ENERGY AWARE GREEDY SUBSET SELECTION FOR SPEECH ENHANCEMENT IN WIRELESS ACOUSTIC SENSOR NETWORKS

Joseph Szurley, Alexander Bertrand*, Marc Moonen*, Peter Ruckebusch†, and Ingrid Moerman†*

∗ KU Leuven, Dept. of Electrical Engineering
ESAT, SCD-SISTA
IBBT Future Health Department
Kasteelpark Arenberg 10
B-3001 Leuven, Belgium
E-mail: joseph.szurley@esat.kuleuven.be,
alexander.bertrand@esat.kuleuven.be,
marc.moonen@esat.kuleuven.be

† Ghent University - IBBT
Dept. of Information Technology (INTEC)
Gaston Crommenlaan 8 Bus 201
9050 Ghent, Belgium
E-mail: peter.ruckebusch@intec.ugent.be,
ingrid.moerman@intec.ugent.be

ABSTRACT

A wireless acoustic sensor network is envisaged that relies on a collection of spatially distributed microphones, which observe a speech signal together with additive background noise. The microphone signals are sent to a fusion center where they are filtered and combined to produce an estimate of the speech signal. In order to save energy and extend network lifetime, it is desired to only have a subset of the microphones active at any one moment. This subset selection unfortunately comes with the adverse effect of decreasing the accuracy of the signal estimation. Since the network now has two competing objectives a trade-off develops that balances the energy consumption to estimation accuracy. We propose a network model that is cast similarly to a 0-1 knapsack problem that uses a greedy method to balance the output signal-to-noise ratio to total transmission energy expended by the wireless microphones. Simulations show that although a greedy approach is used, a relatively small decrease in output signal-to-noise ratio is achieved while there is a marked decrease in energy usage of the system.

Index Terms— wireless sensor networks, acoustic sensor networks, multimedia sensor networks, sensor fusion, sensor subset selection, greedy algorithms

This research 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 P6/04 'Dynamical systems, control and optimization' (DYSCO) 2007-2011, Research Project IBBT, Research Project FWO nr. G.0763.12 'Wireless acoustic sensor networks for extended auditory communication'. Alexander Bertrand is supported by a Postdoctoral Fellowship of the Research Foundation Flanders (FWO). The scientific responsibility is assumed by its authors.

1. INTRODUCTION

Resource allocation is a fundamental design challenge in wireless sensor networks (WSN). This is due, in part, to devices being spatially distributed throughout an area and relying on limited resources to perform a certain predefined task. In order to efficiently allocate network resources, algorithms must be developed that are able to determine which subset of signals benefit the system goal the most while utilizing the fewest number of resources possible. As the usage of WSNs for signal estimation has become more prevalent [1], there has been growing interest on exploring subset selection in regards to resource allocation [2, 3].

A wireless acoustic sensor network (WASN) is a collection of microphones that are interconnected through wireless links [4]. Here, a WASN is envisaged that observes a speech signal together with background noise, where the task is to produce an estimate of the speech signal. In a centralized scenario, the microphone signals are sent to a fusion center (FC) where they are used to derive an optimal filter in the linear minimum mean square error (MMSE) sense which takes the form of the well known multi-channel Wiener filter (MWF) [5]. Since the microphones may be distributed over a large area with limited energy resources it is advantageous to limit the total amount of active microphones to a subset in order to extend the lifetime of the network. Unfortunately limiting the network to a subset of microphones results in a decrease in the output signal-to-noise ratio ($\mathrm{SNR}_{\mathrm{out}}$).

The aim of this paper is to use information derived from the unique properties of an assumed rank-1 speech model in the WASN and the individual energy usage of the wireless sensors in order to determine a subset of microphones that offer an acceptable trade-off between the $\mathrm{SNR}_{\mathrm{out}}$ and total transmission energy (E_T). While previous methods have proposed greedy type algorithms based on the total information gain of individual microphones [6, 7], they fail to couple the energy usage of the network in order to facilitate better network resource management.

The problem is comparable to a 0-1 knapsack problem (0-1 KP) that maximizes the overall SNR_{out} while meeting a predefined energy budget. A similar approach was presented in [8] that minimized the transmission energy while keeping the mean square error (MSE) below a certain bound. While both methods are similar, our problem statement presents particular challenges as the actual contribution of a microphone depends on the other microphones that are in the current subset. Therefore a greedy algorithm is used that removes the signal that has the lowest contribution to the overall SNR_{out} compared to its energy usage until the desired energy consumption of the network is met.

The paper is organized as follows. Section 2 describes the problem formulation and notation of the envisaged WASN. Section 3 describes an efficient way to determine the contribution of each signal to the current estimation in terms of full-bandwidth SNR_{out}. In section 4 a greedy algorithm is proposed similarly to a 0-1 KP using the WASN parameters. Simulations are performed in section 5 which show the effect on SNR_{out} while removing signals to reach the desired system energy. Finally in section 6 conclusions are drawn from the simulation data.

2. DATA MODEL AND NOTATION

We assume a spatially distributed set of microphones that collect and transmit their observations to an FC. A signal impinges on each microphone $k \in \{1 \dots M\}$ in the form of

$$y_k(\omega) = x_k(\omega) + n_k(\omega) \tag{1}$$

where x_k is the desired speech component, n_k is the undesired noise component and ω is the frequency bin. The frequency bin ω will be omitted from the following derivations bearing in mind that the operations take place in the frequency domain.

The FC collects the entire M channel signal in a stacked vector $\mathbf{y} = [y_1 \dots y_M]^T$, where T is the transpose operator. An M channel speech vector, $\mathbf{x}$ and noise vector $\mathbf{n}$ are similarly defined. We assume that there is a single speech source, s, hence the speech component in each microphone is represented as

$$\mathbf{x} = \mathbf{a}s \tag{2}$$

where $\mathbf{a}$ is a steering vector that contains information pertaining to the room acoustic transfer functions from the speech source location to the microphones.

The FC performs an MMSE estimate of the desired speech component in a reference microphone which, without loss of generality, is chosen as the first microphone, x_1. The MSE cost function at the fusion center is represented as

$$\begin{aligned} J(\mathbf{w}) &= E\{|x_1 - \mathbf{w}^H\mathbf{y}|^2\} \\ &= E\{|x_1 - \mathbf{w}^H\mathbf{x}|^2\} + E\{|\mathbf{w}^H\mathbf{n}|^2\} \end{aligned} \tag{3}$$

where $E\{\}$ is the expectation operator, H is the conjugate transpose, and it is assumed that the speech and noise components are statistically independent. Alternatively a tuning parameter μ may be added to (3), i.e.,

$$J(\mathbf{w}) = E\{|x_1 - \mathbf{w}^H\mathbf{x}|^2\} + \mu E\{|\mathbf{w}^H\mathbf{n}|^2\} \tag{4}$$

which controls a trade-off between speech distortion and noise reduction. If a single speech source is assumed, the optimal solution in an MMSE sense to (4) is [9],

$$\hat{\mathbf{w}} = \frac{\mathbf{R}_{\mathbf{nn}}^{-1}\mathbf{R}_{\mathbf{xx}}\mathbf{e}_1}{\mu + \text{Tr}\{\mathbf{R}_{\mathbf{nn}}^{-1}\mathbf{R}_{\mathbf{xx}}\}} \tag{5}$$

where $\text{Tr}\{\mathbf{A}\}$ is the trace of the matrix $\mathbf{A}$, $\mathbf{e}_1$ is a vector containing a one in the first entry (corresponding to the reference microphone) and zero otherwise, $\mathbf{R}_{\mathbf{nn}}^{-1}$ is the inverse of the noise correlation matrix $\mathbf{R}_{\mathbf{nn}} = E\{\mathbf{n}\mathbf{n}^H\}$ and $\mathbf{R}_{\mathbf{xx}} = E\{\mathbf{x}\mathbf{x}^H\}$ is the speech correlation matrix. For the ease of exposition we will represent $\text{Tr}\{\mathbf{R}_{\mathbf{nn}}^{-1}\mathbf{R}_{\mathbf{xx}}\}$ as $\text{Tr}\{\mathbf{D}\}$ unless otherwise stated.

Since the speech and noise are assumed to be uncorrelated $\mathbf{R}_{\mathbf{xx}}$ may be estimated by subtracting a noise+speech correlation matrix $\mathbf{R}_{\mathbf{yy}}$, estimated during speech activity, by the noise correlation matrix $\mathbf{R}_{\mathbf{nn}}$, estimated during speech pauses[1], i.e.,

$$\mathbf{R}_{\mathbf{xx}} = \mathbf{R}_{\mathbf{yy}} - \mathbf{R}_{\mathbf{nn}}. \tag{6}$$

3. SIGNAL REMOVAL AND THE EFFECT ON OUTPUT SNR

The SNR_{out} at the FC evaluated at a given frequency bin, ω, is given by the ratio of the variance of the filtered signal to the variance of filtered noise

$$\begin{aligned} \text{SNR}_{\text{out}} &= \frac{E\{|\hat{\mathbf{w}}^H\mathbf{x}|^2\}}{E\{|\hat{\mathbf{w}}^H\mathbf{n}|^2\}} \\ &= \frac{\hat{\mathbf{w}}^H\mathbf{R}_{\mathbf{xx}}\hat{\mathbf{w}}}{\hat{\mathbf{w}}^H\mathbf{R}_{\mathbf{nn}}\hat{\mathbf{w}}}. \end{aligned} \tag{7}$$

Using the rank-1 speech model it has been shown in [9] that SNR_{out} may also be represented as $\text{Tr}\{\mathbf{R}_{\mathbf{nn}}^{-1}\mathbf{R}_{\mathbf{xx}}\}$ or

$$\begin{aligned} \text{SNR}_{\text{out}} &= \text{Tr}\{\mathbf{R}_{\mathbf{nn}}^{-1}\mathbf{R}_{\mathbf{xx}}\} \\ &= \text{Tr}\{\mathbf{R}_{\mathbf{nn}}^{-1}P_s\mathbf{a}\mathbf{a}^H\} \\ &= P_s\mathbf{a}^H\mathbf{R}_{\mathbf{nn}}^{-1}\mathbf{a} \end{aligned} \tag{8}$$

[1] It should be noted that there are better ways to estimate the $\mathbf{R}_{\mathbf{xx}}$, such as using the dominate eigenvector, the described method is only used for its simplicity as it is not the main topic of the paper.

where P_s is the power of the desired speech signal.

The impact that a microphone has on the SNR_{out}, i.e., the reduction that happens when that signal is removed, is used to determine the importance of a microphone to the current estimation. The decrease in SNR_{out} when a signal k is removed from the system can therefore be calculated by monitoring the change in (8), i.e.,

$$\text{SNR}_{\text{out}-k} = \text{Tr}\{\mathbf{D}_{-k}\} \tag{9}$$

where $\text{Tr}\{\mathbf{D}_{-k}\}$ is the trace with signal k removed.

In [7] a computationally efficient expression, $O(M)$, was derived that simultaneously calculates the difference in the trace at a given ω for all signals in the current estimation,

$$[\text{Tr}\{\mathbf{D}_{-1}\}\ldots\text{Tr}\{\mathbf{D}_{-M}\}]^T = \text{Tr}\{\mathbf{D}\}\mathbb{1} - \mathbf{\Lambda}_{NX}^{-1}|\mathbf{\Lambda}_D|^2\mathbb{1} \tag{10}$$

where $\mathbf{\Lambda}_D$ is a diagonal matrix of elements that define the trace, $\mathbf{\Lambda}_{NX}^{-1}$ is a matrix product of the diagonal elements of the inverse noise and speech correlation matrix and $\mathbb{1}$ is a vector with all entries equal to one.

Due to spectral differences in the desired speech and undesired noise components, the signal contributions to the SNR_{out} may differ greatly throughout the frequency bins which makes the decision on which signal to remove an arduous task. Therefore we extend (8) to the full-bandwidth SNR_{out} ($\text{FB-SNR}_{\text{out}}$) so that the contribution each signal makes to the full estimation of the desired speech signal may be known.

In order to determine the impact of the removal of signal k on the $\text{FB-SNR}_{\text{out}}$, the variance of the filtered speech and filtered noise must first by summed over all frequency bins respectively,

$$\text{FB-SNR}_{\text{out}} = \frac{\sum\limits_{\omega=0}^{L-1} E\{|\hat{\mathbf{w}}^H\mathbf{x}|^2\}}{\sum\limits_{\omega=0}^{L-1} E\{|\hat{\mathbf{w}}^H\mathbf{n}|^2\}} \tag{11}$$

where L is the DFT size. The variance of the filtered speech component, $E\{|\hat{\mathbf{w}}^H\mathbf{x}|^2\}$, in a single frequency bin may be expanded using (5), i.e.,

$$\hat{\mathbf{w}}^H\mathbf{R}_{\mathbf{xx}}\hat{\mathbf{w}} = \frac{\mathbf{e}_1^T\mathbf{R}_{\mathbf{xx}}\mathbf{R}_{\mathbf{nn}}^{-1}\mathbf{R}_{\mathbf{xx}}\mathbf{R}_{\mathbf{nn}}^{-1}\mathbf{R}_{\mathbf{xx}}\mathbf{e}_1}{(\mu + \text{Tr}\{\ \mathbf{D}\})^2} \tag{12}$$

which when using the relationship in (8) reduces to

$$\hat{\mathbf{w}}^H\mathbf{R}_{\mathbf{xx}}\hat{\mathbf{w}} = \frac{P_1\text{Tr}\{\ \mathbf{D}\}^2}{(\mu + \text{Tr}\{\ \mathbf{D}\})^2} \tag{13}$$

where $P_1 = P_s|\mathbf{e}_1^T\mathbf{a}|^2$ denotes the speech signal power in the reference microphone.

Likewise the filtered noise variance, $E\{|\hat{\mathbf{w}}^H\mathbf{n}|^2\}$ can also be represented as

$$\begin{aligned}\hat{\mathbf{w}}^H\mathbf{R}_{\mathbf{nn}}\hat{\mathbf{w}} &= \frac{\mathbf{e}_1^T\mathbf{R}_{\mathbf{xx}}\mathbf{R}_{\mathbf{nn}}\mathbf{R}_{\mathbf{nn}}^{-1}\mathbf{R}_{\mathbf{nn}}\mathbf{R}_{\mathbf{xx}}\mathbf{e}_1}{(\mu + \text{Tr}\{\ \mathbf{D}\})^2} \\ &= \frac{P_1\text{Tr}\{\ \mathbf{D}\}}{(\mu + \text{Tr}\{\ \mathbf{D}\})^2}\end{aligned} \tag{14}$$

which when used with (13) and the definition of SNR_{out} (7), reduces to (8). If instead we wish to determine the $\text{FB-SNR}_{\text{out}}$ it can now efficiently be computed as a sum of the powers in the reference microphone and trace products over all frequency bins,

$$\text{FB-SNR}_{\text{out}} = \frac{\sum\limits_{\omega=0}^{L-1} \frac{P_1\text{Tr}\{\ \mathbf{D}\}^2}{(\mu+\text{Tr}\{\ \mathbf{D}\})^2}}{\sum\limits_{\omega=0}^{L-1} \frac{P_1\text{Tr}\{\ \mathbf{D}\}}{(\mu+\text{Tr}\{\ \mathbf{D}\})^2}}. \tag{15}$$

Furthermore the $\text{FB-SNR}_{\text{out}}$ with a signal k removed may be calculated by using the trace with the signal k removed as

$$\text{FB-SNR}_{\text{out}-k} = \frac{\sum\limits_{\omega=0}^{L-1} \frac{P_1\text{Tr}\{\ \mathbf{D}_{-k}\}^2}{(\mu+\text{Tr}\{\ \mathbf{D}_{-k}\})^2}}{\sum\limits_{\omega=0}^{L-1} \frac{P_1\text{Tr}\{\ \mathbf{D}_{-k}\}}{(\mu+\text{Tr}\{\ \mathbf{D}_{-k}\})^2}}. \tag{16}$$

The difference between the current $\text{FB-SNR}_{\text{out}}$ and $\text{FB-SNR}_{\text{out}-k}$ may then be given as

$$\Delta\text{FB-SNR}_{\text{out}-k} = \text{FB-SNR}_{\text{out}-k} - \text{FB-SNR}_{\text{out}}. \tag{17}$$

Since $\text{Tr}\{\ \mathbf{D}_{-k}\}$ can be found simultaneously for each signal left in the estimation, $\text{FB-SNR}_{\text{out}-k}$ may be found with relatively little increase in computationally complexity. Notice that once a signal is removed from the estimation, $\mathbf{R}_{\mathbf{nn}-k}^{-1}$ and $\hat{\mathbf{w}}_{-k}$ must be re-calculated to perform optimal filtering with the remaining signals. It is noted that both values can also be efficiently computed as shown in [6].

4. GREEDY APPROXIMATION

In order to determine the importance of each microphone to the estimation while meeting the network resource allocation constraints, it is necessary to evaluate the amount of information gain of the individual microphones compared to their usage of network resources.

In the envisaged network scheme, the FC maximizes the $\text{FB-SNR}_{\text{out}}$ while also restricting the combined transmission energy of the individual nodes to below a given energy threshold E_T. Microphones that are not used in the estimation are put into sleep mode to reduce the energy usage of the network. Since microphones either transmit or do not transmit to the FC, the problem of which subset to select is an inherent combinatorial optimization problem. This formulation is similar to a 0-1 KP that maximizes the value of a set of objects while ensuring that the sum of the weights of the objects stays below a certain constraint.

The optimal solution to combinatorial optimization problems may be found by using an exhaustive search that finds all 2^M combinations. In order to reduce the computational burden associated with an exhaustive search, especially when the number of microphones is large, we use a sub-optimal approximation often used in 0-1 KPs, which uses a value per weight ratio and employs a greedy method to add or remove elements from the system [10].

In this context, each microphone k is associated with a value representative of the reduction in FB-SNR$_{\text{out}}$ when it is removed from the estimation, $v_k = \Delta$FB-SNR$_{\text{out}-k}$. The FC places these values in a stacked vector of the form

$$\mathbf{v} = [v_1, \ldots, v_M]^T. \tag{18}$$

The microphones are also associated to a weight that is represented by their transmission energy e_k to communicate with the FC. The values in (18) are divided by their transmission energies to produce a value per energy ratio, i.e.,

$$\mathbf{v}_w = [\frac{v_1}{e_1}, \ldots, \frac{v_M}{e_M}]^T. \tag{19}$$

The FC then begins the sensor selection process by removing the microphone that has the lowest contribution or value per weight ratio, $min\{\mathbf{v}_w\}$. The greedy algorithm repeats the process until the combined energy of the remaining sensors is less than E_T.

4.1. Weighted Greedy Approach

In using the proposed greedy method based on (19), sensors with relatively small energy usage may seem to contribute quite heavily to the estimation thereby eliminating nodes that contribute to a higher FB-SNR$_{\text{out}}$ which is empirically shown in section 5. Conversely a greedy method that relies strictly on (18), maximizing FB-SNR$_{\text{out}}$, may utilize nodes that are at a substantial distance from the FC and consume a large amount of energy.

With the purpose of balancing out these two solutions, a relaxation term θ is introduced to the selection process,

$$\mathbf{v}\theta + (1-\theta)\mathbf{v}_w \;\; 0 \leq \theta \leq 1 \tag{20}$$

where $\theta = 0$ maximizes FB-SNR$_{\text{out}}$ with emphasis on minimizing E and $\theta = 1$ will focus on maximizing FB-SNR$_{\text{out}}$ only, which is equivalent to the standard approach. This allows for a more flexible trade-off between SNR performance and network lifetime.

5. SIMULATIONS

An acoustic scenario was simulated with room dimensions of (5x5x5) m. Figure 1 depicts the room with a white noise

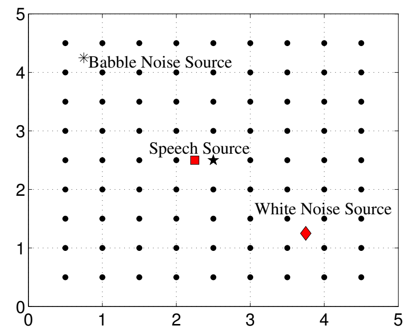

Fig. 1. Simulated room environment.

source ($\Diamond$), a babble noise source ($*$), and a speech source ($\Box$) all placed at a height of 1.5 m. Microphones were placed in a grid pattern 0.5 m away from the walls and every 0.5 m at a height of 1.5 m throughout the room. The reference microphone ($\star$) was at the location (2.5,2.5). The simulation was carried out using a reflection coefficient of 0.4 (T_{60} = 0.16 using Sabine's formula) for all measurements. All processing was done in batch mode on the whole length of the audio signal with a DFT size of $L = 512$. A perfect voice activity detector (VAD) was used so that errors in the estimation in $\mathbf{R_{xx}}$ and $\mathbf{R_{nn}}$ could be neglected.

We used an ideal transmission scheme given in [11] in which the transmission rate is constant for every sensor and delays in the system are ignored. The power required to transmit from sensor k to the FC is then given as

$$P_k(r_k) = Kr_k^{(\alpha)} \tag{21}$$

where K is a constant ($K \approx 10^{-10} J/m^{-\alpha}/bit$), α is a power loss factor (nominally between 2 and 6), and r_k is the distance to the FC. We assume a sensor link capacity, S_k, of 212kbs, which is a typical value for current wireless binaural hearing aid systems [12]. The transmission energy required for each sensor e_k is then given by

$$e_k(r_k, S_k) = KS_kr_k^{(\alpha)}. \tag{22}$$

The FC was placed at the microphone location (0.5,0.5) in figure 1 and the euclidean distance from the fusion center to the other microphones was used for r_k.

The greedy algorithm as described in section 4 was started with a full set of signals and removed the signal that contributed the least to the estimation as defined by (20). The decrease in FB-SNR$_{\text{out}}$ and transmission energy for each microphone were converted to a dB scale. The energies were also scaled by dividing by $min\{e_k\}$. The algorithm terminated the selection process once half of the signals of the total network were removed. Figures 2,3 show the network configuration when half of the nodes have been removed from the system for the limiting scenarios of $\theta = 0$ and $\theta = 1$. As expected $\theta = 0$ weights the network topology

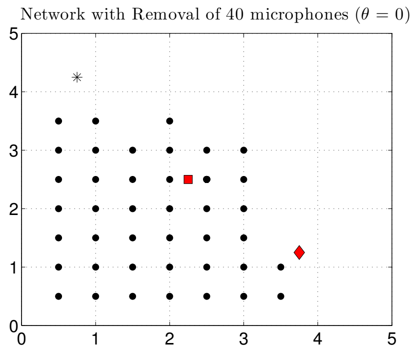

Fig. 2. Network topology with $\theta = 0$.

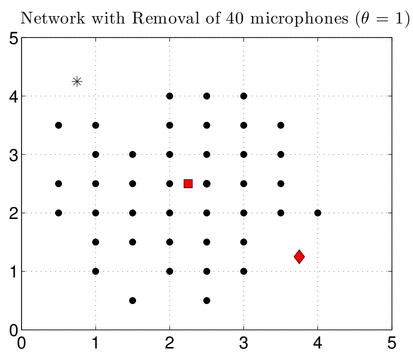

Fig. 3. Network topology with $\theta = 1$.

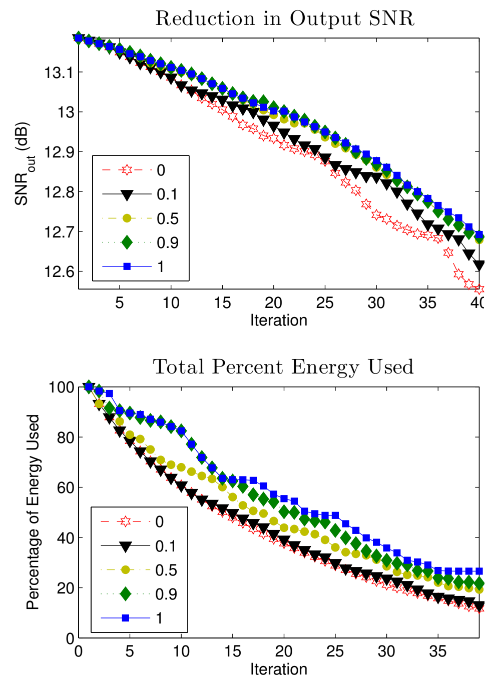

Fig. 4. Difference in FB-SNR_{out} and percentage of energy usage per iteration for $0 \leq \theta \leq 1$.

heavily in favor of a nearest neighbor scenario. This may in fact become a problem if the FC lies somewhere near the noise source as it would effectively remove all nodes that can greatly contribute to the FB-SNR_{out}. On the other extreme $\theta = 1$ the network relies strictly on the FB-SNR_{out} which contains some nodes that are a large distance from the FC.

Figure 4 shows the decrease in SNR_{out} and total percentage of power consumption after each signal removal for varying values of θ. For $\theta = 0.1$ the network achieves a 12% reduction in energy consumption while only losing 0.14 dB in FB-SNR_{out} when compared to $\theta = 1$.

6. CONCLUSIONS

A relaxation term that was related to the energy use was applied to a greedy subset selection algorithm in order to balance output signal-to-noise ratio to energy consumption. A previous method of ranking the signals in terms of their frequency dependent output SNR was extended to a full-bandwidth measurement. This in conjunction with the relaxation term applied to the output SNR/Energy allowed for a noticeable reduction in energy consumption of the network while still maintaining a high level of output SNR.

7. REFERENCES

[1] I.F. Akyildiz, T. Melodia, and K.R. Chowdury, "Wireless multimedia sensor networks: A survey," *Wireless Communications, IEEE*, vol. 14, no. 6, pp. 32 –39, december 2007.

[2] M. Shamaiah, S. Banerjee, and H. Vikalo, "Greedy sensor selection: Leveraging submodularity," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, dec. 2010, pp. 2572 –2577.

[3] H. Rowaihy, S. Eswaran, M. Johnson, D. Verma, A. Bar-noy, and T. Brown, "A survey of sensor selection schemes in wireless sensor networks," in *In SPIE Defense and Security Symposium Conference on Unattended Ground, Sea, and Air Sensor Technologies and Applications IX*, 2007.

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

[5] S.S. Haykin, *Adaptive filter theory*, Prentice-Hall information and system sciences series. Prentice Hall, 2002.

[6] A. Bertrand and M. Moonen, "Efficient sensor subset selection and link failure response for linear MMSE signal estimation in wireless sensor networks," in *Proc. of the European signal processing conference (EUSIPCO)*, Aalborg - Denmark, August 2010, pp. 1092–1096.

[7] J. Szurley, A. Bertrand, and M. Moonen, "Efficient computation of microphone utility in a wireless acoustic sensor network with multi-channel wiener filter based noise reduction," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, March 2012.

[8] H. Godrich, A.P. Petropulu, and H.V. Poor, "Sensor selection in distributed multiple-radar architectures for localization: A knapsack problem formulation," *Signal Processing, IEEE Transactions on*, vol. 60, no. 1, pp. 247 –260, jan. 2012.

[9] M. Souden, J. Benesty, and S. Affes, "On optimal frequency-domain multichannel linear filtering for noise reduction," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 18, no. 2, pp. 260 –276, feb. 2010.

[10] H. Kellerer, U. Pferschy, and D. Pisinger, *Knapsack problems*, Springer, 2004.

[11] D. Ciullo, G. D. Celik, and E. Modiano, "Minimizing transmission energy in sensor networks via trajectory control," in *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2010 Proceedings of the 8th International Symposium on*, 31 2010-june 4 2010, pp. 132 –141.

[12] F. Kuk, B. Crose, T. Kyhn, M. Mrkebjerg, M.L. Rank, M. Nrgaard, and H. Pontoppidan, "Digital wireless hearing aids, part 3: Audiological benefits," *Hearing Review*, vol. 18, no. 8, pp. 48–56, 2011.