Performance Characterization of Linear Arrays with Respect to Robust MVDR Beamforming

Sebastian Gergen  
Institute of Communication Acoustics  
Ruhr-Universität Bochum  
44780 Bochum, Germany  
Email: sebastian.gergen@rub.de

Nilesh Madhu  
ExpORL, Dept. Neurosciences  
Katholieke Universiteit Leuven  
B-3000, Leuven, Belgium  
Email: nilesh.madhu@med.kuleuven.be

Rainer Martin  
Institute of Communication Acoustics  
Ruhr-Universität Bochum  
44780 Bochum, Germany  
Email: rainer.martin@rub.de

Abstract—In this contribution we focus on measures for characterizing array designs in terms of beamformer performance. The specific beamformer considered here is the MVDR beamformer. We first combine two existing performance criteria into an informative measure that qualifies both noise reduction as well as beamformer robustness to uncorrelated noise. Based on the results obtained with the new measure and on simulations of linear array configurations, we introduce the distance histogram as an instructive tool to qualitatively rate an array for beamforming performance. The aim throughout is to generate optimal microphone positioning with wideband audio signal applications in mind.

I. INTRODUCTION

Sensor arrays, along with appropriate algorithms, can be used to exploit the spatial diversity of desired signals and the noise in order to enhance the quality of the desired signal and reduce noise in the output [1], [2]. This process is known as beamforming. The sensor arrays may be arranged in different geometries (i.e., linear, planar or 3-D) and, for a given geometry, in different configurations (i.e., sensor number and positioning).

Such multichannel signal processing for target enhancement and noise cancellation has been the subject of research for a long time [3]–[5] in manifold application areas, of which radar, sonar, speech and audio are just a few well-known examples. Our focus is on the design of microphone arrays and algorithms to the field of audio and speech. Audio signals mostly cover a wide frequency range: the signal may be broadband itself, e.g., in speech and music processing [6], or the signal may be narrowband but occurring in a wide frequency range, e.g., in noise detection systems [7].

For the practical implementation of a microphone array the robustness against uncorrelated sensor noise is highly important, especially where superdirective beamforming is concerned. Not only sensor self noise but also sensor positioning errors and sensor gain mismatch can be modeled as this type of noise. Usually such a robust design of the beamforming algorithm requires a trade-off with directivity. Optimizing the sensor positioning can reduce this effect as the configuration of an array has high impact not only on the directivity properties, but also on parameters affecting the robustness of the beamformer. In this investigation we focus on the minimum variance distortionless response (MVDR) beamformer design [8] and compare different linear array configurations. More general array geometries and beamforming algorithms will be the substance of further research. As the performance measure for the comparison we propose the combination of two criteria indicating, respectively, the residual noise power and the robustness to uncorrelated noise. Based upon this evaluation we introduce the array distance histogram as an informative and easy to use tool to rate array configurations.

The paper is organized as follows: first we introduce the MVDR beamformer and basic performance criteria. Then, we define our measure for array comparison. We next describe our simulation framework for testing linear array configurations. In section V we evaluate the simulated arrays on the basis of the proposed performance measure. Based on these results we introduce the distance histogram in section VII. We conclude with the ramifications of our paper and outline the steps for further research in section VIII.

II. MVDR BEAMFORMING AND PERFORMANCE CRITERIA

In order to exploit spatial diversity of a signal source at \( r_s = [x_s, y_s, z_s] \) and noise sources in an acoustical environment we use \( N \) microphones located in the farfield of the sources at \( r_n = [x_n, y_n, z_n]^T \). The desired signal \( s(t) \) received at each microphone \( n = 1 \ldots N \) may then be written as:

\[
y_n(t) = s(t - \tau_n) + v_n(t),
\]

with \( \tau_n \) being the time delay of arrival (TDOA) relative to a reference point \( r_0 \) which, for instance, can be one of the microphones. The effects of background noise (correlated and uncorrelated) and reflections are subsumed into the definition of \( v(t) \).

Beamforming is usually done by introducing an FIR filter \( w_n = (w_n(0), w_n(1), \ldots, w_n(M))^T \) of order \( M \) in each signal path followed by summing the filtered signals of each microphone. The output \( \hat{S}(\Omega) \) of this filter-and-sum beamformer (FSB) in the frequency domain is traditionally expressed as:

\[
\hat{S}(\Omega) = \sum_{n=1}^{N} H_{n,s}^*(\Omega) Y_n(\Omega),
\]

with \( H_{n,s}^*(\Omega) = W_n(\Omega) \) denoting the frequency response of the \( n \)-th filter and \( Y_n(\Omega) \) being the Fourier transform of the signal \( y_n \) [8].

A widely used approach to FSB design is the MVDR beamformer, which minimizes the output power subject to unity gain along the target direction [8]. The MVDR beamformer provides high directivity even in low frequencies but comes along with a major drawback: uncorrelated noise (e.g., sensor self noise) is amplified in the lower frequencies [9]. This sensitivity is termed susceptibility. By including an additional constraint in the design stage of the MVDR beamformer this susceptibility can be limited to a maximum value \( K_0 \) and therefore a low frequency noise boost can be controlled at cost of a reduced directivity characteristic.

To obtain such a limited-susceptibility MVDR beamformer \( \mathbf{H}_{\text{MVDR-LS}}(r_s) = (H_1(\Omega, r_s), \ldots, H_N(\Omega, r_s))^T \), we take recourse to the method of Lagrange multipliers, which yields

\[
\mathbf{H}_{\text{MVDR-LS}}(\Omega, r_s) \triangleq \frac{\mathbf{A}^H(\Omega, r_s)(\mathbf{Y}(\Omega) + q(\Omega) \mathbf{I})^{-1} \mathbf{A}(\Omega, r_s)}{\mathbf{A}^H(\Omega, r_s)(\mathbf{Y}(\Omega) + q(\Omega) \mathbf{I})^{-1} \mathbf{A}(\Omega, r_s)},
\]

where

\[
\mathbf{Y}(\Omega) = \sum_{n=1}^{N} H_{n,s}^*(\Omega) Y_n(\Omega),
\]

\( q(\Omega) \) denotes the Lagrange multiplier. We emphasize that this expression is the classical MVDR beamformer, i.e., \( \mathbf{H}_{\text{MVDR-LS}}(\Omega, r_s) = \mathbf{H}_{\text{MVDR}}(\Omega, r_s) \) for uncorrelated noise.

for \( \mathbf{A}(\Omega, r_s) = \sum_{n=1}^{N} H_{n,s}^*(\Omega) \mathbf{A}(\Omega, r_s) \),

where

\[
\mathbf{Y}(\Omega) = \sum_{n=1}^{N} H_{n,s}^*(\Omega) Y_n(\Omega),
\]

\( q(\Omega) \) denotes the Lagrange multiplier. We emphasize that this expression is the classical MVDR beamformer, i.e., \( \mathbf{H}_{\text{MVDR-LS}}(\Omega, r_s) = \mathbf{H}_{\text{MVDR}}(\Omega, r_s) \) for uncorrelated noise.
where $\Psi_{NN}(\Omega)$ is the noise power spectral density matrix, $A(\Omega, \mathbf{r}_n) = (e^{-j\Omega d f r_1(\mathbf{r}_n)}, e^{-j\Omega d f r_2(\mathbf{r}_n)}, \ldots, e^{-j\Omega d f r_N(\mathbf{r}_n)})^T$ is the propagation vector along the direction of the source, $\mathbf{I}$ is the identity matrix, and $q(\Omega)$ is the Lagrange multiplier, which can be determined for a given $K_0$ and frequency [8], [10].

All other factors being equal, the geometry and the configuration of the array has strong influence on the MVDR-LS beamformer performance as it affects the condition of $\Psi_{NN}(\Omega)$ in correlated noise fields. Additionally, the array configuration should also be chosen such that grating lobes are avoided. For linear arrays, this requires a minimum distance between two sensors of $d \leq \frac{\lambda_{\text{min}}}{2}$, with $\lambda_{\text{min}}$ being the wavelength of the maximum frequency in the signal [1].

To evaluate the performance of a beamformer several basic criteria are defined. The white noise gain (WNG) $G_W(\Omega)$ presents the gain to uncorrelated (white) noise and is the inverse of the susceptibility

$$K(\Omega) = \frac{1}{G_W(\Omega)} = \mathbf{H}^H(\Omega, \mathbf{r}_s) \mathbf{H}(\Omega, \mathbf{r}_s). \quad (4)$$

The MVDR filters suffer from the limitation of the susceptibility $K(\Omega) \leq K_0$ up to the frequency $\Omega_{\text{desc}}$ where the susceptibility starts to descend (Figure 1). Therefore the beamformer does not provide the maximum directivity in the frequency range $\Omega < \Omega_{\text{desc}}$. At this cut-off frequency the susceptibility drops below $K_0$ and the WNG starts to increase sharply.

The directivity pattern depicts the attenuation of the sound energy from a given direction $\mathbf{r}$ [8].

![Fig. 1: Susceptibility of an MVDR-LS beamformer limited to $K_0$ with the cut-off frequency $\Omega_{\text{desc}}$.](image)

### III. Definition of the New Performance Measure

To evaluate the performance of an MVDR-LS beamformer related to the underlying array geometry and configuration we combine two criteria to a new measure. First we take the cut-off frequency of the susceptibility $\Omega_{\text{desc}}$ into account, which we term susceptibility descending frequency (and its discrete counterpart, the susceptibility descending frequency bin $\text{SDFB}$) and define it as:

Let $\{\Omega\}$ be the ordered set of frequencies, with an associated set of susceptibilities $\{K(\Omega)\}$. Then, $\Omega_{\text{desc}}$ is the maximal element in $\{\Omega\}$ where $K(\Omega) = K_0$.

The smaller $\Omega_{\text{desc}}$, the higher the number of frequencies in which the filter calculation is solved optimally and in which a large WNG is achieved.

Secondly, as the MVDR beamformer should minimize the output noise power, the residual noise power (5), compared to the unfiltered single microphone noise power, can be used as a measure for noise reduction performance.

$$a = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{H}^H(\Omega, \mathbf{r}_s) \Psi_{NN}(\Omega) \mathbf{H}(\Omega, \mathbf{r}_s) d\Omega \quad (5)$$

Combining these values one can characterize an array for its beamforming capability for a fixed $K_0$. Changing $K_0$ for a given configuration leads to a change of $\Omega_{\text{desc}}$ as well as of the noise reduction performance.

### IV. Simulation

In the simulation we used a MVDR-LS beamformer with $K_0 = 10$ steered to $\theta = 0^\circ, 30^\circ, \ldots, 180^\circ$ to evaluate different configurations of linear arrays on the goal of finding an array configuration which is optimal for all steering directions and thus generally applicable. To rate the performance of an array configuration we evaluated the mean SDFB and the mean residual noise power over all simulated steering directions.

The maximum signal frequency was assumed to be $f_{\text{max}} = 8$kHz having speech processing applications in mind. The noise field was specified to be diffuse with unity power in each frequency bin. The sampling frequency was set to $f_s = 16$kHz and we used a DFT size of $M = 512$. We constrained the maximum array length to $d_{\text{max}} = 20$cm with possible microphone positions at $x_{\text{mic}} = 0, 1, 2, 3, \ldots$cm. The minimum distance between a sensor pair was set to $d_{\text{min}} = 2$cm as this fulfills the spatial aliasing condition at $f_{\text{max}}$. The simulation covered all possible array configurations with $N = \{4, 5, 6, 7\}$ microphones, within the constraints of $d_{\text{min}}$ and $d_{\text{max}}$.

### V. Evaluation

Each point in Figure 2 represents one of all possible array configurations in terms of the mean residual noise power over the mean SDFB. In the light of the above discussion we search for the smallest mean SDFB of an array and secondly the lowest mean residual noise power. We find that for four and five microphones the optimal solutions are related to configurations with $d > d_{\text{min}}$ for all inter-element distances, so that the spatial aliasing theorem is violated. Therefore we only consider configurations which fulfill this condition. For arrays with six and seven microphones, the optimal configuration already includes $d_{\text{min}} = 2$cm (Figure 3).

![Fig. 2: Mean residual noise power vs. mean SDFB of different array configurations with four (blue), five (green), six (red) and seven (cyan) microphones, with unity input power](image)
Fig. 3: Optimal array configurations with \( N \) = \{4, 5, 6, 7\} microphones and \( d_{\text{min}} = 2\) cm

Table I presents the mean SDFB and the mean residual noise power of the four optimal array configurations. Also presented are the values obtained for a uniform linear array (ULA) with the same number of microphones and \( d_{\text{max}} = 20\) cm (\( d_{\text{min}} = 5.66\) cm, \( d_{\text{mic}} = 5\) cm, \( d_{\text{mic}} = 4\) cm, \( d_{\text{mic}} = 3.33\) cm). This comparison is included as ULA are commonly used.

### Table I: Mean SDFB and mean residual noise power (RNP), each with standard deviation, of optimal configurations and the ULA, with unity input power

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>34.0±5.0</td>
<td>30.6±6.0</td>
<td>-5.8±0.6</td>
<td>5.4±0.6</td>
</tr>
<tr>
<td>5</td>
<td>48.7±6.7</td>
<td>50.1±6.9</td>
<td>-6.7±0.9</td>
<td>7.0±0.9</td>
</tr>
<tr>
<td>6</td>
<td>59.4±5.3</td>
<td>61.6±7.6</td>
<td>-7.3±1.2</td>
<td>-7.3±1.2</td>
</tr>
<tr>
<td>7</td>
<td>89.6±3.0</td>
<td>95.4±2.6</td>
<td>-7.8±1.5</td>
<td>-8.1±1.8</td>
</tr>
</tbody>
</table>

Fig. 4: (a)-(d) Directivity patterns of optimal array configurations with \( N \) = \{4, 5, 6, 7\} microphones and (e)-(f) ULA with \( N \) = \{6, 7\} microphones, steered to 0\(^\circ\) azimuth, in [dB]

The directivity patterns in (Figures 4(a)-(d)) depict the spatial selectivity of the optimal arrays for a steering angle \( \theta_s = 0\)^\( \circ \). The directivity patterns of the corresponding ULA for \( N \) = \{6, 7\} are also presented in Figures 4(e) and 4(f).

VI. Analysis

In general a trade-off between the SDFB and the residual noise power has to be accepted. Increasing the number of microphones \( N \), the residual noise power decreases, but the SDFB increases. This is to be expected: a larger number of microphones presents a greater spatial diversity, thus allowing for better noise suppression. However when \( d_{\text{max}} \) is constrained, increasing \( N \) implies smaller inter-microphone distances which worsens the conditioning of \( \Psi_{\text{NN}} \) and increases the SDFB.

We also see that the optimal array configurations outperform the ULA with \( N \geq 5 \) microphones. Furthermore the ULA exhibit regions which are affected by strong spatial aliasing, whereas in Figure 4(c) and 4(d) it can explicitly be seen that the optimal configurations provide attenuation in all directions but the steering direction.

A detailed look into the results (Table II) shows the variance of SDFB and noise reduction with the steering direction. Thereby it is conceivable that for applications with single steering directions, optimized arrays with even lower SDFB and noise power can be expected. Also the comparison of the optimal configurations in Table II with the ULA in Table III highlights the fact that only for a single steering direction the ULA perform better in terms of the RNP. When configurations are optimized only for this single steering direction they will outperform the ULA even in these cases.

Lastly, comparing configurations with a small mean SDFB to configurations with a high mean SDFB we can see that arrays with a high mean SDFB suffer under the common occurrence of \( d_{\text{min}} \). The worst results in terms of the mean SDFB are given by ULA with \( d = d_{\text{min}} = 2\) cm, as their SDFB is maximal. This observation is the basis for developing the distance histogram for analyzing the results in the next section.

### Table II: SDFB and residual noise power (RNP) in [dB] of optimal array configurations steering to \( \theta_s \), with unity input power

<table>
<thead>
<tr>
<th>( \theta_s )</th>
<th>( N = 4 )</th>
<th>( N = 5 )</th>
<th>( N = 6 )</th>
<th>( N = 7 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0(^\circ)</td>
<td>34</td>
<td>30</td>
<td>67</td>
<td>87</td>
</tr>
<tr>
<td>30(^\circ)</td>
<td>34</td>
<td>30</td>
<td>67</td>
<td>87</td>
</tr>
<tr>
<td>60(^\circ)</td>
<td>34</td>
<td>30</td>
<td>67</td>
<td>87</td>
</tr>
<tr>
<td>90(^\circ)</td>
<td>34</td>
<td>30</td>
<td>67</td>
<td>87</td>
</tr>
<tr>
<td>120(^\circ)</td>
<td>34</td>
<td>30</td>
<td>67</td>
<td>87</td>
</tr>
<tr>
<td>150(^\circ)</td>
<td>34</td>
<td>30</td>
<td>67</td>
<td>87</td>
</tr>
<tr>
<td>180(^\circ)</td>
<td>34</td>
<td>30</td>
<td>67</td>
<td>87</td>
</tr>
</tbody>
</table>

### Table III: SDFB and residual noise power (RNP) in [dB] of ULA configurations steering to \( \theta_s \), with unity input power

<table>
<thead>
<tr>
<th>( \theta_s )</th>
<th>( N = 4 )</th>
<th>( N = 5 )</th>
<th>( N = 6 )</th>
<th>( N = 7 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0(^\circ)</td>
<td>31</td>
<td>30</td>
<td>67</td>
<td>94</td>
</tr>
<tr>
<td>30(^\circ)</td>
<td>31</td>
<td>30</td>
<td>67</td>
<td>94</td>
</tr>
<tr>
<td>60(^\circ)</td>
<td>31</td>
<td>30</td>
<td>67</td>
<td>94</td>
</tr>
<tr>
<td>90(^\circ)</td>
<td>31</td>
<td>30</td>
<td>67</td>
<td>94</td>
</tr>
<tr>
<td>120(^\circ)</td>
<td>31</td>
<td>30</td>
<td>67</td>
<td>94</td>
</tr>
<tr>
<td>150(^\circ)</td>
<td>31</td>
<td>30</td>
<td>67</td>
<td>94</td>
</tr>
<tr>
<td>180(^\circ)</td>
<td>31</td>
<td>30</td>
<td>67</td>
<td>94</td>
</tr>
</tbody>
</table>
VII. DISTANCE HISTOGRAM

The distance histogram (DH) is designed as an easy-to-use rating tool for complex sensor arrays. Especially arrays geometries like planar arrays with a high number of sensors are difficult to compare. The DH visualizes the distribution of the inter-element distances, for now in linear array configurations. In further work it will be extended to planar arrays. Each color in the DH represents one array configuration. The y-axis value number for $d = 0$ cm indicates the number of microphones in the configuration. Comparing the DHs of the five best (Figure 5(a)-5(d)) and the five worst (Figure 5(e)-5(h)) configurations, a better insight into good array design can be obtained. Without going into the details of each configuration, we see that a ‘good’ array profits on the one hand from the occurrence of the maximum inter-element distance $d_{\text{max}}$ and on the other hand from the occurrence of many different inter-element distances. For instance, the maximum occurrence of an inter-element distance in configurations with $N \leq 6$ is two. In contrast, the ‘bad’ arrays show, firstly, a higher occurrence of small inter-element distances and, secondly, a maximum distance smaller than $d_{\text{max}} = 20$ cm in most cases. Thus, for linear arrays the DH as defined provides a first qualitative estimate of beamforming performance in terms of our newly defined measure.

![Distance Histogram](image)

Fig. 5: Distance Histogram of (a)-(d) five best and (e)-(h) five worst array configurations

We have presented a new performance measure for characterizing array configurations in context of MVDR-LS beamforming. The new measure represents both, the noise reduction capability and the robustness of the array to uncorrelated noise. Array configurations which are optimal in terms of this measure reduce the trade-off between directivity and white noise gain.

The evaluation of our simulation of linear arrays yielded preferable configurations for universal steering directions and wideband beamforming. For applications requiring a restricted range of steering vectors and bandwidth, one can similarly optimize the configuration according to the proposed measure and obtain improved results. In this context our measure can be used as input to an optimization procedure which will generate the optimal configuration. The comparison to ULAs showed that for five or more microphones the optimized configurations outperform the ULAs.

Further, we introduced the distance histogram as a demonstrative tool to estimate the performance of an array configuration. It provides information about the distribution of sensors, for now in a linear array, without the knowledge of the exact positions and allows a qualitative comparison of the ability of arrays for robust MVDR-LS beamforming, based on evaluations with our new measure. In general, a uniform distribution of distances with several different distances occurring over the maximum length of an array is preferable. Redundancy in the occurrence of distances results in worse beamformer performance.

This work lays the foundation for further research on holistic beamformer design that considers both sensor positioning and geometry and algorithmic aspects of beamforming. In our further investigation we shall extend the concepts presented here to different beamformer design algorithms and geometries, i.e. planar arrays. The distance histogram will also be extended to a general multi-dimensional case.

ACKNOWLEDGMENT

This work was supported in parts by the Ministry of Economic Affairs and Energy of the State of North Rhine-Westphalia, grant 31543-022/005-WFBO-009. The work of Nilesh Madhu was supported by the EU project AUDIS – digital signal processing in audiology (http://www.audis-itn.eu).

REFERENCES