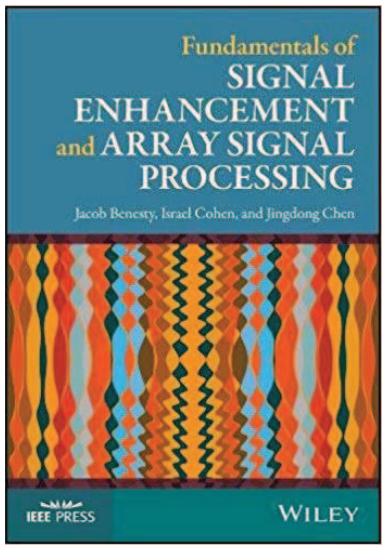
Adaptive Beamforming

J. Benesty, I. Cohen, and J. Chen,
Fundamentals of Signal Enhancement and Array Signal Processing,
Wiley-IEEE Press, 2017.



Outline

Introduction

Fixed beamformers do not depend on the statistics of the array data.

Fixed beamformers use a model for the noise field, and can work pretty well in different scenarios.

However, in very challenging acoustic environments, the performance of these algorithms, in terms of noise reduction, may be limited.

Therefore, in this talk we present adaptive beamformers, i.e., optimal linear filters that take into consideration the statistics of the incoming data.

Signal Model and Problem Formulation

We consider a plane wave, in the farfield that propagates in an anechoic acoustic environment at the speed of sound, i.e., $c = 340$ m/s, and impinges on a uniform linear sensor array consisting of M omnidirectional microphones (see Fig. 1).

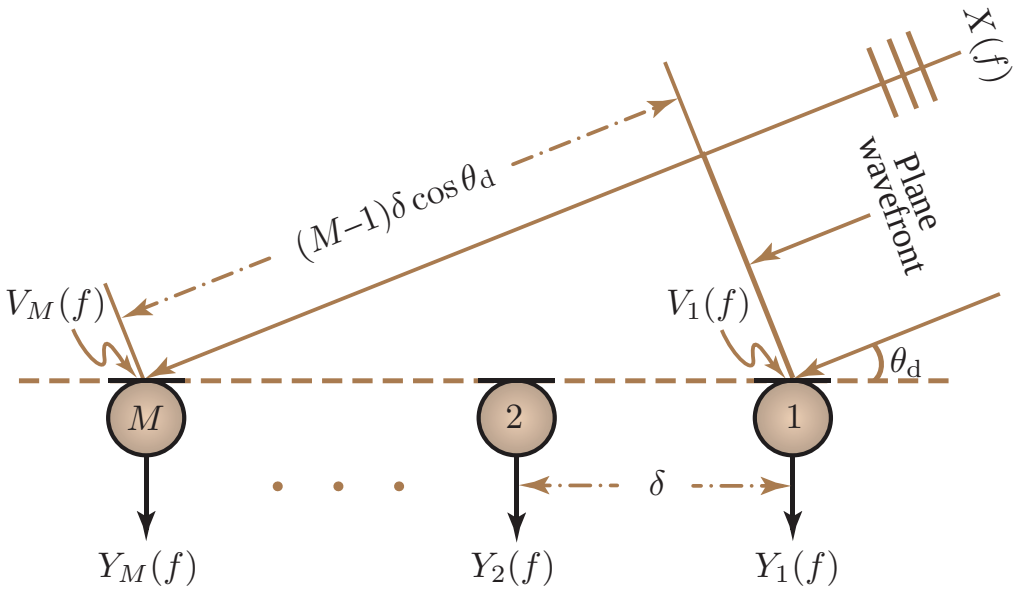


Figure 1: A uniform linear array with M sensors.

The signal model is given by

$$\begin{aligned}\mathbf{y}(f) &= \mathbf{x}(f) + \mathbf{v}(f) \\ &= \mathbf{d}\left(f, \cos\theta_{\mathrm{d}}\right) X(f) + \mathbf{v}(f),\end{aligned} \tag{1}$$

where $\mathbf{y}(f)$ is observation signal vector (of length M), $\mathbf{d}\left(f, \cos\theta_{\mathrm{d}}\right)$ is the steering vector associated with the desired signal, $X(f)$, impinging on the array from the direction θ_{d}, and $\mathbf{v}(f)$ is the noise signal vector.

The correlation matrix of $\mathbf{y}(f)$ is

$$\begin{aligned}\mathbf{\Phi}_{\mathbf{y}}(f) &= \mathbf{\Phi}_{\mathbf{x}}(f) + \mathbf{\Phi}_{\mathbf{v}}(f) \\ &= \phi_X(f)\mathbf{d}\left(f, \cos\theta_{\mathrm{d}}\right)\mathbf{d}^H\left(f, \cos\theta_{\mathrm{d}}\right) + \mathbf{\Phi}_{\mathbf{v}}(f),\end{aligned} \tag{2}$$

where $\mathbf{\Phi}_{\mathbf{x}}(f)$ and $\mathbf{\Phi}_{\mathbf{v}}(f)$ are the correlation matrices of $\mathbf{x}(f)$ and $\mathbf{v}(f)$, respectively, and $\phi_X(f)$ is the variance of $X(f)$.

Beamforming or linear filtering [1] consists of applying a complex-valued linear filter, $\mathbf{h}(f)$, of length M to $\mathbf{y}(f)$, i.e.,

$$\begin{aligned} Z(f) &= \mathbf{h}^H(f)\mathbf{y}(f) \\ &= \mathbf{h}^H(f)\left[\mathbf{x}(f) + \mathbf{v}(f)\right] \\ &= X_{\mathrm{fd}}(f) + V_{\mathrm{rn}}(f), \end{aligned} \tag{3}$$

where $Z(f)$ is, in general, the estimate of the desired signal, and $X_{\mathrm{fd}}(f)$ and $V_{\mathrm{rn}}(f)$ are the filtered desired signal and residual noise, respectively.

Assuming that $X_{\mathrm{fd}}(f)$ and $V_{\mathrm{rn}}(f)$ are uncorrelated, the variance of $Z(f)$ is

$$\begin{aligned} \phi_Z(f) &= \phi_{X_{\mathrm{fd}}}(f) + \phi_{V_{\mathrm{rn}}}(f) \\ &= \phi_X(f)\left|\mathbf{h}^H(f)\mathbf{d}\left(f, \cos\theta_{\mathrm{d}}\right)\right|^2 + \mathbf{h}^H(f)\mathbf{\Phi}_{\mathbf{v}}(f)\mathbf{h}(f). \end{aligned} \tag{4}$$

Our objective is to design and study beamformers that do depend on the statistics of the signals as well as the knowledge of the direction of the desired signal.

These so-called adaptive beamformers can usually adapt pretty quickly to changes in the environments in which they operate and do not rely on some model of the noise field such as in fixed beamformers.

Performance Measures
Signal-to-Noise Ratio

The narrowband and broadband input SNRs are, respectively,

$$\mathrm{iSNR}(f) = \frac{\phi_X(f)}{\phi_{V_1}(f)} \tag{5}$$

and

$$\mathrm{iSNR} = \frac{\int_f \phi_X(f)df}{\int_f \phi_{V_1}(f)df}, \tag{6}$$

where $\phi_{V_1}(f)$ is the variance of $V_1(f)$, which is the first element of the vector $\mathbf{v}(f)$.

From (4), we deduce the narrowband output SNR:

$$\begin{aligned} \mathrm{oSNR}\left[\mathbf{h}(f)\right] &= \frac{\phi_{X_{\mathrm{fd}}}(f)}{\phi_{V_{\mathrm{rn}}}(f)} \\ &= \frac{\phi_X(f)\left|\mathbf{h}^H(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)\right|^2}{\mathbf{h}^H(f)\mathbf{\Phi}_{\mathbf{v}}(f)\mathbf{h}(f)} \end{aligned} \tag{7}$$

and the broadband output SNR:

$$\mathrm{oSNR}\left(\mathbf{h}\right) = \frac{\int_f \phi_X(f)\left|\mathbf{h}^H(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)\right|^2 df}{\int_f \mathbf{h}^H(f)\mathbf{\Phi}_{\mathbf{v}}(f)\mathbf{h}(f)df}. \tag{8}$$

It follows from the definitions of the input and output SNRs that the narrowband and broadband array gains are, respectively,

$$\mathcal{G}\left[\mathbf{h}(f)\right] = \frac{\mathrm{oSNR}\left[\mathbf{h}(f)\right]}{\mathrm{iSNR}(f)}, \tag{9}$$

$$\mathcal{G}\left(\mathbf{h}\right) = \frac{\mathrm{oSNR}\left(\mathbf{h}\right)}{\mathrm{iSNR}}. \tag{10}$$

Adaptive beamformers should be designed in such a way that $\mathcal{G}\left[\mathbf{h}(f)\right] > 1$ and $\mathcal{G}\left(\mathbf{h}\right) > 1$.

Noise Rejection Factor

Other useful definitions to quantify noise reduction are the narrowband noise reduction factor:

$$\xi_{\mathrm{n}}\left[\mathbf{h}(f)\right] = \frac{\phi_{V_1}(f)}{\mathbf{h}^H(f)\mathbf{\Phi}_{\mathbf{v}}(f)\mathbf{h}(f)} \tag{11}$$

and the broadband noise reduction factor:

$$\xi_{\mathrm{n}}\left(\mathbf{h}\right) = \frac{\int_f \phi_{V_1}(f)df}{\int_f \mathbf{h}^H(f)\mathbf{\Phi}_{\mathbf{v}}(f)\mathbf{h}(f)df}. \tag{12}$$

In the distortionless case, i.e.,

$$\mathbf{h}^H(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right) = 1, \tag{13}$$

the noise reduction factor coincides with the array gain for both the narrowband and broadband measures.

Desired-Signal Reduction Factor

In order to quantify distortion of the desired signal due to the beamforming operation, we define the narrowband desired-signal reduction factor:

$$\xi_{\mathrm{d}}\left[\mathbf{h}(f)\right] = \frac{1}{\left|\mathbf{h}^{H}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)\right|^{2}} \tag{14}$$

and the broadband desired-signal reduction factor:

$$\xi_{\mathrm{d}}\left(\mathbf{h}\right) = \frac{\int_{f}\phi_{X}(f)df}{\int_{f}\phi_{X}(f)\left|\mathbf{h}^{H}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)\right|^{2}df}. \tag{15}$$

In the distortionless case, we have $\xi_{\mathrm{d}} = 1$, but when distortion occurs, we have $\xi_{\mathrm{d}} > 1$.

Desired-Signal Distortion Index

An alternative measure to the desired-signal reduction factor is the desired-signal distortion index.

We have the following definitions:

- the narrowband desired-signal distortion index,

$$\upsilon_{\mathrm{d}}\left[\mathbf{h}(f)\right] = \left|\mathbf{h}^H(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right) - 1\right|^2 \tag{16}$$

- and the broadband desired-signal distortion index,

$$\upsilon_{\mathrm{d}}\left(\mathbf{h}\right) = \frac{\int_f \phi_X(f)\left|\mathbf{h}^H(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right) - 1\right|^2 df}{\int_f \phi_X(f) df}. \tag{17}$$

Mean-Squared Error

The error signal between the estimated and desired-signals at the frequency f is given by

$$\begin{aligned}\mathcal{E}(f) &= Z(f) - X(f) \\ &= X_{\mathrm{fd}}(f) + V_{\mathrm{rn}}(f) - X(f) \\ &= \mathcal{E}_{\mathrm{d}}(f) + \mathcal{E}_{\mathrm{n}}(f),\end{aligned} \tag{18}$$

where

$$\mathcal{E}_{\mathrm{d}}(f) = \left[\mathbf{h}^H(f)\mathbf{d}(f, \cos\theta_{\mathrm{d}}) - 1\right] X(f) \tag{19}$$

is the desired-signal distortion due to the beamformer and

$$\mathcal{E}_{\mathrm{n}}(f) = \mathbf{h}^H(f)\mathbf{v}(f) \tag{20}$$

represents the residual noise.

Assuming that $\mathcal{E}_{\mathrm{d}}(f)$ and $\mathcal{E}_{\mathrm{n}}(f)$ are incoherent, the narrowband MSE can be expressed as

$$\begin{aligned} J\left[\mathbf{h}(f)\right] &= E\left[\left|\mathcal{E}(f)\right|^2\right] \\ &= E\left[\left|\mathcal{E}_{\mathrm{d}}(f)\right|^2\right] + E\left[\left|\mathcal{E}_{\mathrm{n}}(f)\right|^2\right] \\ &= J_{\mathrm{d}}\left[\mathbf{h}(f)\right] + J_{\mathrm{n}}\left[\mathbf{h}(f)\right] \\ &= \phi_X(f) + \mathbf{h}^H(f)\mathbf{\Phi}_{\mathbf{y}}(f)\mathbf{h}(f) - \phi_X(f)\mathbf{h}^H(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right) - \\ &\quad \phi_X(f)\mathbf{d}^H\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{h}(f), \end{aligned} \tag{21}$$

where

$$\begin{aligned} J_{\mathrm{d}}\left[\mathbf{h}(f)\right] &= \phi_X(f)\left|\mathbf{h}^H(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right) - 1\right|^2 \\ &= \phi_X(f)\upsilon_{\mathrm{d}}\left[\mathbf{h}(f)\right] \end{aligned} \tag{22}$$

and

$$J_{\mathrm{n}}\left[\mathbf{h}(f)\right] = \mathbf{h}^{H}(f)\mathbf{\Phi}_{\mathbf{v}}(f)\mathbf{h}(f) = \frac{\phi_{V_1}(f)}{\xi_{\mathrm{n}}\left[\mathbf{h}(f)\right]}. \tag{23}$$

We have the following classical relationships:

$$\frac{J_{\mathrm{d}}\left[\mathbf{h}(f)\right]}{J_{\mathrm{n}}\left[\mathbf{h}(f)\right]} = \mathrm{iSNR}(f) \times \xi_{\mathrm{n}}\left[\mathbf{h}(f)\right] \times \upsilon_{\mathrm{d}}\left[\mathbf{h}(f)\right] = \mathrm{oSNR}\left[\mathbf{h}(f)\right] \times \xi_{\mathrm{d}}\left[\mathbf{h}(f)\right] \times \upsilon_{\mathrm{d}}\left[\mathbf{h}(f)\right]. \tag{24}$$

Adaptive Beamformers

Wiener

In this section, we show how to design different kinds of adaptive beamformers.

For each one of them, we give several equivalent formulations, depending on what (second-order) statistics we want or need to estimate.

The Wiener beamformer is found by minimizing the narrowband MSE, $J\left[\mathbf{h}(f)\right]$ [eq. (21)].

We easily obtain

$$\mathbf{h}_{\mathrm{W}}\left(f,\cos\theta_{\mathrm{d}}\right)=\phi_{X}(f)\mathbf{\Phi}_{\mathbf{y}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right). \tag{25}$$

In the Wiener beamformer, we need to estimate $\phi_X(f)$ and $\mathbf{\Phi_y}(f)$.

The latter quantity is easy to estimate from the observations, but the former is not.

Let

$$\mathbf{\Gamma_y}(f) = \frac{\mathbf{\Phi_y}(f)}{\phi_{Y_1}(f)} \tag{26}$$

be the pseudo-coherence matrix of the observations, where $\phi_{Y_1}(f)$ is the variance of $Y_1(f)$, we can rewrite (25) as

$$\begin{aligned}\mathbf{h}_{\mathrm{W}}(f, \cos\theta_{\mathrm{d}}) &= \frac{\mathrm{iSNR}(f)}{1+\mathrm{iSNR}(f)}\mathbf{\Gamma_y}^{-1}(f)\mathbf{d}(f, \cos\theta_{\mathrm{d}}) \\ &= H_{\mathrm{W}}(f)\mathbf{\Gamma_y}^{-1}(f)\mathbf{d}(f, \cos\theta_{\mathrm{d}}),\end{aligned} \tag{27}$$

where

$$H_{\mathrm{W}}(f) = \frac{\mathrm{iSNR}(f)}{1 + \mathrm{iSNR}(f)} \tag{28}$$

is the single-channel Wiener gain.

Now, instead of estimating $\phi_X(f)$ as in (25), we need to estimate the narrowband input SNR, $\mathrm{iSNR}(f)$ or, equivalently, $H_{\mathrm{W}}(f)$.

The Wiener filter can also be expressed as a function of the statistics of the observation and noise signals, i.e.,

$$\mathbf{h}_{\mathrm{W}}\left(f, \cos\theta_{\mathrm{d}}\right) = \left[\mathbf{I}_M - \mathbf{\Phi}_{\mathbf{y}}^{-1}(f)\mathbf{\Phi}_{\mathbf{v}}(f)\right]\mathbf{i}_{\mathrm{i}}, \tag{29}$$

where $\mathbf{i}_{\mathrm{i}}$ is the first column of $\mathbf{I}_M$.

Determining the inverse of $\mathbf{\Phi}_{\mathbf{y}}(f)$ from (2) with the Woodbury's identity, we get

$$\mathbf{\Phi}_{\mathbf{y}}^{-1}(f) = \mathbf{\Phi}_{\mathbf{v}}^{-1}(f) - \frac{\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{d}^{H}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)}{\phi_{X}^{-1}(f)+\mathbf{d}^{H}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)}. \quad (30)$$

Substituting (30) into (25) gives

$$\begin{aligned}\mathbf{h}_{\mathrm{W}}\left(f,\cos\theta_{\mathrm{d}}\right) &= \frac{\phi_{X}(f)\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)}{1+\phi_{X}(f)\mathbf{d}^{H}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)} \\ &= \frac{\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{\Phi}_{\mathbf{y}}(f)-\mathbf{I}_{M}}{1-M+\mathrm{tr}\left[\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{\Phi}_{\mathbf{y}}(f)\right]}\mathbf{i}_{\mathrm{i}}. \quad (31)\end{aligned}$$

In the second equation of (31), $\mathbf{h}_{\mathrm{W}}\left(f,\cos\theta_{\mathrm{d}}\right)$ depends on the statistics of the observation and noise signals and the matrix $\mathbf{\Phi}_{\mathbf{v}}(f)$ is inverted, while in the formulaton given in (29), $\mathbf{h}_{\mathrm{W}}\left(f,\cos\theta_{\mathrm{d}}\right)$ depends on the same statistics but the matrix $\mathbf{\Phi}_{\mathbf{y}}(f)$ is inverted.

The Wiener beamformer maximizes the narrowband array gain but does not necessarily maximize the broadband array gain.

Distortion is obviously expected and is increased when the input SNR is decreased.

However, if we increase the number of sensors, we decrease distortion.

Example 1

Consider a ULA of M sensors.

Suppose that a desired signal impinges on the ULA from the direction θ_{d}, and that two statistically independent interferences impinge on the ULA from directions θ_1 and θ_2.

Assume that the desired signal is a harmonic pulse of T samples:

$$x(t) = \begin{cases} A\sin\left(2\pi f_0 t + \phi\right), & 0 \leq t \leq T-1 \\ 0, & t < 0,\ t \geq T \end{cases},$$

with fixed, but unknown, amplitude A and frequency f_0, and random phase ϕ, uniformly distributed on the interval from 0 to 2π.

Assume that the interferences $u_1(t)$ and $u_2(t)$ are IID white Gaussian noise, i.e., $u_1(t), u_2(t) \sim \mathcal{N}\left(0, \sigma_u^2\right)$, uncorrelated with $x(t)$.

In addition, the sensors contain thermal white Gaussian noise, $w_m(t) \sim \mathcal{N}\left(0, \sigma_w^2\right)$, that are mutually uncorrelated.

The noisy received signals are given by $y_m(t) = x_m(t) + v_m(t),\ m = 1, \ldots, M$, where $v_m(t) = u_m(t) + w_m(t),\ m = 1, \ldots, M$ are the interference-plus-noise signals.

For simplicity, we choose a sampling interval T_{s} that satisfies $T_{\mathrm{s}} = \delta / c$.

The variance of $X(f)$ is given by

$$\phi_X(f) = \frac{A^2}{4} D_T^2 \left[\pi \left(f + f_0\right)\right] + \frac{A^2}{4} D_T^2 \left[\pi \left(f - f_0\right)\right],$$

where

$$D_T(x) = \frac{\sin(Tx)}{\sin(x)}.$$

The correlation matrices of $\mathbf{x}(f)$ and $\mathbf{v}(f)$ are given by

$$\begin{aligned}\mathbf{\Phi_x}(f) &= \phi_X(f)\mathbf{d}\left(f,\cos\theta_\mathrm{d}\right)\mathbf{d}^H\left(f,\cos\theta_\mathrm{d}\right),\\ \mathbf{\Phi_v}(f) &= T\sigma_u^2\mathbf{d}\left(f,\cos\theta_1\right)\mathbf{d}^H\left(f,\cos\theta_1\right)+\\ &\qquad T\sigma_u^2\mathbf{d}\left(f,\cos\theta_2\right)\mathbf{d}^H\left(f,\cos\theta_2\right)+T\sigma_w^2\mathbf{I}_M.\end{aligned}$$

The narrowband and broadband input SNRs are, respectively,

$$\begin{aligned}\mathrm{iSNR}(f) &= \frac{\phi_X(f)}{\phi_{V_1}(f)} \\ &= \frac{A^2}{4T\left(2\sigma_u^2+\sigma_w^2\right)} D_T^2\left[\pi\left(f+f_0\right)\right] + \\ &\quad \frac{A^2}{4T\left(2\sigma_u^2+\sigma_w^2\right)} D_T^2\left[\pi\left(f-f_0\right)\right]\end{aligned}$$

and

$$\mathrm{iSNR} = \frac{\int_f \phi_X(f)df}{\int_f \phi_{V_1}(f)df} = \frac{\sum_t E\left[\left|x_1(t)\right|^2\right]}{\sum_t E\left[\left|v_1(t)\right|^2\right]} = \frac{A^2}{2\left(2\sigma_u^2+\sigma_w^2\right)},$$

where we have used Parseval's identity.

The Wiener beamformer, $\mathbf{h}_{\mathrm{W}}\left(f, \cos\theta_{\mathrm{d}}\right)$, is obtained from (29).

To demonstrate the performance of the Wiener beamformer, we choose $A = 0.5$, $f_0 = 0.1c/\delta$, $T = 500$, $\theta_{\mathrm{d}} = 70°$, $\theta_1 = 30°$, $\theta_2 = 50°$, and $\sigma_w^2 = 0.01\sigma_u^2$.

Figure 2 shows plots of the broadband gain in SNR, $\mathcal{G}\left(\mathbf{h}_{\mathrm{W}}\right)$, the broadband noise reduction factor, $\xi_{\mathrm{n}}\left(\mathbf{h}_{\mathrm{W}}\right)$, the broadband desired-signal reduction factor, $\xi_{\mathrm{d}}\left(\mathbf{h}_{\mathrm{W}}\right)$, and the broadband desired-signal distortion index, $\upsilon_{\mathrm{d}}\left(\mathbf{h}_{\mathrm{W}}\right)$, as a function of the broadband input SNR, for different numbers of sensors, M.

For a given input SNR, as the number of sensors increases, the gain in SNR and the noise reduction factor increase, while the desired-signal reduction factor and the desired-signal distortion index decrease.

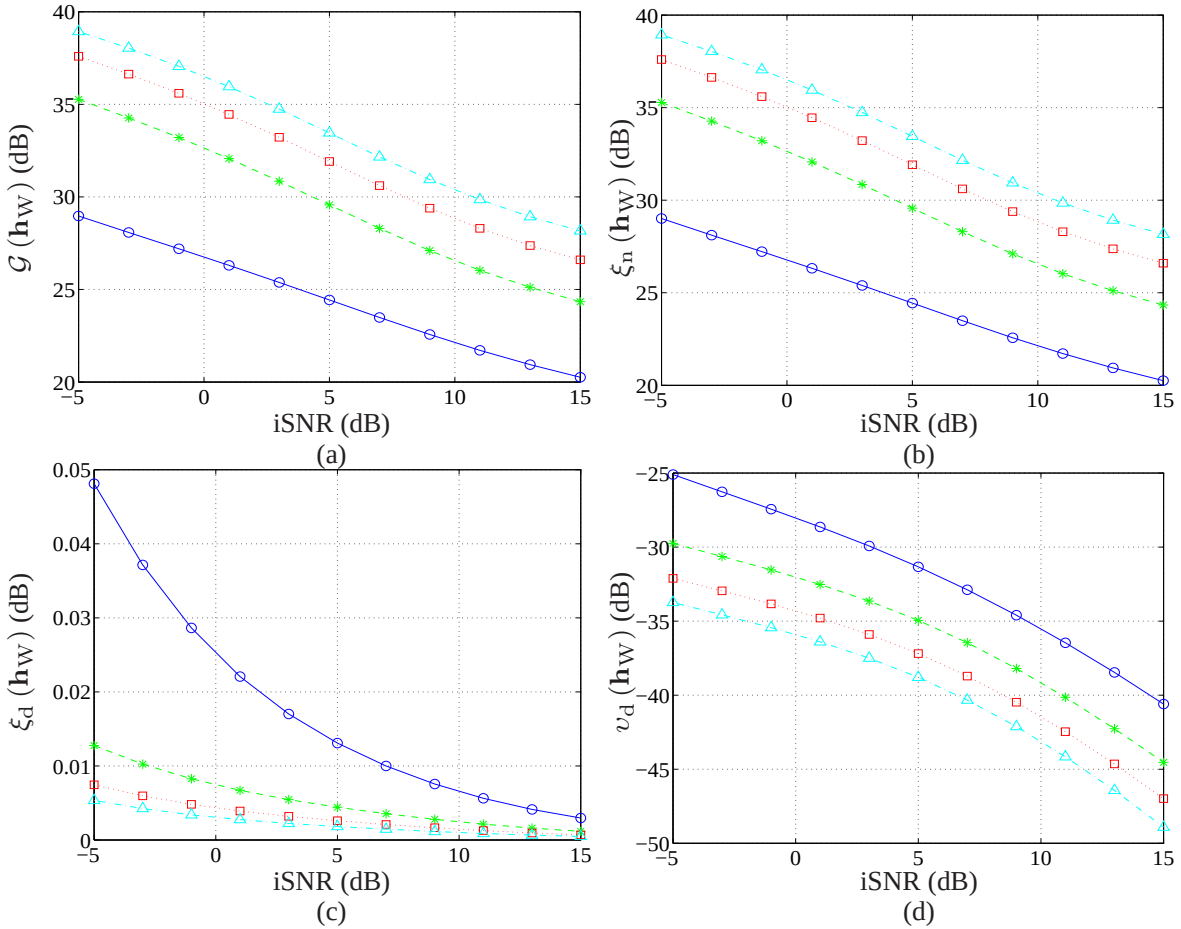


Figure 2: The broadband (a) gain in SNR, (b) noise reduction factor, (c) desired-signal reduction factor, and (d) desired-signal distortion index of the Wiener beamformer for: $M = 10$ (solid line with circles), $M = 20$ (dashed line with asterisks), $M = 30$ (dotted line with squares), and $M = 40$ (dash-dot line with triangles).

Figure 3 shows beampatterns, $|\mathcal{B}\left[\mathbf{h}_{\mathrm{W}}\left(f,\cos\theta_{\mathrm{d}}\right),\cos\theta\right]|$, for $f=f_0$ and different numbers of sensors, M. The main beam is in the direction of the desired signal, i.e., θ_{d}, and there are nulls in the directions of the interferences, i.e., θ_1 and θ_2.

As the number of sensors increases, the width of the main beam decreases, and the nulls in the directions of the interferences become deeper.

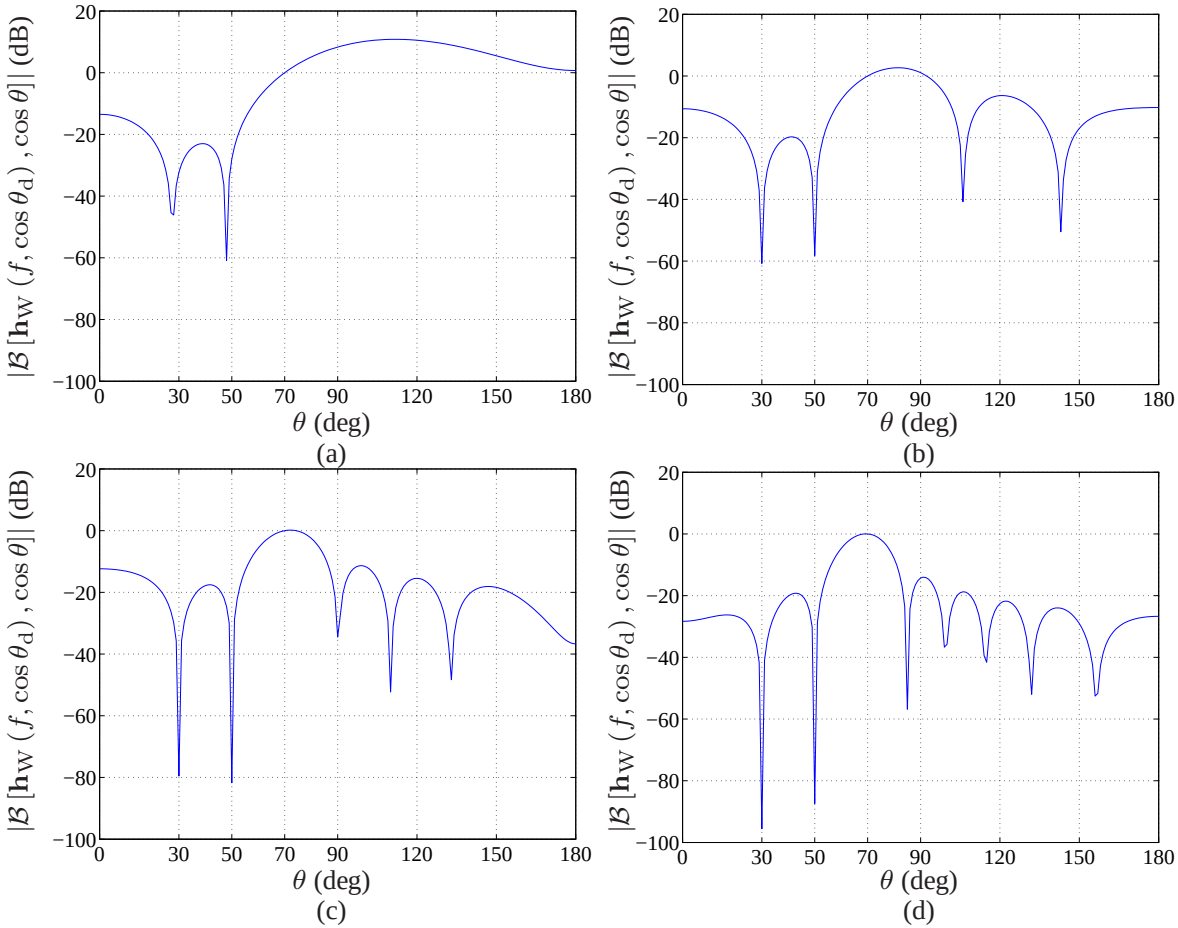


Figure 3: Beampatterns of the Wiener beamformer for $f = f_0$ and different numbers of sensors, M: (a) $M = 10$, (b) $M = 20$, (c) $M = 30$, and (d) $M = 40$.

MVDR

The MVDR beamformer proposed by Capon [2], [3] is obtained by minimizing the narrowband MSE of the residual noise, $J_{\mathrm{r}}\left[\mathbf{h}(f)\right]$, subject to the distortionless constraint, i.e.,

$$\min_{\mathbf{h}(f)} \mathbf{h}^H(f)\mathbf{\Phi}_{\mathbf{v}}(f)\mathbf{h}(f) \quad \text{subject to} \quad \mathbf{h}^H(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)=1. \tag{32}$$

The solution to this optimization problem is

$$\mathbf{h}_{\mathrm{MVDR}}\left(f,\cos\theta_{\mathrm{d}}\right)=\frac{\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)}{\mathbf{d}^H\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)}, \tag{33}$$

which depends on the statistics of the noise only.

Using the Woodbury's identity, it is easy to show that the MVDR beamformer is also

$$\begin{aligned}\mathbf{h}_{\mathrm{MVDR}}\left(f,\cos\theta_{\mathrm{d}}\right) &= \frac{\mathbf{\Phi}_{\mathbf{y}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)}{\mathbf{d}^{H}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{\Phi}_{\mathbf{y}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)} \\ &= \frac{\mathbf{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)}{\mathbf{d}^{H}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)}.\end{aligned} \tag{34}$$

This formulation is really important and practical, since it depends on the statistics of the observations only, which can be easily estimated in practice.

It is clear that the MVDR beamformer maximizes the narrowband array gain, however, for the broadband array gain, we always have

$$1 \leq \mathcal{G}\left(\mathbf{h}_{\mathrm{MVDR}}\right) \leq \mathcal{G}\left(\mathbf{h}_{\mathrm{W}}\right). \tag{35}$$

From a theoretical point of view, it is also clear that we have

$$\upsilon_{\mathrm{d}}\left[\mathbf{h}_{\mathrm{MVDR}}\left(f,\cos\theta_{\mathrm{d}}\right)\right]=0, \tag{36}$$

$$\upsilon_{\mathrm{d}}\left(\mathbf{h}_{\mathrm{MVDR}}\right)=0. \tag{37}$$

However, in practice, this is not true in general because of the reverberation, which is not taken into account in our model.

Example 2

Returning to Example 1, we now employ the MVDR beamformer, $\mathbf{h}_{\mathrm{MVDR}}\left(f, \cos\theta_{\mathrm{d}}\right)$, given in (34).

Figure 4 shows plots of the broadband gain in SNR, $\mathcal{G}\left(\mathbf{h}_{\mathrm{MVDR}}\right)$, the broadband noise reduction factor, $\xi_{\mathrm{n}}\left(\mathbf{h}_{\mathrm{MVDR}}\right)$, the broadband desired-signal reduction factor, $\xi_{\mathrm{d}}\left(\mathbf{h}_{\mathrm{MVDR}}\right)$, and the broadband MSE, $J\left(\mathbf{h}_{\mathrm{MVDR}}\right)$, as a function of the broadband input SNR, for different numbers of sensors, M.

For a given broadband input SNR, as the number of sensors increases, the broadband gain in SNR and the broadband noise reduction factor increase, while the broadband MSE decreases.

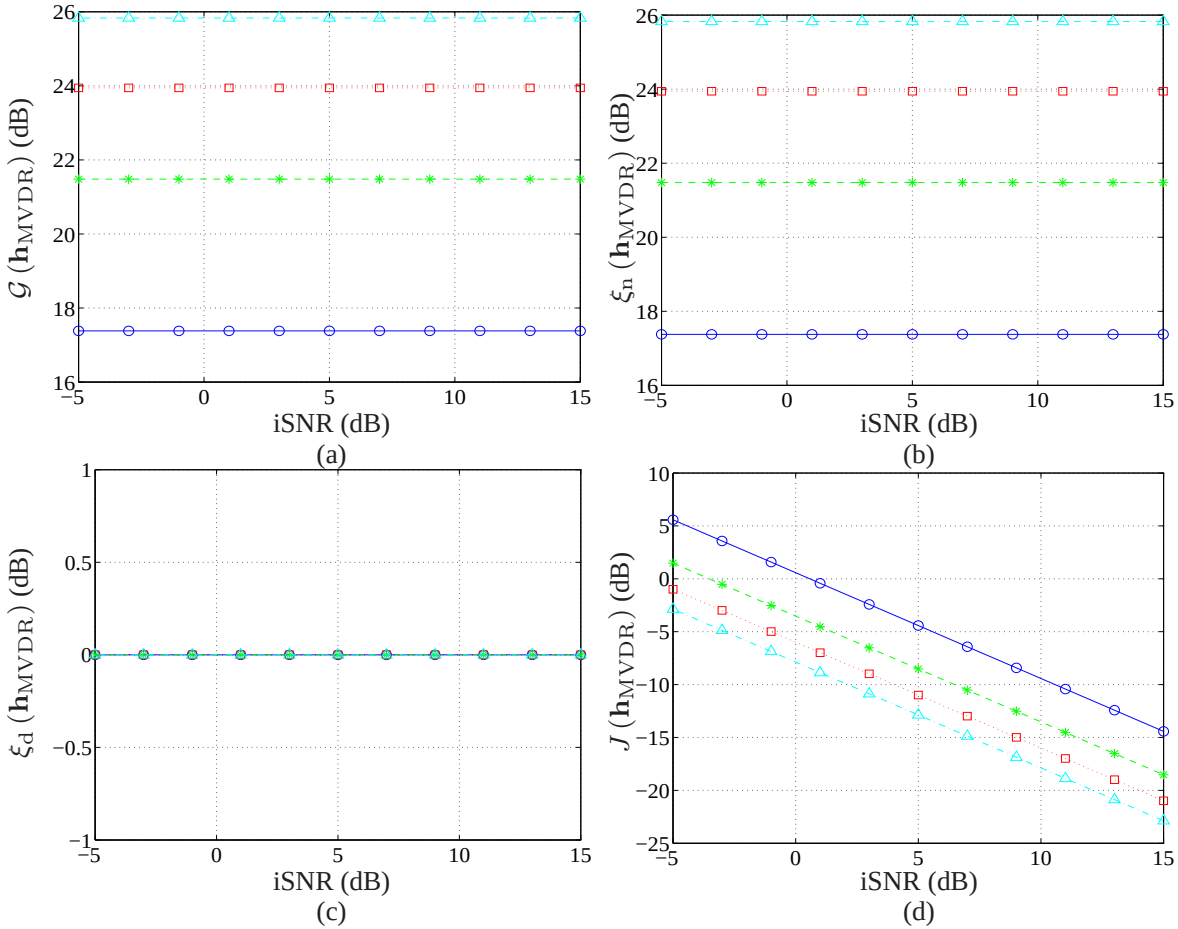


Figure 4: The broadband (a) gain in SNR, (b) noise reduction factor, (c) desired-signal reduction factor, and (d) MSE of the MVDR beamformer for different numbers of sensors, M: $M = 10$ (solid line with circles), $M = 20$ (dashed line with asterisks), $M = 30$ (dotted line with squares), and $M = 40$ (dash-dot line with triangles).

Figure 5 shows beampatterns, $|\mathcal{B}\left[\mathbf{h}_{\mathrm{MVDR}}\left(f,\cos\theta_{\mathrm{d}}\right),\cos\theta\right]|$, for $f=f_0$ and different numbers of sensors, M.

The main beam is in the direction of the desired signal, i.e., θ_{d}, and there are nulls in the directions of the interferences, i.e., θ_1 and θ_2.

In particular, $|\mathcal{B}\left[\mathbf{h}_{\mathrm{MVDR}}\left(f,\cos\theta_{\mathrm{d}}\right),\cos\theta\right]|$ is 1 for $\theta=\theta_{\mathrm{d}}$.

As the number of sensors increases, the width of the main beam decreases, and the nulls in the directions of the interferences become deeper.

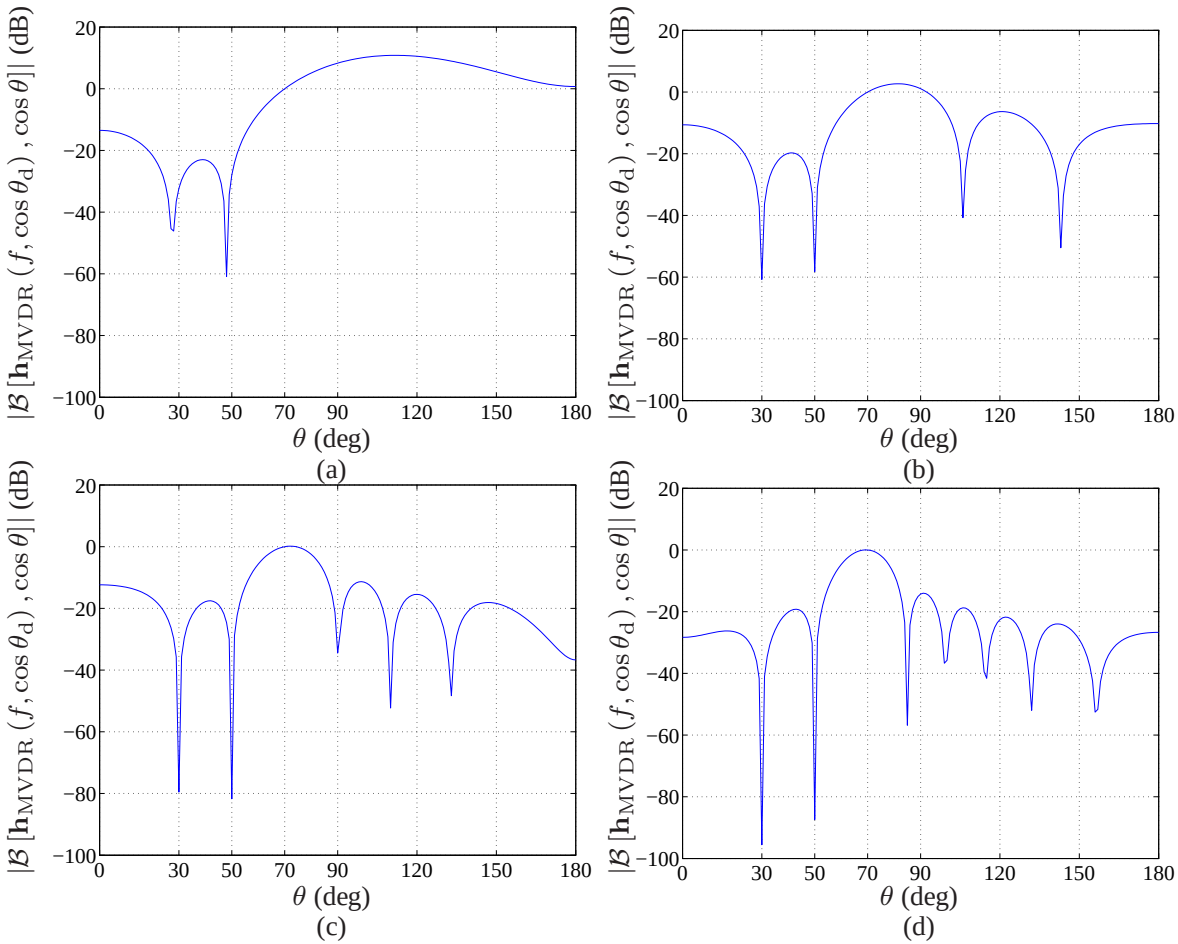


Figure 5: Beampatterns of the MVDR beamformer for $f = f_0$ and different numbers of sensors, M: (a) $M = 10$, (b) $M = 20$, (c) $M = 30$, and (d) $M = 40$.

Tradeoff

In order to better compromise between noise reduction and signal distortion, we constrain the noise reduction factor to be equal to a positive value that is greater than 1, i.e.,

$$\min_{\mathbf{h}(f)} J_{\mathrm{d}}\left[\mathbf{h}(f)\right] \quad \text{subject to} \quad J_{\mathrm{n}}\left[\mathbf{h}(f)\right] = \aleph \phi_{V_1}(f), \tag{38}$$

where $0 < \aleph < 1$ to insure that we get some noise reduction.

By using a Lagrange multiplier, $\mu > 0$, we get the tradeoff beamformer:

$$\begin{aligned}
\mathbf{h}_{\mathrm{T},\mu}\left(f, \cos\theta_{\mathrm{d}}\right) &= \phi_X(f)\left[\mathbf{\Phi}_{\mathbf{x}}(f) + \mu \mathbf{\Phi}_{\mathbf{v}}(f)\right]^{-1} \mathbf{d}\left(f, \cos\theta_{\mathrm{d}}\right) \\
&= \frac{\phi_X(f)\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{d}\left(f, \cos\theta_{\mathrm{d}}\right)}{\mu + \phi_X(f)\mathbf{d}^H\left(f, \cos\theta_{\mathrm{d}}\right)\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{d}\left(f, \cos\theta_{\mathrm{d}}\right)} \\
&= \frac{\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{\Phi}_{\mathbf{y}}(f) - \mathbf{I}_M}{\mu - M + \mathrm{tr}\left[\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{\Phi}_{\mathbf{y}}(f)\right]}\mathbf{i}_{\mathrm{i}\cdot}
\end{aligned} \tag{39}$$

We can see that for

- $\mu = 1$, $\mathbf{h}_{\mathrm{T},1}\left(f, \cos\theta_{\mathrm{d}}\right) = \mathbf{h}_{\mathrm{W}}\left(f, \cos\theta_{\mathrm{d}}\right)$, which is the Wiener beamformer;
- $\mu = 0$, $\mathbf{h}_{\mathrm{T},0}\left(f, \cos\theta_{\mathrm{d}}\right) = \mathbf{h}_{\mathrm{MVDR}}\left(f, \cos\theta_{\mathrm{d}}\right)$, which is the MVDR beamformer;
- $\mu > 1$, results in a beamformer with low residual noise at the expense of high desired-signal distortion (as compared to Wiener); and
- $\mu < 1$, results in a beamformer with high residual noise and low desired-signal distortion (as compared to Wiener).

A more interesting way to express the tradeoff beamformer is

$$\mathbf{h}_{\mathrm{T},\mu}\left(f,\cos\theta_{\mathrm{d}}\right)= \left[(1-\mu)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{d}^{H}\left(f,\cos\theta_{\mathrm{d}}\right)+\mu\frac{1+\mathrm{iSNR}(f)}{\mathrm{iSNR}(f)}\mathbf{\Gamma}_{\mathbf{y}}(f)\right]^{-1}\times \mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right), \tag{40}$$

or, equivalently, with the help of the Woodbury's identity as

$$\mathbf{h}_{\mathrm{T},\mu}\left(f,\cos\theta_{\mathrm{d}}\right)=\frac{\mathrm{iSNR}(f)}{1+\mathrm{iSNR}(f)}\times \frac{\mathbf{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)}{\mu+(1-\mu)\frac{\mathrm{iSNR}(f)}{1+\mathrm{iSNR}(f)}\mathbf{d}^{H}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)}. \tag{41}$$

The previous expression depends only on the estimation of the statistics of the observations as well as the estimation of the narrowband input SNR.

We can simplify (41) by writing it as a function of $H_{\mathrm{W}}(f)$, i.e.,

$$\mathbf{h}_{\mathrm{T},\mu}\left(f,\cos\theta_{\mathrm{d}}\right)=\frac{H_{\mathrm{W}}(f)\mathbf{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)}{\mu+(1-\mu)H_{\mathrm{W}}(f)\mathbf{d}^{H}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)}. \tag{42}$$

Obviously, the tradeoff beamformer also maximizes the narrowband array gain, $\forall\mu\geq 0$.

However, for the broadband array gain, we always have for $\mu \geq 1$,

$$1 \leq \mathcal{G}\left(\mathbf{h}_{\mathrm{MVDR}}\right) \leq \mathcal{G}\left(\mathbf{h}_{\mathrm{W}}\right) \leq \mathcal{G}\left(\mathbf{h}_{\mathrm{T},\mu}\right), \tag{43}$$

and for $0 \leq \mu \leq 1$,

$$1 \leq \mathcal{G}\left(\mathbf{h}_{\mathrm{MVDR}}\right) \leq \mathcal{G}\left(\mathbf{h}_{\mathrm{T},\mu}\right) \leq \mathcal{G}\left(\mathbf{h}_{\mathrm{W}}\right). \tag{44}$$

Distortion of the desired signal, on the other hand, depends quite a lot on the values of μ.

However, the closer is the value of μ to 0, the less distorted is the desired signal.

Example 3

Returning to Example 1, we now employ the tradeoff beamformer, $\mathbf{h}_{\mathrm{T},\mu}\left(f, \cos\theta_{\mathrm{d}}\right)$, given in (39).

Figure 6 shows plots of the broadband gain in SNR, $\mathcal{G}\left(\mathbf{h}_{\mathrm{T},\mu}\right)$, the broadband noise reduction factor, $\xi_{\mathrm{n}}\left(\mathbf{h}_{\mathrm{T},\mu}\right)$, the broadband desired-signal reduction factor, $\xi_{\mathrm{d}}\left(\mathbf{h}_{\mathrm{T},\mu}\right)$, and the broadband desired-signal distortion index, $\upsilon_{\mathrm{d}}\left(\mathbf{h}_{\mathrm{T},\mu}\right)$, as a function of the broadband input SNR, for $M = 30$ and several values of μ.

For a given broadband input SNR, the higher is the value of μ, the higher are the broadband gain in SNR and the broadband noise reduction factor, but at the expense of higher broadband desired-signal reduction factor and higher broadband desired-signal distortion index.

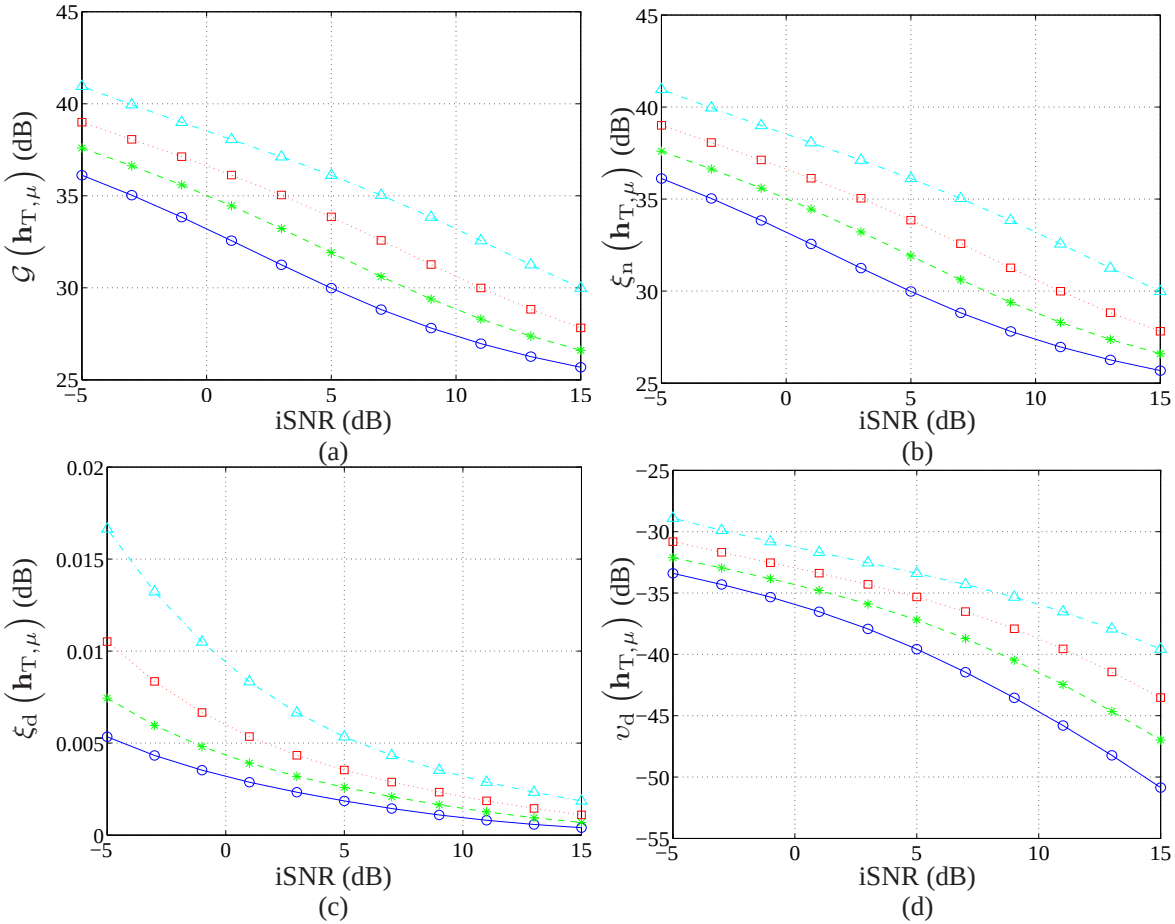


Figure 6: The broadband (a) gain in SNR, (b) noise reduction factor, (c) desired-signal reduction factor, and (d) desired-signal distortion index of the tradeoff beamformer for several values of μ: $\mu = 0.5$ (solid line with circles), $\mu = 1$ (dashed line with asterisks), $\mu = 2$ (dotted line with squares), and $\mu = 5$ (dash-dot line with triangles).

Figure 7 shows beampatterns, $|\mathcal{B}\left[\mathbf{h}_{\mathrm{T},\mu}\left(f,\cos\theta_{\mathrm{d}}\right),\cos\theta\right]|$, for $f=f_0$, $M=30$ and several values of μ.

The main beam is in the direction of the desired signal, i.e., θ_{d}, and there are nulls in the directions of the interferences, i.e., θ_1 and θ_2.

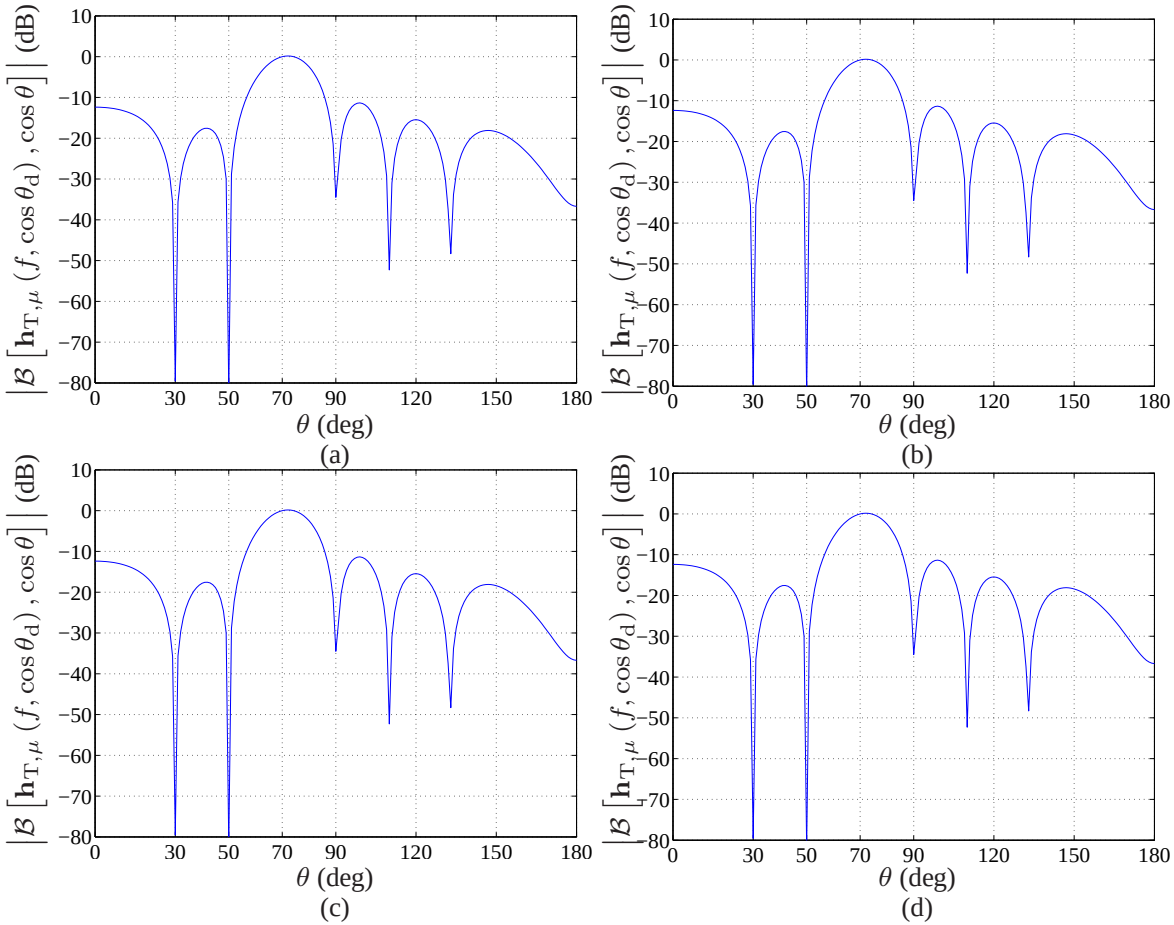


Figure 7: Beampatterns of the tradeoff beamformer for $f = f_0$, $M = 30$ and several values of μ: (a) $\mu = 0.5$, (b) $\mu = 1$, (c) $\mu = 2$, and (d) $\mu = 5$.

Maximum Array Gain

we can express the narrowband array gain as

$$\mathcal{G}\left[\mathbf{h}(f)\right] = \frac{\phi_{V_1}(f)\mathbf{h}^H(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{d}^H\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{h}(f)}{\mathbf{h}^H(f)\mathbf{\Phi}_{\mathbf{v}}(f)\mathbf{h}(f)}. \tag{45}$$

The maximum array gain beamformer, $\mathbf{h}_{\max}\left(f,\cos\theta_{\mathrm{d}}\right)$, is obtained by maximizing the array gain as given above.

In (45), we recognize the generalized Rayleigh quotient [4].

It is well known that this quotient is maximized with the maximum eigenvector of the matrix $\phi_{V_1}(f)\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{d}^H\left(f,\cos\theta_{\mathrm{d}}\right)$.

Let us denote by $\lambda_{\max}\left(f,\cos\theta_{\mathrm{d}}\right)$ the maximum eigenvalue corresponding to this maximum eigenvector.

Since the rank of the mentioned matrix is equal to 1, we have

$$\begin{aligned}\lambda_{\max}\left(f,\cos\theta_{\mathrm{d}}\right) &= \mathrm{tr}\left[\phi_{V_1}\left(f\right)\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{d}^{H}\left(f,\cos\theta_{\mathrm{d}}\right)\right] \\ &= \phi_{V_1}\left(f\right)\mathbf{d}^{H}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right).\end{aligned} \tag{46}$$

As a result,

$$\begin{aligned}\mathcal{G}\left[\mathbf{h}_{\max}\left(f,\cos\theta_{\mathrm{d}}\right)\right] &= \lambda_{\max}\left(f,\cos\theta_{\mathrm{d}}\right) \\ &= \mathcal{G}_{\max}\left(f,\cos\theta_{\mathrm{d}}\right),\end{aligned} \tag{47}$$

which corresponds to the maximum possible narrowband array gain.

Obviously, we also have

$$\mathbf{h}_{\max}\left(f,\cos\theta_{\mathrm{d}}\right) = \varsigma(f)\mathbf{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right), \tag{48}$$

where $\varsigma(f) \neq 0$ is an arbitrary frequency-dependent complex number.

We can observe that all beamformers derived so far are equivalent up to a scaling factor.

LCMV

Assume that we have N interferences, with $N < M$, impinging on the array from the directions $\theta_1 \neq \theta_2 \neq \cdots \neq \theta_N \neq \theta_{\mathrm{d}}$.

We would like to place nulls in the directions $\theta_n,\ n = 1, 2, \ldots, N$, with a beamformer $\mathbf{h}(f)$, and, meanwhile, recover the desired source coming from the direction θ_{d}.

Combining all these constraints together, we get the constraint:

$$\mathbf{C}^H\left(f, \theta_{\mathrm{d}}, \theta_{1:N}\right)\mathbf{h}(f) = \mathbf{i}_{\mathrm{c}}, \tag{49}$$

where

$$\mathbf{C}\left(f, \theta_{\mathrm{d}}, \theta_{1:N}\right) = \left[\begin{array}{cccc} \mathbf{d}\left(f, \theta_{\mathrm{d}}\right) & \mathbf{d}\left(f, \theta_1\right) & \cdots & \mathbf{d}\left(f, \theta_N\right) \end{array}\right] \tag{50}$$

is a matrix of size $M \times (N+1)$ whose columns are linearly independent and $\mathbf{i}_{\mathrm{c}} = \left[\begin{array}{cccc} 1 & 0 & \cdots & 0 \end{array}\right]^T$ is a vector of length $N+1$.

The most convenient way to solve this problem is by minimizing the narrowband MSE of the residual noise, $J_{\mathrm{r}}\left[\mathbf{h}(f)\right]$, subject to (49), i.e.,

$$\min_{\mathbf{h}(f)} \mathbf{h}^H(f)\mathbf{\Phi}_{\mathbf{v}}(f)\mathbf{h}(f) \quad \text{subject to} \quad \mathbf{C}^H\left(f,\theta_{\mathrm{d}},\theta_{1:N}\right)\mathbf{h}(f)=\mathbf{i}_{\mathrm{c}}. \tag{51}$$

The solution to this optimization problem gives the well-known LCMV beamformer [5], [6]:

$$\mathbf{h}_{\mathrm{LCMV}}\left(f,\cos\theta_{\mathrm{d}}\right)=\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{C}\left(f,\theta_{\mathrm{d}},\theta_{1:N}\right)\times \left[\mathbf{C}^H\left(f,\theta_{\mathrm{d}},\theta_{1:N}\right)\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{C}\left(f,\theta_{\mathrm{d}},\theta_{1:N}\right)\right]^{-1}\mathbf{i}_{\mathrm{c}}, \tag{52}$$

which depends on the statistics of the noise only.

It can be shown that a more interesting formulation of the LCMV beamformer is

$$\mathbf{h}_{\mathrm{LCMV}}\left(f,\cos\theta_{\mathrm{d}}\right)=\boldsymbol{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{C}\left(f,\theta_{\mathrm{d}},\theta_{1:N}\right)\times \left[\mathbf{C}^{H}\left(f,\theta_{\mathrm{d}},\theta_{1:N}\right)\boldsymbol{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{C}\left(f,\theta_{\mathrm{d}},\theta_{1:N}\right)\right]^{-1}\mathbf{i}_{\mathrm{c}}. \tag{53}$$

The previous expression depends on the statistics of the observations only, which should be easy to estimate.

Example 4

Returning to Example 1, we now employ the LCMV beamformer, $\mathbf{h}_{\mathrm{LCMV}}\left(f, \cos\theta_{\mathrm{d}}\right)$, given in (53).

Figure 8 shows plots of the broadband gain in SNR, $\mathcal{G}\left(\mathbf{h}_{\mathrm{LCMV}}\right)$, the broadband noise reduction factor, $\xi_{\mathrm{n}}\left(\mathbf{h}_{\mathrm{LCMV}}\right)$, the broadband desired-signal reduction factor, $\xi_{\mathrm{d}}\left(\mathbf{h}_{\mathrm{LCMV}}\right)$, and the broadband MSE, $J\left(\mathbf{h}_{\mathrm{LCMV}}\right)$, as a function of the broadband input SNR, for different numbers of sensors, M.

For a given broadband input SNR, as the number of sensors increases, the broadband gain in SNR and the broadband noise reduction factor increase, while the broadband MSE decreases.

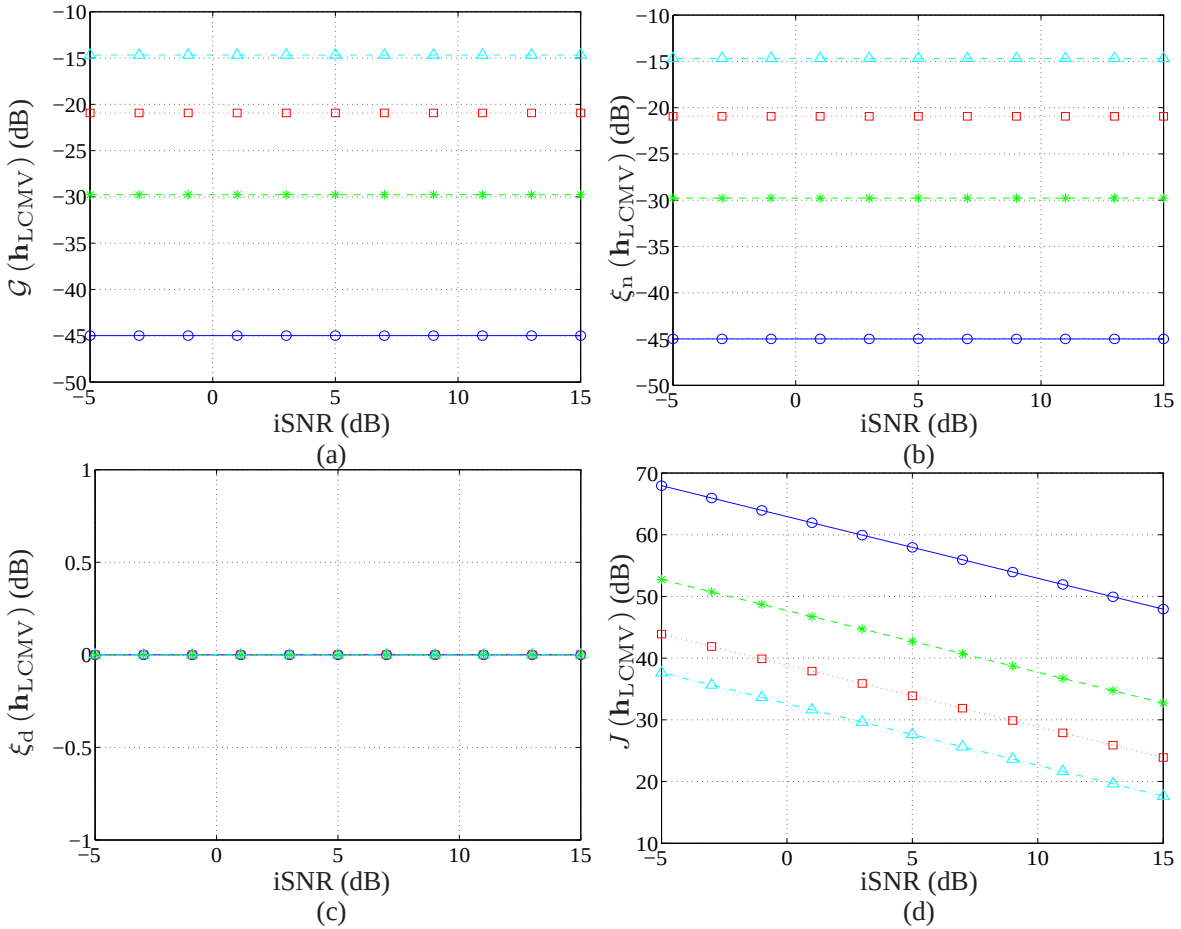


Figure 8: The broadband (a) gain in SNR, (b) noise reduction factor, (c) desired-signal reduction factor, and (d) MSE of the LCMV beamformer for different numbers of sensors, M: $M = 10$ (solid line with circles), $M = 20$ (dashed line with asterisks), $M = 30$ (dotted line with squares), and $M = 40$ (dash-dot line with triangles).

Figure 9 shows beampatterns, $|\mathcal{B}\left[\mathbf{h}_{\mathrm{LCMV}}\left(f,\cos\theta_{\mathrm{d}}\right),\cos\theta\right]|$, for $f=f_0$ and different numbers of sensors, M.

The main beam is in the direction of the desired signal, i.e., θ_{d}, and there are nulls in the directions of the interferences, i.e., θ_1 and θ_2.

In particular, $|\mathcal{B}\left[\mathbf{h}_{\mathrm{LCMV}}\left(f,\cos\theta_{\mathrm{d}}\right),\cos\theta\right]|$ is 1 for $\theta=\theta_{\mathrm{d}}$, and is identically zero for $\theta=\theta_1$ and $\theta=\theta_2$.

As the number of sensors increases, the width of the main beam decreases.

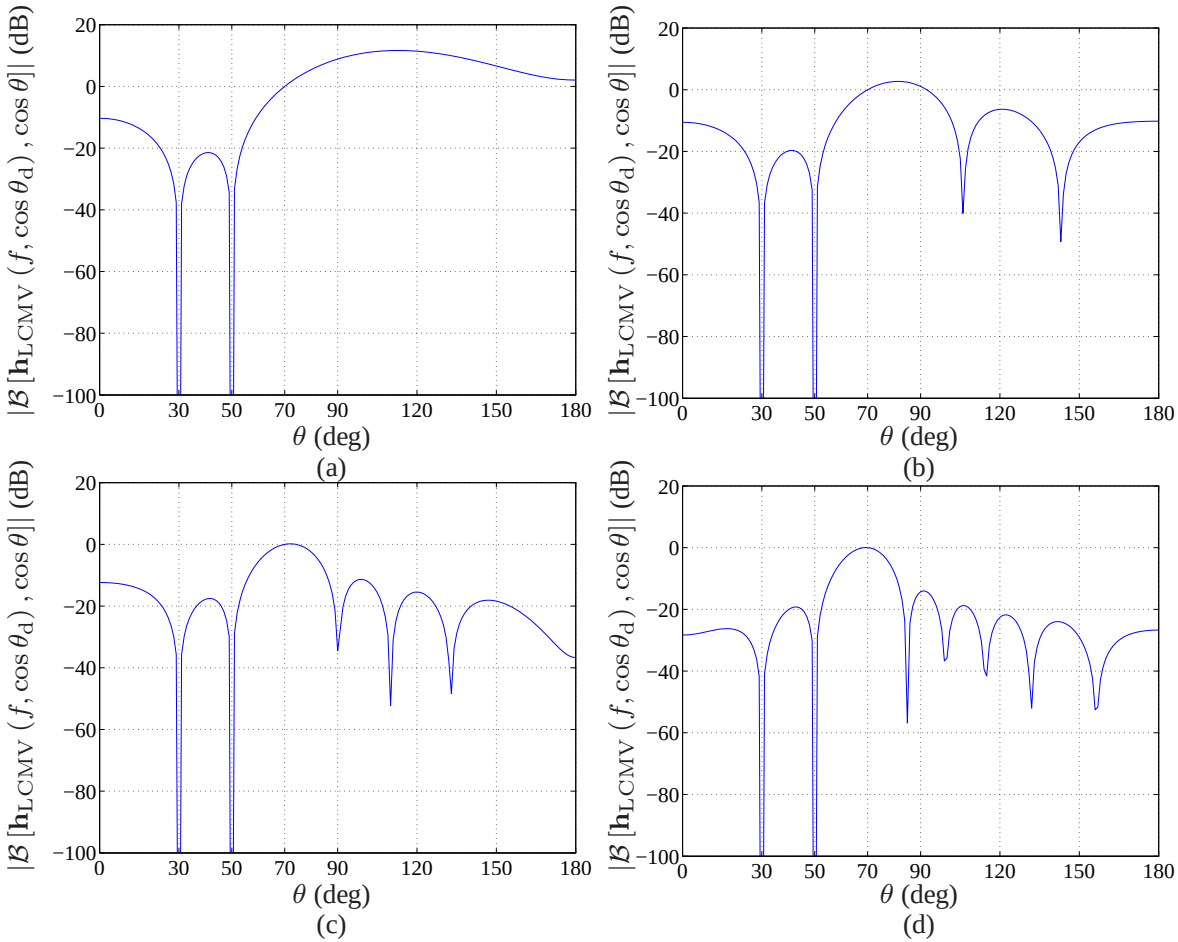


Figure 9: Beampatterns of the LCMV beamformer for $f = f_0$ and different numbers of sensors, M: (a) $M = 10$, (b) $M = 20$, (c) $M = 30$, and (d) $M = 40$.

Table 1 summarizes all the optimal adaptive beamformers studied in this section.

Table 1: Adaptive beamformers.

Wiener:	$\mathbf{h}_{\mathrm{W}}\left(f,\cos\theta_{\mathrm{d}}\right)=H_{\mathrm{W}}(f)\mathbf{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)$
MVDR:	$\mathbf{h}_{\mathrm{MVDR}}\left(f,\cos\theta_{\mathrm{d}}\right)=\dfrac{\mathbf{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)}{\mathbf{d}^{H}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)}$
Tradeoff:	$\mathbf{h}_{\mathrm{T},\mu}\left(f,\cos\theta_{\mathrm{d}}\right)=\dfrac{H_{\mathrm{W}}(f)\mathbf{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)}{\mu+(1-\mu)H_{\mathrm{W}}(f)\mathbf{d}^{H}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)}$
Max. Array Gain:	$\mathbf{h}_{\max}\left(f,\cos\theta_{\mathrm{d}}\right)=\varsigma(f)\mathbf{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right),\ \varsigma(f)\neq 0$
LCMV:	$\mathbf{h}_{\mathrm{LCMV}}\left(f,\cos\theta_{\mathrm{d}}\right)=\mathbf{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{C}\left(f,\theta_{\mathrm{d}},\theta_{1:N}\right)\times\left[\mathbf{C}^{H}\left(f,\theta_{\mathrm{d}},\theta_{1:N}\right)\mathbf{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{C}\left(f,\theta_{\mathrm{d}},\theta_{1:N}\right)\right]^{-1}\mathbf{i}_{\mathrm{c}}$

SNR Estimation

From Table 1, we observe that all the beamformers depend on the statistics of the observations, i.e., $\mathbf{\Gamma_y}(f)$, while some of them depend also on the narrowband input SNR, i.e., $\mathrm{iSNR}(f)$, or, equivalently, on $H_{\mathrm{W}}(f)$.

In practice, while it is easy to get an estimate for $\mathbf{\Gamma_y}(f)$, it is not for $\mathrm{iSNR}(f)$.

In this section, we show one possible way to estimate this SNR.

In fact, it is much more natural to estimate $H_{\mathrm{W}}(f)$ as explained below.

We can express the pseudo-coherence matrix of the observations as

$$\begin{aligned}\mathbf{\Gamma}_{\mathbf{y}}(f) &= \frac{\mathrm{iSNR}(f)}{1+\mathrm{iSNR}(f)}\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{d}^{H}\left(f,\cos\theta_{\mathrm{d}}\right)+\frac{1}{1+\mathrm{iSNR}(f)}\mathbf{\Gamma}_{\mathbf{v}}(f) \\ &= H_{\mathrm{W}}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{d}^{H}\left(f,\cos\theta_{\mathrm{d}}\right)+\left[1-H_{\mathrm{W}}(f)\right]\mathbf{\Gamma}_{\mathbf{v}}(f),\end{aligned} \tag{54}$$

where

$$\mathbf{\Gamma}_{\mathbf{v}}(f)=\frac{\mathbf{\Phi}_{\mathbf{v}}(f)}{\phi_{V_1}(f)}. \tag{55}$$

Let us assume that we are in the presence of the spherically isotropic noise.

In this case, $\mathbf{\Gamma}_{\mathbf{v}}(f)$ coincides with $\mathbf{\Gamma}_{0,\pi}(f)$.

Since $\mathbf{y}(f)$ is observable, it is easy to estimate $\mathbf{\Gamma}_{\mathbf{y}}(f)$.

We denote this estimate by $\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f)$.

By following the approaches developed in [7], [8], [9], [10], we can write the components of the matrix $\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f)$ as

$$\Re\left\{\left[\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f)\right]_{ij}\right\} = \widehat{H}_{\mathrm{W}}(f)\Re\left[D_i\left(f,\cos\theta_{\mathrm{d}}\right)D_j^*\left(f,\cos\theta_{\mathrm{d}}\right)\right]+ \left[1-\widehat{H}_{\mathrm{W}}(f)\right]\left[\mathbf{\Gamma}_{0,\pi}(f)\right]_{ij}, \tag{56}$$

$$\Im\left\{\left[\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f)\right]_{ij}\right\} = \widehat{H}_{\mathrm{W}}(f)\Im\left[D_i\left(f,\cos\theta_{\mathrm{d}}\right)D_j^*\left(f,\cos\theta_{\mathrm{d}}\right)\right], \tag{57}$$

for $i \neq j,\ i,j = 1,2,\ldots,M$, where $\Re[\cdot]$ and $\Im[\cdot]$ are the real part and imaginary part operators, respectively, $\widehat{H}_{\mathrm{W}}(f)$ is the estimate of $H_{\mathrm{W}}(f)$, and $D_m\left(f,\cos\theta_{\mathrm{d}}\right)$ is the mth element of $\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)$.

We deduce from the previous expressions that

$$\widehat{H}_{\mathrm{W}}(f) = \frac{\Re\left\{\left[\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f)\right]_{ij}\right\} - \left[\mathbf{\Gamma}_{0,\pi}(f)\right]_{ij}}{\Re\left[D_i\left(f,\cos\theta_{\mathrm{d}}\right)D_j^*\left(f,\cos\theta_{\mathrm{d}}\right)\right] - \left[\mathbf{\Gamma}_{0,\pi}(f)\right]_{ij}}, \; i \neq j, \tag{58}$$

$$\widehat{H}_{\mathrm{W}}(f) = \frac{\Im\left\{\left[\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f)\right]_{ij}\right\}}{\Im\left[D_i\left(f,\cos\theta_{\mathrm{d}}\right)D_j^*\left(f,\cos\theta_{\mathrm{d}}\right)\right]}, \; i \neq j. \tag{59}$$

To get a much more reliable estimate, it is better to average (58) and (59) over all possible sensor combinations [7], [8], [9], [10], resulting in the estimator:

$$\widehat{H}_{\mathrm{W}}(f) = \frac{1}{M(M-1)} \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \frac{\Re\left\{ \left[\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f) \right]_{ij} \right\} - \left[\mathbf{\Gamma}_{0,\pi}(f) \right]_{ij}}{\Re\left[D_i\left(f, \cos\theta_{\mathrm{d}}\right) D_j^*\left(f, \cos\theta_{\mathrm{d}}\right) \right] - \left[\mathbf{\Gamma}_{0,\pi}(f) \right]_{ij}} + \frac{1}{M(M-1)} \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \frac{\Im\left\{ \left[\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f) \right]_{ij} \right\}}{\Im\left[D_i\left(f, \cos\theta_{\mathrm{d}}\right) D_j^*\left(f, \cos\theta_{\mathrm{d}}\right) \right]}. \quad (60)$$

Obviously, in practice, it is much better to estimate $H_{\mathrm{W}}(f)$ than $\mathrm{iSNR}(f)$ since the former is bounded, i.e., $0 \leq H_{\mathrm{W}}(f) \leq 1$, while the latter is not.

If the estimate of the Wiener gain is greater than 1, then we should force it to 1, and if it is negative, we should put it to 0.

It is possible to estimate the single-channel Wiener gain directly from (54), in a much simpler way, by pre- and post-multiplying both sides of (54) by $\mathbf{d}^H\left(f,\cos\theta_{\mathrm{d}}\right)$ and $\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)$, respectively, and by replacing $\mathbf{\Gamma_y}(f)$ and $\mathbf{\Gamma_v}(f)$ with $\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f)$ and $\mathbf{\Gamma}_{0,\pi}(f)$, respectively.

We easily obtain

$$\widehat{H}_{\mathrm{W}}(f)=\frac{\mathbf{d}^H\left(f,\cos\theta_{\mathrm{d}}\right)\left[\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f)-\mathbf{\Gamma}_{0,\pi}(f)\right]\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)}{M^2-\mathbf{d}^H\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{\Gamma}_{0,\pi}(f)\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)}. \tag{61}$$

Example 5

Consider a ULA of M sensors, and a desired signal that impinges on the ULA from the direction $\theta_{\mathrm{d}} = 70°$.

Assume that the desired signal is a harmonic pulse of T samples:

$$x(t) = \begin{cases} A\sin\left(2\pi f_0 t + \phi\right), & 0 \le t \le T-1 \\ 0, & t < 0,\ t \ge T \end{cases},$$

with fixed, but unknown, amplitude A and frequency f_0, and random phase ϕ, uniformly distributed on the interval from 0 to 2π.

Assume that the interference $u_m(t)$ is a diffuse noise uncorrelated with $x(t)$.

In addition, the sensors contain thermal white Gaussian noise, $w_m(t) \sim \mathcal{N}\left(0, \sigma_w^2\right)$, that are mutually uncorrelated.

The noisy received signals are given by
$y_m(t) = x_m(t) + v_m(t),\ m = 1, \ldots, M$, where
$v_m(t) = u_m(t) + w_m(t),\ m = 1, \ldots, M$ are the interference-plus-noise signals.

The pseudo-coherence matrix of the noise can be written as

$$\mathbf{\Gamma}_{\mathbf{v}}(f) = (1 - \alpha)\, \mathbf{\Gamma}_{0,\pi}(f) + \alpha \mathbf{I}_M, \tag{62}$$

where α ($0 \leq \alpha \leq 1$) is related to the ratio between the powers of the thermal and diffuse noises.

Figures 10 and 11 show plots of the estimators $\widehat{H}_{\mathrm{W}}(f)$, given by (60) and (61), respectively, as a function of the narrowband input SNR for several values of α and different numbers of sensors, M.

The theoretical plot is indicated by thick solid line, and the pseudo-coherence matrix of the observations $\mathbf{\Gamma}_{\mathbf{y}}(f)$ is obtained by

$$\mathbf{\Gamma}_{\mathbf{y}}(f) = \frac{\mathrm{iSNR}(f)}{1+\mathrm{iSNR}(f)}\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{d}^{H}\left(f,\cos\theta_{\mathrm{d}}\right) + \frac{(1-\alpha)\,\mathbf{\Gamma}_{0,\pi}(f)+\alpha\mathbf{I}_{M}}{1+\mathrm{iSNR}(f)}. \tag{63}$$

As the number of sensors is larger and as the value of α is smaller, the estimators (60) and (61) are closer to the theoretical values.

Generally, the estimator (61) produces better results than (60).

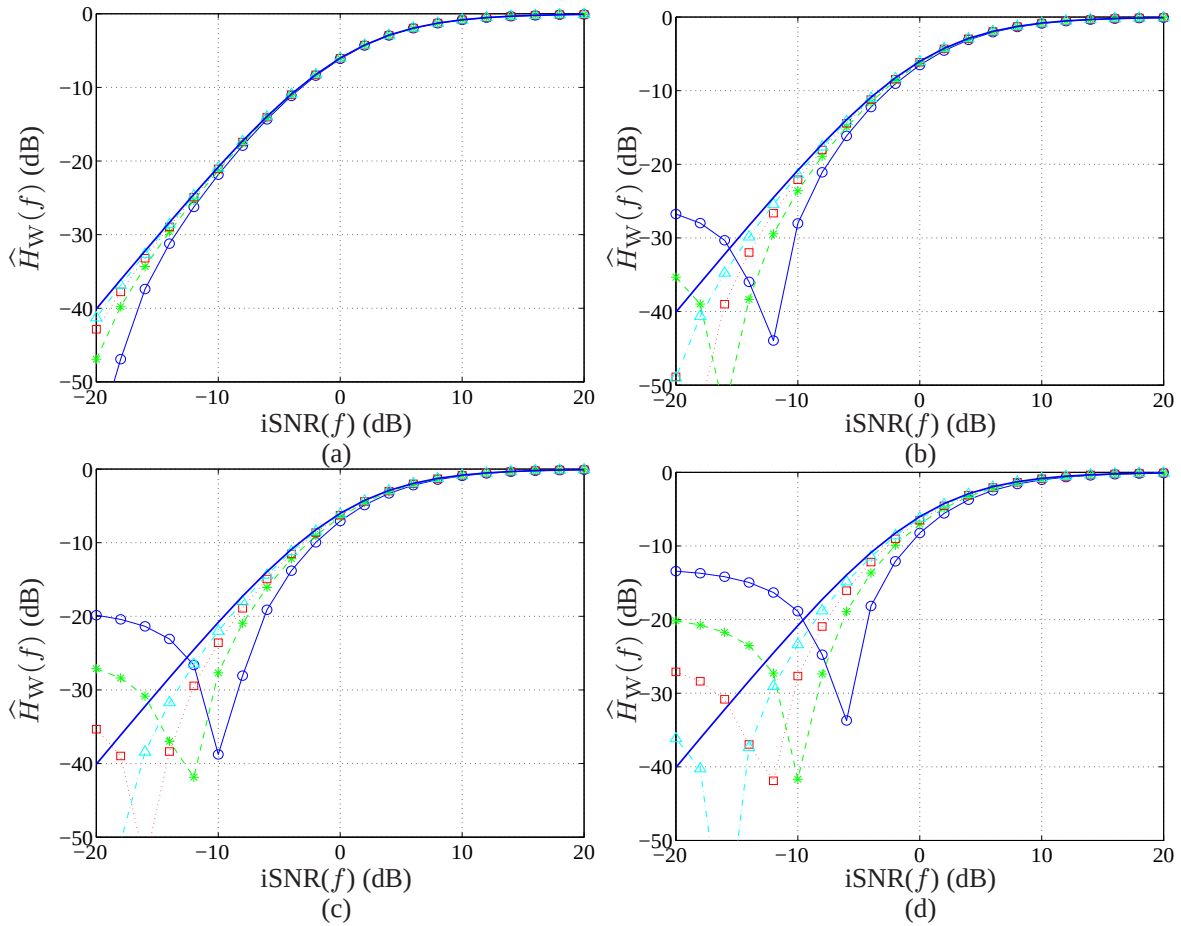


Figure 10: The estimator $\widehat{H}_{\mathrm{W}}(f)$, given by (60) for: (a) $\alpha = 0.001$, (b) $\alpha = 0.005$, (c) $\alpha = 0.01$, and (d) $\alpha = 0.02$. Each figure shows the theoretical plot (thick solid line), and estimates for $M = 2$ (solid line with circles), $M = 5$ (dashed line with asterisks), $M = 10$ (dotted line with squares), and $M = 20$ (dash-dot line with triangles).

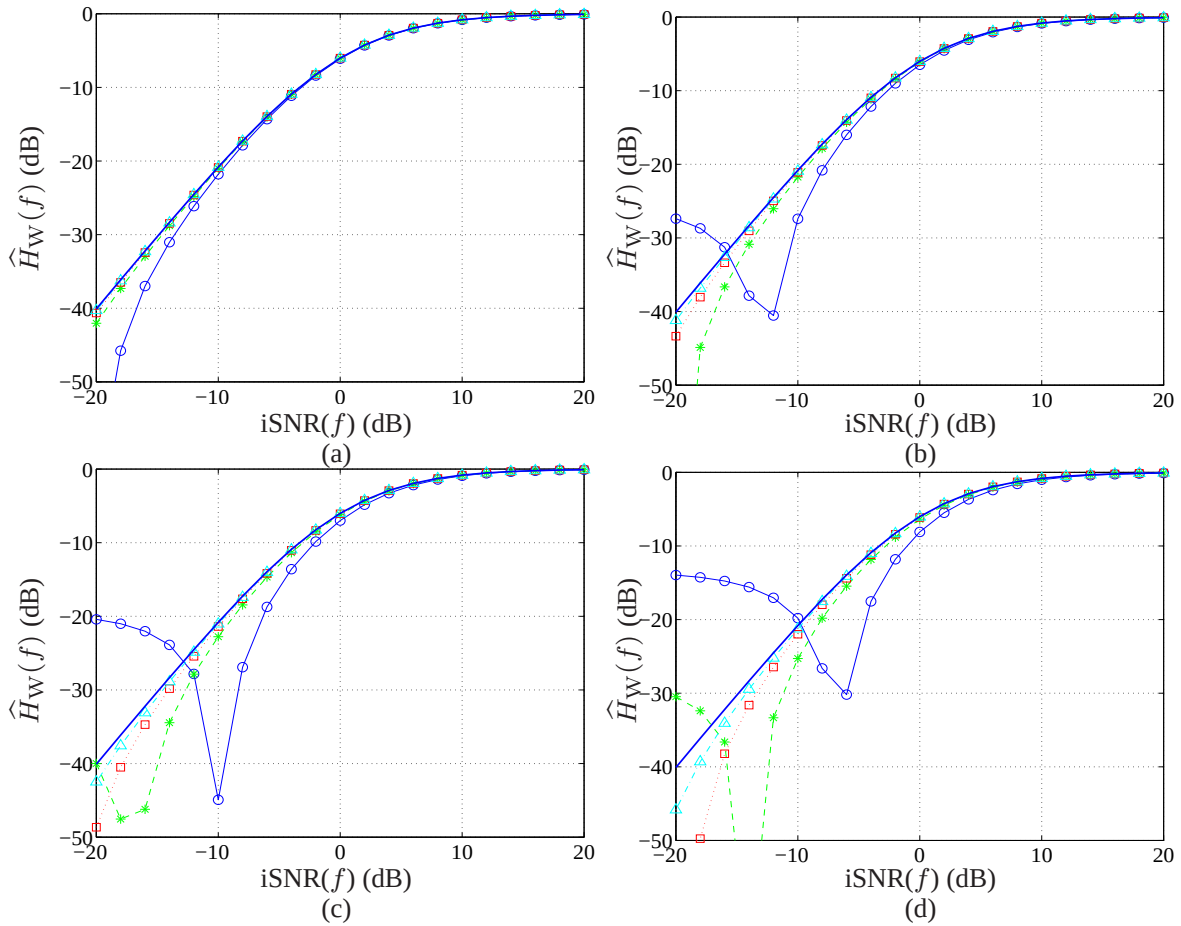


Figure 11: The estimator $\widehat{H}_{\mathrm{W}}(f)$, given by (61) for: (a) $\alpha = 0.001$, (b) $\alpha = 0.005$, (c) $\alpha = 0.01$, and (d) $\alpha = 0.02$. Each figure shows the theoretical plot (thick solid line), and estimates for $M = 2$ (solid line with circles), $M = 5$ (dashed line with asterisks), $M = 10$ (dotted line with squares), and $M = 20$ (dash-dot line with triangles).

Now we set α to 0.001, and estimate the pseudo-coherence matrix of the observations using K random snapshots:

$$\widehat{\mathbf{\Phi}}_{\mathbf{y}}(f) = \frac{1}{K}\sum_{k=1}^{K}\mathbf{y}_k(f)\mathbf{y}_k^H(f), \tag{64}$$

$$\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f) = \frac{\widehat{\mathbf{\Phi}}_{\mathbf{y}}(f)}{\widehat{\phi}_{Y_1}(f)}, \tag{65}$$

where $\mathbf{y}_k(f)$ is a random snapshot of $\mathbf{y}(f)$.

Figures 12 and 13 show plots of the estimators $\widehat{H}_{\mathrm{W}}(f)$, given by (60) and (61), respectively, as a function of the narrowband input SNR for different numbers of snapshots, K, and different numbers of sensors, M.

The theoretical plot is indicated by thick solid line.

As the number of sensors is larger or as the number of snapshots is larger, the estimators (60) and (61) are closer to the theoretical values.

Generally, the estimator (61) produces better results than (60).

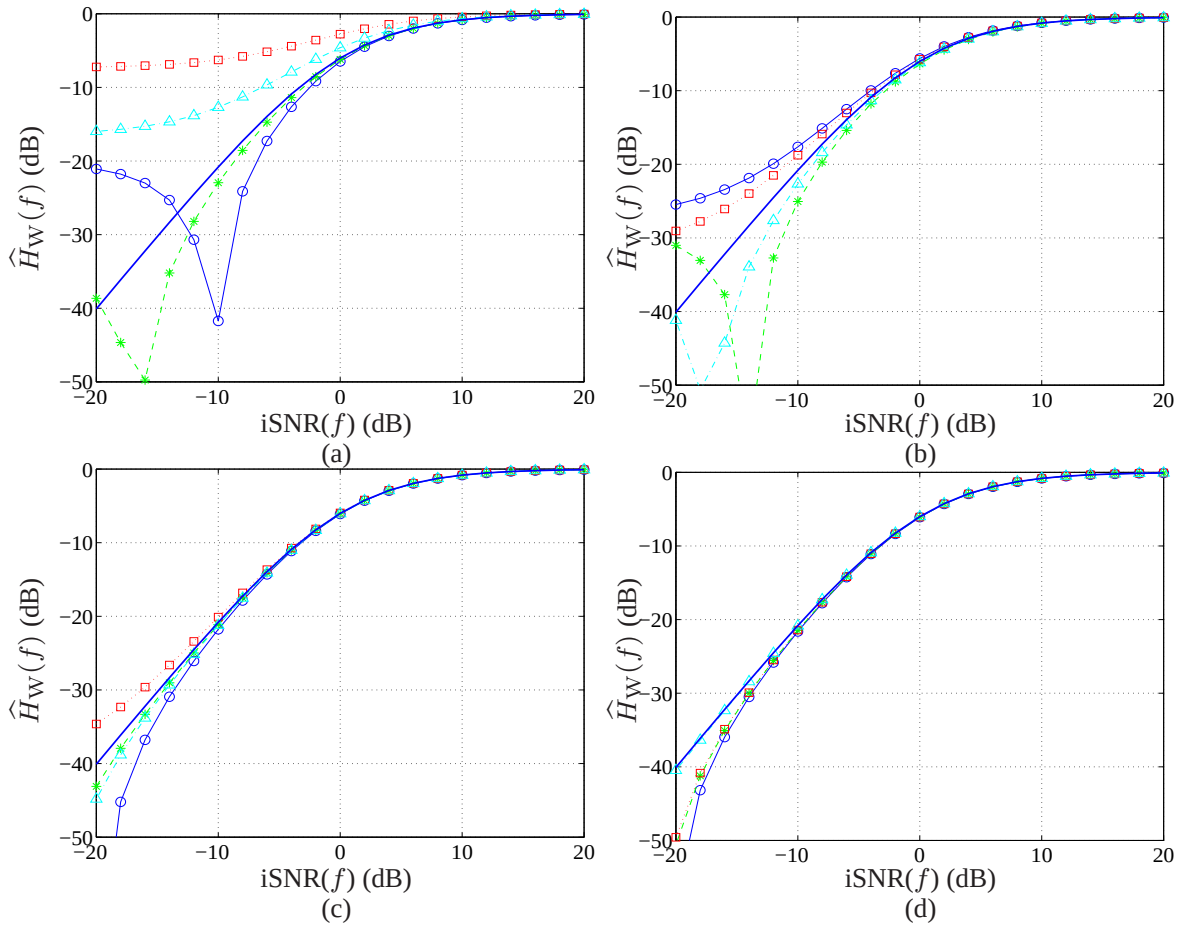


Figure 12: The estimator $\widehat{H}_{\mathrm{W}}(f)$, given by (60) for: (a) $K = 10^3$, (b) $K = 10^4$, (c) $K = 10^5$, and (d) $K = 10^6$. Each figure shows the theoretical plot (thick solid line), and estimates for $M = 2$ (solid line with circles), $M = 5$ (dashed line with asterisks), $M = 10$ (dotted line with squares), and $M = 20$ (dash-dot line with triangles).

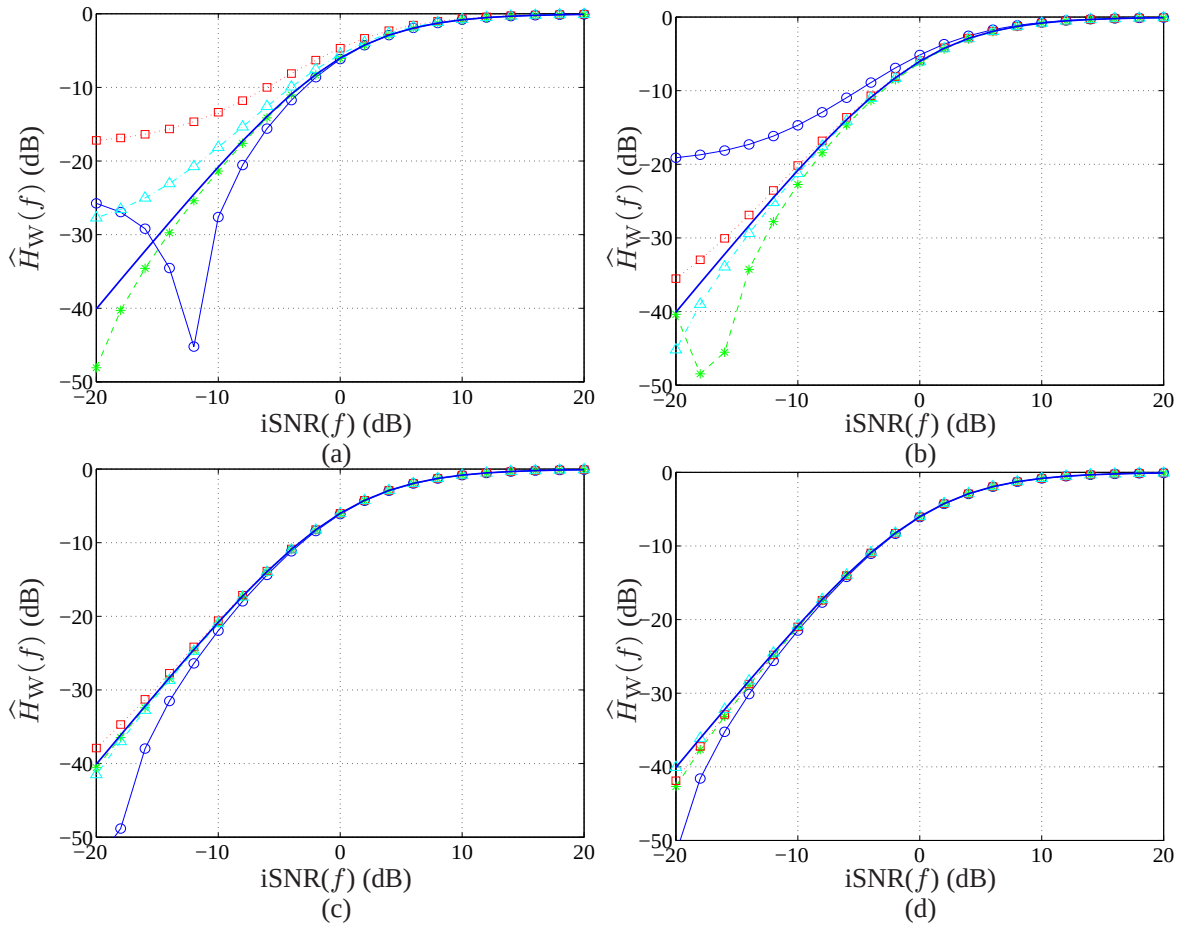


Figure 13: The estimator $\widehat{H}_{\mathrm{W}}(f)$, given by (61) for: (a) $K = 10^3$, (b) $K = 10^4$, (c) $K = 10^5$, and (d) $K = 10^6$. Each figure shows the theoretical plot (thick solid line), and estimates for $M = 2$ (solid line with circles), $M = 5$ (dashed line with asterisks), $M = 10$ (dotted line with squares), and $M = 20$ (dash-dot line with triangles).

DOA Estimation

In practice, the direction-of-arrival (DOA) of the desired signal, i.e., θ_{d}, may not always be known.

Therefore, it is of great interest to be able to estimate this angle.

Obviously, the literature on this subject is extremely rich [11] and many different approaches can be derived depending on several factors.

In this section, we propose a method that naturally flows from the perspective developed above.

Let us assume that an estimate of the pseudo-coherence matrix of the observations, $\mathbf{\Gamma_y}(f)$, is

$$\begin{aligned}\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f) &= \frac{\widehat{\mathbf{\Phi}}_{\mathbf{y}}(f)}{\widehat{\phi}_{Y_1}(f)} \\ &= \frac{\mathrm{iSNR}(f)}{1+\mathrm{iSNR}(f)}\mathbf{d}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{d}^H\left(f,\cos\theta_{\mathrm{d}}\right)+\frac{1}{1+\mathrm{iSNR}(f)}\mathbf{\Gamma}_{0,\pi}(f),\end{aligned} \tag{66}$$

where we explicitly assume that $\mathbf{\Gamma}_{0,\pi}(f)$ is an estimate of $\mathbf{\Gamma_v}(f)$.

The pseudo-coherence matrix corresponding to a source signal coming from the direction θ may be written as

$$\mathbf{\Gamma_x}\left(f,\cos\theta\right)=\mathbf{d}\left(f,\cos\theta\right)\mathbf{d}^H\left(f,\cos\theta\right), \tag{67}$$

which is a rank-1 matrix.

Using the joint diagonalization technique [4], the two matrices $\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f)$ and $\mathbf{\Gamma}_{0,\pi}(f)$ can be decomposed as

$$\mathbf{T}^H(f)\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f)\mathbf{T}(f) = \mathbf{\Lambda}(f), \tag{68}$$

$$\mathbf{T}^H(f)\mathbf{\Gamma}_{0,\pi}(f)\mathbf{T}(f) = \mathbf{I}_M, \tag{69}$$

where

$$\mathbf{T}(f) = \left[\begin{array}{cccc} \mathbf{t}_1(f) & \mathbf{t}_2(f) & \cdots & \mathbf{t}_M(f) \end{array} \right] \tag{70}$$

is a full-rank square matrix and

$$\mathbf{\Lambda}(f) = \mathrm{diag}\left[\lambda_1(f), \lambda_2(f), \ldots, \lambda_M(f)\right] \tag{71}$$

is a diagonal matrix with $\lambda_1(f) \geq \lambda_2(f) \geq \cdots \geq \lambda_M(f) > 0$.

For $\theta = \theta_{\mathrm{d}}$, we have

$$\mathbf{T}^H(f)\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f)\mathbf{T}(f) = \frac{\mathrm{iSNR}(f)}{1+\mathrm{iSNR}(f)}\mathbf{T}^H(f)\mathbf{\Gamma}_{\mathbf{x}}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{T}(f) + \frac{1}{1+\mathrm{iSNR}(f)}\mathbf{I}_M, \tag{72}$$

where $\mathbf{T}^H(f)\mathbf{\Gamma}_{\mathbf{x}}\left(f,\cos\theta_{\mathrm{d}}\right)\mathbf{T}(f)$ is a diagonal matrix whose first diagonal element is

$$\lambda_{\mathbf{x},1}(f) = \frac{\left[1+\mathrm{iSNR}(f)\right]\lambda_1(f)-1}{\mathrm{iSNR}(f)} > 0 \tag{73}$$

and the other diagonal elements are zero.

However, for $\theta \neq \theta_{\mathrm{d}}$, the matrix $\mathbf{T}^H(f)\mathbf{\Gamma}_{\mathbf{x}}(f, \cos\theta)\mathbf{T}(f)$ is no longer diagonal but its rank is still equal to 1.

Consequently, we can take advantage of this property to estimate the DOA.

Indeed, it is easy to observe that

$$\left|\mathbf{t}_i^H(f)\mathbf{d}(f, \cos\theta_{\mathrm{d}})\right|^2 = 0,\ i = 2, 3, \ldots, M \tag{74}$$

but

$$\left|\mathbf{t}_i^H(f)\mathbf{d}(f, \cos\theta)\right|^2 > 0,\ i = 2, 3, \ldots, M,\ \theta \neq \theta_{\mathrm{d}}. \tag{75}$$

As a result, the previous equations may be combined and used as a good criterion for the estimation of θ_{d}, i.e.,

$$\widehat{\theta}_{\mathrm{d}} = \arg\min_{\theta} \sum_{i=2}^{M} \left| \mathbf{t}_i^H(f)\mathbf{d}\left(f, \cos\theta\right) \right|^2 . \tag{76}$$

The previous expression corresponds to a narrowband estimation of the desired angle.

A more reliable estimator can be obtained by integrating the criterion over a range of frequencies, i.e.,

$$\widehat{\theta}_{\mathrm{d}} = \arg\min_{\theta} \int_{f_1}^{f_2} \sum_{i=2}^{M} \left| \mathbf{t}_i^H(f)\mathbf{d}\left(f, \cos\theta\right) \right|^2 df, \tag{77}$$

which corresponds to a broadband estimation of θ_{d}.

This approach is a generalization of the MUSIC (multiple signal classification) algorithm [12], [13], which was originally developed for spatially white noise, to the spherically isotropic noise field.

Obviously, this approach works for more than one desired angle, but less than M desired angles.

A byproduct of this method is that the input SNR can be easily estimated. Indeed, from (72), we deduce that

$$\mathbf{t}_i^H(f)\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f)\mathbf{t}_i(f) = \frac{1}{1+\mathrm{iSNR}(f)},\ i = 2,3,\ldots,M. \tag{78}$$

As a result, an estimate of the input SNR is

$$\widehat{\mathrm{iSNR}}(f) = \frac{1}{M-1}\sum_{i=2}^{M}\frac{1}{\mathbf{t}_i^H(f)\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f)\mathbf{t}_i(f)} - 1. \tag{79}$$

Example 6

Returning to Example 5, we set the narrowband input SNR to $\mathrm{iSNR}(f) = -10$ dB, and compute $\mathbf{\Gamma}_{\mathbf{y}}(f)$ using (63).

We obtain $\mathbf{T}(f)$ using (68) and (69), and calculate the following function:

$$R(f,\theta) = \sum_{i=2}^{M} \left| \mathbf{t}_i^H(f)\mathbf{d}\left(f, \cos\theta\right) \right|^2 . \tag{80}$$

According to (76), the minimum of $R(f,\theta)$ is obtained for $\theta = \widehat{\theta}_{\mathrm{d}}$.

Figure 14 shows plots of $R(f, \theta)$ as a function of θ for several values of α and different numbers of sensors, M.

The value of θ_{d} is better estimated for smaller values of α and larger number of sensors.

As the value of α is smaller, $\widehat{\theta}_{\mathrm{d}}$ can be obtained using a larger number of sensors with a better accuracy than that obtained with less sensors.

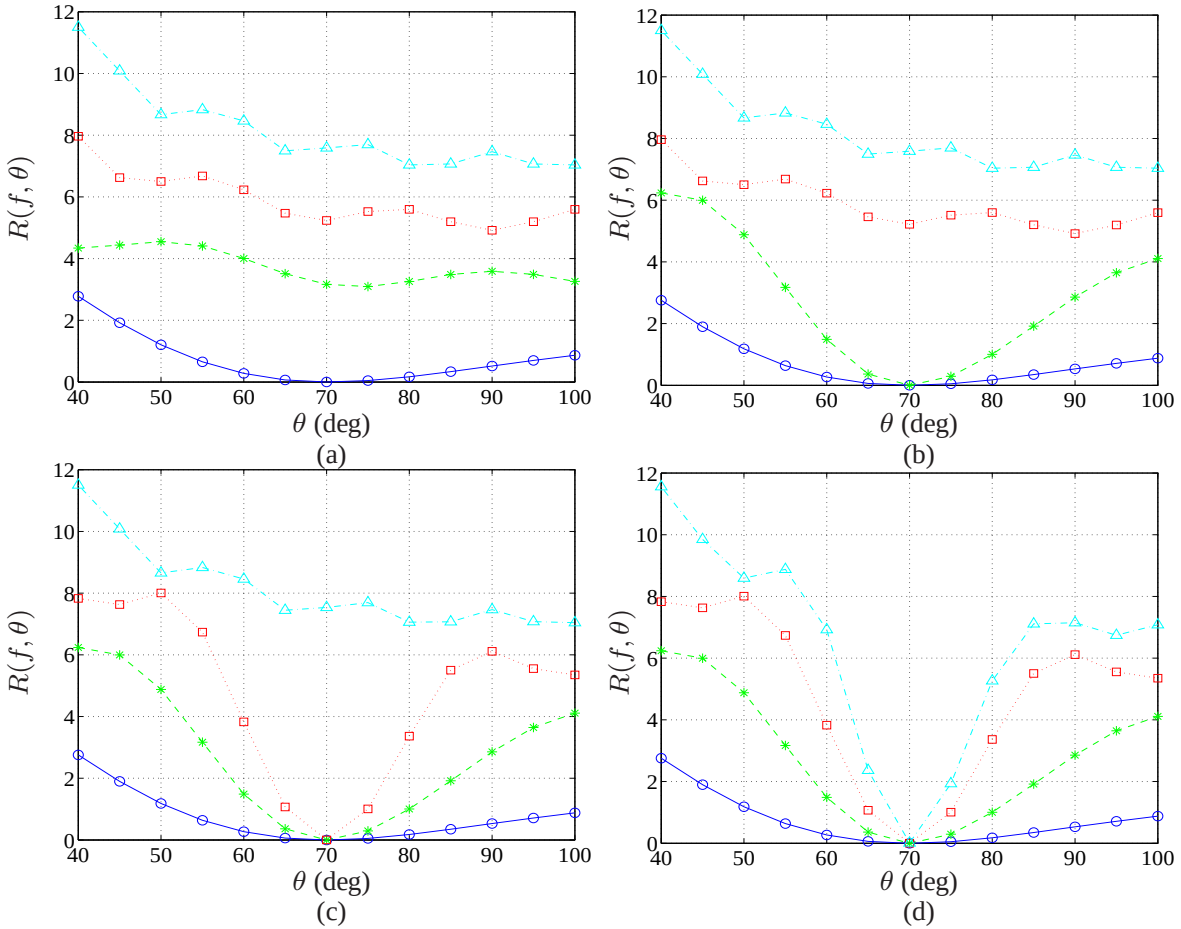


Figure 14: The function $R(f,\theta)$, given by (80) for several values of α and M: (a) $\alpha = 10^{-5}$, (b) $\alpha = 10^{-10}$, (c) $\alpha = 10^{-15}$, and (d) $\alpha = 10^{-20}$. Each figure shows plots for $M = 3$ (solid line with circles), $M = 6$ (dashed line with asterisks), $M = 9$ (dotted line with squares), and $M = 12$ (dash-dot line with triangles).

Now we set α to 10^{-20}, and estimate the pseudo-coherence matrix of the observations using (64) and (65) with K random snapshots.

Figure 15 shows plots of $R(f, \theta)$ as a function of θ for different numbers of snapshots, K, and different numbers of sensors, M.

For a small number of snapshots, a good estimate of θ_{d} requires a larger number of sensors.

As the number of snapshots is larger, we can estimate θ_{d} using a smaller number of sensors, but still a better estimate is obtained using a larger number of sensors.

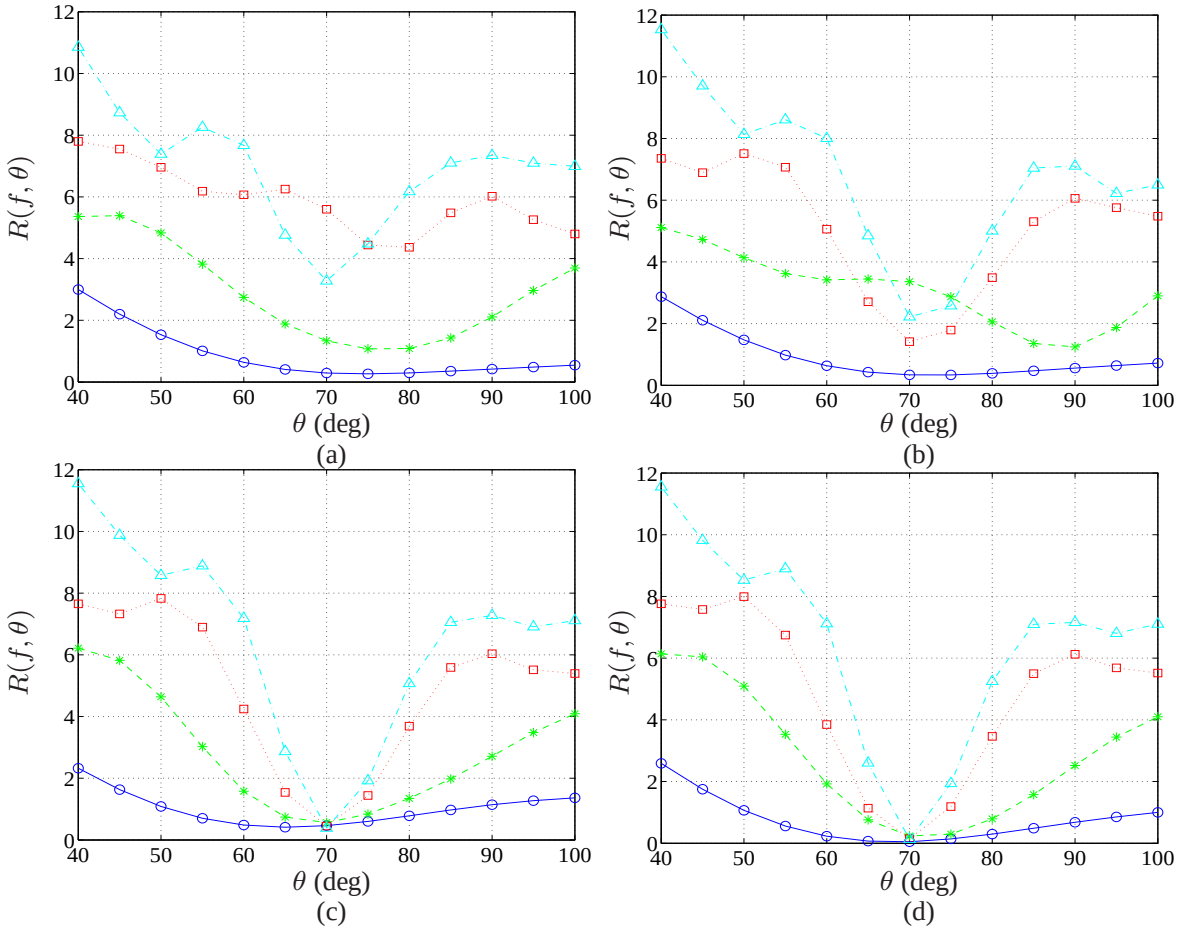


Figure 15: The function $R(f,\theta)$, given by (80) for different K and M values: (a) $K = 30$, (b) $K = 50$, (c) $K = 300$, and (d) $K = 1000$. Each figure shows plots for $M = 3$ (solid line with circles), $M = 6$ (dashed line with asterisks), $M = 9$ (dotted line with squares), and $M = 12$ (dash-dot line with triangles).

[1] J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing*. Berlin, Germany: Springer-Verlag, 2008.

[2] J. Capon, "High resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, pp. 1408–1418, Aug. 1969.

[3] R. T. Lacoss, "Data adaptive spectral analysis methods," *Geophysics*, vol. 36, pp. 661–675, Aug. 1971.

[4] J. N. Franklin, *Matrix Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1968.

[5] A. Booker and C. Y. Ong, "Multiple constraint adaptive filtering," *Geophysics*, vol. 36, pp. 498–509, June 1971.

[6] O. Frost, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, pp. 926–935, Jan. 1972.

[7] R. Zelinski, "A microphone array with adaptive post-filtering for noise reduction in reverberant rooms," in *Proc. IEEE ICASSP*, vol. 5, 1988, pp. 2578–2581.

[8] J. Meyer and K. Uwe Simmer, "Multi-channel speech enhancement in a car environment using Wiener filtering and spectral subtraction," in *Proc. IEEE ICASSP*, vol. 2, 1997, pp. 1167–1170.

[9] I. A. McCowan and H. Bourlard, “Microphone array post-filter based on noise field coherence,” *IEEE Trans. Speech, Audio Process.*, vol. 11, pp. 709–716, Nov. 2003.

[10] S. Lefkimmiatis and P. Maragos, “A generalized estimation approach for linear and nonlinear microphone array post-filters,” *Speech Communication*, vol. 49, pp. 657–666, 2007.

[11] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. New York, NY: John Wiley & Sons, Inc., 2002.

[12] G. Bienvenu and L. Kopp, “Adaptivity to background noise spatial coherence for high resolution passive methods,” in *Proc. IEEE ICASSP*, 1980, pp. 307–310.

[13] R. O. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propag.*, vol. AP-34, pp. 276–280, Mar. 1986.