DISTRIBUTED NODE-SPECIFIC DIRECTION-OF-ARRIVAL ESTIMATION IN WIRELESS ACOUSTIC SENSOR NETWORKS

Amin Hassani[*,†], *Alexander Bertrand*[*,†], *Marc Moonen*[*,†]

∗ KU Leuven, Dept. of Electrical Engineering-ESAT, SCD-SISTA / † iMinds Future Health Department
Address: Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
E-mail: amin.hassani@esat.kuleuven.be
alexander.bertrand@esat.kuleuven.be
marc.moonen@esat.kuleuven.be

ABSTRACT

In this paper, we study the effect of collaboration between nodes for direction of arrival (DOA) estimation in a full connected wireless acoustic sensor network (WASN) where the position of the nodes is unknown. Each node is equipped with a linear microphone array which defines a node-specific DOA with respect to a single common target speech source. We assume that the DOA estimation algorithm is operated in conjunction with a distributed noise reduction algorithm, referred to as the distributed adaptive node-specific signal estimation (DANSE) algorithm. To avoid additional data exchange between the nodes, the goal is to exploit the shared signals used in the DANSE algorithm to also improve the node-specific DOA estimation. The DOA estimation is based on the multiple signal classification (MUSIC) algorithm (if sufficient computing power is available), or a least-squares (LS) method based on a locally estimated steering vector which allows to eliminate the exhaustive search in MUSIC and results in a significantly lower computational complexity. Simulation results demonstrate that collaboration between nodes improves the performance of the DOA estimation compared to the case where the nodes operate individually, i.e. do not collaborate.

Index Terms— Direction-of-arrival estimation, wireless sensor networks, distributed estimation

1. INTRODUCTION

Microphone arrays allow one to exploit spatial information in acoustic signal processing applications. For example, they allow for localization of target sound sources, such as speech sources, as well as cancellation of undesired sound waves impinging on the array from certain directions other than the directions from which the target sound signals come [1]. Wireless acoustic sensor networks (WASNs) are an emerging technology in the field of multi-microphone acoustic signal processing. A WASN consists of a collection of nodes which are equipped with a local microphone array, a signal processing unit and wireless communication facilities. The nodes can then cooperate to solve certain acoustic signal processing tasks by exchanging relevant information amongst each other. In this paper, we assume that each node is equipped with a uniform linear array (ULA) of microphones which means that the microphones are placed on a single line with uniform spacing.

Considering the fact that sensor nodes are often battery-powered and have a limited computational capacity, WASN algorithms that need less communication and computational power are generally desired. To reach this goal, unlike centralized processing in which all the nodes send their raw observations to a central unit for processing, we consider here a distributed approach. The processing task is then spread over all the nodes and they exchange signals, possibly after a local in-node processing, among each other.

In this paper, node-specific direction of arrival (DOA) estimation in a fully connected WASN is addressed where the position of the nodes is unknown. By the term node-specific, we refer to the case in which each of the nodes estimates its own specific DOA for a single common target speech source. We assume that the DOA estimation is operated in conjunction with a distributed noise reduction algorithm. One application could be a video conferencing in which on top of the noise reduction for speech enhancement, we are also interested in steering each node's built-in camera towards the loca-

Acknowledgements : This work was carried out at the ESAT Laboratory of KU Leuven, in the frame of KU Leuven Research Council CoE EF/05/006 Optimization in Engineering (OPTEC) and PFV/10/002 (OPTEC), Concerted Research Action GOA-MaNet, the Belgian Programme on Interuniversity Attraction Poles initiated by the Belgian Federal Science Policy Office IUAP P7/19 (DYSCO, Dynamical systems, control and optimization, 2007-2011), and Research Project FWO nr. G.0763.12 (Wireless acoustic sensor networks for extended auditory communication'). The work of A. Bertrand was supported by a Postdoctoral Fellowship of the Research Foundation - Flanders (FWO). The scientic responsibility is assumed by its authors.

tion of the speaker. We will investigate whether the DOA estimation can benefit from using the signals that are exchanged between nodes within the distributed noise reduction algorithm that is already in place, and we will compare this to the case where the nodes operate individually, i.e. use only their own microphone signals for DOA estimation.

For the sake of this assessment, we use the distributed adaptive node-specific signal estimation (DANSE) algorithm [2], [3] for noise reduction. To avoid additional data exchange between the nodes, the goal is to exploit the shared signals used in the DANSE algorithm to also improve the node-specific DOA estimation. With DANSE in place, we estimate the node-specific speech correlation matrix in each node given the node's own microphone signals together with the compressed-and-filtered broadcast signals of the other nodes. Given the node-specific speech correlation matrix, the node-specific steering vector corresponding to the target speech source is extracted, from which the node-specific DOA is estimated based on either MUSIC or a simpler least-squares (LS) method. The main noticeable advantage of the proposed LS method over MUSIC is that it eliminates the need for an exhaustive search over all possible angles which could be problematic taking the limited available computational power into account. It is reiterated that the proposed approach is blind in that the position of the nodes is unknown, and only the local array geometry at the individual nodes is exploited.

The paper is organized as follows. The data model and problem statement are presented in section 2. In section 3, the LS method and MUSIC are explained for DOA estimation. Section 4 presents the proposed collaborative steering vector estimation and DANSE. Simulation results are presented in section 5 and the conclusions are drawn in section 6.

2. PROBLEM STATEMENT AND DATA MODEL

We consider a WASN with J nodes in which each node k has direct access to its own M_k microphones forming a ULA. The signal of microphone m at node k in the frequency domain can be decomposed as:

$$y_{km}(\omega_n) = s_{km}(\omega_n) + n_{km}(\omega_n) \tag{1}$$

where $s_{km}(\omega_n)$ and $n_{km}(\omega_n)$ are the target speech signal and undesired noise signal in microphone m of node k and $\omega_n = 2\pi n/L$ is the discrete frequency domain variable in which the resolution is defined by the discrete Fourier transform (DFT) of size L and and $n = 0, 1, \dots (L/2+1)$. By stacking (1) for $m = 1, \dots, M_k$, we obtain $\mathbf{y}_k(\omega_n) = [y_{k1}(\omega_n) \dots y_{kM_k}(\omega_n)]^T = \mathbf{s}_k(\omega_n) + \mathbf{n}_k(\omega_n)$. All the $\mathbf{y}_k(\omega_n)$'s are stacked in the full M-channel signal $\mathbf{y}(\omega_n) = [\mathbf{y}_1(\omega_n)^T \dots \mathbf{y}_J(\omega_n)^T]^T$ in which $M = \sum_{k=1}^{J} M_k$. Considering $s(\omega_n)$ as the target speech source signal propagated from a single source, we have $\mathbf{s}_k(\omega_n) = \mathbf{a}_k(\omega_n)s(\omega_n)$ in which $\mathbf{a}_k(\omega_n)$ is the node-specific M_k-dimensional steering vector. In general, $\mathbf{a}_k(\omega_n)$ is composed of the acoustic transfer functions (including room acoustics and microphone characteristics) from the target speech source to each microphone of node k. In a sufficiently large room with negligible reverberations, we can write $\mathbf{a}_k(\omega_n)$ as a function of the array manifold vector $\mathbf{g}_k(\omega_n, \theta_k)$ which expresses the phase shifts with respect to the first microphone of node k:

$$\mathbf{a}_k(\omega_n) = a_{k1}(\omega_n)\mathbf{g}_k(\omega_n, \theta_k) \tag{2}$$

$$= a_{k1}(\omega_n) \begin{bmatrix} 1 \\ e^{-j\omega_n d\, cos(\theta_k) F_s/c} \\ \vdots \\ e^{-j\omega_n (M_k-1) d\, cos(\theta_k) F_s/c} \end{bmatrix} \tag{3}$$

where θ_k is the DOA at node k, $a_{k1}(\omega_n)$ is the acoustic transfer function from the target speech source to the first microphone of node k, F_s is the sampling frequency, c is the speed of sound, and d is the distance between microphones at the node. For the sake of an easy exposition, the relative attenuation factors are neglected here, i.e., we assume that $\frac{|a_{km}(\omega_n)|}{|a_{k1}(\omega_n)|} = 1$ (this is without loss of generality, as we will only extract the phase information from (2)).

In this paper, we study the effect of collaboration between nodes on the performance of the node-specific DOA estimation in a WASN where the position of the nodes is unknown.

3. DOA ESTIMATION

This section describes two methods to extract the DOA from a given steering vector estimate $\bar{\mathbf{a}}_k(\omega_n)$. The proposed cooperative procedure for this estimation will be explained in section 4. As the steering vector can only be estimated up to an unknown scaling, we assume in the sequel that $\bar{\mathbf{a}}_k(\omega_n)$ is normalized with respect to its first entry, i.e., each element is divided by the first entry such that $\bar{a}_{k1}(\omega_n) = 1$.

3.1. Least Squares DOA estimator

We define $\mathbf{p}_k(\omega_n, \theta_k)$ as the absolute phase of the generic array manifold $\mathbf{g}_k(\omega_n, \theta_k)$ at node k, i.e.:

$$\begin{bmatrix} 0 \\ \frac{\omega_n d cos(\theta_k) F_s}{c} \\ \vdots \\ \frac{\omega_n (M_k-1) d cos(\theta_k) F_s}{c} \end{bmatrix} = \mathbf{p}_k(\omega_n, \theta_k). \tag{4}$$

Defining $x_k = cos(\theta_k)$, node k then estimates the x_k from an overdetermined set of equations which finally leads to solving a least square minimization problem with the following cost function:

$$\min_{x_k} \sum_{\omega_n} \|\mathbf{p}_{k_t}(\omega_n, \theta_k) - \bar{\mathbf{p}}_{k_t}(\omega_n)\|^2 \tag{5}$$

where subscript t denotes the truncated $\mathbf{p}_k$ in which the first row is omitted (because it is always zero) and $\bar{\mathbf{p}}_{k_t}$ is the corresponding truncated phase vector of $\bar{\mathbf{a}}_k(\omega_n)$. Stacking these variables for the different discrete frequencies yields

$$\mathbf{p}_{k_s} = [\mathbf{p}_{k_t}^T(\omega_1) \dots \mathbf{p}_{k_t}^T(\omega_{L/2+1}))]^T \tag{6}$$

$$\bar{\mathbf{p}}_{k_s} = [\bar{\mathbf{p}}_{k_t}^T(\omega_1) \dots \bar{\mathbf{p}}_{k_t}^T(\omega_{L/2+1}))]^T. \tag{7}$$

To correct the jumps in phase angles, phase unwrapping must be performed first. The solution of (5) is then given by:

$$\bar{x}_k = \frac{\mathbf{p}_{k_s}^T \bar{\mathbf{p}}_{k_s}}{\mathbf{p}_{k_s}^T \mathbf{p}_{k_s}} \tag{8}$$

and finally we compute the node-specific DOA as $\bar{\theta}_k = cos^{-1}(\bar{x}_k)$

3.2. MUSIC

MUSIC is one of the well-known high resolution algorithms for DOA estimation. MUSIC decomposes the speech correlation matrix into a signal and noise subspace which are orthogonal to each other. In the case of a single target speech source, the signal subspace is defined by the eigenvector corresponding to the largest eigenvalue of the speech correlation matrix, and the noise subspace can be constructed as the $(M_k - 1)$-dimensional subspace orthogonal to this signal subspace, e.g., using Gramm-Schmidt orthogonalization. The matrices containing the basis vectors of the signal and noise subspaces are then denoted as:

$$\mathbf{E}_{\mathbf{s}_k}(\omega_n) = [\mathbf{q}_1(\omega_n)] \tag{9}$$

$$\mathbf{E}_{\mathbf{n}_k}(\omega_n) = [\mathbf{q}_2(\omega_n)| \dots |\mathbf{q}_{M_k}(\omega_n)] \tag{10}$$

where $\mathbf{q}_1(\omega_n)$ is the eigenvector defining the signal subspace, and $\mathbf{E}_{\mathbf{n}_k}^H(\omega_n)\mathbf{q}_1(\omega_n) = \mathbf{0}$. An exhaustive search over all possible θ_k is performed, each of them yielding a different generic array manifold vector $\mathbf{g}_k(\omega_n, \theta_k)$. Merging all frequency-dependent DOA estimations can be performed by using one of the three available methods: arithmetic (used in this paper), geometric and harmonic averaging [4]. The θ_k for which the so-called MUSIC pseudospectrum [5] is maximized, will be the estimated DOA, i.e.,

$$\bar{\theta}_k = \arg\max_{\theta_k} \sum_{\omega_n} \frac{1}{\mathbf{g}_k^H(\omega_n, \theta_k)\mathbf{E}_{\mathbf{n}_k}(\omega_n)\mathbf{E}_{\mathbf{n}_k}^H(\omega_n)\mathbf{g}_k(\omega_n, \theta_k)}. \tag{11}$$

4. COLLABORATIVE STEERING VECTOR ESTIMATION USING DANSE

In this section we propose a collaborative approach to estimate each node-specific steering vector using the shared broadcast signals of DANSE together with each node's own signals. The first step is then to estimate the speech correlation matrix at each node. A complete block diagram of the scheme is illustrated in Figure 1. The correlation matrix of the target speech signal component of the microphone signals, $\mathbf{s}_k$, can be written as:

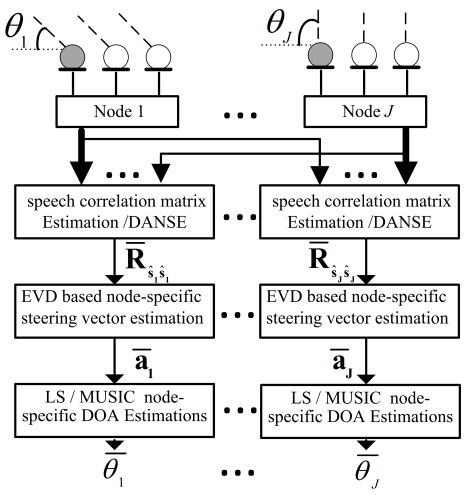

Fig. 1. The complete procedure for the proposed collaborative estimation of node-specific DOAs.

$$\mathbf{R}_{\mathbf{s_k s_k}}(\omega_n) = E\{\mathbf{s}_k(\omega_n)\mathbf{s}_k(\omega_n)^H\} = P_s(\omega_n)\mathbf{a}_k(\omega_n)\mathbf{a}_k(\omega_n)^H \tag{12}$$

where $P_s(\omega_n) = E\{|s(\omega_n)|^2\}$ is the power of the target speech signal, $E\{\cdots\}$ denotes the expected value operator, and the superscript H indicates the conjugate transpose operator. In general, $\mathbf{R}_{\mathbf{s_k s_k}}(\omega_n)$ is unknown and has to be estimated from the collected signal observations. In the rest of the paper, we use the hat superscript ($\hat{\cdot}$) as a case-dependent notation which will be defined later. With the assumption of uncorrelated $\hat{\mathbf{s}}_k(\omega_n)$ and $\hat{\mathbf{n}}_k(\omega_n)$, we have (we use an overline (bar) to denote an estimate):

$$\bar{\mathbf{R}}_{\hat{\mathbf{s}}_\mathbf{k}\hat{\mathbf{s}}_\mathbf{k}}(\omega_n) = \bar{\mathbf{R}}_{\hat{\mathbf{y}}_\mathbf{k}\hat{\mathbf{y}}_\mathbf{k}}(\omega_n) - \bar{\mathbf{R}}_{\hat{\mathbf{n}}_\mathbf{k}\hat{\mathbf{n}}_\mathbf{k}}(\omega_n) \tag{13}$$

where $\bar{\mathbf{R}}_{\hat{\mathbf{n}}_\mathbf{k}\hat{\mathbf{n}}_\mathbf{k}}(\omega_n) \approx E\{\hat{\mathbf{n}}_k(\omega_n)\hat{\mathbf{n}}_k(\omega_n)\}$ and $\bar{\mathbf{R}}_{\hat{\mathbf{y}}_\mathbf{k}\hat{\mathbf{y}}_\mathbf{k}}(\omega_n) \approx E\{\hat{\mathbf{y}}_k(\omega_n)\hat{\mathbf{y}}_k(\omega_n)\}$ which can be estimated during "noise-only" and "speech-and-noise" periods, respectively. To distinguish between "noise-only" and "speech-and-noise" periods, a voice activity detection (VAD) mechanism must be applied. Estimation of the second order signal statistics ($\mathbf{R}$ matrices) can be done by time averaging in the short-time-Fourier-transform (STFT) domain. In theory, $\mathbf{R}_{\hat{\mathbf{s}}_\mathbf{k}\hat{\mathbf{s}}_\mathbf{k}}(\omega_n)$ is a rank-1 matrix for a single target speech source. In practice, however, due to the finite DFT size in the STFT analysis, non stationarity of the noise and finite observation window (which leads to estimation errors), the rank of $\bar{\mathbf{R}}_{\hat{\mathbf{s}}_\mathbf{k}\hat{\mathbf{s}}_\mathbf{k}}(\omega_n)$ will be greater than one. Therefore, we should use a method for rank-1 approximation to extract the steering vector based on (12).

To evaluate the performance of the collaborative DOA estimation, it will be compared with 2 cases: namely the isolated case and the centralized case; in the first, each node has only access to its own observations, i.e. $\hat{\mathbf{y}}_k(\omega_n) = \mathbf{y}_k(\omega_n)$, whereas in the second each node has access to all M observations throughout the entire network, i.e. $\hat{\mathbf{y}}_k(\omega_n) = \mathbf{y}(\omega_n)$.

4.1. DANSE algorithm

For the purpose of noise reduction, node k then employs the multi-channel Wiener filter (MWF) [6] in which the filter coefficients $\mathbf{w}_k(\omega_n)$ are computed such that the following MSE cost function, taking the target speech signal component of the first microphone signal of node k as the desired signal, is minimized:

$$\min_{\mathbf{w}_k(\omega_n)} E\{\left|s_{k1}(\omega_n) - \mathbf{w}_k(\omega_n)^H \hat{\mathbf{y}}_k(\omega_n)\right|^2\}. \qquad (14)$$

The solution is then given by [6]:

$$\hat{\mathbf{w}}_k(\omega_n) = \left(\mathbf{R}_{\hat{\mathbf{y}}_\mathbf{k}\hat{\mathbf{y}}_\mathbf{k}}(\omega_n)\right)^{-1} \mathbf{R}_{\hat{\mathbf{s}}_\mathbf{k}\hat{\mathbf{s}}_\mathbf{k}}(\omega_n)\mathbf{e}_1 \qquad (15)$$

where $\mathbf{e}_1 = [1\ 0 \ldots 0]^T$.

In this paper, we apply DANSE which is a distributed adaptive noise reduction algorithm [2], [3]. This algorithm is designed primarily for a fully connected sensor network in which all the nodes broadcast pre-processed microphone signals to all other nodes. The main objective of DANSE is to generate a node-specific estimate of the target speech signal as it impinges on the first microphone of each individual node. To reach this goal, DANSE compresses the individual microphone signals of each node into a single-channel signal z_k which is then broadcast to the other nodes. Surprisingly, without accessing all the observations in the network, the optimal estimation for each node can be obtained [2], [3]. For the case of DANSE, $\hat{\mathbf{y}}_k(\omega_n)$ in (13)-(15) will be a $M_k + J - 1$ dimensional vector such that:

$$\hat{\mathbf{y}}_k(\omega_n) = \begin{bmatrix} \mathbf{y}_k(\omega_n) \\ \mathbf{z}_{-k}(\omega_n) \end{bmatrix} \qquad (16)$$

where, considering $\mathbf{z}(\omega_n) = [z_1(\omega_n)\ \ldots\ z_J(\omega_n)]^T$, $\mathbf{z}_{-k}(\omega_n)$ denotes the vector $\mathbf{z}(\omega_n)$ in which $z_k(\omega_n)$ is excluded. These signals are generated based on the following filter-and-sum process:

$$z_k(\omega_n) = \mathbf{w}_{\mathbf{y}_k}(\omega_n)^H \mathbf{y}_k(\omega_n) \qquad (17)$$

where $\mathbf{w}_{\mathbf{y}_k}(\omega_n)$ is the part of the $\hat{\mathbf{w}}_k(\omega_n)$ that is only applied to node k's own M_k microphone signals $\mathbf{y}_k(\omega_n)$. Note that (17) compresses the M_k-channel signal $\mathbf{y}_k(\omega_n)$ into a single-channel signal $z_k(\omega_n)$, hence DANSE considerably reduces the required communication bandwidth, as well as the per-node computational complexity when compared to the centralized case. As a result, DANSE can considerably reduce the required communication bandwidth as well as the local computational requirement. In DANSE, the nodes update their $\hat{\mathbf{w}}_k(\omega_n)$ with (15) either sequentially [2] or simultaneously [3] (rS-DANSE). Here we consider a case in which nodes update their filters in a sequential round robin fashion. For further reading, we refer to [2] and [3].

4.2. Collaborative steering vector estimation

In order to estimate the steering vector from $\bar{\mathbf{R}}_{\hat{\mathbf{s}}_\mathbf{k}\hat{\mathbf{s}}_\mathbf{k}}(\omega_n)$ with exploiting the effect of collaboration between nodes, the following procedure is proposed. We perform eigenvalue decomposition (EVD) for rank-1 approximation of the speech correlation matrix, i.e.:

$$\bar{\mathbf{R}}_{\hat{\mathbf{s}}_\mathbf{k}\hat{\mathbf{s}}_\mathbf{k}}(\omega_n) \approx\ \hat{\mathbf{v}}_{max_k}(\omega_n)\hat{\mathbf{v}}^H_{max_k}(\omega_n)\lambda_{max_k}(\omega_n) \qquad (18)$$

where λ_{max_k} is the largest eigenvalue and $\hat{\mathbf{v}}_{max_k}$ is its corresponding normalized eigenvector. We define $\mathbf{v}_{max_k}$ as the first M_k entries of $\hat{\mathbf{v}}_{max_k}$, only containing the part corresponding to node k's own microphone signals and ignoring the signals obtained from the other nodes. Although this means that we throw away information, there is still implicit collaboration between the nodes as the EVD-based computation of first $\hat{\mathbf{v}}_{max_k}$ and then of $\mathbf{v}_{max_k}$ indeed also relies on the signals from other nodes, which will (hopefully) result in a better estimate of the steering vector. The reason why we only proceed with $\mathbf{v}_{max_k}$ rather than $\hat{\mathbf{v}}_{max_k}$, is because the relative geometry between the microphone arrays of different nodes is assumed to be unknown. According to (12), $\mathbf{v}_{max_k}$ can be treated as a normalized estimate of the steering vector, i.e. $\bar{\mathbf{a}}_k \approx \beta \mathbf{v}_{max_k}$, where β is a complex number. Consequently, for the LS algorithm $\bar{\mathbf{p}}_{k_t}$ in (5) will be the phase of $\mathbf{v}_{max_{k_t}}/v_1$ where v_1 is the first element of $\mathbf{v}_{max_k}$ and $\mathbf{v}_{max_{k_t}}$ is $\mathbf{v}_{max_k}$ with the first element removed. For the case of MUSIC subspace decomposition in (9) we will have $\mathbf{q}_1 = \mathbf{v}_{max_k}$. To evaluate how the collaboration between the nodes impacts the DOA estimation, the acoustic scenario illustrated in Figure 2 is simulated. In a non-reverberant room, we assume a symmetric arrangement to create a scenario in which the input signal to noise ratio (iSNR) will be identical for all the nodes. We consider 4 nodes, each having 3 microphones that form a ULA, as well as 4 uncorrelated multi-talker noise sources with equal noise power which will be varied to manipulate the input SNR. A speech source in the center of the room produces the target signal. Consequently, the true values of the specific DOAs to be estimated are 90 degrees for each node. A sampling frequency F_s = 8kHZ and DFT size of $L = 1024$ with 50% Hann-windowed overlaps are used. To model the precise microphone signals of each ULA, fractional delay filters are applied. Moreover, an ideal VAD is used to exclude the effect of VAD errors. DANSE is performed in batch mode, which means that all iterations are done on the full signal length, and the nodes are updated in a sequential round robin fashion. The speech source produces short English sentences with a silence period between each

two consecutive sentences. Sensor noise is modeled as an uncorrelated white Gaussian noise signal with 5% of the power of the target speech signal as observed at the microphones.

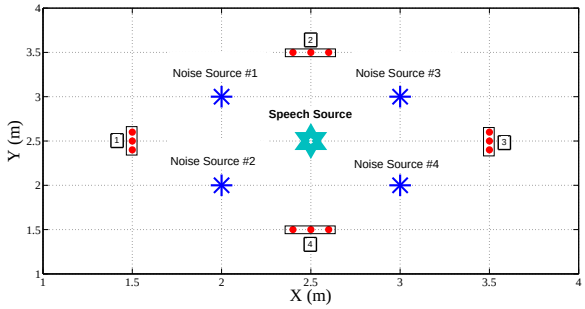

Fig. 2. Acoustic scenario

5. SIMULATION RESULTS

To investigate the performance of the distributed DOA estimation, it is compared with the centralized and isolated case. A database with 28 speech signals of 68 seconds each is used to simulate 28 Monte-Carlo (MC) runs with uncorrelated multi-talker noise for each independent run. Figure 3 shows the averaged absolute values of the DOA estimation errors using the proposed LS method. As can be seen in this figure, collaboration between nodes leads to a better performance compared to the isolated estimation. Although the difference is more clear for lower input SNRs, it should be mentioned that in higher input SNR levels, collaborative estimation still outperforms the isolated estimation. In the case of MUSIC, more MC runs were required to obtain an intelligible figure. Therefore, we have performed with 56 speech signals each 34 seconds. A resolution of 1 degree is used for the exhaustive search. Again, it can be observed in Figure 4 that collaboration between the nodes significantly improves the DOA estimation performance. As can be seen from the figures,the DOA estimator based on MUSIC performs better at lower SNRs than the estimations based on the LS method. This comes at a cost of a significantly higher computational complexity due to an exhaustive search over all possible DOA's, which may be impractical in WASNs with limited power supply.

6. CONCLUSIONS

In this paper, we have studied the benefits of cooperation between nodes in a node-specific DOA estimation task in a WASN. The nodes use the broadcast signals generated by the DANSE algorithm to improve the estimation of the local node-specific steering vectors. To keep the effects of cooperation on the local steering vector extraction, we have used an EVD-based rank-1 approximation. A LS method has been

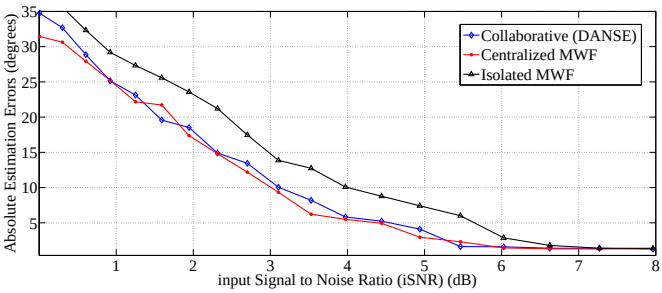

Fig. 3. Absolute errors based on LS for distributed, centralized and isolated case

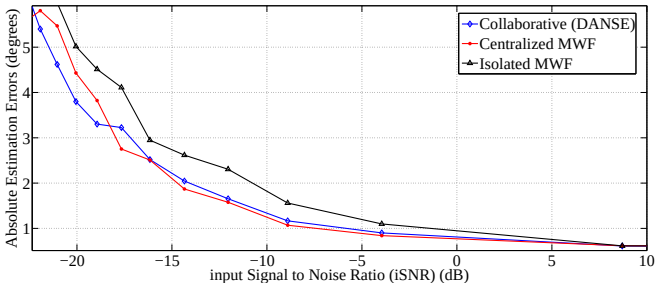

Fig. 4. Absolute errors based on MUSIC for distributed, centralized and isolated case

utilized for the estimation of node-specific DOAs. In addition, the MUSIC algorithm has also been employed to further evaluate the performance of the estimation with collaborative nodes. It has been demonstrated that the collaborative estimation of DOAs with exploiting the shared signals used in the DANSE algorithm, yields better results compared to the isolated case.

7. REFERENCES

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

[2] A. Bertrand and M. Moonen, "Distributed adaptive node-specific signal estimation in fully connected sensor networks part I: sequential node updating," in *IEEE Trans. Signal Processing*, 2010, vol. 58, pp. 5277–5291.

[3] A. Bertrand and M. Moonen, "Distributed adaptive node-specific signal estimation in fully connected sensor networks part II: simultaneous and asynchronous node updating," in *IEEE Trans. Signal Processing*, 2010, vol. 58, pp. 5292–5306.

[4] L. L. Scharf M. R. Azimi-Sadjadi, A. Pezeshki and M. Hohil, "Wideband DOA estimation algorithms for multiple target detection and tracking using unattended acoustic sensors," in *Proc. of SPIE04 Defense and Security Symposium*, 2004, vol. 5417, pp. 1–11.

[5] R. Schmidt, "Multiple emitter location and signal parameter estimation," in *IEEE Trans. on Antennas and Propagation*, 1986, vol. 34, pp. 276–280.

[6] S. Doclo and M. Moonen, "GSVD-based optimal filtering for single and multimicrophone speech enhancement," in *IEEE Trans. Signal Processing*, 2002, vol. 50, pp. 2230–2244.