GENERALISED SIDELOBE CANCELER — FEATURES AND REALIZATION —

Jan Ingerle
Department of Circuit Theory, Faculty of Electrical Engineering
Czech Technical University in Prague
Technická 2, 166 27 Praha 6, Czech Republic
e-mail: `xingerle@fel.cvut.cz`

Abstract

One of the most often used multi-channel noise suppression structure is Generalised Sidelobe Canceler (GSC). The structure seems to be promising fundament for extended systems with much more satisfying performance than others systems. The detail study of the GSC structure necessary to design new algorithms is summarized in this paper and results of performed experiments and solution of some discovered problems are presented here. The direction of next work is proposed as the conclusion of the paper.
Keywords: beamforming, generalised sidelobe canceler, multi-channel speech enhancement.

1 Introduction

Many multi-channel noise suppression systems developed during last years are based on the Generalised Sidelobe Canceler (GSC) proposed by Griffiths and Jim in [8] as an alternative structure to the Adaptive Beamformer (Frost [4]). The structure was being studied by Simmer in his earlier work (see [5]) and the structure with ability to suppress incoherent as well as coherent noise was developed.

Later studies (e.g. [3]) showed that original GSC has no sufficient ability to reduce coherent noise and new systems based on the GSC and postfiltering were developed. Significant works use several ideas: a combination of GSC version of superdirective beamformer and a postfiltration [2], extension of beamformer by coherence filter [6] or extension by cepstral lifter [7]. Another interesting work is Hoshuyama's and Sugiyama's improvements of the GSC direction dependence behaviour in paper [9].

To develop new efficient multi-channel noise suppression algorithm based on the GSC, it is necessary to make detail study of the system. Therefore the first part of this paper is focused to algorithm description and analysis. The second part of the paper contains comments to realization and some performed experiments results. The direction of next work is outlined in the conclusion of the paper.

2 Algorithm description

The block diagram of the GSC algorithm is in figure 1(a). The GSC consists of three basic part: delay and sum beamformer (DAS) — M microphones with summation block (S), Wiener filter (WF) and sidelobe canceler (SC) — blocking matrix (BM) with adaptive canceler (ANC) composed of L adaptive filters H_i.

The system works in following way: the DAS at the input of the system performs primary amplification of the signal coming from the look-direction (direction orthogonal to microphone array). The DAS output $X(k)$ is processed by the WF. The weights of the WF are counted

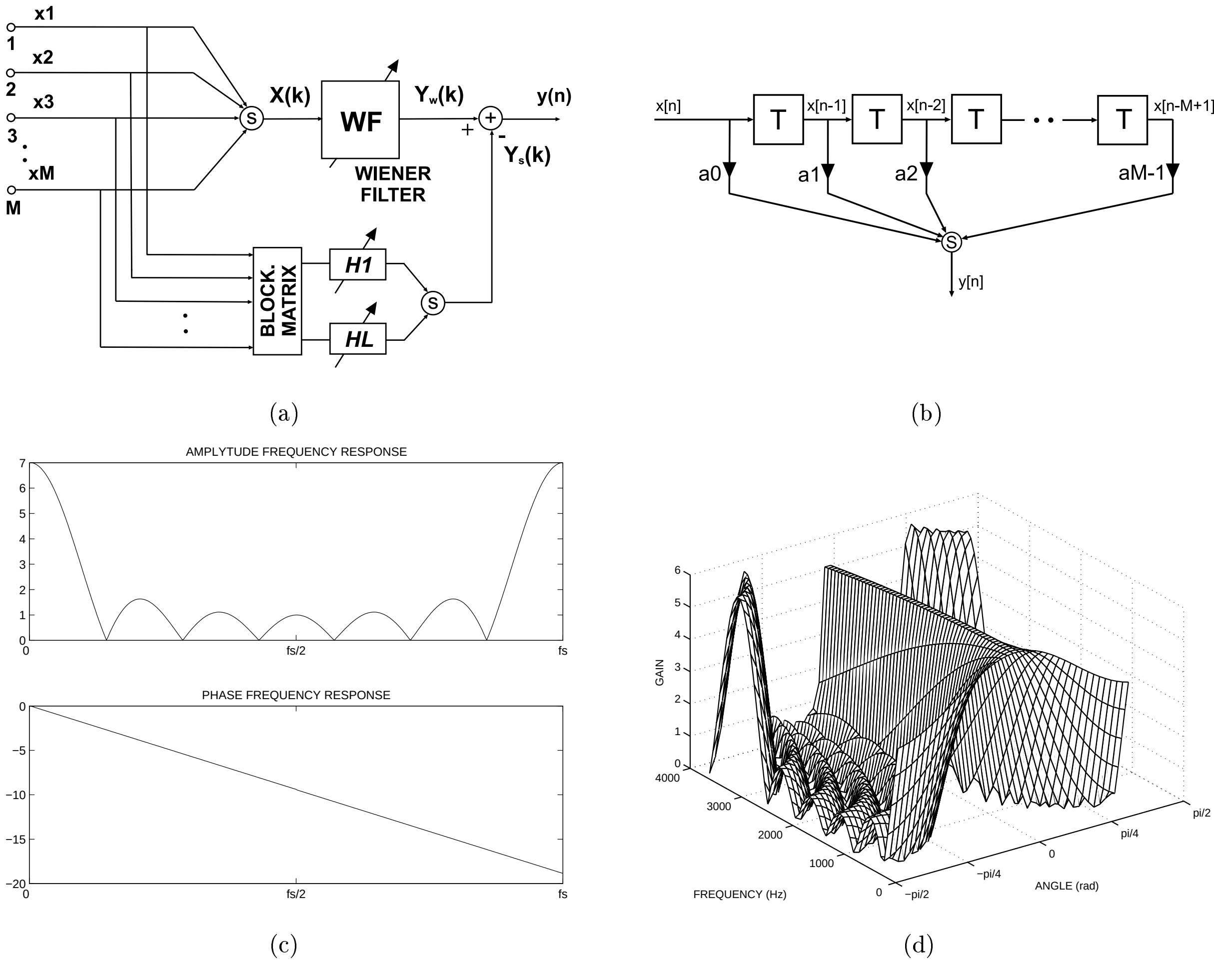


Figure 1: *(a) block diagram of the GSC structure, (b) model of the input part, (c) amplitude and phase frequency response for $\tau = const.$, (d) gain of the input part*

from the WF output $Y_W(k)$ to minimize power of the input signal. The second branch — the BM and the ANC, cancels coherent noise: the BM separate the non look-direction signal from the input signal. This signal is the input of the ANC part. The weights of the ANC are set from the GSC output signal $y(n)$ to model coherent non-look direction noise. Finally the ANC output $Y_s(k)$ is subtracted from the WF output signal to get the GSC output signal $y(n)$.

3 Analysis and realization

The analyses of each part of the system are given in this section. Each analysis contains practice notice to realization of the relevant part.

Delay and sum beamformer

The function of the delay and sum beamformer — DAS, can be, at first approximation, modeled by the finite impulse response filter depicted in figure 1(b). The DAS output can be written as:

$$y[n] = \sum_{l=0}^{M-1} a_j \, x[n - j\,\tau] \tag{1}$$

where a_j are filter weights and τ is time delay of the filter dependent on the incoming signal direction. An example of a frequency response of the filter described by equation 1 for $\tau = \text{const.}$,

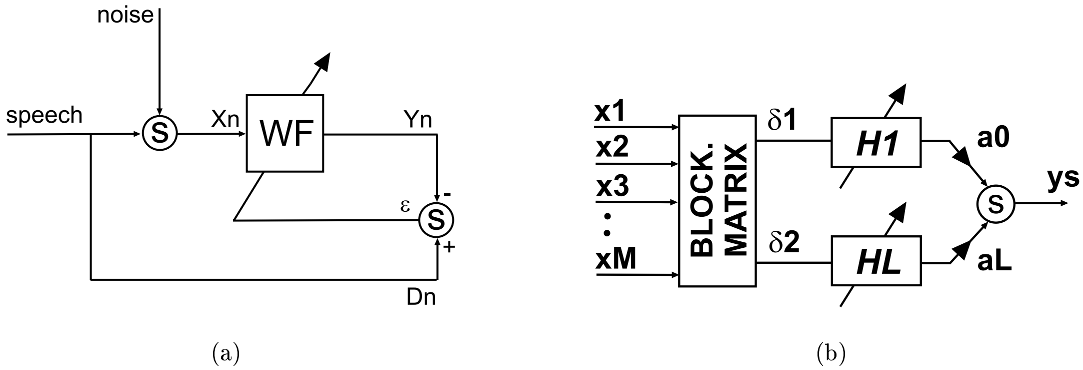


Figure 2: *(a) the Wiener filter block diagram, (b) the SC part of the system*

$a_j = 1$ and $M = 7$ is in the figure 1(c). Another example of a frequency response — as a function of the time delay is in figure 1(d).

It is noticeable from the examples that the input part of the GSC is M-order low-pass filter with incoming signal direction dependent frequency response. The weighs of the filter can be practically set to 1 or to $1/M$. The second variant — $a_j = 1/M$, should be preferred to constrain maximal gain of the filter to 1 what ensures equality of the WF and the ANC branches gains.

Wiener filter

The Wiener filter — WF, realizes MMSE estimation of $X(k)$. The equation for the WF weights can be derived as following (see figure 2(a)):

The power of the error signal:

$$\begin{aligned} \mathrm{E}[\varepsilon^2] &= \mathrm{E}[(d_n - y_n)^2] = \Phi_{dd}(0) + \Phi_{yy}(0) - 2\Phi_{dy}(0) = \\ &= \Phi_{dd}(0) + \mathcal{Z}^{-1}\left\{[W(z^{-1})\Phi_{xx}(z) - 2\Phi_{dx}(z)]W(z)\right\} = \left| W(z) = \sum_{m=0}^{M} w_m z^{-m} \right| = \\ &= \Phi_{dd}(0) + \sum_{l=0}^{M}\sum_{m=0}^{M} w_l w_m \Phi_{xx}(l-m) - 2\sum_{m=0}^{M} w_m \Phi_{dx}(-m), \end{aligned} \tag{2}$$

where E[] is the operator of mean expected value, ε the error signal, $\Phi_{xx}(k)$ the power spectrum density (PSD) of signal $x(n)$, $\Phi_{xy}(k)$ the cross-PSD of signals $x(n)$ and $y(n)$, $W(z)$ the transfer function of the filter, w_m are the weights of the filter and $\mathcal{Z}\{\}$ is operator of z-transformation.

The minimum of the error signal can be counted from derivation of the expression 2:

$$\frac{\partial \varepsilon}{\partial w_k} = 2\sum_{l=-\infty}^{\infty} w_l \Phi_{xx}(k-l) - 2\Phi_{dx}(-k) = 0 \quad \Rightarrow \quad \sum_{l=-\infty}^{\infty} w_l^* \Phi_{xx}(k-l) = 2\Phi_{xd}(k), \tag{3}$$

so that

$$W(z) = \mathcal{Z}\{w_k^*\} = \frac{\Phi_{xd}(z)}{\Phi_{xx}(z)}. \tag{4}$$

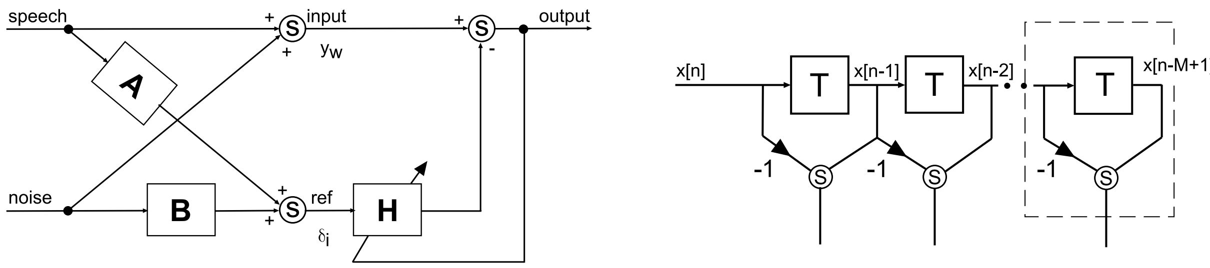
To realize the expression 4 is necessary to assume uncorrelated input noise and no correlation between input noise and useful signal:

$$\Phi_{xd}(z) = \Phi_{dx}(z) = \Phi_{dd}(z) \quad \text{and} \quad \Phi_{xx}(z) = \Phi_{dd}(z) = \Phi_{nn}(z), \tag{5}$$

where d means useful signal, n input noise and x input signal.

The final expression for WF weights can be obtained using equation 4 with assumption 5:

$$W(z) = \frac{|\Phi_{dd}(z)|}{\Phi_{dd}(z) + \Phi_{nn}(z)}. \tag{6}$$



(a) (b)

Figure 3: *(a) basic structure of the ANC, (b) model of the BM considered in text*

The PSDs in the expression 6 have to be procured from the microphone array at the input of GSC. One way how to do it showed Fisher and Simmer in [5] by recursive formulas:

$$\hat{\Phi}^{l}_{dd}(k) = \alpha\,\hat{\Phi}^{l-1}_{dd}(k) + \hat{\Phi}'_{dd}(k) \quad \text{and} \quad \hat{\Phi}^{l}_{xx}(k) = \alpha\,\hat{\Phi}^{l-1}_{xx}(k) + \hat{\Phi}'_{xx}(k), \tag{7}$$

where superscript l denotes number of segment,

$$\hat{\Phi}'_{dd}(k) = \frac{2}{M(M-1)} \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} X_i^*(k) X_j(k) \quad \text{and} \quad \hat{\Phi}'_{xx}(k) = \left| \frac{1}{M} \sum_{j=1}^{M} X_j(k) \right|^2. \tag{8}$$

To improve the WF estimation method proposed by Nuttall and Carter [12] based on smoothing the PSD estimation can be used: The PSD estimation obtained from expression 8 is transferred to time domain, weighted by window

$$w_{lag}(n) = \frac{w_d(n)\ r_{ww}(0)}{r_{ww}(n)}, \tag{9}$$

where w_d is desired window — e.g. Bartlett window, and r_{ww} is auto-correlation function of data-window, and transferred back to frequency domain. This procedure improves the WF filter estimation so that musical tones are reduced.

The necessity of conditions 5 accomplishment leads to the fact that the WF suppresses only non-coherent noise and coherent noise is not touched so no distortion of the input speech is possible in this section.

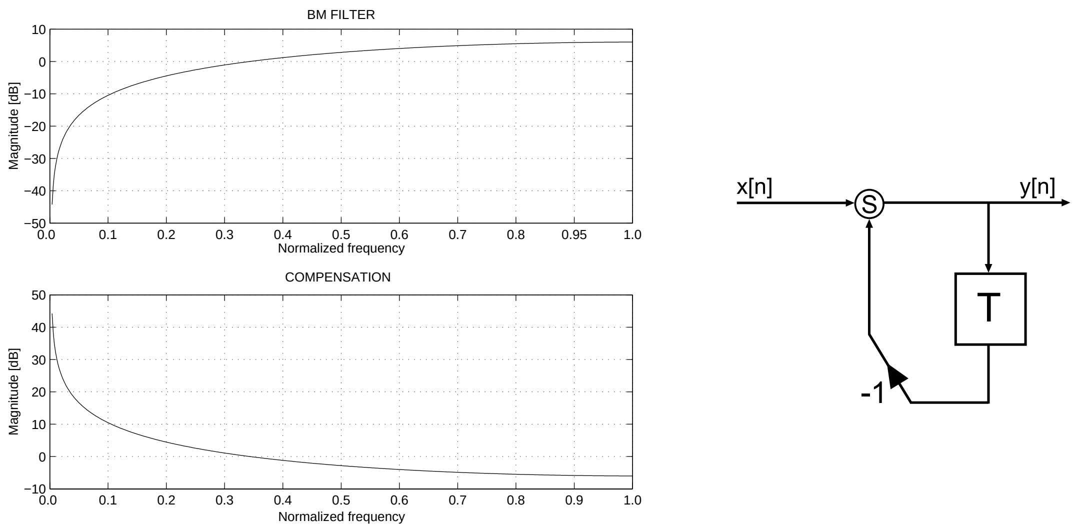
A notice to the WF realisation should be written: the frequency response value of the filter should be in interval $\langle 0; 1 \rangle$. The practical value can be out of this interval due to estimation error of the PDS so it is necessary to constrain it to the relevant values.

Adaptive noise canceler

The ANC branch is in the figure 2(b). The function of this part can be described by the ANC diagram in figure 3(a). The filter H is the Wiener filter and its weights are set from the output by formula 4 completed by expressions for PSD from the diagram:

$$H_i(z) = \frac{\Phi_{\delta_i y_w}(z)}{\Phi_{\delta_i \delta_i}(z)} = \frac{\Phi_{ss}(z)A(z^{-1}) + \Phi_{nn}(z)B(z^{-1})}{\Phi_{ss}(z)\,|A(z)|^2 + \Phi_{nn}(z)\,|B(z)|^2}, \tag{10}$$

where meanings of the symbols are obvious from the diagram. The adaptive filter models noise included in the input part from signal obtained at the reference point. If there is no desired signal at the reference point ($A(z) = 0$) then $H(z) = B^{-1}(z)$ and the modeled noise will be subtracted at the output. If the desired signal appears at the reference point ($A(z) \neq 0$) then model will fail and the desired signal at the output will be distorted. The range of the noise



(a) (b)

Figure 4: *(a) frequency response of the BM and its compensation, (b) compensation filter*

suppression and useful signal distortion can be expressed by signal to noise ratios (SNR) and signal distortion ($D(z)$):

if the SNR at the input is denoted as SNR_{in} and SNR at the reference point as SNR_{ref}:

$$SNR_{in}(z) = \frac{\Phi_{ss}(z)}{\Phi_{nn}(z)} \quad \text{and} \quad SNR_{ref}(z) = \frac{\Phi_{ss}(z)\,|A(z)|^2}{\Phi_{nn}(z)\,|B(z)|^2}, \tag{11}$$

then SNR at the output can be written as:

$$\begin{aligned} SNR_{out}(z) &= \frac{\Phi_{ss}(z)\,|1 - A(z)H(z)|^2}{\Phi_{nn}(z)\,|1 - B(z)H(z)|^2} = \frac{\Phi_{ss}(z)}{\Phi_{nn}(z)} \left| \frac{\Phi_{nn}(z)B(z^{-1})}{\Phi_{ss}(z)A(z^{-1})} \right|^2 = \\ &= \frac{\Phi_{nn}(z)\,|B(z)|^2}{\Phi_{ss}(z)\,|A(z)|^2} = \frac{1}{SNR_{ref}(z)}, \end{aligned} \tag{12}$$

what shows that SNR_{out} nad SNR_{out} are reciprocal.

The signal distortion $D(z)$ defined as the ratio of the useful signal at the output of the Wiener filter to the useful signal at the input can be expressed as

$$D(z) = \frac{\Phi_{ss}(z)\,|A(z)H(z)|^2}{\Phi_{ss}(z)} = \frac{\Phi_{ss}(z)\,|A(z)/B(z)|^2}{\Phi_{ss}(z)} = \left| \frac{A(z)}{B(z)} \right|^2 = \frac{SNR_{ref}(z)}{SNR_{in}(z)}, \tag{13}$$

what indicates that low signal distortion results from high SNR_{in} and low SNR_{ref}.

The accomplishment of the conditions 12 and 13 is realised by the BM separating the non-look direction signal from the input signal so that behind the BM just a noise appears.

Several realisations of the BM were proposed in papers dealing with the GSC. The realisations are based on mutual combination of particular channels (e.g. [8]) or adaptive filtering (e.g. [9]). The BM realised by subtraction of adjacent channels is frequently used due to easy realisation. The mathematical notation of this solution is:

$$\mathbf{BM} = \begin{pmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -1 \end{pmatrix}. \tag{14}$$

	non-correlated wide-band noise			correlated noise 3000Hz			correlated noise 1200 -1300 Hz		
$\mathbf{SNR_{in}}$	$\mathbf{LAR_{in}}$	$\mathbf{LAR_{out}}$	**SD**	$\mathbf{LAR_{in}}$	$\mathbf{LAR_{out}}$	**SD**	$\mathbf{LAR_{in}}$	$\mathbf{LAR_{out}}$	**SD**
20	4.52	4.33	1.79	4.66	4.32	1.78	4.52	4.33	1.81
10	4.61	4.47	1.95	5.24	4.42	1.78	4.55	4.32	1.83
0	4.82	4.46	1.81	6.06	4.68	1.91	4.65	4.30	1.75
-10	5.25	4.72	2.05	7.06	5.55	1.89	4.96	4.46	1.81
-20	5.55	4.83	2.69	7.77	6.64	2.12	5.54	4.74	2.05

Table 1: *LAR of input and output signals and SD as a function of input SNR*

The BM expressed by 14 works well in ideal case. In real environment when the signal wave form is no plane and the look-direction is no right orthogonal to the microphone array the leakage of useful signal through the BM can be observed and the useful signal at the output of the GSC will be distorted.

Realisation of the BM from expression 14 leads to finite response filters model at figure 3(b). This filters represent first order high-pass filters with frequency response at figure 4(a) upper and cause undesired suppression of low frequency in the ANC part. The compensation with frequency response showed in figure 2(b) lower has to be used. It can be realised by inverse filter to the first order FIR filter — first order infinite impulse response filter showed at figure 4(b). The weight of stable filter realisation has to be bigger then -1, e.g. -0.95.

If the DAS summation is weighted (see delay and sum beamformer part), then summation at the output of the ANC can be also realised as weighted summation with weights $a_i = 1/L$ (see figure 2(b)), where L is a number of the ANC channels, to compensate different number of channels in the DAS and the ANC parts. $L = M - 1$ for the BM realised according to formula 14 so that the compensation is not precise but it is sufficient.

4 Simulation and results

The simulations to test the GSC structure were performed with following parameters: number of microphones $M = 7$, sampling frequency $f_s = 8$kHz, segmentation length $L_{seg} = 256$ samples, overlap $L_{over} = 128$ samples and FFT size $N_{FFT} = 512$.

The diffuse character of the input signal was modeled by method described by Allen and Berkley in [1]. This method models impulse response of rooms by image model:

$$p(t, X, X') = \sum_{p=0}^{1} \sum_{r=-\infty}^{\infty} \beta_{x1}^{|n-q|} \beta_{x2}^{|n|} \beta_{y1}^{|l-j|} \beta_{y2}^{|l|} \beta_{z1}^{|m-k|} \beta_{z2}^{|m|} \frac{\delta[t - (|R_p + R_T|/c)]}{4\pi |R_P + R_T|}, \tag{15}$$

where c is speed of sound, β_i are pressure reflex coefficients of relevant walls, $R_p = (x - x' + 2qx'; y - y' + 2jy'; z - z' + 2kz')$, $R_T = 2(nL_x; lL_y; mL_z)$, $L = (L_x; L_y; L_z)$, $X = (x; y; z)$ is talker location, $X' = (x'; y'; z')$ is microphone location and p = (q; j; k) and r = (n; l; m).

The impulse response (IR) for parameters $L = (7.0; 3.5; 2.5)$ meters, $X = (5.0; 0.5; 1.0)$ meters, $X' = (5.0; 3.5; 1.0)$ meters with 0.05m microphones spacing, $\beta = (0.9; 0.9; 0.9; 0.9; 0.7; 0.7)$ and $c = 340\text{ms}^{-1}$ and $f_s = 80$kHz were used to simulate input signal. The input signal of each microphone was composed of speech, wide-band non-correlated noise and narrow-band correlated noise. Each of this components was obtained as a convolution of appropriate signal with IR obtained according to the previous paragraph.

Several criteria to study the GSC behaviour were used. Influence of a noise type and a level of SNR_{in} was watched by spectrograms, Log Area Raio — LAR, and signal distortion — $D(z)$ (see [14]). Also convergence of the adaptive filters were watched.

Some results of the simulations are in table 1. The improvement of speech quality can be noticed from LAR comparison in this table. The signal improvement measured by LAR grows with decreasing SNR_{in}. There is necessary to consider also SD in this case — the signal distortion also grows (see line for $SNR_{in} = -20$dB) and the distortion of useful signal have to be checked by reception tests. The problems with the weights convergence appear if SNR_{in} is low.

The experiments shown following important fact:

At first, there is no sufficient separation of noise and useful signal in the BM so that conditions for good work of the ANC part are not accomplished and the distortion of the output signal can be observed.

At second, problem results from zero-gain points of the DAS frequency response (see figure 1(d))— this points lead to missadjustment of the adaptive filters weights.

The last fact is, that a correlated high power signal in the input signal causes divergence of the ANC weights and destruction of the output signal.

5 Conclusion

The study described in this paper showed features and problems of the GSC. The elimination of mentioned problems and reduction of microphone number M without decrease of performance are goal of next work. Because the main source of the useful signal distortion is the BM failure and consecutively the ANC filters divergence, special attention will be paid to this subject.

A solutions of the goals mentioned in previous paragraph will be approached by

- study of new possibilities of BM implementation. This study will go out from Nordebo, Claesson and Nordholm work [13],
- considering the ANC implementation using coherence filter. The promising result of the coherence method used in beamforming was published by Le Bouguin in [11] and Gunzalez-Rodrigues and Ortega-Garcia in [6],
- possibility of microphone number reduction will be test using method published by Houston in [10].

Acknowledgement

This work is supported by the project: Trans-disciplinary Research in Biomedical Engineering, no. J04/98:210000012 of the Czech Technical University and it is a part of research of the Biological Signal Laboratory at Department of Circuit Theory, FEE, CTU.

References

[1] Allen, J.B., Berkley, D.A.: *Image method for efficiently simulating small-room acoustics*, Journal of Acoustic Society of America, vol.65, no.4, April 1979

[2] Bitzer, J., Simmer, K.U., Kammeyer, K.D.: *Multi-microphone Noise Reduction by Post-Filter and Superdirective Beamformer*, Proceeding of the IWANEC-99

[3] Bitzer, J., Simmer, K.U., Kammeyer, K.D.: *Theoretical Noise Reduction Limits of the Generalized Sidelobe Canceler (GSC) for Speech Enhancement*, Proceedings of the ICASSP-99, vol. 5, pp. 2965-2968, Phoenix, Arizona, March 1999

[4] Frost, III, O.L.: *An Algorithm for Linearly Constrained Adaptive Array Processing*, Proceedings of the IEEE, vol.60, No.8, 1972

[5] Fischer, S., Simmer, K.U.: *Beamforming microphone arrays for speech acquisition in noisy environments*, Speech communication, vol. 20, pp. 215-227, 1996

[6] Gonzalez-Rodrigues, J., Ortega-Garcia, J.: *Coherence-based Decomposition for Efficient Reverberation and Noise Removal in Enclosed Sound Field*, 5th International Conference on Spoken Language Processing ICSLP'98, Sydney, 1998

[7] Gonzalez-Rodrigues, J., Sanchez-Bote, J.L., Ortega-Garcia, J.: *Speech Dereverberation and Noise Reduction with a Combined Microphone Array Approach*, IEEE International Conference on Acoustic, Speech and Signal Processing, ICASSP-00, Istambul, June 2000

[8] Griffiths, L.J., Jim, Ch.W.: *An Alternative Approach to Linearly Constrained Adaptive Beamforming*, IEEE Transaction on Antennas and Propagation, vol. AP-30, No.1, January 1982

[9] Hoshuyama, O., Sugiyama, A.: *Robust Adaptive Beamforming*, Brandstein, M., Ward, D. edt.: Microphone Arrays, Signal Processing Techniques and Applications, Springer, Berlin 2001, ISBN 3-540-41953-5

[10] Houston, K.M.: *A Fast Beamforming Algorithm*, Proceedings of the 1994 IEEE Oceans Conference, Part 1, pp. 211-216, 1994

[11] Le Bouquin, R.: *Enhancement of noisy speech signals: Application to mobile radio communications*, Speech Communication, vol.18, pp. 3-19, 1996

[12] Nuttall, A.H., Carter, G.C.: *Spectral Estimation Using Combined Time and Lag Weighting*, Proceedings of the IEEE, vol.70, no.9, September 1982

[13] Nordebo, S., Claesson, I., Nordholm, S.: *Adaptive Beamforming: Spatial Filter Designed Blocking Matrix*, IEEE Journal of Oceanic Engineering, vol.19, no.4, October 1994

[14] Widrow, B., Stearns, S.D.: *Adaptive Signal Processing*, A.V. Oppenheim (eds.), Prentice-Hall, 1985, ISBN 0-13-004029-0