A BROADBAND ADAPTIVE BEAMFORMER USING NESTED ARRAYS AND MULTIRATE TECHNIQUES

Yahong R. Zheng, Rafik A. Goubran and Mohamed El-Tanany

Dept. of Systems and Computer Engineering
Carleton University, Ottawa, K1S 5B6, Canada
E-mail: {yzheng, goubran, tanany}@sce.carleton.ca

ABSTRACT

A broadband adaptive array beamformer is proposed using harmonic nesting arrays and multirate sampling techniques. An harmonic nesting microphone array is designed to have several uniform linear subarrays, each covering an octave frequency band. This is analogous to time domain subbanding. An adaptive beamformer following each subarray is then implemented using a Generalized Sidelobe Canceler (GSC) structure. By combining time domain multirate sampling in the subarrays, we obtain an identical normalized frequency band for all subarrays. Thus a single GSC design using eigenvector constraint method is applied to all subarray beamformers. Multirate sampling also provides the convenience on designing the analysis and synthesis filters for perfect reconstruction of the subbanded signals. A specific example is presented to show the simplified design procedures and the improved beamformer performance.

1. INTRODUCTION

Microphone array processing has been widely studied for hands-free audio communication. Its applications include teleconferencing, computer telephony, speech recognition, speech enhancement and hearing aids, etc. Adaptive microphone arrays are often difficult to design because speech and audio are broadband signals with high degrees of non-stationarity.

The problem of broadbanding a sensor array has been a challenge for beamformer design due to the frequency dependent array properties. Several approaches are found in the literature. A straight forward approach [1] is to use a frequency domain beamformer which consists of a large number of independent narrow band beamformers. Another approach [2] is to use multidimensional optimization method to find the optimal sensor spacing and gains. The frequency invariant approach proposed in [3] is to factor the frequency response of a continuous sensor into two parts and then approximate the response by an unequally spaced discrete array. Yet another approach is "harmonic nesting" [4], [5]. With this approach, a broadband array is composed of several equal-spaced linear subarrays, each designed for a smaller frequency range. The subarray beamformers are implemented separately and their outputs are summed.

The harmonic nesting approach has become favorable among other broadband beamforming methods, especially in microphone arrays dealing with speech and audio frequency bands. A nested array is more economical than a single Uniform Linear Array (ULA) because some elements in the subarrays are superimposed, reducing the total number of elements required. A nested array is also easy to design and install, since each subarray is designed and processed separately and the well established beamforming methods for ULAs may be employed directly. Furthermore, a nested array is in fact a subbanding scheme using spatial sampling. When employed by adaptive beamforming scheme, it can have good tracking performance and fast convergence similar to that of other subbanded adaptive systems [6], especially in non-stationary environment.

In this paper, we propose a broadband adaptive microphone array beamformer, incorporating harmonic nesting and multirate sampling. With harmonic nesting, the voice band beamforming is spatially subbanded and the extent of beamwidth variation is reduced to that which occurs within a single octave. With multirate sampling, the subarrays can have identical frequency response and adaptive structure. An identical structure of adaptive beamforming is then designed using a Generalized Sidelobe Canceler (GSC) and eigenvector constraint method. We show that the proposed scheme simplifies the design and implementation of the beamformer. It also improves the performance of the broadband beamformer.

This work was partially supported by Communications and Information Technology Ontario (CITO), Canada

2. HARMONIC NESTING ARRAYS USING MULTIRATE TECHNIQUES

An harmonic nesting array is composed of several equal-spaced subarrays. Let the first subarray be an M-element Uniform Linear Array (ULA), designed for a frequency range $[f_1/2, f_1]$. To avoid grating lobes, the intersensor spacing d_1 is at most half the wavelength of the highest frequency within the band of interest, i.e. $d_1 = c/(2f_1)$, where c is the speed of propagation. The second subarray is then designed for frequency range $[f_1, 2f_1]$ with intersensor spacing being $d_1/2$. The second subarray is nested with the first subarray with $(M+1)/2$ superimposed elements. The third and more subarrays are designed similarly until the highest frequency f_2 is covered or the sensor spacing limit is reached. The total number of elements required is $M + (M-1)(\log_2(f_2/f_1 - 1)/2$. On contrast, a single ULA requires Mf_2/f_1 elements to achieve the same aperture for all frequencies.

Fig. 1. shows the sensor geometry of the harmonic nesting array implemented in this paper. The microphone array is composed of 11 elements, with 5 elements in each subarray. They form 4 nested subarrays and cover the frequency ranges $B_1 = [212.5, 425]Hz$, $B_2 = [425, 850]Hz$, $B_3 = [850, 1700]Hz$, and $B_4 = [1700, 3400]Hz$, respectively, as shown in Fig. 2.

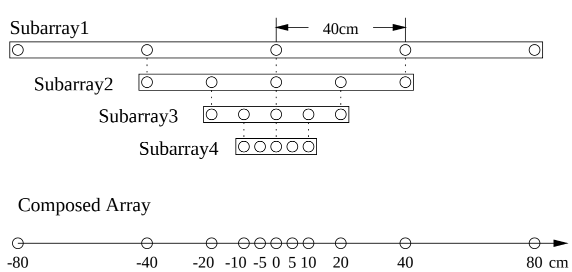


Figure 1: Configuration of harmonic nested array

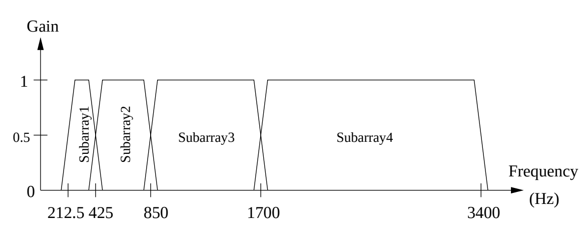


Figure 2: Frequency bands covered by subarrays

With the octave nesting configuration above, it is very convenient to use a common method of multirate sampling for the beamformer, as shown in Fig. 3. Every element is first sampled at F_s, the highest rate required. Then each subarray uses a proper rate of decimation $D_i (i = 1, 2, 3, 4)$ to achieve the desired sampling rate F_i. An analysis filter $H_i(z)$ is placed before each downsampler to avoid aliasing. After down sampling, the four subarrays will have an identical normalized frequency range B_n. Then an identical adaptive beamformer is designed and applied to all the subarrays. The outputs of the adaptive beamformers are up sampled to F_s and then summed. The synthesis filters $G_i(z)$ are needed before the summation to remove the images generated by upsampling. The optional bandpass filter may be also needed to achieve better frequency responses.

In our paper, F_s is chosen to be $16kHz$, and the sampling frequencies of subarrays are $F_1 = 2kHz, F_2 = 4kHz, F_3 = 8kHz$ and $F_4 = 16kHz$. The downsamplers are $D_1 = 8$, $D_2 = 4$, $D_3 = 2$, and $D_4 = 1$. The analysis filters $H_i(z)$ are designed to have the desired frequency responses for the corresponding subarrays, as in Fig. 2. The identical normalized frequency range for the four subarrays is $B_n = [0.10625, 0.2125]$ after the downsamplers.

The use of multirate sampling may appear to be complicated at the first glance, but it in fact simplifies the design and implementation of the broadband beamformer. The added synthesis filters may be paired with the analysis filters for perfect reconstruction [7], while the requirements for these filters may be relaxed. Without multirate sampling, the analysis filters are still needed for each element in each subarray, and more stringent filter specifications are required to avoid aliasing. Multirate sampling also produces an identical normalized frequency band for all subarray beamformers. So we need to design only one adaptive beamformer and apply it to each subarray. This simplifies the broadband beamforming problem to a single ULA beamformer design.

The identical adaptive beamformer used in the subarrays is implemented by a GSC, as depicted in Fig. 4. It consists of a fixed beamformer $\mathbf{W_q}$, a signal blocking matrix $\mathbf{C_a}$ and an unconstrained adaptive weight vector $\mathbf{W_a}$. The design of this beamformer is discussed in the next section.

3. ADAPTIVE BEAMFORMER DESIGN

Let M be the number of elements in a subarray and $\{x_m; m = 1, 2, \ldots, M\}$ be the location of the elements on the array axis. Each element attaches a tapped delay line of length K. There are p plane waves impinging on the array from directions $\Theta = [\theta_1, \theta_2, \ldots, \theta_p]$, measured relative to the array axis. The desired signal is from a known look direction θ_1, and the remaining $p-1$

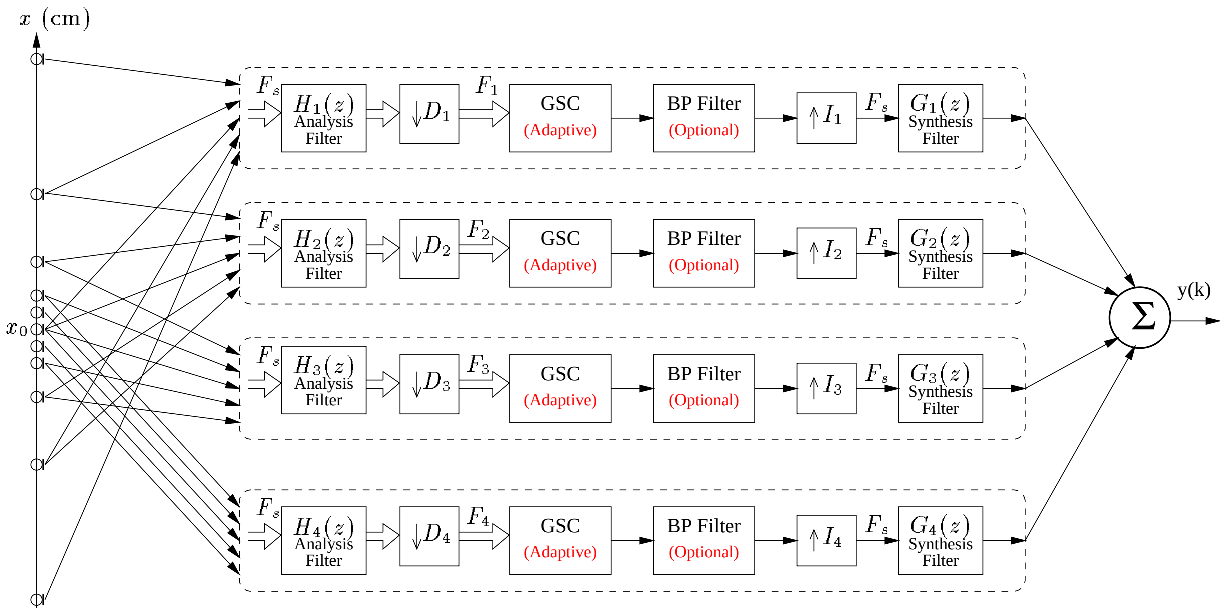


Figure 3: Nested array using multirate sampling

signals are treated as interference with unknown DOA and statistical properties. All the signals are assumed to be band limited to the normalized angular frequency range $[\omega_a, \omega_b] = 2\pi B_n$. The signal received at the mth sensor is

$$
\begin{aligned}
u_m(t) &= \sum_{i=1}^{p} s_i(t - \tau_{i,m}) + n_m(t) \qquad (1)\\
\tau_{i,m} &= \frac{x_m}{c}\cos(\theta_i)
\end{aligned}
$$

where $c = 343m/s$ is the speed of sound. $\tau_{i,m}$ is the propagation delay of the ith source at the mth sensor.

Let the M-dimensional snapshot vector of the signal obtained at the GSC input be

$$\mathbf{u}(k) = [u_1(k), u_2(k), \ldots, u_M(k)]^T \qquad (2)$$

and the $N(= MK)$-dimensional vector of the concatenated snapshot samples be

$$\mathbf{U} = [\mathbf{u}^T(k), \mathbf{u}^T(k-1), \ldots, \mathbf{u}^T(k-K+1)]^T \qquad (3)$$

then the beamformer output can be expressed in matrix form as

$$y(k) = \mathbf{W}^H \mathbf{U}(k) \qquad (4)$$

where $\mathbf{W}$ is the concatenated weight vector and can be determined by solving

$$\min_{\mathbf{W}} \mathbf{W}^H \mathbf{R} \mathbf{W} \quad \text{subject to} \quad \mathbf{C}^H \mathbf{W} = \mathbf{f} \qquad (5)$$

where $\mathbf{R} = E\{\mathbf{U}(k)\mathbf{U}^H(k)\}$ is the $N \times N$ data covariance matrix, $\mathbf{C}$ is the constraint matrix, and $\mathbf{f}$ is the response vector. Each column of $\mathbf{C}$ imposes a linear constraint on the weight vector $\mathbf{W}$. Assuming L constraints, $\mathbf{C}$ is $N \times L$ and $\mathbf{f}$ is L-dimensional.

The fixed beamformer $\mathbf{W_q}$ in the GSC is an N-dimensional vector satisfying the constraints $\mathbf{C}^H \mathbf{W_q} = \mathbf{f}$. Hence

$$\mathbf{W_q} = \mathbf{C}(\mathbf{C}^H \mathbf{C})^{-1} \mathbf{f} \qquad (6)$$

The signal blocking matrix $\mathbf{C_a}$ is orthogonal to the constraint matrix $\mathbf{C}$, $\mathbf{C}^H \mathbf{C_a} = \mathbf{0}$. The unconstrained adaptive weight vector $\mathbf{W_a}$ is updated iteratively using any of the adaptive algorithms such as the Least Mean Square (LMS), the Recursive Least Square (RLS) [11] or the Fast Affine Projection (FAP) [12] algorithms.

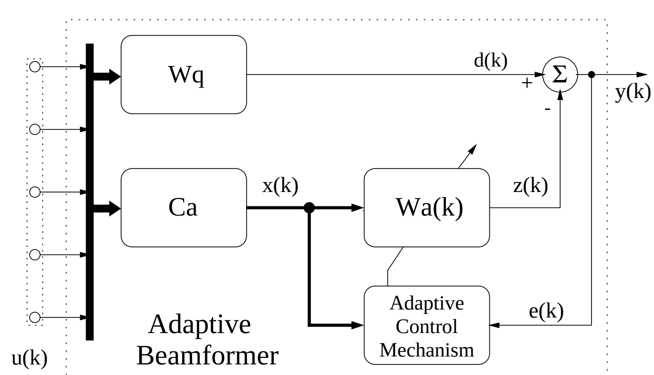


Figure 4: Generalized Sidelobe Canceler

Several different approaches can be employed for designing the constraint matrix $\mathbf{C}$ and response vector $\mathbf{f}$, namely point constraints [8], derivative constraints [9] and eigenvector constraints [10]. In this paper, we use the eigenvector constraint method since it is very efficient for broadband beamformers.

The idea of eigenvector constraint design is to approximate the desired beamformer response at a large number of frequency points, in a least squares sense, using only a small number of constraints. The eigenvector constraints which ensure the real valued implementation of $\mathbf{W}$ are designed as follows:

First of all, select a large number of frequency points $\{\omega_j, j = 1, 2, \ldots, J\}$ within the band of interest $[\omega_a, \omega_b]$. Let the desired array response at these frequencies and the look direction θ_1 be of unit gain and group delay τ_0, thus form the equation

$$\mathbf{A}^H \mathbf{W} = \mathbf{d} \tag{7}$$

where

$$\mathbf{A} = [\mathbf{c}(\omega_1), \ldots, \mathbf{c}(\omega_J) \mid \mathbf{s}(\omega_1), \ldots, \mathbf{s}(\omega_J)] \tag{8}$$

$$\mathbf{d} = [d_1 \cos(\omega_1 \tau_0), \ldots, d_J \cos(\omega_J \tau_0) \mid d_1 \sin(\omega_1 \tau_0), \ldots, d_J \sin(\omega_J \tau_0)]^T \tag{9}$$

$\mathbf{c}(\omega_j)$ and $\mathbf{s}(\omega_j)$ are, respectively, the real and imaginary parts of the concatenated steering vector $\mathbf{a}(\theta_1, \omega_j)$ defined as

$$\mathbf{a}(\theta_1, \omega_j) = [e^{j\omega_j \tau_{1,m}}, e^{j\omega_j \tau_{1,2}}, \ldots, e^{j\omega_j(\tau_{1,M} - K+1)}]^H \tag{10}$$

Then perform Singular Value Decomposition (SVD) of $\mathbf{A}^H$ to obtain

$$\mathbf{A}^H = \mathbf{P} \mathbf{\Sigma} \mathbf{Q}^H \tag{11}$$

where

$$\mathbf{P} = [\mathbf{p}_1, \ldots, \mathbf{p}_L, \ldots, \mathbf{p}_N],$$
$$\mathbf{Q} = [\mathbf{q}_1, \ldots, \mathbf{q}_L, \ldots, \mathbf{q}_{2J}],$$

The columns of $\mathbf{P}$ and $\mathbf{Q}$ are, respectively, the left and right singular vectors of $\mathbf{A}^H$, and $\mathbf{\Sigma}$ is the $2J \times N$ dimensional diagonal singular value matrix.

Select the L largest singular values of $\mathbf{A}^H$ and their corresponding singular vectors to form $\mathbf{\Sigma}_L$, $\mathbf{P}_L$ and $\mathbf{Q}_L$. We obtain the rank L approximation of $\mathbf{A}$

$$\mathbf{A}^H \approx \mathbf{A}_L^H = \mathbf{P}_L \mathbf{\Sigma}_L \mathbf{Q}_L^H \tag{12}$$

Replacing $\mathbf{A}$ in (7) by $\mathbf{A}_L$ yields

$$\mathbf{P}_L^H \mathbf{W} = \mathbf{\Sigma}_L^{-1} \mathbf{Q}_L^{\dagger} \mathbf{d} \tag{13}$$

where $\mathbf{Q}_L^{\dagger}$ is the pseudo inverse of $\mathbf{Q}_L$.

The equation (13) has the same form as $\mathbf{C}^H \mathbf{W} = \mathbf{f}$. So the designed eigenvector constraints are $\mathbf{C} = \mathbf{P}_L$ and $\mathbf{f} = \mathbf{\Sigma}_L^{-1} \mathbf{Q}_L^{\dagger} \mathbf{d}$.

The group delay τ_0 is chosen to be the temporal center of the desired signal source across the array and the tapped delay line,

$$\tau_0 = \frac{1}{2}[\frac{(M-1)\pi}{\omega_b} |\cos(\theta_1)| + K] \tag{14}$$

And L is generally selected in the vicinity of double the observation *Time Band Width Product* (TBWP), where the TBWP is defined as $B_n[\frac{(M-1)\pi}{\omega_b} + K]$.

4. SIMULATION AND PERFORMANCE

The adaptive beamformer described in Fig. 3 is designed as follows. $K = 30$ taps are attached to the $M = 5$ elements in each subarray. $L = 9$ eigenvector constraints are designed using $J = 80$ frequency points uniformly distributed within the normalized passband $B_n = [0.10625, 0.2125]$.

To illustrate the reduction of frequency dependent variation obtained by harmonic nesting, Fig. 5 shows the beam patterns at the five frequency points $300Hz$, $800Hz$, $1300Hz$, $2300Hz$ and $3300Hz$. Fig. 5(a) is the beam pattern for the nested array described in Fig. 3. Fig. 5(b) is obtained by an 11-element ULA spaced at half-wavelength of the highest frequency $3400Hz$. Each element in the ULA also has $K = 30$ taps. Eigenvector constraints are designed similarly with $L = 30$.

The mainlobe beamwidth in Fig. 5(a) varies slightly within 10°. The beamwidth of lower frequencies $300Hz$, $800Hz$ and $1300Hz$ is actually narrower than that of $2300Hz$, thanks to the nested subarrays. But the beamwidth in Fig. 5(b) widens linearly as the frequency decreases. The mainlobe beamwidth at $3300Hz$ and $300Hz$ is approximately 15° and 170°, respectively. The beamwidth variation is more than 150°. So the four-octave nested array reduces the frequency dependent variation to 10% of the ULA's.

For the 11-element ULA, the effective aperture at the low frequencies is less than one half-wavelength. In other words, the 11 elements is equivalent to 2 elements spaced at 0.88 half-wavelength of $300Hz$, or 0.625 half-wavelength of $212.5Hz$. The low efficiency of single ULAs for broadband beamforming is obvious. With harnomic nesting, however, the effective aperture is 2 to 4 half-wavelength for all frequencies. So the four-octave nested array provides more than 200% increase of aperture for the lowest frequency band.

Fig. 6. shows the adaptive beamformer performance of one subarray. Three signal sources impinge on the array from directions 75°, 90° and 125°. The signal

sources are assumed to be uncorrelated. Each signal has a SNR=20dB and is band limited within the range $[200, 3400] Hz$. The array is steered to $125°$. The beam patterns in Fig. 6(a) are obtained at the GSC output after convergence. It shows the effective nulling at the interfering directions $75°$ and $90°$. The frequency response in Fig. 6(b) shows that the look direction has a flat response in the passband, while the responses at the interference directions are attenuated by at least 38dB.

On the other hand, the stopband attenuation of the look direction response is only -15dB which may not be satisfactory for background noise reduction. To achieve better suppression of the sidelobes is desired, one can increase the number of taps attached to each element in the adaptive beamformer. But this will increase the computational complexity and decrease the convergence speed. A better choice we find is to use a fixed bandpass filter at the output of each GSC.

5. CONCLUSION

A broadband adaptive microphone array, incorporating harmonic nesting and multirate sampling, is proposed in this paper. The array is composed of 4 octave nested subarrays, each having 5 elements. With superimposed elements, the total number of elements is only 11. Compared with 11-element equal-spaced array, the nested array reduces the frequency dependent variation of the beam pattern by 90% and increases the aperture for lower frequency end by 200%. The nested array is also easy to be scaled, implemented and maintained.

The octave nesting approach is combined with multirate sampling techniques to obtain an identical normalized frequency response for each subarray. Thus the adaptive beamformers in the subarrays can be implemented identically by a Generalized Sidelobe Canceler structure, simplifying the broadband beamforming design to a single Uniform Linear Array problem. The analysis and synthesis filters associated with the downsamplers and upsamplers may be implemented by the established subbanding techniques and provide perfect reconstruction.

REFERENCES

[1] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Magazine*, pp.4–24, April 1988.

[2] J.H. Doles, III and F.D. Benedict, "Broadband array design using the asymptotic theory of unequally spaced arrays," *IEEE Trans. Antennas Propag.*, Vol. 36, No.1, pp.27–33, Jan. 1988.

[3] D.B. Ward, R.A. Kennedy and R.C. Williamson, "Theory and design of broadband sensor arrays with frequency invariant far-field beam patterns," *J. Acoust. Soc. Amer.*, Vol.97, No.2, pp.1023–1034, Feb. 1995.

[4] T. Chou, "Frequency-independent beamformer with low response error," *Proc. IEEE ICASSP-95*, pp.2995–2998, May 1995.

[5] C. Marro, Y. Mahieux and K.U. Simmer, "Analysis of noise reduction and dereverberation techniques based on microphone arrays with postfiltering," *IEEE Trans. on Speech and Audio Proc.*, Vol.6 No.3, pp. 240–259, May 1998.

[6] R.B. Wallace and R.A. Goubran, "Improved tracking adaptive noise canceler for nonstationary environments," *IEEE Trans. on SP*, Vol.40, No.3, pp.700–703, Mar. 1992.

[7] P.P. Vaidyanathan, "Multirate Digital filters, filter banks, polyphase networks, and applications: a tutorial," *Proc. IEEE*, Vol.78, No.1, pp.56–93, Jan. 1990.

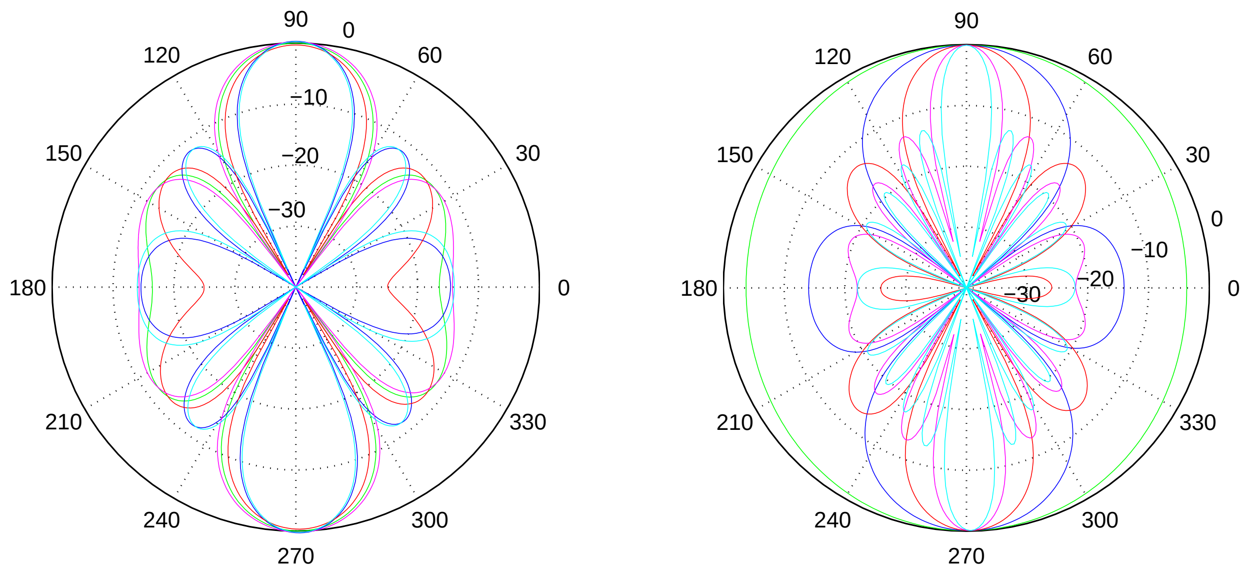
[8] E. L. Kelly, Jr. and M.L. Levin, "Signal parameter estimation for seismometer arrays," *Mass. Inst. Tech., Lincoln Lab., Tech. Report 339*, 1964.

[9] M.H. Er and A. Cantoni, "Derivative constraints for broad-band element space antenna array processors," *IEEE Trans. on ASSP*, Vol.ASSP-31, No.12, pp.1378–1393, Dec. 1983.

[10] K.M. Buckley, "Spatial/spectral filtering with linearly-constrained minimum variance beamformers," *IEEE Trans. on ASSP*, Vol.ASSP-35, No.3, pp.249–266, Mar. 1987.

[11] S. Haykin, *Adaptive Filter Theory*, 3rd ed., Upper Saddle River, N.J., Prentice Hall, 1996.

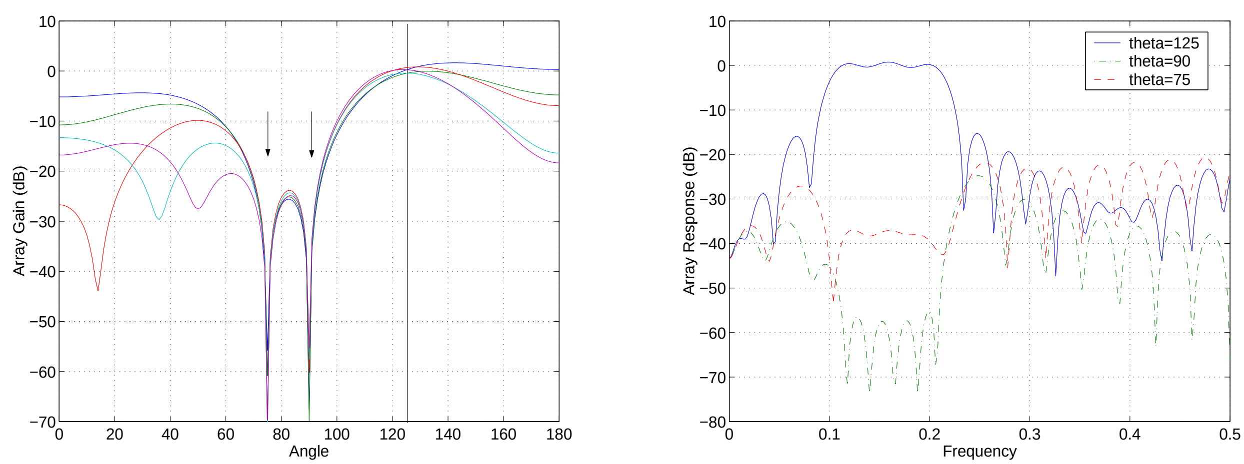
[12] M. Tanaka, Y. Kaneda, S. Makino, and J. Kojima, "Fast projection algorithm and its step size control," *IEEE ICASSP-95*, pp.945–948, 1995.



(a) Harmonically nested array

(b) 11-element uniform linear array

Figure 5: Beam pattern at frequency points $0.3kHz, 0.8kHz, 1.3kHz, 2.3kHz, 3.3kHz$



(a) Beam patterns at normalized frequency points 0.12, 0.14, 0.16, 0.18 and 0.20, vertical line indicates the look direction, arrows indicate the interference directions

(b) Frequency response for $\theta_1 = 125°$ (solid), $\theta_2 = 90°$ (dashed) and $\theta_3 = 75°$ (dash-dot)

Figure 6: Adaptive beamforming performance of nested subarrays