocation.

II. Fundamentals

A. Signal model

We consider a WASN consisting of K microphones that monitor the sound field. We assume there is one target source and several interferers[1]. In the short-time Fourier transfer (STFT) domain, let l and ω denote the time-frame index and

[1]As will be shown in the experimental results, the proposed methods are also validated when multiple target sources are present. Without loss of generalization, we stick here to the case of one target source.

978-9-0827-9703-9/19/$31.00

the angular frequency, respectively. The noisy and transmitted STFT coefficient that is received by the FC from the kth microphone, say $Y_k(\omega, l)$, is then given by

$$Y_k(\omega, l) = X_k(\omega, l) + N_k(\omega, l), k = 1, 2, \cdots, K, \quad (1)$$

where $X_k(\omega, l) = a_k(\omega)S(\omega, l)$ with $a_k(\omega)$ the acoustic transfer function (ATF) of the target source from the source location to the kth microphone and $S(\omega, l)$ the STFT coefficient of the target source at its original location. In (1), the total noise (including interferers) received by the FC is denoted by $N_k(\omega, l)$, which is given by

$$N_k(\omega, l) = Z_k(\omega, l) + Q_k(\omega, l), \quad (2)$$

where $Z_k(\omega, l)$ denotes the total received acoustic noise by the kth microphone (including correlated noise sources and uncorrelated noise) and $Q_k(\omega, l)$ denotes the quantization noise introduced by representing the data at a low communication rate.[2] For notational convenience, the frequency variable ω and the frame index l will be omitted taking into account that all processing takes place per time frame and per frequency band. We use vector notation and stack the quantized signals from K microphones in a vector $\mathbf{y} = [Y_1, ..., Y_K]^T \in \mathbb{C}^K$. Similarly, we define K dimensional vectors $\mathbf{a}$, $\mathbf{x}$, $\mathbf{n}$, $\mathbf{z}$, and $\mathbf{q}$ for the ATFs, the target speech component, the received noises by the FC, the acoustic noise and the quantization noise, respectively, such that (1) can be rewritten as

$$\mathbf{y} = \mathbf{a}S + \mathbf{z} + \mathbf{q}, \quad (3)$$

where $\mathbf{x} = \mathbf{a}S$ and $\mathbf{n} = \mathbf{z} + \mathbf{q}$. Assuming that all sources are mutually uncorrelated, the correlation matrix of the received signals is given by [10]

$$\mathbf{R_{yy}} = \mathbb{E}\{\mathbf{y}\mathbf{y}^H\} = \mathbf{R_{xx}} + \underbrace{\mathbf{R_{zz}} + \mathbf{R_{qq}}}_{\mathbf{R_{nn}}} \in \mathbb{C}^{K \times K}, \quad (4)$$

where $\mathbb{E}\{\cdot\}$ denotes mathematical expectation. In case the sensors utilize uniform quantizers to quantize their recordings, the correlation matrix of the quantization noise $\mathbf{R_{qq}}$ is given by [9], [10], [12]

$$\mathbf{R_{qq}} = \frac{1}{12}\mathrm{diag}\left(\left[\frac{\mathcal{A}_1^2}{4^{b_1}}, \frac{\mathcal{A}_2^2}{4^{b_2}}, ..., \frac{\mathcal{A}_K^2}{4^{b_K}}\right]\right), \quad (5)$$

where $\mathcal{A}_k = \max\{|Y_k|\}$ and b_k denotes the number of bits used by the kth microphone node. In this work, we assume that the ATFs of the target sources are known.

B. LCMV beamforming

The LCMV beamformer is widely used in acoustic array processing. It is formulated by minimizing the noise power subject to a number of linear constraints, resulting in the following constrained problem formulation [13]–[16]

$$\hat{\mathbf{w}}_{\mathrm{LCMV}} = \arg\min_{\mathbf{w}} \ \mathbf{w}^H\mathbf{R_{nn}}\mathbf{w}, \quad \text{s.t.} \quad \mathbf{\Lambda}^H\mathbf{w} = \mathbf{f}, \quad (6)$$

where $\mathbf{f} = [f_1, \cdots, f_{\mathcal{U}}]^T \in \mathbb{C}^{\mathcal{U}}$ and $\mathbf{\Lambda} \in \mathbb{C}^{\mathcal{U} \times K}$ with $\mathcal{U}$ denoting the number of equality constraints. The closed-form solution to (6) is given by [13]–[16]

$$\hat{\mathbf{w}}_{\mathrm{LCMV}} = \mathbf{R}_{\mathbf{nn}}^{-1}\mathbf{\Lambda}(\mathbf{\Lambda}^H\mathbf{R}_{\mathbf{n+q}}^{-1}\mathbf{\Lambda})^{-1}\mathbf{f}, \quad (7)$$

resulting in an output noise power [16]

$$\hat{\mathbf{w}}^H\mathbf{R_{nn}}\hat{\mathbf{w}} = \mathbf{f}^H(\mathbf{\Lambda}^H\mathbf{R}_{\mathbf{nn}}^{-1}\mathbf{\Lambda})^{-1}\mathbf{f}. \quad (8)$$

The linear constraints in (6) can be used to preserve sources or cancel interferers by specifying the matrix $\mathbf{\Lambda}$ and vector $\mathbf{f}$. In binaural applications [11], [17]–[19], LCMV beamforming can also be used to preserve certain interaural relations (e.g., spatial cue preservation).

III. Proposed framework

A. Problem formulation

Using Shannon's channel capacity theorem [20], the power needed to transmit data from microphone k to the FC for a single time-frequency bin is given by [21]–[23]

$$E_k = d_k^2 V_k(4^{b_k} - 1), \quad (9)$$

where d_k denotes the transmission distance and V_k the power spectral density (PSD) of the channel noise.

In this paper, our focus is on reducing the transmission costs between sensors and the FC, while reaching a prescribed noise reduction performance. The transmission costs are reduced by utilizing rate distribution or sensor selection strategies, and noise reduction is performed using LCMV beamforming. The initial problem is formulated by minimizing the total power usage, which consists of the sum of the transmission powers over all sensors, and the power that is needed to keep the transmitting devices activated, subject to a constraint on the noise reduction performance. That is,

$$\begin{aligned} \min_{\mathbf{w},\mathbf{b}} \quad & g(\mathbf{b}) = \sum_{k=1}^{K} d_k^2 V_k(4^{b_k} - 1) + c_0||\mathbf{b}||_0 \\ \text{s.t.} \quad & \mathbf{w}^H\mathbf{R_{nn}}\mathbf{w} \le \frac{\beta}{\alpha} \\ & \mathbf{\Lambda}^H\mathbf{w} = \mathbf{f}, \quad b_k \in \mathbb{Z}_+, b_k \le b_0, \forall k, \end{aligned} \quad \text{(P1)}$$

where β denotes the output noise power that is achieved if all sensors are activated at the maximum rate, $\alpha \in (0, 1]$ is to control a certain desired performance compared to β, $\mathbb{Z}_+$ denotes a non-negative integer set, b_0 the maximum bit-rate per sample, and c_0 is the cost for having a sensor activated. The ℓ_0-(quasi) norm of $\mathbf{b}$ refers to the number of non-zero entries in $\mathbf{b}$, i.e., $||\mathbf{b}||_0 := |\{k : b_k \neq 0\}|$. Note that here we assume that the sensors are homogeneous without loss of generality, i.e., their self-costs are the same. Notice that the output noise power $\mathbf{w}^H\mathbf{R_{nn}}\mathbf{w}$ depends on $\mathbf{b}$ via the noise correlation matrix $\mathbf{R_{nn}}$. The general form in (P1) is built from the rate distribution point of view, yet it can easily be extended to the context of sensor selection, which will be shown in Sec. 3.3. Notice that (P1) differs from the one presented in [10] as now the power to keep sensors activated is also taken into account. This might change the bit distribution, depending on the value of c_0. It might lead to situations where instead of a small number of bits, it is preferred to allocate no bits to a particular device in order to save the power for keeping that device active. By solving (P1), one obtains the optimal rate

[2] In practice, the microphone measurements are already quantized by the analog-to-digital converter before being transmitted to the FC. In this case, $Q_k(\omega, l)$ indicates the error caused by resampling the microphone measurements at a lower rate.

distribution that the microphone nodes should use to quantize their measurements before transmission, such that a certain specified noise-reduction performance is guaranteed. Solving (P1) using an exhaustive search is intractable unless b_0 and/or K are very small, as this offers $(b_0+1)^K$ choices. Next, we will solve (P1) using convex optimization techniques.

B. Rate-distributed LCMV beamforming

Considering the Lagrangian function of (P1), it can easily be verified that the LCMV beamformer in (7) is a candidate solution. Hence, we can substitute (7) into (P1), such that (P1) can be simplified by leaving out the linear constraints, resulting in the following optimization problem:

$$\begin{aligned} \min_{\mathbf{b}} \quad & g_1(\mathbf{b}) = \sum_{k=1}^{K} d_k^2 V_k (4^{b_k}-1) + c_0 ||\mathbf{b}||_0 \\ \text{s.t.} \quad & \mathbf{f}^H (\boldsymbol{\Lambda}^H \mathbf{R}_{\mathbf{nn}}^{-1} \boldsymbol{\Lambda})^{-1} \mathbf{f} \leq \frac{\beta}{\alpha} \\ & b_k \in \mathbb{Z}_+, \quad b_k \leq b_0, \forall k, \end{aligned} \tag{P2}$$

which is clearly non-linear and non-convex in $\mathbf{b}$. Note that the unknowns $\mathbf{b}$ are implicit in the term $\mathbf{f}^H(\boldsymbol{\Lambda}^H\mathbf{R}_{\mathbf{nn}}^{-1}\boldsymbol{\Lambda})^{-1}\mathbf{f}$ through $\mathbf{R}_{\mathbf{nn}}$. We therefore need to split $\mathbf{b}$ out from $\mathbf{f}^H(\boldsymbol{\Lambda}^H\mathbf{R}_{\mathbf{nn}}^{-1}\boldsymbol{\Lambda})^{-1}\mathbf{f}$. To do this, we introduce a Hermitian positive definite matrix $\mathbf{G} \in \mathbb{S}_{++}^{\mathcal{U}}$ with $\mathbb{S}_{++}^{\mathcal{U}}$ denoting a set for Hermitian positive definite matrices of dimension $\mathcal{U} \times \mathcal{U}$, such that the first inequality constraint in (P2) can equivalently be written as the following two constraints:

$$\boldsymbol{\Lambda}^H \mathbf{R}_{\mathbf{nn}}^{-1} \boldsymbol{\Lambda} = \mathbf{G}, \tag{10}$$

$$\mathbf{f}^H \mathbf{G}^{-1} \mathbf{f} \leq \frac{\beta}{\alpha}, \tag{11}$$

where (10) can be expressed as a linear matrix inequality (LMI) using the Schur complement [24, p.650], i.e.,

$$\begin{bmatrix} \mathbf{G} & \mathbf{f} \\ \mathbf{f}^H & \frac{\beta}{\alpha} \end{bmatrix} \succeq \mathbf{O}_{\mathcal{U}+1}. \tag{12}$$

As the equality constraint in (10) is still non-convex as a function of $\mathbf{b}$, it can be relaxed to

$$\boldsymbol{\Lambda}^H \mathbf{R}_{\mathbf{nn}}^{-1} \boldsymbol{\Lambda} \succeq \mathbf{G}. \tag{13}$$

In order to linearize (13) in $\mathbf{b}$, we calculate $\mathbf{R}_{\mathbf{nn}}^{-1}$ as

$$\begin{aligned} \mathbf{R}_{\mathbf{nn}}^{-1} &\overset{(a)}{=} \left(\mathbf{R}_{\mathbf{zz}} + \mathbf{R}_{\mathbf{qq}}\right)^{-1} \\ &\overset{(b)}{=} \mathbf{R}_{\mathbf{zz}}^{-1} - \mathbf{R}_{\mathbf{zz}}^{-1}\left(\mathbf{R}_{\mathbf{zz}}^{-1} + \mathbf{R}_{\mathbf{qq}}^{-1}\right)^{-1}\mathbf{R}_{\mathbf{zz}}^{-1}, \end{aligned} \tag{14}$$

where (b) is derived from the matrix inversion lemma [25, p.18][3]. Substituting $\mathbf{R}_{\mathbf{nn}}^{-1}$ from (14) into (13), we obtain

$$\boldsymbol{\Lambda}^H\mathbf{R}_{\mathbf{zz}}^{-1}\boldsymbol{\Lambda} - \mathbf{G} \succeq \boldsymbol{\Lambda}^H\mathbf{R}_{\mathbf{zz}}^{-1}\left(\mathbf{R}_{\mathbf{zz}}^{-1} + \mathbf{R}_{\mathbf{qq}}^{-1}\right)^{-1}\mathbf{R}_{\mathbf{zz}}^{-1}\boldsymbol{\Lambda}. \tag{15}$$

Using the Schur complement, we obtain the following LMI

$$\begin{bmatrix} \mathbf{R}_{\mathbf{zz}}^{-1} + \mathbf{R}_{\mathbf{qq}}^{-1} & \mathbf{R}_{\mathbf{zz}}^{-1}\boldsymbol{\Lambda} \\ \boldsymbol{\Lambda}^H\mathbf{R}_{\mathbf{zz}}^{-1} & \boldsymbol{\Lambda}^H\mathbf{R}_{\mathbf{zz}}^{-1}\boldsymbol{\Lambda} - \mathbf{G} \end{bmatrix} \succeq \mathbf{O}_{K+\mathcal{U}}, \tag{16}$$

where $\mathbf{R}_{\mathbf{qq}}^{-1}$ can be computed from (5) as

$$\begin{aligned} \mathbf{R}_{\mathbf{qq}}^{-1} &\overset{(a)}{=} 12\text{diag}\left(\left[\frac{4^{b_1}}{\mathcal{A}_1^2}, \frac{4^{b_2}}{\mathcal{A}_2^2}, ..., \frac{4^{b_K}}{\mathcal{A}_K^2}\right]\right) \\ &\overset{(b)}{=} \text{diag}\left(\mathbf{e} \odot (\mathbf{t} + \mathbf{1}_K)\right), \end{aligned} \tag{17}$$

[3] $(\mathbf{A}+\mathbf{CBC}^T)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{C}(\mathbf{B}^{-1}+\mathbf{C}^T\mathbf{A}^{-1}\mathbf{C})^{-1}\mathbf{C}^T\mathbf{A}^{-1}$.

where $\mathbf{1}_K$ denotes an all-ones column vector, $\odot$ is Hadamard product, and (b) is obtained by using the variable change $t_k = 4^{b_k} - 1 \in \mathbb{Z}, \forall k$ and defining a constant vector $\mathbf{e} = [\frac{12}{\mathcal{A}_1^2}, \cdots, \frac{12}{\mathcal{A}_K^2}]$, such that $\mathbf{R}_{\mathbf{qq}}^{-1}$ and (16) are both linear in $\mathbf{t}$. Note that $||\mathbf{b}||_0 = ||\mathbf{t}||_0$. For the non-convex integer constraint $b_k \in \mathbb{Z}_+, \forall k$, we relax it to $b_k \in \mathbb{R}_+$, i.e., $t_k \in \mathbb{R}_+$. In addition, the non-convex $||\mathbf{b}||_0$ will be relaxed to $||\mathbf{b}||_1$. Due to the fact that $b_k = \log_4(t_k+1) \approx t_k/\ln 4$ using the first-order Taylor expansion of the function $\log_4(t_k+1)$, we have $||\mathbf{b}||_1 \approx ||\mathbf{t}||_1/\ln 4$. Altogether, we arrive at the following standard semi-definite programming problem:

$$\min_{\mathbf{t},\mathbf{G}} \quad g_1(\mathbf{b}) = \sum_{k=1}^{K} d_k^2 V_k t_k + \frac{c_0}{\ln 4}||\mathbf{t}||_1 \tag{18}$$

$$\text{s.t.} \quad \begin{bmatrix} \mathbf{G} & \mathbf{f} \\ \mathbf{f}^H & \frac{\beta}{\alpha} \end{bmatrix} \succeq \mathbf{O}_{\mathcal{U}+1}, \tag{18a}$$

$$\begin{bmatrix} \mathbf{R}_{\mathbf{zz}}^{-1} + \mathbf{R}_{\mathbf{qq}}^{-1} & \mathbf{R}_{\mathbf{zz}}^{-1}\boldsymbol{\Lambda} \\ \boldsymbol{\Lambda}^H\mathbf{R}_{\mathbf{zz}}^{-1} & \boldsymbol{\Lambda}^H\mathbf{R}_{\mathbf{zz}}^{-1}\boldsymbol{\Lambda} - \mathbf{G} \end{bmatrix} \succeq \mathbf{O}_{K+\mathcal{U}}, \tag{18b}$$

$$0 \leq t_k \leq 4^{b_0} - 1, \quad \forall k, \tag{18c}$$

which can be solved in polynomial time using standard solvers, like CVX [26]. After (18) is solved, the rate distribution can be resolved by $b_k = \log_4(t_k+1), \forall k$. As these are continuous values, they need to be post-processed by randomized rounding [5], [6], [10] to obtain the final integer solution. Note that due to relaxing $||\mathbf{t}||_0$ to $||\mathbf{t}||_1/\ln 4$, the actual power usage of the rate allocation method will always be lower than the cost function in (18).

C. Sensor selection based LCMV beamforming

In [10], it was shown that sensor selection is a special case of rate distribution. Based on the rate distribution problem in (P2), we will reformulate it for sensor selection based LCMV beamforming. To do so, we represent (18) from the perspective of Boolean optimization. Using the variable change $p_k = t_k/(4^{b_0}-1)$ in (18), we can obtain an equivalent optimization problem, given by

$$\min_{\mathbf{p},\mathbf{G}} \quad g_2(\mathbf{p}) = \left(4^{b_0}-1\right)\sum_{k=1}^{K} p_k d_k^2 V_k + c_0||\mathbf{p}||_1 \tag{19}$$

$$\text{s.t.} \quad \begin{bmatrix} \mathbf{G} & \mathbf{f} \\ \mathbf{f}^H & \frac{\beta}{\alpha} \end{bmatrix} \succeq \mathbf{O}_{\mathcal{U}+1}, \tag{19a}$$

$$\begin{bmatrix} \mathbf{R}_{\mathbf{zz}}^{-1} + \mathbf{R}_{\mathbf{qq}}^{-1} & \mathbf{R}_{\mathbf{zz}}^{-1}\boldsymbol{\Lambda} \\ \boldsymbol{\Lambda}^H\mathbf{R}_{\mathbf{zz}}^{-1} & \boldsymbol{\Lambda}^H\mathbf{R}_{\mathbf{zz}}^{-1}\boldsymbol{\Lambda} - \mathbf{G} \end{bmatrix} \succeq \mathbf{O}_{K+\mathcal{U}}, \tag{19b}$$

$$0 \leq p_k \leq 1, \forall k, \tag{19c}$$

where $\mathbf{R}_{\mathbf{qq}}^{-1} = \text{diag}[\mathbf{e} \odot ((4^{b_0}-1)\mathbf{p} + \mathbf{1}_K)]$. The box constraint $0 \leq p_k \leq 1$ can be regarded as the relaxation of the Boolean constraint, i.e., $p_k \in \{0,1\}, \forall k$. The cost function in (19) can be interpreted as each selected sensor quantizing at the maximum rate b_0, i.e., with transmission power $d_k^2 V_k (4^{b_0}-1)$. Given the solution of (19), the rates to be allocated can be resolved by $b_k = \log_4[p_k(4^{b_0}-1)+1]$.

Strictly speaking, (18) and (19) are two equivalent optimization problems with the same computational complexity. From

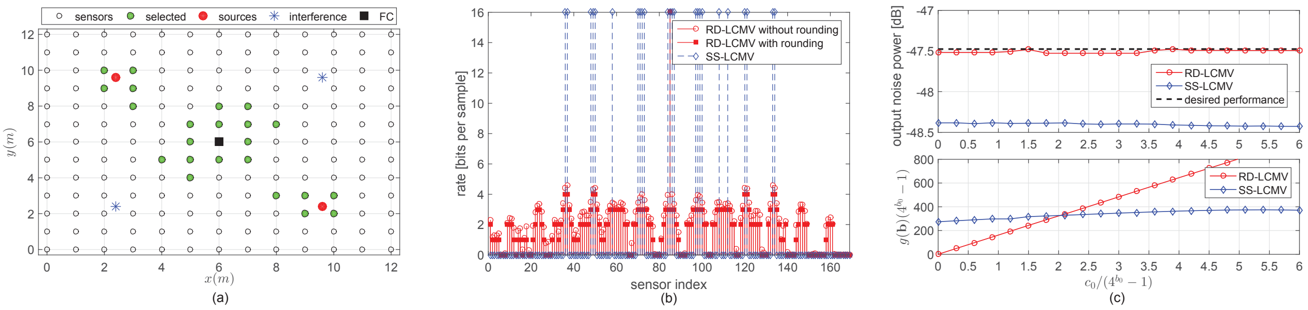

Figure 1. (a) The experimental setup, where the green sensors are selected by the sensor selection method with $\alpha = 0.6$, (b) an example for rate distribution with $\alpha = 0.6$, (c) noise reduction performance and energy usage in terms of the normalized microphone self-cost with $\alpha = 0.6$.

either problem, we can resolve the optimal rate distribution. Due to the fact that (19) is built from the perspective of Boolean optimization, the optimization variable p_k can indicate whether sensor k is selected or not. Hence, (19) can also perform sensor selection, and we therefore call it sensor selection based LCMV beamforming.

IV. Comparison in terms of power usage

In this section, we present an informal metric to compare the power usage of the two approaches. Let $\mathcal{S}_1$ denote the subset of microphones that are activated by the rate allocation algorithm and $\mathcal{S}_2$ the subset of microphones that are selected by the sensor selection method. In [10], it was shown that in general $|\mathcal{S}_1| \geq |\mathcal{S}_2|$, and the sensors in $\mathcal{S}_1$ are allocated with a much lower rate than the maximum rate b_0. For the power usage we thus have from (18) and (19) that

$$g_1(\mathbf{b}) = \sum_{k=1}^{K} d_k^2 V_k (4^{b_k} - 1) + c_0 |\mathcal{S}_1|, \tag{20}$$

$$g_2(\mathbf{p}) = P_0 \sum_{k=1}^{K} p_k d_k^2 V_k + c_0 |\mathcal{S}_2|, \tag{21}$$

where $||\mathbf{b}||_0 = |\mathcal{S}_1|$, $||\mathbf{p}||_0 = |\mathcal{S}_2|$ and $P_0 = 4^{b_0} - 1$. Further, the difference between the power usage of the two approaches can be calculated as

$$\begin{aligned} \Delta g &= g_1(\mathbf{b}) - g_2(\mathbf{p}) \qquad (22) \\ &= \sum_{k=1}^{K} d_k^2 V_k (4^{b_k} - 1) - P_0 \sum_{k \in \mathcal{S}_2} d_k^2 V_k + c_0 \left(|\mathcal{S}_1| - |\mathcal{S}_2|\right) \\ &= P_0 \Big(\sum_{k=1}^{K} d_k^2 V_k \frac{4^{b_k} - 1}{4^{b_0} - 1} - \sum_{k \in \mathcal{S}_2} d_k^2 V_k + \frac{c_0 \left(|\mathcal{S}_1| - |\mathcal{S}_2|\right)}{4^{b_0} - 1} \Big) \\ &\approx P_0 \left(\frac{c_0 \left(|\mathcal{S}_1| - |\mathcal{S}_2|\right)}{4^{b_0} - 1} - \sum_{k \in \mathcal{S}_2} d_k^2 V_k \right), \end{aligned}$$

where the approximation $\approx$ is due to the fact that $\frac{4^{b_k}-1}{4^{b_0}-1} \ll 1$ with $b_k < b_0$. To make rate allocation more efficient in terms of power usage, i.e., $\Delta g < 0$, the power c_0 to keep the sensor activated needs to satisfy the following inequality condition:

$$\frac{c_0}{P_0} < \frac{\sum_{k \in \mathcal{S}_2} d_k^2 V_k}{|\mathcal{S}_1| - |\mathcal{S}_2|} = \eta, \tag{23}$$

where the right-hand side only depends on the activated sets of the two approaches. In other words, given the same noise reduction performance, if c_0 is high, the sensor selection method is cheaper in power usage. Otherwise the rate allocation method is cheaper in terms of power consumption.

V. Simulations

Fig. 1(a) shows the simulation setup, where $K = 169$ microphones are placed uniformly in a 2D room with dimensions (12×12) m. We consider two target point sources (red solid circles) at $(2.4, 9.6)$ m and $(9.6, 2.4)$ m, respectively. Two interfering point sources (blue stars) are located at $(2.4, 2.4)$ m and $(9.6, 9.6)$ m, respectively. We thus have $\mathcal{U} = 2$. The FC (black solid square) is placed at $(6, 6)$ m in the center of the room. The target speech signals originate from the TIMIT database [27]. The interferers are stationary Gaussian speech shaped noise sources. The uncorrelated noise is modeled as sensor noise at an SNR of 50 dB. The signal-to-interferer ratio is set to be 0 dB. The sampling frequency is set to be 16 kHz. We use a square-root-Hann window of 20 ms for framing with 50% overlap. The ATFs are generated using [28] with reverberation time $T_{60} = 200$ ms. In addition, we assume that all the sensors are synchronized. The channel noise PSD is assumed to be the same for all sensors, e.g., $V_k = 1, \forall k$ without loss of generality.

Fig. 1(b) shows an example of rate distributions that are obtained by the rate allocation method (referred to as RD-LCMV) and the sensor selection method (referred to as SS-LCMV) for $\alpha = 0.6$. The maximum rate is set to be $b_0 = 16$ bits per sample, and the selected sensors of the SS-LCMV method quantize their measurements at the maximum rate. As an example, the power needed to keep a sensor activated is set at $c_0 = 4^{b_0} - 1$. We can see that for the RD-LCMV method, the sensors that are close to the target sources and the FC are distributed with higher rate (still much lower than the maximum rate), since they have higher SNR and less transmission cost. Fig. 1(a) shows an example of sensor selection by the SS-LCMV method for $\alpha = 0.6$. Clearly, the sensors that are close to the target sources and the FC are more likely to be selected due to the reasons stated before.

Fig. 1(c) compares the two approaches in terms of the power c_0. Observing the output noise power in the top plot, we can

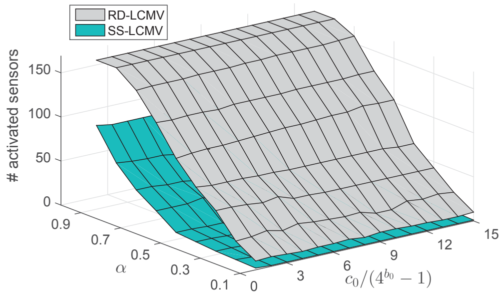

Figure 2. The cardinality of the activated sets of the two approaches in terms of α and sensor self-cost.

conclude that with the power c_0 being taken into account, both methods still satisfy the performance requirement, but the rate distribution method is closer to the desired performance. Further, observing the energy usage in the bottom plot, it is clear that when c_0 is small, the rate allocation method consumes less energy; when it is large, the sensor selection method is cheaper. For the setup in Fig. 1(a) where $|\mathcal{S}_1| = 124$ and $|\mathcal{S}_2| = 24$, we can calculate from (23) that $\eta = \frac{\sum_{k\in\mathcal{S}_2} d_k^2 V_k}{|\mathcal{S}_1|-|\mathcal{S}_2|} \approx 2.2$. This also validates our theoretical analysis in Sec. IV.

Furthermore, in Fig. 2 we show the number of the activated sensors for the two approaches in terms of α and the power c_0. Clearly, the rate allocation method always needs to activate more sensors than the sensor selection method. For both methods, more sensors need to be activated when α is large.

VI. Concluding Remarks

In this paper, we presented a comparison between sensor selection and rate allocation for beamforming in WASNs, taking into account also the power that it takes to keep sensors activated. The problem was formulated by minimizing the total power usage and constraining the noise reduction performance. By taking into account the power to keep sensors activated, the sensor selection and rate allocation methods could be compared in a fair way. Both theoretical analysis and simulation results showed that given a pre-defined noise reduction performance, it depends on the actual cost to keep sensors activated whether sensor selection or bit-rate allocation is more efficient in terms of power consumption.

References

[1] A. Bertrand, "Applications and trends in wireless acoustic sensor networks: a signal processing perspective," in *IEEE Symposium on Communications and Vehicular Technology in the Benelux (SCVT)*, 2011, pp. 1–6.

[2] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 451–462, 2009.

[3] A. Bertrand and M. Moonen, "Efficient sensor subset selection and link failure response for linear mmse signal estimation in wireless sensor networks," in *EURASIP Europ. Signal Process. Conf. (EUSIPCO)*, 2010, pp. 1092–1096.

[4] J. Szurley, A. Bertrand, M. Moonen, P. Ruckebusch, and I. Moerman, "Energy aware greedy subset selection for speech enhancement in wireless acoustic sensor networks," in *EURASIP Europ. Signal Process. Conf. (EUSIPCO)*, 2012, pp. 789–793.

[5] S. P. Chepuri and G. Leus, "Sparsity-promoting sensor selection for non-linear measurement models," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 684–698, 2015.

[6] J. Zhang, S. P. Chepuri, R. C. Hendriks, and R. Heusdens, "Microphone subset selection for MVDR beamformer based noise reduction," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 26, no. 3, pp. 550–563, 2018.

[7] O. Roy and M. Vetterli, "Rate-constrained collaborative noise reduction for wireless hearing aids," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 645–657, 2009.

[8] J. Amini, R. C. Hendriks, R. Heusdens, M. Guo, and J. Jensen, "On the impact of quantization on binaural MVDR beamforming," in *12th ITG Symposium of Speech Communication*. VDE, 2016, pp. 1–5.

[9] F. de la Hucha Arce, F. Rosas, M. Moonen, M. Verhelst, and A. Bertrand, "Generalized signal utility for LMMSE signal estimation with application to greedy quantization in wireless sensor networks," *IEEE Signal Process. Lett.*, vol. 23, no. 9, pp. 1202–1206, 2016.

[10] J. Zhang, R. Heusdens, and R. C. Hendriks, "Rate-distributed spatial filtering based noise reduction in wireless acoustic sensor networks," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 26, no. 11, pp. 2015–2026, 2018.

[11] J. Zhang, R. Heusdens, and R. C. Hendriks, "Rate-distributed binaural LCMV beamforming for assistive hearing in wireless acoustic sensor networks," in *IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2018.

[12] R. M. Gray, "Quantization noise spectra," *IEEE Trans. Information Theory*, vol. 36, no. 6, pp. 1220–1244, 1990.

[13] O. Lamont Frost III, "An algorithm for linearly constrained adaptive array processing," *Proceedings of the IEEE*, vol. 60, no. 8, pp. 926–935, 1972.

[14] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Signal Process. Mag.*, vol. 5, no. 2, pp. 4–24, 1988.

[15] J. Benesty, S. Makino, and J. Chen, *Speech enhancement*, Springer Science & Business Media, 2005.

[16] M. Souden, J. Benesty, and S. Affes, "A study of the LCMV and MVDR noise reduction filters," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4925–4935, 2010.

[17] A. I. Koutrouvelis, R. C. Hendriks, R. Heusdens, and J. Jensen, "Relaxed binaural LCMV beamforming," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 25, no. 1, pp. 137–152, 2017.

[18] E. Hadad, S. Doclo, and S. Gannot, "The binaural LCMV beamformer and its performance analysis," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 24, no. 3, pp. 543–558, 2016.

[19] E. Hadad, D. Marquardt, S. Doclo, and S. Gannot, "Theoretical analysis of binaural transfer function MVDR beamformers with interference cue preservation constraints," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 23, no. 12, pp. 2449–2464, 2015.

[20] C. E. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, 1949.

[21] S. Shah and B. Beferull-Lozano, "Adaptive quantization for multihop progressive estimation in wireless sensor networks," in *EURASIP Europ. Signal Process. Conf. (EUSIPCO)*. IEEE, 2013, pp. 1–5.

[22] Y. Huang and Y. Hua, "Multihop progressive decentralized estimation in wireless sensor networks," *IEEE Signal Process. Lett.*, vol. 14, no. 12, pp. 1004–1007, 2007.

[23] Y. Huang and Y. Hua, "Energy planning for progressive estimation in multihop sensor networks," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 4052–4065, 2009.

[24] S. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge university press, 2004.

[25] K. B. Petersen, M. S. Pedersen, et al., "The matrix cookbook," *Technical University of Denmark*, vol. 7, pp. 15, 2008.

[26] M. Grant, S. Boyd, and Y. Ye, "CVX: Matlab software for disciplined convex programming," 2008.

[27] J. S. Garofolo, "DARPA TIMIT acoustic-phonetic speech database," *National Institute of Standards and Technology (NIST)*, vol. 15, pp. 29–50, 1988.

[28] E. A. P. Habets, "Room impulse response generator," *Technische Universiteit Eindhoven, Tech. Rep*, vol. 2, no. 2.4, pp. 1, 2006.