IEEE International Symposium on Circuits and Systems ISCAS, Monterey, CA, June, 1998, vol. 5, pp. 166-169.

ON THE DESIGN OF THE TARGET-SIGNAL FILTER IN ADAPTIVE BEAMFORMING

Marcel Joho, *George S. Moschytz*

Signal and Information Processing Laboratory,
Swiss Federal Institute of Technology Zürich, Switzerland
joho@isi.ee.ethz.ch, tel.:+41 (0) 1 632 27 71, *moschytz@isi.ee.ethz.ch, tel.:+41 (0) 1 632 27 63*

ABSTRACT

This paper deals with the improvement of jammer suppression for adaptive broadband beamforming. The analysis is carried out with a 2 microphone Griffiths-Jim Beamformer [1] and one jammer signal. By examination of the derived optimal filter, a design strategy for the constant target signal filter in the main channel of the beamformer is given. This leads to a substantial reduction in the required filter length of the adaptive filter while maintaining the same jammer suppression. Furthermore, a faster rate of convergence can be achieved, as there are less filter coefficients that have to be adapted. Finally, it is shown that the design strategy also applies if more than two microphones are used.

1. INTRODUCTION

Adaptive broadband beamforming is becoming increasingly important in acoustical applications for spatial filtering [2]. Beamforming can be applied whenever multiple sound sources are present which can be subdivided into two groups: target and jammer signals. The purpose of beamforming is to amplify a target signal while attenuating the jammer signals. Typical acoustical applications are in hearing aids and for videoconferencing, where one speaker is the target signal and the other speakers or noise sources located in the same room are considered as jammers.

In [3] an adaptive beamformer is presented in which the filter coefficients are updated by an LMS-algorithm in the time-domain. After every adaptation step the coefficient vector is adjusted such that the specified target signal transfer function in the frontal-direction is preserved.

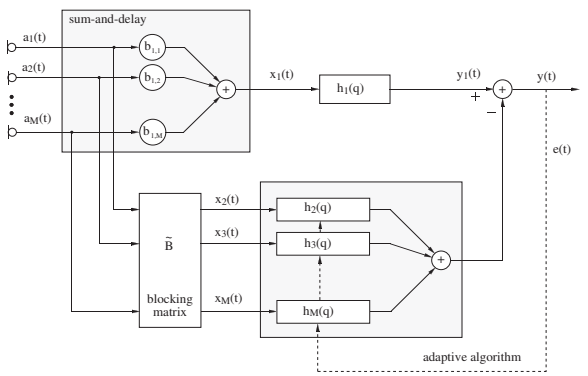

Figure 1: Griffiths-Jim beamformer with blocking matrix

This additionally required operation after every coefficient update is elegantly circumvented in the Griffiths-Jim Beamformer [1] by the introduction of a *blocking matrix* for preprocessing the input signals (Fig. 1). This leads to a simple update rule with no constraints on the adapted filter coefficients. Therefore any adaptive MISO-algorithm (multiple input single output) can be applied in a Griffiths-Jim beamformer [7].

Little attention has been paid to the influence of the constant *target signal filter* $h_1(q)$ located in the main channel. Usually, a simple delay of l samples is chosen for this filter (i.e., $h_1(q) = q^{-l}$) to let the target signal pass through to the beamformer output without any distortion [1, 4].

This paper focuses on the role of the constant target signal filter $h_1(q)$ in the Griffiths-Jim beamformer. It can be shown that the steady-state adaptation error strongly depends on the choice of $h_1(q)$. As a result, a design rule for the target signal filter $h_1(q)$ is derived, which leads to shorter adaptive filters in the auxiliary channels and therefore to a faster adaptation towards the optimal filter coefficients. The drawback in this method is that the target signal is high-pass filtered. In an acoustical application, e.g. hearing aids, where improving speech intelligibility is the main objective, this is no real limitation, because low frequencies are of little importance. In section 2 the conventional Griffiths-Jim beamformer is described. In section 3 we analyze the two-microphone Griffiths-Jim beamformer and derive a design strategy for the target-signal filter. Simulation results are presented in section 4.

2. ADAPTIVE BEAMFORMING

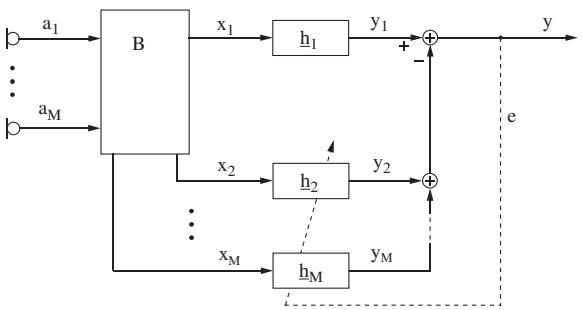

Figure 2: Griffiths-Jim beamformer

2.1. Griffiths-Jim Beamformer

The system studied here is based on the adaptive beamformer described by Griffiths and Jim [1]. As seen in Fig. 2 the microphone array consists of M sensors. Its input signals a_m are transformed with a matrix B into a *main channel* x_1 and $M-1$ *auxiliary channels* $x_2..x_M$. B is an $M \times M$ matrix organized such

that $x_m[t] = \sum_{i=1}^{M} b_{m,i} \cdot a_i[t]$ or, in vector notation, $\underline{x}[t] = B\,\underline{a}[t]$, where $\underline{a}[t] = (a_1[t], \ldots, a_M[t])^T$ and $\underline{x}[t] = (x_1[t], \ldots, x_M[t])^T$.

The matrix B is set up such that it prevents a *target* (in form of a plane wave impinging perpendicularly on the array), from passing through to the auxiliary channels, while letting it pass unimpeded through to the main channel ($x_m[t] = 0$ for $m = 2..M$ if $a_i[t] = a[t]$ for $i = 1..M$). The matrix B without the first row is known as the *blocking matrix*. One possible realization of B is (as described in [1] for $M = 3$):

$$B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}. \tag{1}$$

2.2. Time-Domain Filtering and Adaptation

Each channel x_m of the beamformer contains an FIR-filter $\underline{h}_m$ with N taps. The filter in the main channel, $\underline{h}_1$, is assumed to be time-invariant and designed to shape the target spectrum. $\underline{h}_2..\underline{h}_M$ are adaptive filters, updated after every time sample, such as to minimize the power of the beamformer output:

$$y[t] = y_1[t] - \sum_{m=2}^{M} y_m[t]\,. \tag{2}$$

If perfect adaptation occurs, only the target signal filtered with $\underline{h}_1[t]$ remains at the beamformer output $y[t]$, whereas signal components from other spatial directions (*jammers*) vanish.

For the adaptation it is assumed that target and jammers are uncorrelated, therefore the beamformer output is taken as the adaptation error as well: $y[t] = e[t]$. The filter LMS update equations for real-valued input signals are:

$$\underline{h}_m[t+1] = \underline{h}_m[t] + \mu_0 \underline{x}_m[t]\,e[t] \tag{3}$$

$$\underline{x}_m[t] = (x_m[t], \ldots, x_m[t-N+1])^T \tag{4}$$

where μ_0 is the step-size controlling the rate of convergence and stability of the adaptation.

3. TWO-MICROPHONE ANALYSIS

For the sake of simplicity the Griffiths-Jim beamformer is analyzed with $M = 2$ microphones with one target and one jammer signal as shown in Fig. 3. As will be seen, the same conclusions hold

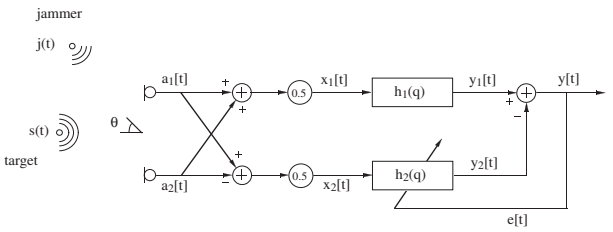

Figure 3: Griffiths-Jim beamformer with 2 microphones.

for $M > 2$ microphones. Furthermore it is assumed that the beamformer is located in an anechoic chamber, such that no reflections from the target and jammer signal appear at the microphones of the beamformer. The corresponding matrix B for $M = 2$ microphones is:

$$B = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{5}$$

The target signal is assumed to impinge perpendicularly on the array ($\theta = 0°$) and therefore simultaneously at both microphones, whereas the jammer signal arrives with an intersensor delay τ at the two sensors, which is dependent on the angle of incidence θ:

$$a_1(t) = s(t) + j(t) \tag{6}$$

$$a_2(t) = s(t) + j(t-\tau)\,. \tag{7}$$

In a sampled-data system the input signals are

$$a_1[t] = s[t] + j[t] \tag{8}$$

$$a_2[t] = s[t] + j[t-d] = s[t] + q^{-d} j[t] \tag{9}$$

where $d = \tau / T_s = \tau f_s$ is the intersensor delay normalized by the sampling time T_s, q is the delay-operator and [.] denotes the sample index. Note that with this notation d can be a rational number.[1] This means that the sequence of the jammer signal $j(.)$ is sampled twice, both with the same sampling frequency but with a constant relative time-shift d: $j(.)$, $j(.-\tau)$ and $j[.]$, $j[.-d]$, respectively.

The beamformer output $y[t]$, which is also used as the adaptation error $e[t]$, is obtained from Fig. 3:

$$y[t] = h_1(q)s[t] + \frac{1}{2}\left(h_1(q)(1+q^{-d}) - h_2(q)(1-q^{-d})\right) j[t]\,. \tag{10}$$

As can be seen from (10), the beamformer transfer function of the target signal is just $h_1(q)$, which is why it is named *target signal filter*. As $h_1(q)$ is an arbitrary design parameter, a simple delay q^{-l} is normally chosen for it. The second term of (10) is the transfer function of the jammer signal. For its complete cancelation at the beamformer output, the transfer function must be equal to zero, which is true for:

$$h_1(q)(1+q^{-d}) = h_2(q)(1-q^{-d})\,. \tag{11}$$

The optimal filter $h_2^o(q)$ and its z-transform becomes

$$h_2^o(q) = \frac{(1+q^{-d})}{(1-q^{-d})} h_1(q) \tag{12}$$

$$H_2^o(z) = \frac{(1+z^{-d})}{(1-z^{-d})} H_1(z)\,. \tag{13}$$

The optimal filter $H_2^o(z)$ has a pole at $z = 1$ because of:

$$1 - z^{-d}\Big|_{z=1} = 0 \qquad \forall d \tag{14}$$

and therefore has an infinite impulse response (IIR). If FIR filters are used, which is usually the case in adaptive beamforming, the optimum solution can only be approximated, and an adaptation error remains in the steady-state. To avoid this error from occuring, several methods have been proposed (e.g., [5, 6]).

One solution proposed in [5] is to insert a filter into the auxiliary channel $x_2[t]$ with a fixed or an adaptive pole. The disadvantage of this method is that low-frequency sensor noise or a DC-offset at the output of the microphone preamplifiers are strongly amplified in the auxiliary channels. Even in the case of a small adaptation error $H_2(z) - H_2^o(z)$, this type of signal appears as disturbing low-frequency noise at the output of the beamformer $y[t]$. As the beamformer output is also used for the adaptation of the

[1] To prevent spatial aliasing $-1 \le d \le 1$ must hold.

filter coefficients, these are also affected by this noise, resulting in a low-frequency fluctuation.

The solution to the IIR-filter problem (see eq. (13)) proposed in this paper is *to explicitly include a zero at unity in the target signal filter, to prevent low-frequency signal components appearing at the output of the main channel* $y_1[t]$:

$$H_1(z) = (1-z^{-1})\overline{H}_1(z)\,. \tag{15}$$

The filter $\overline{H}_1(z)$ remains a design parameter. It can still be chosen to specify the target transfer function $H_1(z)$ for higher frequencies and may be of linear phase. The optimal adaptive filter now results as

$$H_2^o(z) = \frac{(1+z^{-d})}{(1-z^{-d})}(1-z^{-1})\overline{H}_1(z)\,. \tag{16}$$

By substituting $z = e^{j2\pi f T_s}$ in $H_2^o(z)$ and evaluating the gain at $f = 0$ it is seen that this is now finite, in contrast to (13) for a general $H_1(z)$. It is difficult, in the general case, to show mathematically that

$$\|\tilde{H}_2(z)\| = \|H_2^\#(z) - H_2^o(z)\| \qquad \text{or} \tag{17}$$

$$\|\tilde{h}_2(q)\| = \|h_2^\#(q) - h_2^o(q)\| \tag{18}$$

is smaller for $H_2^o(z)$ for (16) than for (13). The optimal FIR filter with N taps is denoted by $H_2^\#(z)$ and $h_2^\#(q)$ in the z-domain and time-domain, respectively. But what is obvious from (13) is that $H_2^\#(z)$ needs to approximate a pole at $z = 1$, which (16) does not. The fact that FIR filters can realize only sharp notches but not sharp peaks leads to the conclusion that the approximation error (17) or (18) is smaller with $H_2^o(z)$ from (16) than from (13).

In the special case when $d = 1$, which happens when the jammer impinges from *endfire* ($\theta = 90°$) on the array, the optimal filter in (16) results in

$$H_2^o(z) = (1+z^{-1})\overline{H}_1(z) \tag{19}$$

and is therefore FIR. This leads to the fact that $H_2^\#(z) = H_2^o(z)$ and therefore the jammer signal can be perfectly suppressed at the beamformer output. Note that this is never the case with (13) when (15) does not hold for $H_1(z)$, because in that case $H_2^o(z)$ must be IIR.

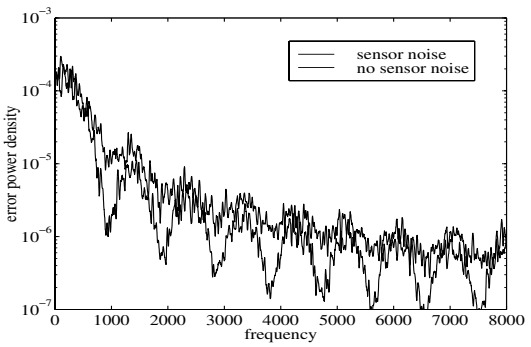

Figure 4: Beamformer output power spectrum in the steady-state with and without sensor noise.

Note that placing the first order filter $(1-z^{-1})$ after each microphone input $a_m[t]$, or after the beamformer output $y[t]$ is *not equivalent to including it in the target signal filter* $H_1(z)$. It makes no difference for the target transfer function, but as seen from (10), it certainly does for the jammer transfer function!

4. SIMULATION

In this section, simulation results are presented to back up the theoretical investigation of the target signal filter $H_1(z)$. The beamformer of Fig. 3 with a jammer signal impinging from 90° ($d = 1$) is examined. No target signal is present in the following simulations. The sampling frequency is $f_s = 16$ kHz, the jammer is a band-limited ($f_s/2$) white Gaussian noise source with power $\sigma_j^2 = 1$. The adaptive filter has $N = 16$ taps. The speed of sound being 330 m/s, the intersensor distance of 2 cm is just close enough to prevent spatial aliasing. The system is located inside an anechoic chamber and additional white Gaussian sensor noise of power 30 dB below the jammer power is added to every microphone. Furthermore, the system is examined in the steady-state, after convergence of the adaptive algorithm.

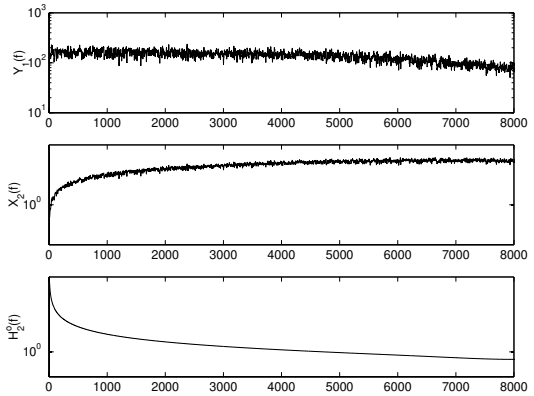

Figure 5: Power spectra for $h_1(q) = q^{-l}$. top: y_1; middle: x_2; bottom: transfer function of the optimal filter $|H_2^o(f)|^2$.

First a simple delay of $l = 7$ samples is chosen for $h_1(q) = q^{-l}$. The power spectrum of the error signal is shown in Fig. 4. Clearly the main contribution arises in the low frequencies below 1 kHz. This is to be expected because of the small aperture of the array.

In Fig. 5 the power spectra of y_1 and x_2 are shown. Note, that almost no power appears in x_2 at the low frequencies, in spite of the white spectrum of the jammer signal. This can be explained by inspecting the transfer function of the jammer signal to x_2, namely $(1-z^{-1})$, which has a zero at unity and therefore corresponds to a time-discrete differentiator. The transfer function of the optimal filter $H_2^o(z)$ is obtained by spectral division of y_1 by x_2 as shown at the bottom of Fig. 5. This causes a sharp peak at 0 Hz as predicted by (13) with $d = 1$:

$$H_2^o(z) = \frac{(1+z^{-1})}{(1-z^{-1})}H_1(z)\,. \tag{20}$$

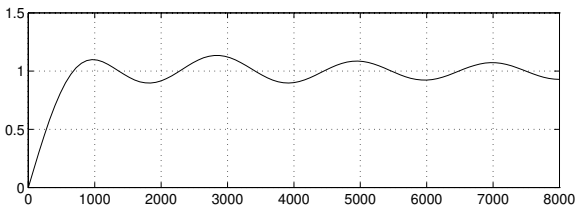

Figure 6: Transfer function of the target signal filter $H_1(z)$ with a zero at unity.

Next, the same set up is chosen but with a different target signal filter as shown in Fig. 6. It was tuned to be flat over a large frequency range (0.5-8 kHz) while fulfilling (15). It has $N = 16$ taps, i.e., the same as the adaptive filter $H_2(z)$. The new spectra of y_1 and x_2 are shown in Fig. 7. Note that the power spectrum of y_1 is now high-pass filtered. The sharp peak of the optimal filter at 0 Hz has now disappeared as can be seen in enlarged detail of Fig. 8.

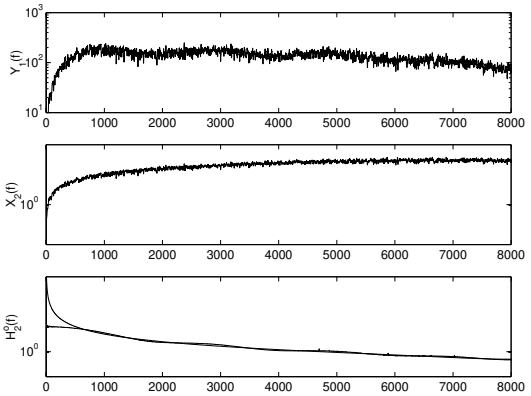

Figure 7: Power spectra for $h_1(q) = (1 - q^{-1})\overline{h}_1(q)$. top: y_1; middle: x_2; bottom: transfer function of the optimal filter $|H_2^o(f)|^2$.

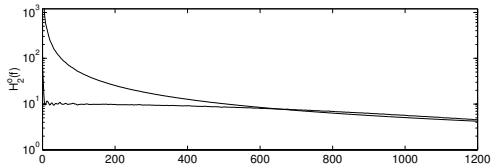

Figure 8: Detail of Fig. 7 bottom.

In Fig. 9 the error power spectrum is shown, where again sensor noise was added at each microphone, power 30 dB below the jammer. In the lower frequency range a remarkable reduction of approximately 10 dB is noticed, although the number of adaptive filter coefficients remains the same. It is noteworthy that there is a substantial reduction of error power at high frequencies as well. An even greater improvement occurs when the sensor noise is removed (Fig. 10). In this case a perfect adaptation of (19) is possible: $H_2^{\#}(z) = H_2^o(z)$. The remaining error power spectrum at the beamformer output is due to the excess mean-squared error of the LMS adaptation in the steady-state.

5. SUMMARY

A design strategy for the target signal filter in a Griffiths-Jim Beamformer is presented. It is shown that by a proper choice of this filter, namely high-pass characteristics with an explicit zero at unity, the pole of the optimal filter vanishes, resulting in a smoother transfer function. As the adaptive FIR filters can model only steep notches but no sharp peaks, the task of adapting towards the optimal solution can thereby be simplified, and in a special case perfect adaptation can even be achieved. The design strategy was derived by analysing a two-microphone beamformer, but it is valid also when multiple microphones are considered. The paper concludes with simulations which back up the theoretical investigations.

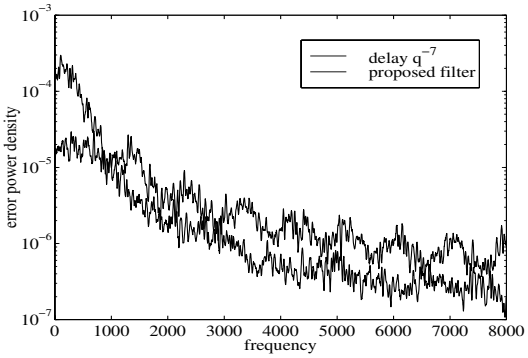

Figure 9: Comparison of beamformer error power spectrum in the steady-state (-30dB sensor noise).

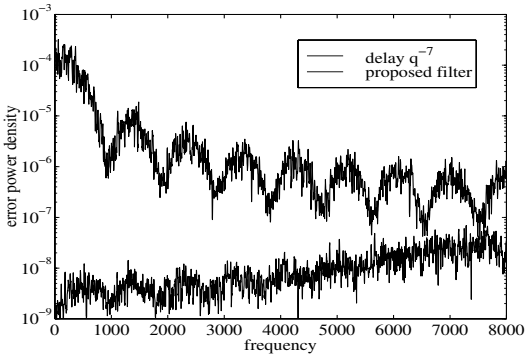

Figure 10: Comparison of beamformer error power spectrum in the steady-state (no sensor noise).

6. REFERENCES

[1] L.J. Griffiths and C.W. Jim, "An Alternative Approach to Linearly Constrained Adaptive Beamforming," *IEEE Trans. Ant. Prop.*, AP-30, Jan. 1982, pp. 27–34.

[2] J.M. Kates, "A comparison of hearing-aid array-processing techniques," J. Acoust. Soc. Amer., Vol. 99, 1996, pp. 3138–48.

[3] O.L. Frost III, "An Algorithm for Linearly Constraind Adaptive Array Processing," Proc. IEEE, vol 60, pp. 926-935, Aug. 1972.

[4] J.E. Greenberg, P.M. Zurek, " Evaluation of an Adaptive Beamforming Method for Hearing Aids," Journal of the Acoustical Society of America, 91(3), pp. 1662-76, March 1992.

[5] R.P. Gooch, "Adaptive Pole-Zero Array Processing," Proc. 16th Asilomar Conf. Circuits Syst. Comput., Nov. 1982.

[6] B. Widrow, S.D. Stearns, "Adaptive Signal Processing," Prentice-Hall, Englewood Cliffs, NJ,1985.

[7] M. Joho, G.S. Moschytz, "Adaptive Beamforming with Partitioned Frequency-Domain Filters," IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Palz, NY, USA, 19-22 Oct. 1997.