Analysis and Simulation of Frost's Beamformer

Miroslav ŠTRUPL, Pavel SOVKA

Dept. of Circuit Theory, Czech Technical University, Technická 2, 166 27 Praha 6, Czech Republic

struplm@feld.cvut.cz, sovka@feld.cvut.cz

Abstract. *Sensor arrays are often used for a signal separation from noises using the information about the direction of arrival. The aim of this paper is to analyze Frost's beamformer with respect to the speech preprocessing for the hearing impaired people. The frequency response of the system including the background noise attenuation are derived as functions of the direction of arrival. The derivation supposes a uniform linear array of sensors and plane waves. It is shown that the number of possible configurations can be decreased by using some symmetries. The impact of the used algorithm constraint on the frequency response and subsequently on the directional noise suppression is derived analytically.*

Keywords

Adaptive arrays, Frost's beamformer, speech enhancement, noise reduction.

1. Introduction

This article deals with the application of Frost's beamformer for the background noise reduction. Frost's beamformer belongs to the broad field of beamforming systems, see. e.g. [7], [8], [9]. The application of this beamforming system discussed in this paper is the speech preprocessing for hearing impaired people. Other possible applications of this system can be speech processing for conferencing and a mobile telephony (direct sequence CDMA) [1], where an array of sensors helps to discriminate between signal and noises sources. For example, a hearing impaired person using a small microphone array mounted on glasses is able to discriminate between different speakers. In these applications the broadband rather than narrowband input signals should be assumed. Thus the frequency response and SNRE (Signal to Noise Ratio Enhancement) evaluated over the whole frequency band have to be used.

The following restrictions about an environment, a geometry of array and input signals are made.

- In order to obtain the frequency response easily the environment should be homogeneous, isotropic, lossless, nondispersive with respect to the wave propagation (phase) velocity.
- The array of sensors are assumed to be a linear with uniformly spaced sensors. Sensors are assumed to have omnidirectional characteristics.
- Both the directional signals and noises are assumed to propagate with the plane wave front. This assumption is not too restrictive because spherical waves generated in a far field from the array can be approximated with plane waves.
- Noises are supposed to be either directional (forming the plane wave - see preceding item) or uncorrelated between sensors.
- Frequency response is derived under assumption of linearity and time invariance - LTI (no coefficients update) of the filter.

The paper is organized as follows. In section 2 Frost's beamformer is described including a linear constraint. In section 3 its frequency response for a plane wave input is derived. In section 4 two symmetries of configurations are proved. In section 5 properties of the frequency response caused by the constraint are discussed. In section 6 the influence of a noise uncorrelated between sensors and samples is shown. Finally, the conclusion is given.

2. Frost's Beamformer

Frost's beamformer [2] (see Fig. 1**a**) consists of an array with K sensors, where each sensor is followed by a transversal filter with J weights. The number of weights is equal for all transversal filters. The sum of the filter outputs is the beamformer output. Weights are updated by Frost's constrained least mean square (CLMS) algorithm which minimizes the mean square error of the output signal while satisfying a constraint [2]. In order the input signal $s(t)$ (arriving to all sensors with the same delay) to be passed without any distortion, the impulse response of the whole system must be equal to the unit impulse. This impulse response represents the constraint for the weights of all filters. The whole system can be replaced by one transversal FIR filter for the signal $s(t)$. The replacement is shown in Fig. 1b, where $f_1, f_2, \ldots, f_J$ is the impulse response for the signal. Constraint equations (see Fig. 1b) can be written also in matrix form

$$\mathbf{W}\begin{bmatrix}1\\1\\\vdots\\1\end{bmatrix}=\begin{bmatrix}f_1\\f_2\\\vdots\\f_J\end{bmatrix}, \tag{1}$$

where $\mathbf{W}$ stands for weight matrix with real elements

$$\mathbf{W}=\begin{bmatrix}w_1 & w_2 & \dots & w_K\\ w_{K+1} & w_{K+2} & \dots & w_{2K}\\ \vdots & \vdots & \ddots & \vdots\\ w_{JK-K+1} & w_{JK-K+2} & \dots & w_{JK}\end{bmatrix}. \tag{2}$$

To discuss the Frost's beamformer behavior in details, let us define some terms needed. The digitized input noisy signals $x_i[n], i = 1, 2, \dots, JK$ are formed by components of both clean signal $s(t)$ and noise $n(t)$. The vector $\tilde{\mathbf{x}}[n]$ represents noisy signals on taps, the vector $\mathbf{w}$ consists of weights value, the vector $\mathcal{F}$ represents the constrained impulse response and the matrix $\mathbf{C}$ will be used in constraint formulation

$$\begin{aligned}\tilde{\mathbf{x}}^T[n] &= [\ x_1[n]\quad x_2[n]\quad \dots\quad x_{JK}[n]\ ],\\ \mathbf{w}^T &= [\ w_1\quad w_2\quad \dots\quad w_{JK}\ ],\\ \mathcal{F}^T &= [\ f_1\quad f_2\quad \dots\quad f_J\ ],\\ \mathbf{C} &= [\ \mathbf{c}_1\quad \mathbf{c}_2\quad \dots\quad \mathbf{c}_J\ ].\end{aligned} \tag{3}$$

Elements $\mathbf{c}_i$ are the column vectors of length JK with $(i-1)K$ zeros followed by K ones and $(J-i)K$ zeros

$$\mathbf{c}_i^T = [\underbrace{\ 0\quad 0\quad \dots\quad 0\ }_{(i-1)K\ zeros}\underbrace{\ 1\quad 1\quad \dots\quad 1\ }_{K\ ones}\underbrace{\ 0\quad 0\quad \dots\quad 0\ }_{(J-i)K\ zeros}]. \tag{4}$$

Now the problem of finding the optimum weight vector for a stationary signal $\mathbf{w}_{opt}$ (Wiener solution) can be formulated. The weight vector minimizing $E[y^2[n]] = \mathbf{w}^T E[\tilde{\mathbf{x}}[n]\tilde{\mathbf{x}}[n]^T]\mathbf{w} = \mathbf{w}^T\mathbf{R}_{xx}\mathbf{w}$ and satisfying the constraint $\mathbf{C}^T\mathbf{w} = \mathcal{F}$ have to be found. $\mathbf{R}_{xx}$ stands for the autocorrelation matrix. In [2] the method of Lagrange multipliers was used to obtain the Wiener solution

$$\mathbf{w}_{opt} = \mathbf{R}_{xx}^{-1}\mathbf{C}(\mathbf{C}^T\mathbf{R}_{xx}^{-1}\mathbf{C})^{-1}\mathcal{F} \tag{5}$$

and the adaptive CLMS algorithm

$$\begin{aligned}\mathbf{w}[0] &= \mathbf{f},\\ \mathbf{w}[n+1] &= \mathbf{P}(\mathbf{w}[n] - \mu y[n]\tilde{\mathbf{x}}[n]) + \mathbf{f}.\end{aligned} \tag{6}$$

The vector $\mathbf{f}$ and the projection matrix $\mathbf{P}$ are defined as

$$\begin{aligned}\mathbf{f} &= \mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1}\mathcal{F},\\ \mathbf{P} &= \mathbf{E} - \mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T.\end{aligned} \tag{7}$$

Positive scalar μ is a step-size parameter. The choice of μ is the tradeof between the convergence time and the missadjustment of weights from Wiener solution. An easily computable upper bound for μ is given by $\mu < 2/(3E[\tilde{\mathbf{x}}^T\tilde{\mathbf{x}}])$.

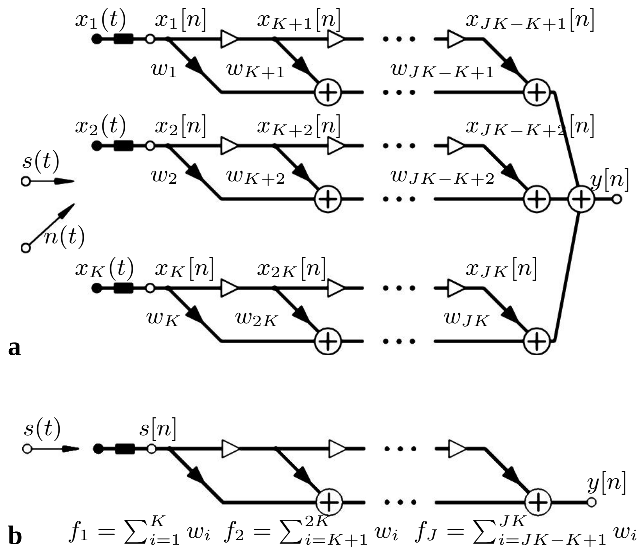

Fig. 1. **a** Frost's beamformer structure, **b** Frost's beamformer from $s(t)$ view - constraint formulation

The convergence behavior and the choice of μ is deeply discussed in [2].

The alternative form of equation (6) for the implementation is

$$w_i[n+1] = w_i[n] - \mu y[n]x_i[n] - \frac{1}{K}\sum_{j=\lfloor\frac{i}{K}\rfloor K+1}^{(\lfloor\frac{i}{K}\rfloor+1)K}(w_j[n] - \mu y[n]x_j[n]) + \frac{f_{\lfloor\frac{i}{K}\rfloor+1}}{K}. \tag{8}$$

3. Frequency Response Analysis

Now the frequency response for the LTI case (fixed weights, fixed geometry, ...) and plane wave fronts will be derived. Since there are digital filters behind sensors, it will be more convenient for the analysis to replace the transfer paths from the source to sensors by a discrete time model. Thus the variable t will be replaced by the nT_s, where T_s is the sampling period used for filters behind sensors. The main reason is to understand the constraint influence on the system behavior.

Now the coordinate system used for the sources description will be given (see Fig. 2). Note that the origin is placed in the center of the array and *y*-axis goes through all the sensors of the array, which are spaced uniformly with the distance d and are numbered from 0 to $K-1$. The sensors are assumed to have omnidirectional characteristics. Due to the rotation symmetry of the array (about *y*-axis) every three-dimensional scenario with arbitrary plane wave sources can be described in $x \in < 0, +\infty)$, $y \in (-\infty, +\infty)$ half-plane (signals at sensors will be the same). Thus without the loose of generality only wave fronts generated by sources lying in this half-plane and having normal vectors in *xy*-plane can be considered. Let us assume one source generating a plane wave with a waveform $u(t)$. Because the normal vector of a wave front at the source lies in *xy*-plane, the whole wave front is projected onto a single line (depicted by the thick line labeled by $u(t)$). Thus for the source description it is sufficient to give the position $\mathbf{r}$ of the nearest point of this wave front to the origin (note that the vector $\mathbf{r}$ is perpendicular to the wave front). It seems to be the most appropriate to express $\mathbf{r}$ in polar coordinates by giving angle $\varphi \in < -\frac{\pi}{2}, \frac{\pi}{2} >$ and distance $r \in \mathcal{R}^+$. The distance r can be measured using delay between the source and the center of the array $R = \frac{r}{T_s v}$. The direction of arrival can be also measured using delay C (in samples) between adjacent sensors. C relates to φ according to

$$C = \frac{d \sin\varphi}{T_s v}, \tag{9}$$

where d denotes the distance between adjacent sensors, v the speed of sound. Due to the assumed environment (homogeneous, isotropic, lossless, nondispersive) $u(t)$ reaches the sensor m with the delay $T_s[R + C(m - \frac{K-1}{2})]$ without any distortion. Thus the frequency response (measured from the source to a sensor m) is $\hat{P}_m(\omega) = e^{-j\omega T_s[R+C(m-\frac{K-1}{2})]}$, where ω stands for the angular frequency. To obtain corresponding frequency response P_m of the discrete time model $\hat{P}_m$ has to be restricted to the range $(-\frac{\omega_s}{2}, \frac{\omega_s}{2})$ (antialiasing filters must be placed behind sensors). Using the normalized frequency $\Theta = \omega T_s$ for $\omega \in (-\frac{\omega_s}{2}, \frac{\omega_s}{2})$ and the periodic extension of the frequency interval we obtain

$$\forall k \in \mathcal{Z}, \forall \Theta \in < -\pi, \pi >: \quad P_m(\Theta + 2k\pi) = e^{-j\Theta[R+C(m-\frac{K-1}{2})]}. \tag{10}$$

It is convenient to use the normalized frequency when describing a digital system. This approach assures the independence of the system description on a sample period T_s (or on a sample frequency $f_s = \frac{1}{T_s}$). The frequency response for the whole system (see Fig. 1**a**) can be derived using frequency responses of transversal filters and (10)

$$\begin{aligned}
&\forall k \in \mathcal{Z}, \forall \Theta \in < -\pi, \pi >: \quad H(\Theta + 2k\pi) = \\
&= \sum_{m=0}^{K-1} e^{-j\Theta[R+C(m-\frac{K-1}{2})]} \Big(w_m + w_{K+m} e^{-j\Theta} + \dots \\
&\quad + w_{K(J-1)+m} e^{-j(J-1)\Theta} \Big) = \\
&= e^{-j\Theta R} e^{j\Theta C(\frac{K-1}{2})} \\
&\quad \cdot \begin{bmatrix} 1 & e^{-j\Theta} & \dots & e^{-j(J-1)\Theta} \end{bmatrix} \mathbf{W} \begin{bmatrix} 1 \\ e^{-jC\Theta} \\ \vdots \\ e^{-j(K-1)C\Theta} \end{bmatrix},
\end{aligned} \tag{11}$$

where the matrix $\mathbf{W}$ is defined in (2). Now the power gain AP will be derived for a system with the frequency response given by (11). The input waveform $u[n]$ of the system is assumed to be the white stationary process with zero mean and variance σ_u. Let us define AP as the ratio of expectations of squared output and input

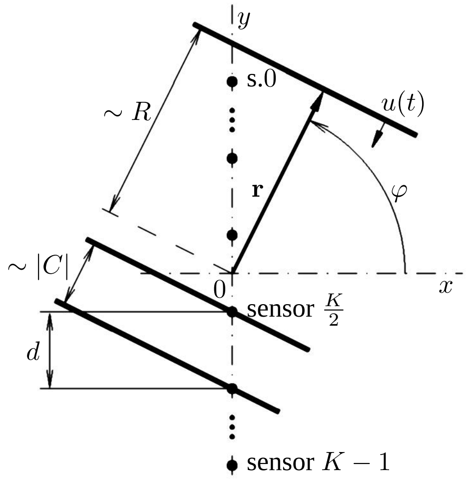

Fig. 2. Configuration used to derive frequency response (depicted for even number of sensors)

$$\begin{aligned}
AP &= \frac{E[v^2[n]]}{E[u^2[n]]} = \\
&= \frac{1}{\sigma_u^2} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} h[k]h[l]E[u[n-k]u[n-l]] = \\
&= \sum_{k=-\infty}^{\infty} h^2[k] = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H(\Theta)|^2 \mathrm{d}\Theta,
\end{aligned} \tag{12}$$

where $h[n]$ and $v[n]$ are the impulse response of the system and the system output respectively (Parceval's theorem was used in the last step). The $H(\Theta)$ should be written in more suitable form than (11) to proceed with the integration

$$\forall \Theta \in < -\pi, \pi >: \quad H(\Theta) = \sum_{i=1}^{JK} w_i e^{-jl(i)\Theta}. \tag{13}$$

The $l(i)$ is a new function expressing the delay connected with the weight w_i [1]

$$l(i) = C\big((i-1) - K\lfloor\frac{i-1}{K}\rfloor\big) + \lfloor\frac{i-1}{K}\rfloor. \tag{14}$$

We can continue with the integration of equation (12)

$$\begin{aligned}
AP &= \frac{1}{4\pi} \int_{-\pi}^{\pi} \big(H(\Theta)H(\Theta)^* + (H(\Theta)H(\Theta)^*)^*\big) = \\
&= \frac{1}{2\pi} \sum_{r=1}^{JK} \sum_{s=1}^{JK} w_r w_s \int_{-\pi}^{\pi} \cos([l(r) - l(s)]\Theta) \mathrm{d}\Theta = \\
&= \sum_{r=1}^{JK} \sum_{s=1}^{JK} w_r w_s \frac{\sin([l(r) - l(s)]\pi)}{[l(r) - l(s)]\pi}.
\end{aligned} \tag{15}$$

4. Configurations

The verification of the system behavior by the simulation cannot be carried out for all the possible configurations.

[1] An arbitrary sensor placement could be expressed by changing $l(i)$ definition

Thus the number of configurations have to be decreased. This can be done by using some types of equivalences. The most convenient choice seems to be the equivalence with regard to generated output, because it seems to be invaluable to simulate set of configurations giving the same output. Of course, the precise definition depends on what we are interested in. Such the equivalence is moreover conditioned by the equality of input signals usually. This means, the system is considered to be a black box with inputs and one output. Needed definitions will be introduced now.

Definition 1 *The configuration is an ordered triplet of input signals used, transfer functions from sources to sensors (a function of an environment, a placement of sources and sensors and generated wave fronts), the filter behind sensors (the filter is described by the filter order and weight values, filtration and weight update equations).*

Definition 2 *Configurations K and K' are equivalent $K \equiv K'$ if and only if K and K' having the same input signals generate the same output* $\forall n \in \mathcal{Z}^+ : \; y[n] = y'[n]$.

4.1 Equivalent Configurations with Different Distance between Sensors

The frequency response from the plane wave source to the chosen sensor (10) depends only on the delay C and neither on the angle φ nor on the distance between adjacent sensors d. Consider the configuration K with the adjacent sensor distance d, s plane wave sources with directions of arrival given by angles $\varphi_1, \varphi_2, \ldots \varphi_s$ leaving delays $C_1, C_2, \ldots C_s$. Let us construct the configuration K' from K by changing d to d' and choosing $\varphi_1', \varphi_2', \ldots \varphi_s'$ under the condition $C_1' = C_1, C_2' = C_2, \ldots C_s' = C_s$. Thus the frequency response (10) will be the same. Therefore the filter behind sensors gets the same input and generates the same output and consequently $K \equiv K'$ according Def. 2. But how do we chose φ_i' in order to C_i remain the same? Equation (9) must hold for such φ_i: $C_i = \frac{d' \sin \varphi_i'}{T_s v}$. Since φ_i is restricted to $< -\frac{\pi}{2}, \frac{\pi}{2} >$, preceding relation between C_i and φ_i' can be rewritten in the form

$$\varphi_i' = \arcsin \frac{v C_i T_s}{d'} \qquad i = 1, 2, \ldots s. \tag{16}$$

4.2 Symmetry of Filters behind Sensors

Consider two configurations K, K' which are identical except for directions of arrival of plane waves $\varphi_1' = -\varphi_1$, $\varphi_2' = -\varphi_2$, $\ldots$ $\varphi_s' = -\varphi_s$. Thus from (9) $C_1' = -C_1$, $C_2' = -C_2$, $\ldots C_s' = -C_s$. Then equation (10) must hold for spectra of waveforms on sensors

$$\begin{aligned}
\forall k \in \mathcal{Z}, \forall \Theta \in < -\pi, \pi >: \quad & X_{K-1-m}(\Theta + 2k\pi) = \\
&= \sum_{i=1}^{s} e^{-j\Theta C_i (K-1-m)} e^{j\Theta C_i \frac{K-1}{2}} e^{-j\Theta R_i} U_i(\Theta) \\
&= \sum_{i=1}^{s} e^{-j\Theta (-C_i) m} e^{j\Theta (-C_i) \frac{K-1}{2}} e^{-j\Theta R_i} U_i(\Theta) \\
&= X_m'(\Theta + 2k\pi).
\end{aligned} \tag{17}$$

It can be concluded, that waveforms on sensors are the same except for the permutation

$$x_0'[n] = x_{K-1}[n], x_1'[n] = x_{K-2}[n], \ldots x_{K-1}'[n] = x_0[n]. \tag{18}$$

If a noise uncorrelated between sensors and samples is present, its components on sensors should satisfy (18) in order to waveforms on sensors satisfy (18).

If there is the pure summation behind sensors (delay and sum beamformer), the permutation takes no effect due to the commutativity of addition. Thus generated outputs are the same in both configurations and $K \equiv K'$ from Def. 2.

The case of the filter driven by Frost's CLMS algorithm will be discussed now. It can be stated after looking at the filter structure Fig. 1**a** and considering the property (18), that $\tilde{\mathbf{x}}'[n]$ is the permutated version of $\tilde{\mathbf{x}}[n]$ (at most after $J-1$ iterations). The permutation occurs within each column. The permutation vector $\boldsymbol{\pi}$ and projection π performing that permutation are defined for easier manipulation

$$\begin{aligned}
\boldsymbol{\pi} = [\pi_1, \ldots \pi_{JK}]^T = [&K, K{-}1, \ldots 1, 2K, 2K{-}1, \ldots K{+}1 \\
&\ldots JK, JK-1, \ldots JK-K+1]^T,
\end{aligned}$$

$$\pi([x_1 \;\; x_2 \;\; \ldots \;\; x_{JK}]^T) = [x_{\pi_1} \;\; x_{\pi_2} \;\; \ldots \;\; x_{\pi_{JK}}]^T. \tag{19}$$

Let us assume that K and K' initial conditions fulfill the described property

$$\tilde{\mathbf{x}}'[0] = \pi(\tilde{\mathbf{x}}[0]), \qquad \mathbf{w}'[0] = \pi(\mathbf{w}[0]). \tag{20}$$

The usual initial conditions (zero $x_i[0]$ values and weight setup $\mathbf{f}$) satisfy (20), because these initial conditions have the same values within each column. The observation made about $\tilde{\mathbf{x}}'[n]$ and $\tilde{\mathbf{x}}[n]$ components now holds for every iteration

$$\forall n \in \mathcal{Z}^+ : \qquad \tilde{\mathbf{x}}'[n] = \pi(\tilde{\mathbf{x}}[n]). \tag{21}$$

In order to show $K \equiv K'$, the following conditions remain to be proved

$$\forall n \in \mathcal{Z}^+ : \qquad y'[n] = y[n], \qquad \mathbf{w}'[n] = \pi(\mathbf{w}[n]). \tag{22}$$

The induction scheme will be followed here. The desired equation for the weight vector is included in (20) for $n = 0$. Using (20) following equation can be written for outputs

$$y'[0] = \mathbf{w}'[0]^T \tilde{\mathbf{x}}'[0] = \pi(\mathbf{w}[0]^T) \pi(\tilde{\mathbf{x}}[0]) = y[0]. \tag{23}$$

The last equality follows from the fact that if components of both vectors are equally permutated, then scalar multiplication leaves the same result due to the commutativity of addition. Now suppose that the property holds for some $n \geq 0$

$$y'[n] = y[n], \quad w_i'[n] = w_{\pi_i}[n]. \tag{24}$$

Notice that any column number (from 0 to $J-1$) for a weight w_i could be expressed by $\lfloor \frac{i}{K} \rfloor$. The updating equation (8) for a weight w'_i can be written

$$w'_i[n+1] = w'_i[n] - \mu y'[n]x'_i[n] - \frac{1}{K} \sum_{j=\lfloor \frac{i}{K} \rfloor K+1}^{(\lfloor \frac{i}{K} \rfloor +1)K} (w'_j[n] - \mu y'[n]x'_j[n]) + \frac{f_{\lfloor \frac{i}{K} \rfloor +1}}{K}. \quad (25)$$

Using the substitution from (24) and (21) into (25) leads to (notice that $\lfloor \frac{i}{K} \rfloor = \lfloor \frac{\pi_i}{K} \rfloor$)

$$w'_i[n+1] = w_{\pi_i}[n] - \mu y[n]x_{\pi_i}[n] - \frac{1}{K} \sum_{j=\lfloor \frac{\pi_i}{K} \rfloor K+1}^{(\lfloor \frac{\pi_i}{K} \rfloor +1)K} (w_{\pi_j}[n] - \mu y[n]x_{\pi_j}[n]) + \frac{f_{\lfloor \frac{\pi_i}{K} \rfloor +1}}{K}. \quad (26)$$

Because the permutation occurs only within one column and the summation goes through the whole column, summing non permutated version leaves the same result

$$w'_i[n+1] = w_{\pi_i}[n] - \mu y[n]x_{\pi_i}[n] - \frac{1}{K} \sum_{j=\lfloor \frac{\pi_i}{K} \rfloor K+1}^{(\lfloor \frac{\pi_i}{K} \rfloor +1)K} (w_j[n] - \mu y[n]x_j[n]) + \frac{f_{\lfloor \frac{\pi_i}{K} \rfloor +1}}{K} = w_{\pi_i}[n+1]. \quad (27)$$

The last equality was obtained using (8). This can be rewritten using vector notation

$$\mathbf{w}'[n+1] = \pi(\mathbf{w}[n+1]). \quad (28)$$

Using (21),(28) and following the same reasoning as in obtaining (23) the equation for $y'[n+1]$ can be written

$$y'[n+1] = \mathbf{w}'[n+1]^T \tilde{\mathbf{x}}'[n+1] = = \pi(\mathbf{w}[n+1]^T)\pi(\tilde{\mathbf{x}}[n+1]) = y[n+1]. \quad (29)$$

This completes the induction step. Equations (22) are proved and through Def. 2 $K \equiv K'$.

The consequences of introduced properties are discussed now for the uniform linear array, plan waves, the filter behind sensors driven by CLMS and the signal coming from $\varphi = 0°$. If one directional noise source is present (noise uncorrelated between sensors and samples may be present too), then it is sufficient to simulate configurations with noise directions φ from $0°$ to $90°$, because other configurations are equivalent with them due to 4.2 (see sect. Symmetry of Filters behind Sensors). Obtained results can be further generalized to different distances between sensors according to 4.1 (see sect. Equivalent Configurations with Different Distance between Sensors). If there are an arbitrary number of noise sources, then it is sufficient to simulate system for directions φ between $-90°$ and $90°$. Obtained results can be generalized by using 4.1 similarly as in previous case. Non zero distance r (proportional to delay R) causes only superposition of the linear trend to the phase of frequency response (see (11)). Thus it seems sometimes to be uninteresting to consider different r.

5. Constraint Influence

Now some properties of the frequency response for plane wave inputs (11) will be discussed. If the elements of the rightmost vector (11) became ones for some Θ, then the resulting expression simplifies greatly. But for which normalized frequency Θ this simplification can be made? This is satisfied for

$$\Theta = \frac{2\pi}{C} l \quad l \in \mathcal{Z}, |l| \leq \lfloor \frac{C}{2} \rfloor. \quad (30)$$

Now substitution from (1) to (11) for Θ given by (30) can be made

$$H(\frac{2\pi}{C} l + 2k\pi) = = e^{j\frac{2\pi}{C} l C(\frac{K-1}{2})} e^{-j\frac{2\pi}{C} l R} \begin{bmatrix} 1 & e^{-j\frac{2\pi}{C} l} & \dots & e^{-j\frac{2\pi}{C} l(J-1)} \end{bmatrix} \cdot \begin{bmatrix} f_1 & f_2 & \dots & f_J \end{bmatrix}^T = e^{j\frac{2\pi}{C} l \left(C(\frac{K-1}{2}) - R\right)} \sum_{i=0}^{J-1} f_{i+1} e^{-j\frac{2\pi}{C} li} \quad l \in \mathcal{Z}, |l| \leq \lfloor \frac{C}{2} \rfloor. \quad (31)$$

If f_i is the unit impulse, then this expression further simplifies to

$$H(\frac{2\pi}{C} l + 2k\pi) = e^{j\frac{2\pi}{C} l \left(C(\frac{K-1}{2}) - R\right)} \quad l \in \mathcal{Z}, |l| \leq \lfloor \frac{C}{2} \rfloor. \quad (32)$$

It can be concluded that due to the imposed constraint the frequency response is fixed at some normalized frequencies. These normalized frequencies are determined by the direction C (30). Fig. 3 illustrates this for some directions C when the unit impulse constraint is assumed (black dots stands for fixed points). Using (30) it can be concluded about normalized frequencies fixed by the constraint (assuming one half period of frequency response):

1) The zero normalized frequency is always included.

2) Fixed normalized frequencies are uniformly spaced with spaces $\frac{2\pi}{C}$. Thus as C increases the spaces between fixed normalized frequencies decreases.

3) The number of fixed normalized frequencies increases with increasing C even.

The values of frequency response at these normalized frequencies are determined by (31). Following can be noted:

4) The value of frequency response at zero normalized frequency ($l = 0$) is the same as the value of frequency response at zero normalized frequency for the signal $s(t)$ coming through the path with impulse response f_i.

5) The unit impulse constraint causes all magnitudes of fixed normalized frequencies to be one (32).

If the signal should be passed without any distortion then the unit impulse constraint is the most convenient choice. Assume that the plane wave of the white stationary noise is arriving from the direction C. Then the output power minimization could occur only at bands excluding fixed normalized frequencies. This leads to sharp local maxima at these normalized frequencies (see Fig. 3, dashed curves). Thus the power gain for the white stationary noise is almost determined by the number of these maxima. Therefore the dependence of the power gain on the direction C is similar to the stair function where the rapid increasing occurs at C even (see item **3)**). The existence of local maxima has also the impact on $SNRE$ (Signal to Noise Ratio Enhancement) which therefore falls with every C even (see Fig. 4). It can be concluded that the delay range should be restricted to C between 0 and 2 samples.

In the case of an input colour noise the situation is more complicated. The behavior of the system depends, on whether the noise spectrum hits or miss the normalized frequencies fixed by the constraint. In general, the delay range should be restricted similarly as in the previous case, because there is only the zero normalized frequency to hit (see Tab. 1 for examples). The car noise is the worst example. Most power of its spectra is spread around the zero normalized frequency, which is fixed by the unit impulse constraint to one (see **5)**), thus most of the noise passes through the system. It motivates us to abandon the unit impulse constraint and turn rather to the high-pass constraint. However it excludes the zero normalized frequency from the signal spectrum and causes some distortion, it fixes the frequency response at the zero normalized frequency to zero value for directional noises according **4)**. If some distortion is present in the signal $s(t)$ (due to the small length of transversal filters) it could be equalized by a subsequent filter. But this equalization must be done carefully, because this subsequent filtration affects the noise too. The proposed frequency response of Frost's beamformer for the signal, the equalization filter and the resulting frequency response, respectively, are shown in Fig. 9. Results for this constraint are summarized also in Tab. 1. The choice of the frequency response for the signal assuming Griffiths-Jim beamformer [3] is discussed also in [4], leading to the same result.

The examples in Tab. 1 are obtained by simulating following configuration (The same array and filter setup was used to obtain curves in Fig. 4). The array consists of three uniformly spaced sensors with the adjacent sensor distance $d = 76\mathrm{mm}$ (the sound speed $v = 330\mathrm{ms}^{-1}$). The noise arrives from the direction $\varphi = 30°$, which corresponds to delay $C = 0.92$. The adaptive filter driven by Frost's CLMS algorithm used behind sensors has the convergence constant $\mu = 0.001$, 63 weights and the sampling frequency $f_s = 8000\mathrm{Hz}$. A speech (words: "0,1,0,1") is used as the input signal. Three types of input noise are used (see Tab. 1). Input waveforms are scaled to have the unit average power. Therefore the global SNR is $0\mathrm{dB}$. The steady state $SSNRE$ is computed from last "0,1" segments by time averaging from one realization. The length of used waveforms is 2.47 s. The case of the white stationary noise and the unit impulse constraint is further illustrated by showing the signal and the noise vs. time in Fig. 5**a**, $SNRE$ vs. time in Fig. 5**b**, and spectrograms Fig. 6. The magnitude frequency response and the power gain for the white stationary noise vs. the direction of arrival are shown in Fig. 7 and Fig. 8.

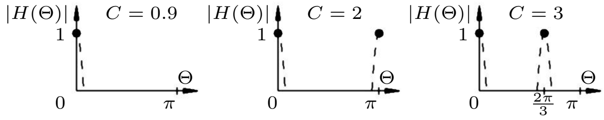

Fig. 3. Impact of constraint on frequency response - examples for different C, unit impulse constraint is assumed, black dots denotes fixed points

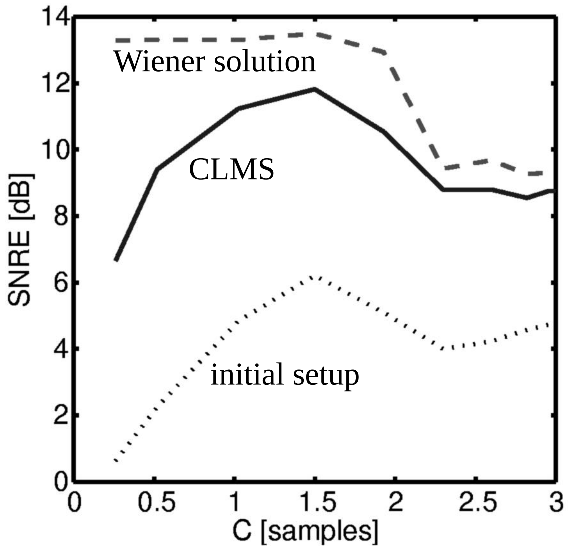

Fig. 4. $SNRE$ in dependence on direction C for unit impulse constraint and zero mean white stationary unit variance stochastic processes used for both signal and noise

			noise		
			white st.	fan	car
constraint type	unit imp.	$\|H\|$			
		s.s. $SNRE$	12.36 dB	12.10 dB	6.57 dB
	high-pass	$\|H\|$			
		s.s. $SNRE$	15.18 dB	12.82 dB	22.22 dB

Tab. 1 Reached frequency responses and s.s. $SNRE$ for some directional noises arriving from direction (delay) $C = 0.9$ [samples] and constraint type.

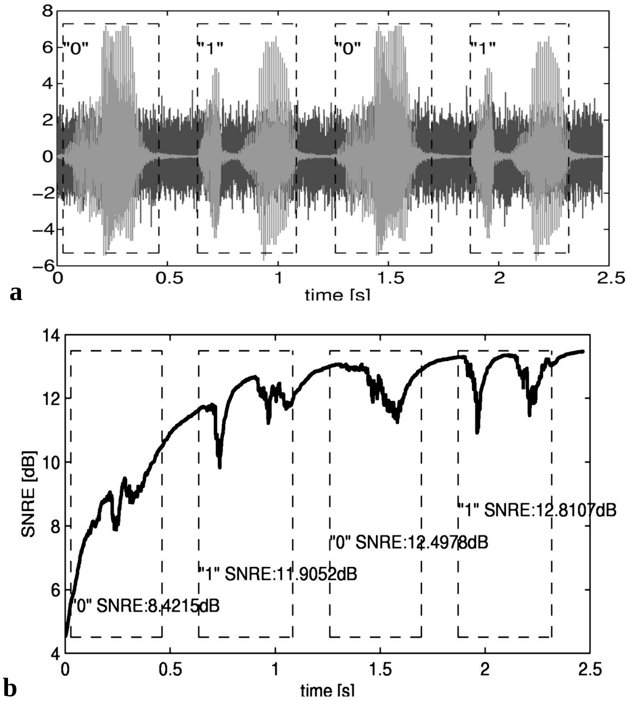

Fig. 5. **a** Signals vs. time (desired speech - light gray, white stationary noise - dark gray), **b** $SNRE$ vs. time ($SNRE$ in each segment is shown)

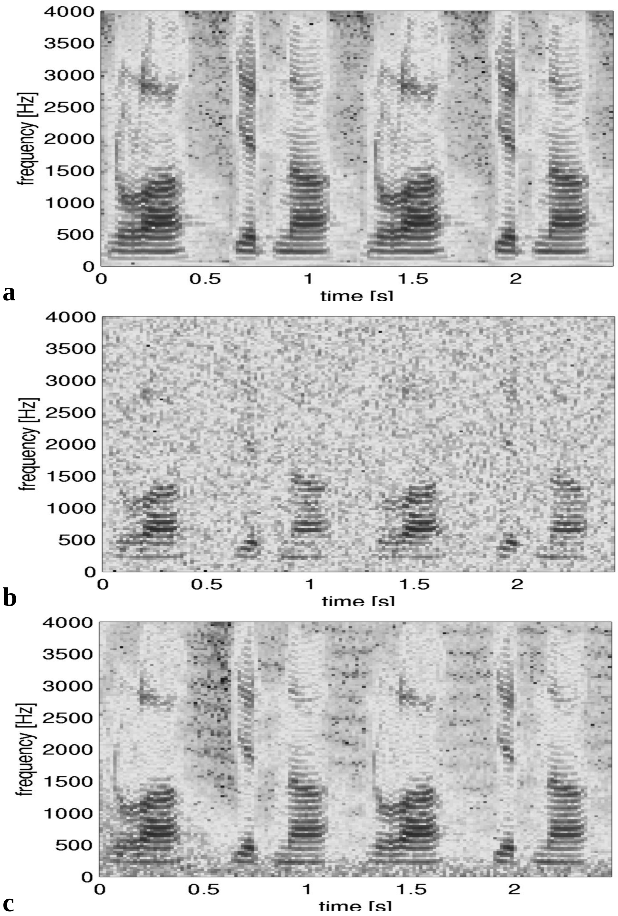

Fig. 6. Spectrograms: **a** desired speech,**b** mixture on sensor, **c** filter output

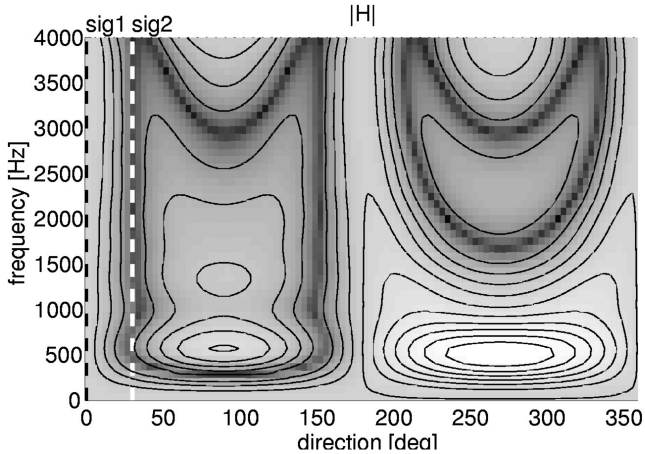

Fig. 7. Magnitude frequency response on direction for weights sampled at time instant 1.75 s. Darker color corresponds to lower magnitude. Directions of desired signal and noise are labeled sig1 and sig2 respectively and marked with vertical lines.

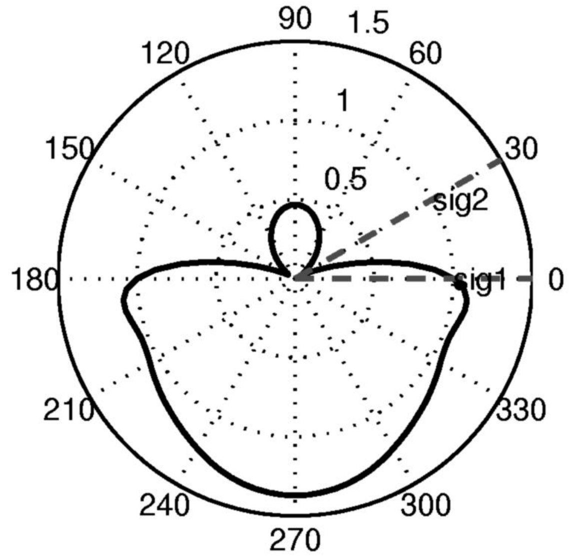

Fig. 8. Power gain on direction for weights sampled at time instant 1.75 s. Directions of desired signal and noise are labeled sig1 and sig2 respectively and marked with rays.

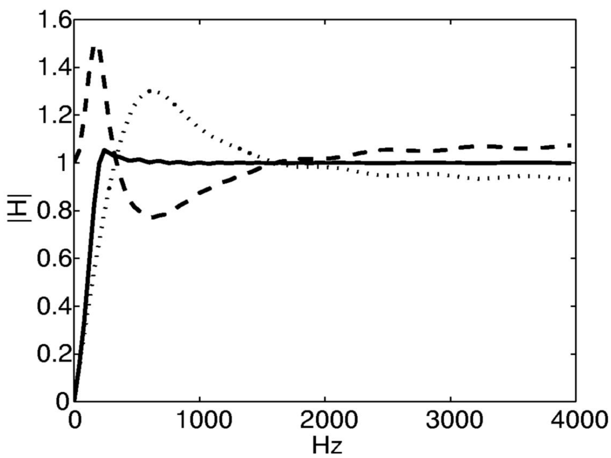

Fig. 9. Frequency response for: desired signal (constraint) - dotted, equalization filter - dashed, result - solid

They are computed from weights using (11) and (15) sampled at the time instant 1.75 s (after the convergence is reached). The decreasing of the system gain close to the noise direction is evident in both cases.

6. Noise Uncorrelated between Sensors and Samples

This section deals with the influence of the noise uncorrelated between sensors and the samples on the system. Noise moments are assumed to be the same across both samples (time stationarity) and sensors (spatial stationarity). The following trend can be observed. As the this type of noise dominates against the directional one, the weights approach to the initial ones. This could be clarified by finding the Wiener solution. The autocorrelation matrix takes the form $\mathbf{R_{xx}} = \sigma_o^2\mathbf{E}$, where σ_o^2 denotes the variance of the this uncorrelated noise. Using (7) the Wiener solution (5) takes the following form

$$\mathbf{w}_{opt} = (\sigma_o^2\mathbf{E})^{-1}\mathbf{C}\left(\mathbf{C}^T\left(\sigma_o^2\mathbf{E}\right)^{-1}\mathbf{C}\right)^{-1}\mathcal{F} = \\ = \mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1}\mathcal{F} = \mathbf{f}. \quad (33)$$

Note that $\mathbf{f}$ stands for the initial weight vector. Thus it can be concluded that the adaptation process takes no effect and the filter remains in the initial state. The noise power at the output is then

$$E[y^2[n]] = \mathbf{w}_{opt}^T\mathbf{R}_{xx}\mathbf{w}_{opt} = \mathbf{f}^T\sigma_o^2\mathbf{E}\mathbf{f} = \\ \sigma_o^2\mathcal{F}^T[(\mathbf{C}^T\mathbf{C})^{-1}]^T\mathcal{F} = \sigma_o^2\mathcal{F}^T[(K\mathbf{E})^{-1}]^T\mathcal{F} = \frac{\sigma_o^2}{K}\sum_{i=1}^{J} f_i^2. \quad (34)$$

While the unit impulse constraint gives the power gain for the signal $AP_s = 1$, the power gain for the noise is

$$AP_n = \frac{E[y^2[n]]}{\sigma_o^2} = \frac{1}{K} \quad (35)$$

and $SNRE$

$$SNRE = \frac{AP_s}{AP_n} = K. \quad (36)$$

Thus for the unit impulse constraint, $SNRE$ equals to the number of sensors used. It is the same as for the delay and sum beamformer. It is not so surprising, because the initial weights $\mathbf{f}$ for the unit impulse constraint are non zero only in one column and their values are equal. In order to compare the $SNRE$ (for the dominating uncorrelated noise between sensors and samples) with results summarized in Tab. 1 (for the directional noise), the equation (36) must be evaluated for the same number of sensors $K = 3$. This evaluation yields 4.8 dB.

7. Estimating $SNRE$

The estimation of $SNRE$ computed directly from waveforms require many realizations to be sufficiently smooth. In the case of limited small number of signal realizations the batch or recursive estimates can be used with the assumption of ergodicity. But what one can do when only one realization is at hand (e.g. speech), and the use of mentioned estimates are rather devastating to give readable results? In this paragraph one alternative method is described which is derived under following assumptions:

1) The system under consideration is Frost's beamformer driven by CLMS algorithm with the unit impulse constraint so the power gain for the signal is $AP_s[n] = 1$.

2) Input noises are mutually independent white stationary noises.

3) The current effective window of input noises (see below) is independent on current weights representing the instantaneous impulse responses.

Of course non-LTI system has not any impulse response in general. But something like the impulse response can be assigned to each iteration. The current output sample is then computed by multiplying input waveform with current impulse response. Assumption **3)** means that the effective window of the input waveform which goes into the multiplication is independent on the current impulse response. This is called the independence assumption and is used for example in [5] for LMS MSE derivation. Reasoning behind this is as follows. When the step size parameter μ is very small then current weights are almost determined by past input samples and the current effective window has negligible impact. When the input samples are mutually independent then the current effective window can be thought to be independent of current weights (or current impulse response).

Let us assume N directional noises $n_{11}[n], \ldots, n_{1N}[n]$ at inputs and noise uncorrelated between sensors and samples $n_{1o}[n]$. Let standard deviations of these noises are $\sigma_{11}, \ldots, \sigma_{1N}, \sigma_{1o}$, impulse responses $h_{1n}[k], \ldots, h_{Nn}[k]$, $h_{on}[k]$ (for iteration n) and power gains $AP_1[n], \ldots, AP_N[n], AP_o[n]$ (see (15), (35)) for iteration n. Let us define the sum of noises at the input $n_1[n]$ and the sum of the noises at the output $n_2[n]$.

$$n_1[n] = \sum_{i=1}^{N} n_{1i}[n] + n_{1o}[n] \quad (37)$$

$$n_2[n] = \sum_{i=1}^{N}\sum_{k_i=-\infty}^{\infty} h_{in}[k_i]n_{1i}[n-k_i] + \\ \sum_{k_o=-\infty}^{\infty} h_{on}[k_o]n_{1o}[n-k_o] \quad (38)$$

Let $AP_n[n] = \frac{E[n_2^2[n]]}{E[n_1^2[n]]}$ to be the power gain for noises in iteration n. And finally let us define the current $SNRE[n]$

$$SNRE[n] = \frac{AP_s[n]}{AP_n[n]} = \frac{E[n_1^2[n]]}{E[n_2^2[n]]} = \frac{\sum_{i=1}^{N}\sigma_{1i}^2 + \sigma_{1o}^2}{E[n_2^2[n]]}. \quad (39)$$

During the last two steps in (39) the assumptions **1)**, **2)** and equation (37) were used. When assumptions **3)**, **2)** and equation (38) are used for the denominator of (39) it can be written

$$\begin{aligned} E[n_2^2[n]] = E\big[(\sum_{i=1}^{N}\sum_{k_i=-\infty}^{\infty} h_{in}[k_i]n_{1i}[n-k_i]+ \\ \sum_{k_o=-\infty}^{\infty} h_{on}[k_o]n_{1o}[n-k_o])^2\big] = \\ = \sum_{i=1}^{N} E\big[\sum_{k_i=-\infty}^{\infty} h_{in}^2[k_i]\big]\sigma_i^2 + E\big[\sum_{k_o=-\infty}^{\infty} h_{on}^2[k_o]\big]\sigma_{1o}^2 \\ = \sum_{i=1}^{N} E\big[AP_i[n]\big]\sigma_i^2 + E\big[AP_o[n]\big]\sigma_{1o}^2. \end{aligned} \tag{40}$$

The substitution (40) to (39) gives

$$SNRE[n] = \frac{\sum_{i=1}^{N}\sigma_{1i}^2 + \sigma_{1o}^2}{\sum_{i=1}^{N} E\big[AP_i[n]\big]\sigma_i^2 + E\big[AP_o[n]\big]\sigma_{1o}^2}. \tag{41}$$

Because weights are smoothed by CLMS itself and second order moments of noises can be computed by time averaging due to their ergodicity, this estimate is much more smoother then previous mentioned batch and recursive estimates. The described approach was used to obtain time dependence of $SNRE$ in Fig. 5.

8. Conclusion

Uniform linear arrays and plane waves propagating in homogeneous, isotropic lossless nondispersive environment were assumed throughout this paper.

Frequency response (11) and white stationary process power gain (15) were derived as functions of the direction of arrival.

Some symmetries were shown using the configuration concept. It was found (4.2) that for the desired signal, one directional noise (noise uncorrelated between sensors and samples may be also present) and Frost's CLMS algorithm behind sensors, it is sufficient to consider noise directions φ between $0°$ and $90°$. For greater number of directional noises it is sufficient to consider noise directions φ between $-90°$ and $90°$. Obtained results can be further extended to different adjacent sensor distances using (4.1).

The impact of constraints for the directional noise suppression was described. Discussion of this issue leads to the restriction of possible delays C between adjacent sensors to the delay range between 0 and 2 samples. This can be done by restricting possible direction of arrival φ or by changing adjacent sensor distance d according to (9). The choice of the constraint was also discussed. The high-pass constraint seems to fit most applications.

Bad performance for noise uncorrelated between sensors and samples was clarified by finding the Wiener solution giving the initial weights. This means that when the noise uncorrelated between sensors and samples dominates then the filter remains in its initial state. Thus adaptation process takes no effect.

The problem of estimating $SNRE$ was considered and the particular solution was given (41).

9. Acknowledgment

This research was supported by grant No. MSM:210000012 Transdisciplinary research in biomedical engineering and grant GACR 102/02/0124 Voice Technologies for Support of Information Society.

References

[1] WERNER, S., APOLINARIO, J., Laakso, T. Multiple-antenna CDMA mobile reception using constrained normalized adaptive algorithms. In *Proceedings SBT/IEEE International Telecommunications Symposium ITS'98*, Sao Paulo (Brazil), 1998, pp. 354-358.

[2] FROST, O. L. An algorithm for linearly constrained adaptive array processing. In *Proceedings of IEEE*, 1972, vol. 60, no. 8, p. 926-934.

[3] GRIFFITHS, L. J., JIM, C. W. An alternative approach to linearly constrained adaptive beamforming. *IEEE Transactions on Anntenas and Propagation*, 1982, vol. AP-30, p. 27-34.

[4] JOHO, M., MOSCHYTZ, G. S. On the design of the target-signal filter in adaptive beamforming. In *Proceedings of IEEE International Symposium on Circuits and Systems*, 1998, vol. 5, p. 166-169.

[5] FAHRANG-BOROUJENY, B. *Adaptive Filters, Theory and Applications.* Wiley and Sons, 1996.

[6] ŠTRUPL, M. *Analysis and simulation of Frost beamformer algorithm.* Master's thesis, FEE CTU in Prague, 2002.

[7] WIDROW, B., DUVALL, K. M., GOOCH, R. P., NEUMAN, W. C. Signal cancelation phenomena in adaptive antenas: causes and cures. *IEEE Transactions on Antenas and Propagation*, 1982, vol. AP-30, no. 3, p. 469-478.

[8] GANNOT, S.,BRUSHTEIN, D., WEINSTEIN, E. Beamforming for multichannel speech enhancement. In *Proceedings of IWAENC'99*, 1999, p. 96-99.

[9] LASENBY, J., FITZGERALD, W. J. A Bayesian approach to high resolution beamforming. In *IEE Proceedings-F*, Radar and Signal Processing, 1991, vol. 138, no. 6.