

Distributed MAP Estimators for Noise Reduction in Fully Connected Wireless Acoustic Sensor Networks

Raziyeh Ranjbaryan¹, Simon Doclo², and Hamid Reza Abutalebi¹

¹ Electrical Engineering Department, Yazd University, Yazd, Iran

² Dept. of Medical Physics and Acoustics and Cluster of Excellence Hearing4All, University of Oldenburg, Germany
Email: ranjbaryan@stu.yazd.ac.ir, simon.doclo@uni-oldenburg.de, habutalebi@yazd.ac.ir

Abstract

Several noise reduction algorithms have been proposed for wireless acoustic sensor networks, which consist of spatially distributed nodes that are connected via a wireless link. To decrease the required bandwidth and computational complexity, in this paper we propose two iterative distributed maximum a posteriori (MAP) estimators. In the first scheme, each node sequentially updates its estimate, whereas in the second scheme, all nodes simultaneously update their estimates. Based on simulations in a reverberant room with three nodes, we have compared the noise reduction performance of the proposed distributed MAP estimators with the centralized MAP estimator, where each node has access to all signals, and the local MAP estimator, where each node only has access to its own signals. The simulation results show that the proposed distributed estimators result in a good noise reduction performance, while decreasing the computational complexity compared to the centralized estimator.

1 Introduction

In order to improve speech quality and intelligibility, several noise reduction algorithms have been proposed for wireless acoustic sensor networks (WASNs) [1–10]. WASNs consist of several spatially distributed nodes, where each node contains one or more microphones. Since the nodes are connected via a wireless link, they are able to cover a larger area and utilize more spatial information than traditional microphone arrays, which typically have a predefined geometry.

In the context of WASNs, noise reduction algorithms can be categorized into two main classes: centralized and distributed algorithms [2, 3, 7]. In centralized algorithms, all (noisy) microphone signals of the nodes are transmitted to a fusion center (FC). Since the FC has direct access to all recorded signals, typically a large noise reduction performance can be achieved. However, this requires a large bandwidth of the wireless link and results in a large computational complexity; in addition, the performance of the WASN is rather sensitive to the FC. In order to enable a trade-off between noise reduction performance and required bandwidth and computational complexity, several distributed beamforming algorithms have been developed for WASNs, which share signals directly between the nodes without requiring an FC. A comprehensive survey on distributed beamforming algorithms in WASNs has been presented in [7], indicating that the main idea among most algorithms is similar: to reduce the number of transmitted signals, each node broadcasts only filtered versions of its recorded signals instead of sending all signals to the other

nodes. It has been shown that under some conditions, distributed algorithms converge to the centralized solution [7].

All aforementioned distributed algorithms have focused on beamforming and to the best of our knowledge, there has been no investigation on multi-microphone Bayesian estimators in WASNs, which consider the statistical properties of the signals to perform noise reduction. In [11], a multi-microphone maximum a posteriori (MAP) estimator was proposed to estimate the clean speech amplitude from the noisy (observed) signals. Since this estimator does not require any assumption about the position of the nodes and the desired speaker and relies only on second-order statistical properties, it can be considered as a good choice to be applied for noise reduction in WASNs.

The paper is organized as follows. After formulating the problem in Section 2, we review the centralized MAP estimator in Section 3. In Section 4, we derive two distributed MAP (DMAP) estimators (sequential and simultaneous update) and show similarities and differences between the proposed estimators and distributed beamforming algorithms. In Section 5 the noise reduction performance of the proposed DMAP estimators is compared with the centralized and local MAP estimators.

2 Problem Formulation

Consider a WASN consisting of K nodes, where the k -th node contains M_k microphones and the total number of microphones is equal to $M = \sum_{k=1}^K M_k$. In the short-time Fourier transform (STFT) domain, the received signals at the m -th microphone of the k -th node can be expressed as

$$Y_{m,k}(l, n) = X_{m,k}(l, n) + V_{m,k}(l, n), \quad m = 1, \dots, M_k \quad (1)$$

where $Y_{m,k}(l, n)$, $X_{m,k}(l, n)$ and $V_{m,k}(l, n)$ denote the noisy signal, the speech signal and the additive noise, respectively, with l the frame index and n the discrete frequency index. For conciseness we omit the frame and frequency indices in the remainder of the paper wherever possible.

In vector notation, (1) is given by $\mathbf{y}_k = \mathbf{x}_k + \mathbf{v}_k$, with $\mathbf{y}_k = [Y_{1,k}, \dots, Y_{M_k,k}]^T$, $\mathbf{x}_k = [X_{1,k}, \dots, X_{M_k,k}]^T$ and $\mathbf{v}_k = [V_{1,k}, \dots, V_{M_k,k}]^T$, where T denotes the transpose operation.

Assuming that the speech and the noise signals are uncorrelated, the noisy correlation matrix of the k -th node can be expressed as

$$\Phi_{\mathbf{y}_k \mathbf{y}_k} = \mathbb{E} \{ \mathbf{y}_k \mathbf{y}_k^H \} = \Phi_{\mathbf{x}_k \mathbf{x}_k} + \Phi_{\mathbf{v}_k \mathbf{v}_k}, \quad (2)$$

where $\mathbb{E} \{ \cdot \}$ denotes the expectation operator, and $\Phi_{\mathbf{x}_k \mathbf{x}_k}$ and $\Phi_{\mathbf{v}_k \mathbf{v}_k}$ are the speech and noise correlation matrices of the k -th node, respectively.

In a centralized algorithm, the FC has access to all noisy signals, and the M -dimensional centralized noisy vector is given by

$$\mathbf{y} = \mathbf{x} + \mathbf{v}, \quad (3)$$

This work was supported in part by the Iran National Science Foundation (INSF) and the German Academic Exchange Service (DAAD).

with $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_K^T]^T$ and \mathbf{x} and \mathbf{v} defined similarly. The centralized noisy correlation matrix can be written as

$$\Phi_{\mathbf{y}\mathbf{y}} = \mathbb{E}\{\mathbf{y}\mathbf{y}^H\} = \Phi_{\mathbf{x}\mathbf{x}} + \Phi_{\mathbf{v}\mathbf{v}}, \quad (4)$$

where $\Phi_{\mathbf{x}\mathbf{x}}$ and $\Phi_{\mathbf{v}\mathbf{v}}$ denote the centralized speech and noise correlation matrices, respectively.

In practice, the centralized noisy and noise correlation matrices can be recursively estimated based on the noisy signals as [12]

$$\hat{\Phi}_{\mathbf{y}\mathbf{y}}(l, n) = \lambda_y \hat{\Phi}_{\mathbf{y}\mathbf{y}}(l-1, n) + (1 - \lambda_y) \mathbf{y}(l, n) \mathbf{y}^H(l, n), \quad (5)$$

$$\hat{\Phi}_{\mathbf{v}\mathbf{v}}(l, n) = \lambda_v \hat{\Phi}_{\mathbf{v}\mathbf{v}}(l-1, n) + (1 - \lambda_v) \mathbf{y}(l, n) \mathbf{y}^H(l, n), \quad (6)$$

with $\lambda_v = \lambda_n + (1 - \lambda_n) \text{SPP}(l, n)$, where $\text{SPP}(l, n)$ denotes the speech presence probability [13] and λ_y and λ_n denote forgetting factors for speech and noise, respectively. The centralized speech correlation matrix can be estimated as

$$\hat{\Phi}_{\mathbf{x}\mathbf{x}}(l, n) = \hat{\Phi}_{\mathbf{y}\mathbf{y}}(l, n) - \hat{\Phi}_{\mathbf{v}\mathbf{v}}(l, n). \quad (7)$$

Considering estimation errors in the centralized noisy and noise correlation matrices, negative eigenvalues of $\hat{\Phi}_{\mathbf{x}\mathbf{x}}(l, n)$ are set to zero to ensure that the resulting speech correlation matrix is positive semidefinite.

2.1 Polar Representation

In polar representation, (1) can be written as

$$Y_{m,k} = R_{m,k} e^{j\vartheta_{m,k}} = A_{m,k} e^{j\alpha_{m,k}} + V_{m,k}, \quad (8)$$

where $R_{m,k}$, $\vartheta_{m,k}$, $A_{m,k}$ and $\alpha_{m,k}$ denote the spectral amplitude and phase of the noisy signal and the speech signal of the m -th microphone of the k -th node, respectively.

Assuming the real and imaginary parts of the speech signal to be zero-mean independent Gaussian variables, the amplitude and the phase of the speech signal are modeled by a Rayleigh and a uniform distribution, respectively [11]. Assuming the noise signal to be Gaussian distributed, the following probability density functions are obtained

$$\begin{aligned} p(A_{m,k}) &= \frac{2A_{m,k}}{\sigma_x^2(m,k)} \exp\left(\frac{-A_{m,k}^2}{\sigma_x^2(m,k)}\right), \\ p(\alpha_{m,k}) &= \frac{1}{2\pi} \quad -\pi < \alpha_{m,k} < \pi, \\ p(A_{m,k}, \alpha_{m,k}) &= \frac{A_{m,k}}{\pi\sigma_x^2(m,k)} \exp\left(\frac{-A_{m,k}^2}{\sigma_x^2(m,k)}\right), \quad (9) \\ p(Y_{m,k} | A_{m,k}, \alpha_{m,k}) &= \\ &= \frac{1}{\pi\sigma_v^2(m,k)} \exp\left(-\frac{|Y_{m,k} - A_{m,k} e^{j\alpha_{m,k}}|^2}{\sigma_v^2(m,k)}\right), \end{aligned}$$

where $\sigma_x^2(m, k)$ and $\sigma_v^2(m, k)$ denote the variances of the speech and the noise signal of the m -th microphone of the k -th node, respectively.

3 Centralized MAP Estimator

In [11], a (direction-independent) MAP estimator was proposed that estimates the speech amplitude from the amplitudes of the noisy signals. In the case of a single speech source, this estimator aims at maximizing the posterior distribution of the amplitude of the reference speech signal

conditioned on the amplitudes of the noisy signals. Without loss of generality, in each node the first microphone is considered as the reference microphone. In a fully connected WASN, where each node has access to the all noisy signals (i.e., each node acts as an FC), the optimization problem for the k -th node can be expressed as

$$\hat{A}_{1,k} = \arg \max_{A_{1,k}} p(A_{1,k} | [\mathbf{r}_1^T, \dots, \mathbf{r}_k^T, \dots, \mathbf{r}_K^T]^T), \quad (10)$$

with $\mathbf{r}_k = [R_{1,k}, \dots, R_{M_k,k}]^T$. The solution of (10) is $\hat{A}_{1,k} = G_{1,k} R_{1,k}$, where the gain $G_{1,k}$ is equal to [11]

$$\begin{aligned} G_{1,k} &= \frac{\sqrt{\frac{\zeta_{1,k}}{\gamma_{1,k}}}}{2 + 2 \sum_{k=1}^K \sum_{m=1}^{M_k} \zeta_{m,k}} \text{Re} \left\{ \sum_{k=1}^K \sum_{m=1}^{M_k} \sqrt{\zeta_{m,k} \gamma_{m,k}} + \right. \\ &\left. \sqrt{\left(\sum_{k=1}^K \sum_{m=1}^{M_k} \sqrt{\zeta_{m,k} \gamma_{m,k}} \right)^2 + (2-M) \left(1 + \sum_{k=1}^K \sum_{m=1}^{M_k} \zeta_{m,k} \right)} \right\}, \quad (11) \end{aligned}$$

where the *a priori* and *a posteriori* signal-to-noise ratios (SNRs) are given by

$$\zeta_{m,k} = \frac{\sigma_x^2(m,k)}{\sigma_v^2(m,k)}, \quad \gamma_{m,k} = \frac{R_{m,k}^2}{\sigma_v^2(m,k)}. \quad (12)$$

The variances $\sigma_x^2(m, k)$ and $\sigma_v^2(m, k)$ can be computed as the diagonal elements of the centralized speech and noise correlation matrices in (7) and (6), respectively, and $R_{m,k}$ can be directly computed from the noisy (observed) signals. The enhanced speech signal of the k -th node is estimated using the noisy phase as

$$\hat{X}_{1,k} = G_{1,k} R_{1,k} e^{j\vartheta_{1,k}} = \hat{A}_{1,k} e^{j\vartheta_{1,k}}. \quad (13)$$

4 Distributed MAP Estimators

In the centralized MAP estimator, all noisy signals of the nodes need to be transmitted, introducing a large computational complexity and requiring a large bandwidth. In addition, some nodes may be located at a large distance from the desired speaker, leading to signals with a low SNR. To decrease the required bandwidth and to increase the SNR of the received signals in each node, in this paper we propose two distributed MAP estimators, either with sequential update or with simultaneous update. Similar to distributed beamforming algorithms, in the proposed MAP estimators each node broadcasts a filtered signal instead of all signals.

For the distributed MAP estimators, we introduce the K -dimensional vector $\hat{\mathbf{x}} = [\hat{X}_{1,1}, \hat{X}_{1,2}, \dots, \hat{X}_{1,K}]^T$, containing the estimated speech signals of all nodes. In addition, we define the $(K-1)$ -dimensional vector $\hat{\mathbf{x}}_{-k}$ by excluding $\hat{X}_{1,k}$ from the vector $\hat{\mathbf{x}}$. Since the proposed distributed estimators require iterative updates, we will use the superscript i to indicate the iteration index. For each node k , we define the distributed noisy vector, consisting of its own local noisy signals \mathbf{y}_k^i and the received signals $\hat{\mathbf{x}}_{-k}^i$ from the other nodes, as

$$\tilde{\mathbf{y}}_k^i = \begin{bmatrix} \mathbf{y}_k^i \\ \hat{\mathbf{x}}_{-k}^i \end{bmatrix} \quad (14)$$

Each node hence has access to $M_k + K - 1$ signals, i.e., M_k local noisy signals and $K - 1$ compressed signals from the other nodes. The distributed noisy correlation matrix is given by

$$\Phi_{\tilde{\mathbf{y}}_k \tilde{\mathbf{y}}_k}^i = \mathbb{E} \{ \tilde{\mathbf{y}}_k^i \tilde{\mathbf{y}}_k^{iH} \} = \Phi_{\tilde{\mathbf{x}}_k \tilde{\mathbf{x}}_k}^i + \Phi_{\tilde{\mathbf{v}}_k \tilde{\mathbf{v}}_k}^i, \quad (15)$$

where $\Phi_{\tilde{\mathbf{x}}_k \tilde{\mathbf{x}}_k}^i$ and $\Phi_{\tilde{\mathbf{v}}_k \tilde{\mathbf{v}}_k}^i$ denote the distributed speech and noise correlation matrices, respectively. Similarly to (5)-(7), the distributed correlation matrices can be recursively estimated in practice. Compared to the centralized case, for the distributed MAP estimators the number of transmitted signals is decreased from $M(K - 1)$ to $K(K - 1)$.

The aim of distributed MAP estimators is to maximize the posterior distribution of the amplitude of the reference speech signal at each node conditioned on the received distributed amplitude vector, i.e.

$$\hat{A}_{1,k}^i = \arg \max_{A_{1,k}^i} p(A_{1,k}^i | \tilde{\mathbf{r}}_k^i), \quad (16)$$

where $\tilde{\mathbf{r}}_k^i$ corresponds to the amplitude of the distributed noisy vector. Similarly to (11), the solution of the distributed MAP estimator is $\hat{A}_{1,k}^i = R_{1,k}^i G_{1,k}^i$, where the gain $G_{1,k}^i$ is equal to

$$G_{1,k}^i = \frac{\sqrt{\frac{\tilde{\zeta}_{1,k}^i}{\tilde{\gamma}_{1,k}^i}}}{2 + 2 \sum_{m=1}^{\tilde{M}} \tilde{\zeta}_{m,k}^i} \operatorname{Re} \left\{ \sum_{m=1}^{\tilde{M}} \sqrt{\tilde{\zeta}_{m,k}^i \tilde{\gamma}_{m,k}^i} + \sqrt{\left(\sum_{m=1}^{\tilde{M}} \sqrt{\tilde{\zeta}_{m,k}^i \tilde{\gamma}_{m,k}^i} \right)^2 + (2 - \tilde{M}) \left(1 + \sum_{m=1}^{\tilde{M}} \tilde{\zeta}_{m,k}^i \right)} \right\}, \quad (17)$$

where $\tilde{M} = M_k + K - 1$. The distributed *a priori* and *a posteriori* SNRs are given by

$$\tilde{\zeta}_{m,k}^i = \frac{(\tilde{\sigma}_x^i(m,k))^2}{(\tilde{\sigma}_v^i(m,k))^2}, \quad \tilde{\gamma}_{m,k}^i = \frac{(\tilde{R}_{m,k}^i)^2}{(\tilde{\sigma}_v^i(m,k))^2}, \quad (18)$$

where the distributed variances can be computed as the diagonal elements of the distributed speech and noise correlation matrices.

Motivated by [2]-[3], we now introduce a sequential update and a simultaneous update DMAP estimator.

4.1 Sequential Update DMAP Estimator

In the first scheme, at each iteration one specific node updates its estimate. The proposed sequential distributed MAP estimator hence runs as follows (cf. block diagram depicted in Fig. 1):

1. The algorithm is initialized with iteration index $i = 0$ and node index $u = 1$. $G_{1,k}^i$ is initialized with a random positive number between 0 and 1 $\forall k \in K$.
2. For each node $k \in K$, the following steps are performed:
 - Compute the amplitude $\hat{A}_{1,k}^i = G_{1,k}^i R_{1,k}^i$.
 - Broadcast $\hat{X}_{1,k}^i = \hat{A}_{1,k}^i e^{j\theta_{1,k}^i}$ to the other nodes.
 - Collect the vector $\hat{\mathbf{x}}_{-k}^i$.
 - Construct the vector $\tilde{\mathbf{y}}_k^i$ using (14).

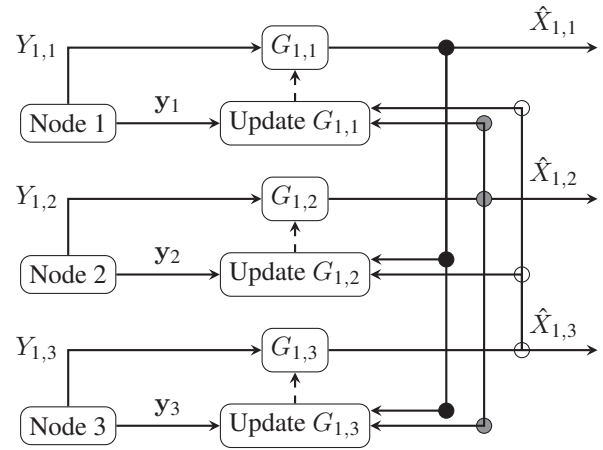


Figure 1: Proposed DMAP estimator with three nodes. Each node uses its own noisy signals and the compressed signals from the other nodes to update a gain.

- Update the distributed correlation matrices for $\tilde{\mathbf{y}}_k^i$, $\tilde{\mathbf{v}}_k^i$ and $\tilde{\mathbf{x}}_k^i$.
- Compute the *a priori* and *a posteriori* SNRs using (18).
- Update the node-specific gain

$$G_{1,k}^{i+1} = \begin{cases} \text{Eq. (17)} & \text{if } k = u \\ G_{1,k}^i & \text{if } k \neq u \end{cases} \quad (19)$$

3. $i \leftarrow i + 1$
4. $u \leftarrow (u \bmod K) + 1$
5. return to step 2.

4.2 Simultaneous Update DMAP Estimator

Although applying the sequential update DMAP estimator reduces the required bandwidth and computational complexity, it may increase the convergence time, especially in the case of a large number of nodes. To increase the convergence speed, motivated by the algorithm presented in [3], we propose a simultaneous update DMAP estimator, where all nodes simultaneously update their estimates. The only difference between the sequential and the simultaneous update DMAP estimators is the step to update the node-specific gain

$$G_{1,k}^{i+1} = \frac{1}{i} G_{1,k}^i + \left(1 - \frac{1}{i}\right) \text{Eq. (17)}. \quad (20)$$

5 Simulation Results

In this section, we evaluate the noise reduction performance of the proposed DMAP estimators with two baseline methods: 1) the local MAP estimator, where each node only has access to its own local noisy signals, and 2) the centralized MAP estimator, where all nodes have access to all noisy signals in a fully connected network (see Section 3).

All signals were recorded in an acoustic laboratory located at the University of Oldenburg with dimensions ($x = 7$ m, $y = 6$ m, $z = 2.7$ m) and reverberation time $RT_{60} \simeq 350$ ms. For the simulations we consider a WASN consisting of three nodes ($K = 3$). The first node corresponds to a hearing aid with 2 microphones, mounted on a dummy head placed approximately in the middle of the laboratory.

The distance between the microphones is about 7.6 mm. The second node consists of two microphones located at $(x = 4.64 \text{ m}, y = 4.13 \text{ m}, z = 2 \text{ m})$ and $(x = 4.64 \text{ m}, y = 2.63 \text{ m}, z = 2 \text{ m})$, respectively, and the third node consists of one microphone located at $(x = 2.36 \text{ m}, y = 4.13 \text{ m}, z = 2 \text{ m})$. The desired source was a male English speaker of duration 20 seconds, played back by a loudspeaker located at a distance of 2 m at the same height and at an angle of 35° on the right side of the dummy head. To generate background noise, four loudspeakers facing the corners of the laboratory were used, playing back different realizations of operation room noise. All signals were recorded at a sampling frequency $f_s = 16 \text{ kHz}$. The STFT processing is implemented using NFFT=512 with half-overlapping frames and using a Hamming window as the analysis window. The forgetting factors to update the noisy and noise correlation matrices are $\lambda_y = \lambda_n = 0.92$. For each node, the SPP was computed using the method proposed in [13]. We used a threshold $G_{1,k,max} = 1$ to ensure that the estimated signal sounds natural.

In this paper, we don't consider the batch implementation, where iterations are performed on the complete signal, but only the recursive implementation where the iteration index is replaced by the frame index.

Fig. 2 depicts the performance of the considered MAP estimators in terms of the PESQ [14] improvement between the enhanced signal and the noisy reference signal at the first node for several input SNRs. The speech signal at the first microphone of the first node has been used as the reference signal. As expected, it can be observed that the PESQ improvement for the centralized MAP estimator is considerably larger than the PESQ improvement for the local MAP estimator, especially at large input SNRs. In addition, it can be observed that the PESQ improvement for the proposed DMAP estimators is similar to the PESQ improvement for the centralized MAP estimator, where the simultaneous update DMAP estimator consistently outperforms the sequential update DMAP estimator.

Fig. 3 depicts the performance of the considered MAP estimators in terms of the segmental noise reduction (segNR), the segmental speech SNR (segSSNR) and the segmental SNR (segSNR) improvement, as proposed in [13]. The segSSNR measure has been defined as a measure for speech distortion, where larger segSSNR values indicate lower speech distortion. As expected, it can be observed from

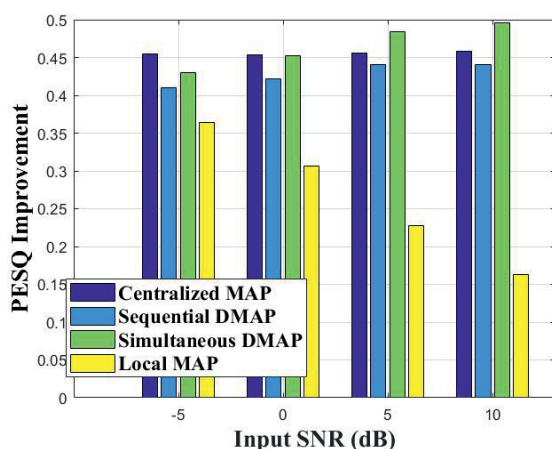


Figure 2: PESQ improvement for the proposed DMAP estimators, the local MAP estimator and the centralized MAP estimator between the enhanced signal and the noisy reference signal at the first node for several input SNRs.

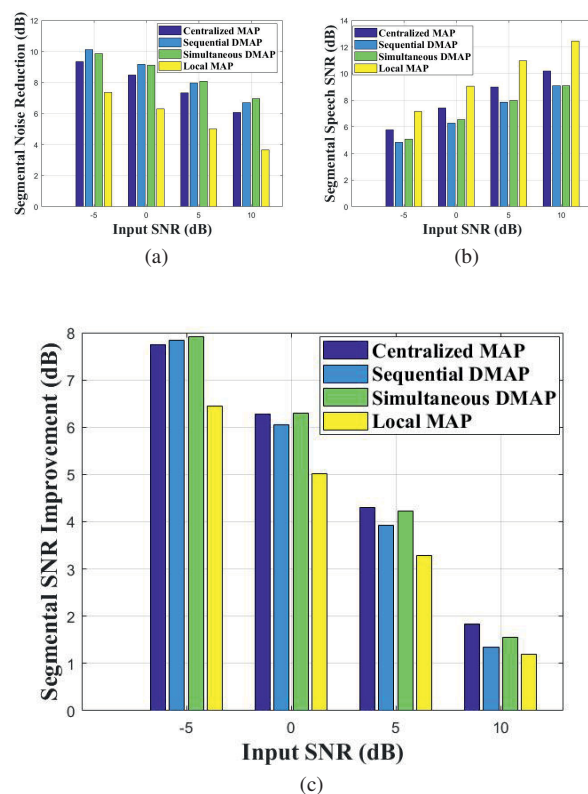


Figure 3: Performance of the proposed DMAP estimators, the local MAP estimator and the centralized MAP estimator in terms of (a) segmental noise reduction, (b) segmental speech SNR and (c) segmental SNR improvement for several input SNRs.

Fig. 3 that the centralized MAP estimator outperforms the local MAP estimator in terms of segSNR improvement, which considers both noise reduction as well as speech distortion. This can be mainly explained by the lower segNR values for the local MAP estimator. Moreover, it can be observed that both proposed DMAP estimators result in a similar segSNR improvement as the centralized MAP estimator, where the proposed DMAP estimators seem to result in slightly larger segNR values (i.e. more noise reduction) and slightly lower segSSNR values (i.e. more speech distortion) than the centralized MAP estimator.

In conclusion, the experimental results show that the performance of the proposed DMAP estimators is similar to the performance of the centralized MAP estimator at a lower computational complexity.

6 Conclusion

In order to decrease the computational complexity and required bandwidth, in this paper we proposed two distributed MAP estimators (sequential and simultaneous update). Each node uses a vector consisting of its local signals and compressed signals from the other nodes, where the compressed signals estimate the speech signal at these nodes. We compared the performance of the proposed distributed MAP estimators with the local and the centralized MAP estimators for a network with 3 nodes, showing that the performance of the DMAP estimators is similar to the performance of the centralized MAP estimator at a lower computational complexity.

References

- [1] S. Doclo, M. Moonen, T. van den Bogaert, and J. Wouters, “Reduced-bandwidth and distributed MWF-based noise reduction algorithms for binaural hearing aids,” *IEEE Transaction on Audio, Speech, and Language Processing*, vol. 17, pp. 38–51, Jan. 2009.
- [2] A. Bertrand and M. Moonen, “Distributed adaptive node-specific signal estimation in fully connected sensor networks – part I: Sequential node updating,” *IEEE Transaction on Signal Processing*, vol. 58, pp. 5277–5291, Oct. 2010.
- [3] A. Bertrand and M. Moonen, “Distributed adaptive node-specific signal estimation in fully connected sensor networks – part II: Simultaneous and asynchronous node updating,” *IEEE Transaction on Signal Processing*, vol. 58, pp. 5292–5306, Oct. 2010.
- [4] M. Taseska and E. A. P. Habets, “Informed spatial filtering for sound extraction using distributed microphone arrays,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, pp. 1195–1207, July 2014.
- [5] T. C. Lawin-Ore, S. Stenzel, J. Freudenberger, and S. Doclo, “Alternative formulation and robustness analysis of the multichannel Wiener filter for spatially distributed microphones,” in *Proc. International Workshop on Acoustic Signal Enhancement (IWAENC)*, Juan les Pins, France, Sept. 2014.
- [6] T. C. Lawin-Ore and S. Doclo, “Analysis of the average performance of the multi-channel Wiener filter for distributed microphone arrays using statistical room acoustics,” *Signal Processing*, vol. 107, pp. 96–108, Feb. 2015.
- [7] S. Markovich-Golan, A. Bertrand, M. Moonen, and S. Gannt, “Optimal distributed minimum-variance beamforming approaches for speech enhancement in wireless acoustic sensor networks,” *Signal Processing*, vol. 107, pp. 4–20, 2015.
- [8] A. Hassani, A. Bertrand, and M. Moonen, “GEVD-based low-rank approximation for distributed adaptive node-specific signal estimation in wireless sensor networks,” *IEEE Transactions on Signal Processing*, vol. 64, pp. 2557–2572, May 2016.
- [9] V. M. Tavakoli, J. R. Jensen, M. G. Christensen, and J. Benesty, “A framework for speech enhancement with ad hoc microphone arrays,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, pp. 1038–1051, June 2016.
- [10] A. I. Koutrouvelis, T. W. Sherson, R. Heusdens, and R. C. Hendriks, “A low-cost robust distributed linearly constrained beamformer for wireless acoustic sensor networks with arbitrary topology,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, pp. 1434–1448, Aug. 2018.
- [11] T. Lotter, C. Benien, and P. Vary, “Multi channel direction independent speech enhancement using spectral amplitude estimation,” *EURASIP Journal on Applied Signal Processing*, pp. 1147–1156, 2003.
- [12] M. Souden, J. Chen, J. Benesty, and S. Affes, “An integrated solution for online multichannel noise tracking and reduction,” *IEEE Transaction on Audio, Speech, and Language Processing*, vol. 19, pp. 2159–2169, Sept. 2011.
- [13] T. Gerkmann and R. C. Hendriks, “Unbiased MMSE-based noise power estimation with low complexity and low tracking delay,” *IEEE Transaction on Audio, Speech, and Language Processing*, vol. 20, pp. 1383–1393, May 2012.
- [14] P. Loizou, *Speech Enhancement: Theory and Practice*. CRC press, 2007.