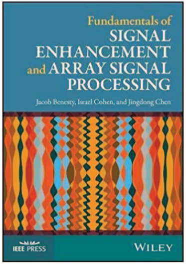
# Multichannel Signal Enhancement in the Time Domain

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



# Outline

# Introduction

We study the signal enhancement problem in the time domain with multiple sensors.

Compared to the single-channel case, better filters can be derived in terms of reduction of the additive noise and distortion of the desired signal.

Specifically, we have much more flexibility to compromise between noise reduction and desired-signal distortion thanks to the space-time processing.

# Signal Model and Problem Formulation

We consider the conventional signal model in which an array of $M$ sensors with an arbitrary geometry captures a convolved desired source signal in some noise field.

The received signals, at the discrete-time index $t$, are expressed as

$$\begin{aligned} y_m(t) &= g_m(t) * x(t) + v_m(t) \\ &= x_m(t) + v_m(t),\ m = 1, 2, \ldots, M, \end{aligned} \tag{1}$$

where $g_m(t)$ is the acoustic impulse response from the unknown desired source, $x(t)$, location to the $m$th sensor, $*$ stands for linear convolution, and $v_m(t)$ is the additive noise at sensor $m$.

We assume that the signals $x_m(t) = g_m(t) * x(t)$ and $v_m(t)$ are uncorrelated, zero mean, stationary, real, and broadband.

By definition, the convolved signals, $x_m(t),\ m = 1, 2, \ldots, M$, are coherent across the array while the noise terms, $v_m(t),\ m = 1, 2, \ldots, M$, are typically only partially coherent across the array.

By processing the data by blocks of $L$ successive time samples, the signal model can be put into a vector form as

$$\mathbf{y}_m(t) = \mathbf{x}_m(t) + \mathbf{v}_m(t),\ m = 1, 2, \ldots, M, \tag{2}$$

where

$$\mathbf{y}_m(t) = \begin{bmatrix} y_m(t) & y_m(t-1) & \cdots & y_m(t-L+1) \end{bmatrix}^T \tag{3}$$

is a vector of length $L$, and $\mathbf{x}_m(t)$ and $\mathbf{v}_m(t)$ are defined similarly to $\mathbf{y}_m(t)$.

It is more convenient to concatenate the $M$ vectors $\{\mathbf{y}_m(t)\}$ together

$$\begin{aligned} \underline{\mathbf{y}}(t) &= \left[ \begin{array}{cccc} \mathbf{y}_1^T(t) & \mathbf{y}_2^T(t) & \cdots & \mathbf{y}_M^T(t) \end{array} \right]^T \\ &= \underline{\mathbf{x}}(t) + \underline{\mathbf{v}}(t), \end{aligned} \quad (4)$$

where the vectors $\underline{\mathbf{x}}(t)$ and $\underline{\mathbf{v}}(t)$ of length $ML$ are defined in a similar way to $\underline{\mathbf{y}}(t)$.

Since $x_m(t)$ and $v_m(t)$ are uncorrelated by assumption, the correlation matrix (of size $ML \times ML$) of the observations is

$$\begin{aligned} \mathbf{R}_{\underline{\mathbf{y}}} &= E\left[\underline{\mathbf{y}}(t)\underline{\mathbf{y}}^T(t)\right] \\ &= \mathbf{R}_{\underline{\mathbf{x}}} + \mathbf{R}_{\underline{\mathbf{v}}}, \end{aligned} \quad (5)$$

where $\mathbf{R}_{\underline{\mathbf{x}}} = E\left[\underline{\mathbf{x}}(t)\underline{\mathbf{x}}^T(t)\right]$ and $\mathbf{R}_{\underline{\mathbf{v}}} = E\left[\underline{\mathbf{v}}(t)\underline{\mathbf{v}}^T(t)\right]$ are the correlation matrices of $\underline{\mathbf{x}}(t)$ and $\underline{\mathbf{v}}(t)$, respectively.

In the rest, unless stated otherwise, it is assumed that $\mathrm{rank}\left(\mathbf{R}_{\underline{\mathbf{x}}}\right) = P < ML$ while $\mathrm{rank}\left(\mathbf{R}_{\underline{\mathbf{v}}}\right) = ML$. In other words, $\mathbf{R}_{\underline{\mathbf{x}}}$ is rank deficient while $\mathbf{R}_{\underline{\mathbf{v}}}$ is full rank.

We consider the first sensor as the reference, so everything will be defined with respect to this sensor.

In this case, the desired signal is the whole vector $\mathbf{x}_1(t)$ of length $L$.

Our problem then may be stated as follows: given $M$ mixtures of two uncorrelated signals $x_m(t)$ and $v_m(t)$, our aim is to preserve $\mathbf{x}_1(t)$ while minimizing the contribution of the noise signal vector, $\underline{\mathbf{v}}(t)$, at the array output.

# Joint Diagonalization

Since $\mathbf{R}_{\underline{\mathbf{v}}}$ has full rank, the two symmetric matrices $\mathbf{R}_{\underline{\mathbf{x}}}$ and $\mathbf{R}_{\underline{\mathbf{v}}}$ can be jointly diagonalized as follows [4]:

$$\underline{\mathbf{T}}^T \mathbf{R}_{\underline{\mathbf{x}}} \underline{\mathbf{T}} = \underline{\mathbf{\Lambda}}, \tag{6}$$

$$\underline{\mathbf{T}}^T \mathbf{R}_{\underline{\mathbf{v}}} \underline{\mathbf{T}} = \mathbf{I}_{ML}, \tag{7}$$

where $\underline{\mathbf{T}}$ is a full-rank square matrix (of size $ML \times ML$), $\underline{\mathbf{\Lambda}}$ is a diagonal matrix whose main elements are real and nonnegative, and $\mathbf{I}_{ML}$ is the $ML \times ML$ identity matrix.

Furthermore, $\underline{\mathbf{\Lambda}}$ and $\underline{\mathbf{T}}$ are the eigenvalue and eigenvector matrices, respectively, of $\mathbf{R}_{\underline{\mathbf{v}}}^{-1}\mathbf{R}_{\underline{\mathbf{x}}}$, i.e.,

$$\mathbf{R}_{\underline{\mathbf{v}}}^{-1}\mathbf{R}_{\underline{\mathbf{x}}}\underline{\mathbf{T}} = \underline{\mathbf{T}}\,\underline{\mathbf{\Lambda}}. \tag{8}$$

The eigenvalues of $\mathbf{R}_{\underline{\mathbf{v}}}^{-1}\mathbf{R}_{\underline{\mathbf{x}}}$ can be ordered as $\underline{\lambda}_1 \geq \underline{\lambda}_2 \geq \cdots \geq \underline{\lambda}_P > \underline{\lambda}_{P+1} = \cdots = \underline{\lambda}_{ML} = 0$. We denote by $\underline{\mathbf{t}}_1, \underline{\mathbf{t}}_2, \ldots, \underline{\mathbf{t}}_{ML}$, the corresponding eigenvectors.

Therefore, the noisy signal correlation matrix can also be diagonalized as

$$\underline{\mathbf{T}}^T \mathbf{R}_{\underline{\mathbf{y}}} \underline{\mathbf{T}} = \underline{\mathbf{\Lambda}} + \mathbf{I}_{ML}. \tag{9}$$

It can be verified from (6) and (7) that

$$\underline{\mathbf{t}}_i^T \underline{\mathbf{x}}(t) = 0, \; i = P+1, P+2, \ldots, ML \tag{10}$$

and

$$\mathbf{R}_{\underline{\mathbf{v}}}^{-1} = \sum_{i=1}^{ML} \underline{\mathbf{t}}_i \underline{\mathbf{t}}_i^T. \tag{11}$$

# Linear Filtering

Since we want to estimate the desired-signal vector, $\mathbf{x}_1(t)$, of length $L$, from the observation signal vector, $\underline{\mathbf{y}}(t)$, of length $ML$, a real-valued rectangular filtering matrix, $\underline{\mathbf{H}}$, of size $L \times ML$ should be used as follows:

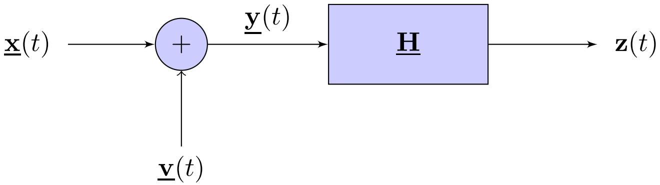



Figure 1: Block diagram of multichannel linear filtering in the time domain.

where $\mathbf{z}(t)$, a vector of length $L$, is the estimate of $\mathbf{x}_1(t)$,

$$\begin{aligned} \mathbf{z}(t) &= \underline{\mathbf{H}}\ \underline{\mathbf{y}}(t) \\ &= \mathbf{x}_{\mathrm{fd}}(t) + \mathbf{v}_{\mathrm{rn}}(t), \end{aligned} \tag{12}$$

the filtered desired signal is given by

$$\mathbf{x}_{\mathrm{fd}}(t) = \underline{\mathbf{H}}\ \underline{\mathbf{x}}(t) \tag{13}$$

and the residual noise is given by

$$\mathbf{v}_{\mathrm{rn}}(t) = \underline{\mathbf{H}}\ \underline{\mathbf{v}}(t)\,. \tag{14}$$

We can always express $\underline{\mathbf{H}}$ as

$$\underline{\mathbf{H}} = \underline{\mathbf{A}}\,\underline{\mathbf{T}}^T, \tag{15}$$

where $\underline{\mathbf{A}}$ is the transformed rectangular filtering matrix also of size $L \times ML$.

Instead of manipulating $\underline{\mathbf{H}}$ directly, we can, equivalently, manipulate $\underline{\mathbf{A}}$, since $\underline{\mathbf{T}}$ (or $\underline{\mathbf{T}}^T$) is a full-rank square matrix.

So when $\underline{\mathbf{A}}$ is estimated, we can easily find $\underline{\mathbf{H}}$ from (15).

We can write (12) as

$$\mathbf{z}(t) = \underline{\mathbf{A}}\,\underline{\mathbf{T}}^T\underline{\mathbf{y}}(t). \tag{16}$$

We deduce that the correlation matrix of $\mathbf{z}(t)$ is

$$\begin{aligned} \mathbf{R}_{\mathbf{z}} &= E\left[\mathbf{z}(t)\mathbf{z}^T(t)\right] \\ &= \underline{\mathbf{A}}\left(\underline{\mathbf{\Lambda}} + \mathbf{I}_{ML}\right)\underline{\mathbf{A}}^T \\ &= \mathbf{R}_{\mathbf{x}_{\mathrm{fd}}} + \mathbf{R}_{\mathbf{v}_{\mathrm{rn}}}, \end{aligned} \tag{17}$$

where

$$\mathbf{R}_{\mathbf{x}_{\mathrm{fd}}} = \underline{\mathbf{A}}\ \underline{\mathbf{\Lambda}}\ \underline{\mathbf{A}}^T \tag{18}$$

is the correlation matrix of the filtered desired signal and

$$\mathbf{R}_{\mathbf{v}_{\mathrm{rn}}} = \underline{\mathbf{A}}\ \underline{\mathbf{A}}^T \tag{19}$$

is the correlation matrix of the residual noise.

# Performance Measures
## Signal-to-Noise Ratio

The input SNR is defined as

$$\mathrm{iSNR} = \frac{\mathrm{tr}\left(\mathbf{R}_{\mathbf{x}_1}\right)}{\mathrm{tr}\left(\mathbf{R}_{\mathbf{v}_1}\right)}, \tag{20}$$

where $\mathbf{R}_{\mathbf{x}_1} = E\left[\mathbf{x}_1(t)\mathbf{x}_1^T(t)\right]$ and $\mathbf{R}_{\mathbf{v}_1} = E\left[\mathbf{v}_1(t)\mathbf{v}_1^T(t)\right]$ are the correlation matrices of $\mathbf{x}_1(t)$ and $\mathbf{v}_1(t)$, respectively.

The output SNR is given by

$$\begin{aligned} \mathrm{oSNR}\left(\underline{\mathbf{H}}\right) &= \frac{\mathrm{tr}\left(\mathbf{R}_{\mathbf{x}_{\mathrm{fd}}}\right)}{\mathrm{tr}\left(\mathbf{R}_{\mathbf{v}_{\mathrm{rn}}}\right)} \\ &= \frac{\mathrm{tr}\left(\underline{\mathbf{A}}\ \underline{\mathbf{\Lambda}}\ \underline{\mathbf{A}}^T\right)}{\mathrm{tr}\left(\underline{\mathbf{A}}\ \underline{\mathbf{A}}^T\right)} \\ &= \mathrm{oSNR}\left(\underline{\mathbf{A}}\right). \end{aligned} \tag{21}$$

It is clear that we always have

$$\text{oSNR}\left(\underline{\mathbf{A}}\right) \leq \underline{\lambda}_1, \tag{22}$$

showing how the output SNR is always upper bounded as long as $\mathbf{R}_{\underline{\mathbf{v}}}$ has full rank.

# Noise Reduction Factor

The noise reduction factor is given by

$$\begin{aligned} \xi_{\mathrm{n}}\left(\underline{\mathbf{H}}\right) &= \frac{\mathrm{tr}\left(\mathbf{R}_{\mathbf{v}_1}\right)}{\mathrm{tr}\left(\underline{\mathbf{A}}\ \underline{\mathbf{A}}^T\right)} \\ &= \xi_{\mathrm{n}}\left(\underline{\mathbf{A}}\right). \end{aligned} \tag{23}$$

For optimal filtering matrices, we should have $\xi_{\mathrm{n}}\left(\underline{\mathbf{A}}\right) \geq 1$.

# Desired-Signal Reduction Factor

Since the desired signal may be distorted by the filtering matrix, we define the desired-signal reduction factor as

$$\xi_{\mathrm{d}}\left(\underline{\mathbf{H}}\right) = \frac{\mathrm{tr}\left(\mathbf{R}_{\mathbf{x}_1}\right)}{\mathrm{tr}\left(\underline{\mathbf{A}}\ \underline{\mathbf{\Lambda}}\ \underline{\mathbf{A}}^T\right)} = \xi_{\mathrm{d}}\left(\underline{\mathbf{A}}\right). \tag{24}$$

For optimal filtering matrices, we generally have $\xi_{\mathrm{d}}\left(\underline{\mathbf{A}}\right) \geq 1$. The closer the value of $\xi_{\mathrm{d}}\left(\underline{\mathbf{A}}\right)$ is to $1$, the less distorted is the desired signal.

Obviously, we have the fundamental relationship:

$$\frac{\mathrm{oSNR}\left(\underline{\mathbf{A}}\right)}{\mathrm{iSNR}} = \frac{\xi_{\mathrm{n}}\left(\underline{\mathbf{A}}\right)}{\xi_{\mathrm{d}}\left(\underline{\mathbf{A}}\right)}, \tag{25}$$

which, basically, states that nothing comes for free.

# Desired-Signal Distortion Index

We can also evaluate distortion via the desired-signal distortion index:

$$
\begin{aligned}
\upsilon_{\mathrm{d}}\left(\underline{\mathbf{H}}\right) &= \frac{E\left\{\left[\mathbf{x}_{\mathrm{fd}}(t)-\mathbf{x}_1(t)\right]^T\left[\mathbf{x}_{\mathrm{fd}}(t)-\mathbf{x}_1(t)\right]\right\}}{\mathrm{tr}\left(\mathbf{R}_{\mathbf{x}_1}\right)} \\
&= \upsilon_{\mathrm{d}}\left(\underline{\mathbf{A}}\right).
\end{aligned} \tag{26}
$$

For optimal filtering matrices, we should have $\upsilon_{\mathrm{d}}\left(\underline{\mathbf{A}}\right) \leq 1$.

# Mean-Squared Error

We define the error signal vector between the estimated and desired signals as

$$\begin{aligned}\mathbf{e}(t) &= \mathbf{z}(t) - \mathbf{x}_1(t) \\ &= \underline{\mathbf{A}}\ \underline{\mathbf{T}}^T \underline{\mathbf{y}}(t) - \mathbf{x}_1(t) \\ &= \mathbf{e}_\mathrm{d}(t) + \mathbf{e}_\mathrm{n}(t),\end{aligned} \tag{27}$$

where

$$\mathbf{e}_\mathrm{d}(t) = \mathbf{x}_\mathrm{fd}(t) - \mathbf{x}_1(t) = \left(\underline{\mathbf{A}}\ \underline{\mathbf{T}}^T - \underline{\mathbf{I}}_\mathrm{i}\right) \underline{\mathbf{x}}(t) \tag{28}$$

is the desired-signal distortion due to the filtering matrix with

$$\underline{\mathbf{I}}_\mathrm{i} = \left[\begin{array}{cc} \mathbf{I}_L & \mathbf{0}_{L\times(M-1)L} \end{array}\right] \tag{29}$$

being the identity filtering matrix of size $L \times ML$, $\underline{\mathbf{I}}_\mathrm{i}\underline{\mathbf{x}}(t) = \mathbf{x}_1(t)$, and

$$\mathbf{e}_\mathrm{n}(t) = \mathbf{v}_\mathrm{rn}(t) = \underline{\mathbf{A}}\ \underline{\mathbf{T}}^T \underline{\mathbf{v}}(t) \tag{30}$$

is the residual noise.

We deduce that the MSE criterion is

$$
\begin{aligned}
J\left(\underline{\mathbf{A}}\right) &= \mathrm{tr}\left\{E\left[\mathbf{e}(t)\mathbf{e}^T(t)\right]\right\} \\
&= \mathrm{tr}\left[\mathbf{R}_{\mathbf{x}_1} - 2\underline{\mathbf{A}}\ \underline{\mathbf{T}}^T\mathbf{R}_{\underline{\mathbf{x}}}\underline{\mathbf{I}}_{\mathrm{i}}^T + \underline{\mathbf{A}}\left(\underline{\mathbf{\Lambda}} + \mathbf{I}_{ML}\right)\underline{\mathbf{A}}^T\right] \\
&= J_{\mathrm{d}}\left(\underline{\mathbf{A}}\right) + J_{\mathrm{n}}\left(\underline{\mathbf{A}}\right),
\end{aligned} \tag{31}
$$

where

$$
\begin{aligned}
J_{\mathrm{d}}\left(\underline{\mathbf{A}}\right) &= \mathrm{tr}\left\{E\left[\mathbf{e}_{\mathrm{d}}(t)\mathbf{e}_{\mathrm{d}}^T(t)\right]\right\} \\
&= \mathrm{tr}\left(\mathbf{R}_{\mathbf{x}_1} - 2\underline{\mathbf{A}}\ \underline{\mathbf{T}}^T\mathbf{R}_{\underline{\mathbf{x}}}\underline{\mathbf{I}}_{\mathrm{i}}^T + \underline{\mathbf{A}}\ \underline{\mathbf{\Lambda}}\ \underline{\mathbf{A}}^T\right) \\
&= \mathrm{tr}\left(\mathbf{R}_{\mathbf{x}_1}\right)\upsilon_{\mathrm{d}}\left(\underline{\mathbf{A}}\right)
\end{aligned} \tag{32}
$$

and

$$J_{\mathrm{n}}\left(\underline{\mathbf{A}}\right) = \mathrm{tr}\left\{E\left[\mathbf{e}_{\mathrm{n}}(t)\mathbf{e}_{\mathrm{n}}^T(t)\right]\right\} = \mathrm{tr}\left(\underline{\mathbf{A}}\ \underline{\mathbf{A}}^T\right) = \frac{\mathrm{tr}\left(\mathbf{R}_{\mathbf{v}_1}\right)}{\xi_{\mathrm{n}}\left(\underline{\mathbf{A}}\right)}. \quad (33)$$

As a result, we have

$$\frac{J_{\mathrm{d}}\left(\underline{\mathbf{A}}\right)}{J_{\mathrm{n}}\left(\underline{\mathbf{A}}\right)} = \mathrm{iSNR} \times \xi_{\mathrm{n}}\left(\underline{\mathbf{A}}\right) \times \upsilon_{\mathrm{d}}\left(\underline{\mathbf{A}}\right) = \mathrm{oSNR}\left(\underline{\mathbf{A}}\right) \times \xi_{\mathrm{d}}\left(\underline{\mathbf{A}}\right) \times \upsilon_{\mathrm{d}}\left(\underline{\mathbf{A}}\right), \quad (34)$$

showing how the different performance measures are related to the MSEs.

# Optimal Filtering Matrices

## Wiener Filtering Matrix

The Wiener filtering matrix is derived from the minimization of the MSE criterion, $J\left(\underline{\mathbf{A}}\right)$.

From this optimization, we obtain

$$\begin{aligned} \underline{\mathbf{A}}_{\mathrm{W}} &= \underline{\mathbf{I}}_{\mathrm{i}}\mathbf{R}_{\underline{\mathbf{x}}}\underline{\mathbf{T}}\left(\underline{\mathbf{\Lambda}}+\mathbf{I}_{ML}\right)^{-1} \\ &= \underline{\mathbf{I}}_{\mathrm{i}}\underline{\mathbf{T}}^{-T}\underline{\mathbf{\Lambda}}\left(\underline{\mathbf{\Lambda}}+\mathbf{I}_{ML}\right)^{-1}. \end{aligned} \quad (35)$$

We deduce that the Wiener filtering matrix is

$$\begin{aligned} \underline{\mathbf{H}}_{\mathrm{W}} &= \underline{\mathbf{A}}_{\mathrm{W}}\underline{\mathbf{T}}^{T} \\ &= \underline{\mathbf{I}}_{\mathrm{i}}\mathbf{R}_{\underline{\mathbf{x}}}\sum_{i=1}^{ML}\frac{\underline{\mathbf{t}}_i\underline{\mathbf{t}}_i^T}{1+\underline{\lambda}_i} \\ &= \underline{\mathbf{I}}_{\mathrm{i}}\mathbf{R}_{\underline{\mathbf{v}}}\sum_{i=1}^{ML}\frac{\underline{\lambda}_i}{1+\underline{\lambda}_i}\underline{\mathbf{t}}_i\underline{\mathbf{t}}_i^T. \end{aligned} \quad (36)$$

Obviously, we can also express $\underline{\mathbf{H}}_{\mathrm{W}}$ as

$$\underline{\mathbf{H}}_{\mathrm{W}} = \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{x}}} \mathbf{R}_{\underline{\mathbf{y}}}^{-1}. \tag{37}$$

**Property**

*With the optimal Wiener filtering matrix given in (37), the output SNR is always greater than or equal to the input SNR, i.e.,* $\mathrm{oSNR}\left(\underline{\mathbf{H}}_{\mathrm{W}}\right) \geq \mathrm{iSNR}$.

# Example 1

Consider an array of $M$ sensors located on a line with a uniform spacing $d$, as shown in Fig. 2. Such an array is known as a uniform linear array (ULA).

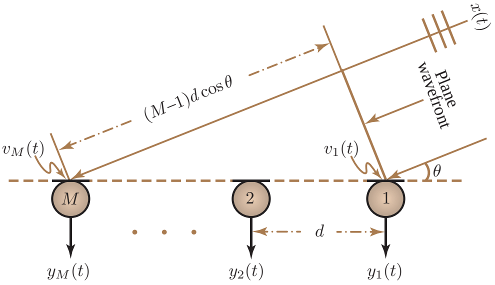



Figure 2: Illustration of a uniform linear array for signal capture in the farfield.

Suppose that a desired signal impinges on the ULA from the broadside direction ($\theta = 90°$), and that an interference impinges on the ULA from the endfire direction ($\theta = 0°$).

Assume that the desired signal is a harmonic random process:

$$x(t) = A \cos(2\pi f_0 t + \phi),$$

with fixed amplitude $A$ and frequency $f_0$, and random phase $\phi$, uniformly distributed on the interval from $0$ to $2\pi$.

Assume that the interference $u(t)$ is white Gaussian noise, i.e., $u(t) \sim \mathcal{N}\left(0, \sigma_u^2\right)$, 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 desired signal needs to be recovered from the noisy received signals, $y_m(t) = x_m(t) + v_m(t),\ m = 1, \ldots, M$, where $v_m(t) = u_m(t) + w_m(t),\ m = 1, \ldots, M$ are the interference-plus-noise signals.

Since the desired source is at the broadside direction and the interference source is at the endfire direction, we have for $i = 2, \ldots, M$:

$$x_i(t) = x_1(t), \tag{38}$$

$$u_i(t) = u_1\left(t - \tau_i\right), \tag{39}$$

where

$$\tau_i = \frac{(i-1)d}{cT_\mathrm{s}} \tag{40}$$

is the relative time delay in samples between the $i$th sensor and the first sensor for an endfire source, $c$ is the speed of wave propagation, and $T_{\rm s}$ is the sampling interval.

Assuming that the sampling interval satisfies $T_{\rm s} = \frac{d}{c}$, then the delay $\tau_i = i - 1$ becomes an integer and, therefore, (38) and (39) can be written as

$$[\underline{\mathbf{x}}(t)]_{l+(m-1)L} = [\underline{\mathbf{x}}(t)]_l \,, \tag{41}$$

$$[\underline{\mathbf{u}}(t)]_{l+(m-1)L} = [\underline{\mathbf{u}}(t)]_{l+m-1} \,, \tag{42}$$

for $l = 1, \ldots, L$, $m = 1, \ldots, M$, and $l + m - 1 \leq L$.

Hence, the correlation matrix of $\underline{\mathbf{x}}(t)$ is

$$\mathbf{R}_{\underline{\mathbf{x}}} = \mathbf{1}_M \otimes \mathbf{R}_{\mathbf{x}_1} ,$$

where $\otimes$ is the Kronecker product, $\mathbf{1}_M$ is an $M \times M$ matrix of all ones, and the elements of the correlation matrix of $\mathbf{x}_1(t)$ are $[\mathbf{R}_{\mathbf{x}_1}]_{i,j} = \frac{1}{2}A^2 \cos[2\pi f_0(i-j)]$.

The correlation matrix of $\underline{\mathbf{v}}(t)$ is $\mathbf{R}_{\underline{\mathbf{v}}} = \mathbf{R}_{\underline{\mathbf{u}}} + \sigma_w^2 \mathbf{I}_{LM}$, where the elements of the $LM \times LM$ matrix $\mathbf{R}_{\underline{\mathbf{u}}}$ are

$$\left[\mathbf{R}_{\underline{\mathbf{u}}}\right]_{i+(m_1-1)L, j+(m_2-1)L} = \sigma_u^2 \, \delta\left(i + m_1 - j - m_2\right),$$
$$i, j = 1, \ldots, L, \; m_1, m_2 = 1, \ldots, M.$$

The input SNR is

$$\mathrm{iSNR} = 10 \log \frac{A^2/2}{\sigma_u^2 + \sigma_w^2} \quad (\mathrm{dB}).$$

The optimal filter $\underline{\mathbf{H}}_{\mathrm{W}}$ is obtained from (36).

To demonstrate the performance of the Wiener filtering matrix, we choose $A = 0.5$, $f_0 = 0.1$, $\sigma_u^2 = 0.5$, and $\sigma_w^2 = 0.01\sigma_u^2$. The input SNR is $-6.06$ dB.

Figure 3 shows the effect of the filter length, $L$, and the number of sensors, $M$, on the gain in SNR, i.e., $\mathcal{G}\left(\underline{\mathbf{H}}_{\mathrm{W}}\right) = \mathrm{oSNR}\left(\underline{\mathbf{H}}_{\mathrm{W}}\right)/\mathrm{iSNR}$, and the MMSE per sample, $J\left(\underline{\mathbf{H}}_{\mathrm{W}}\right)/L$.

As the length of the filter increases, or as the number of sensors increases, the Wiener filtering matrix better enhances the harmonic signal, in terms of higher gain in SNR and lower MMSE per sample.

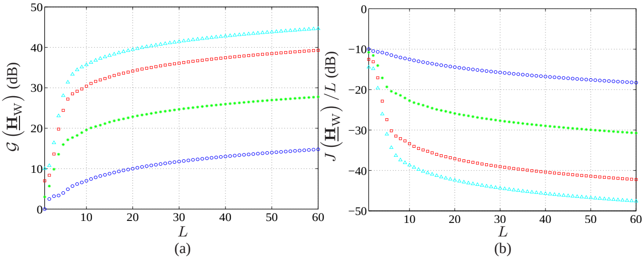



Figure 3: (a) The gain in SNR and (b) the MMSE per sample of the Wiener filtering matrix as a function of the filter length, $L$, for different numbers of sensors, $M$: $M = 1$ (circles), $M = 2$ (asterisks), $M = 5$ (squares), and $M = 10$ (triangles).

If we choose a fixed filter length, $L = 30$, and change $\sigma_u^2$ so that iSNR varies from $-5$ to $15$ dB, then Fig. 4 shows plots of the gain in SNR, the MMSE, the noise reduction factor, and the desired-signal reduction factor, as a function of the 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 MMSE and the desired-signal reduction factor decrease.

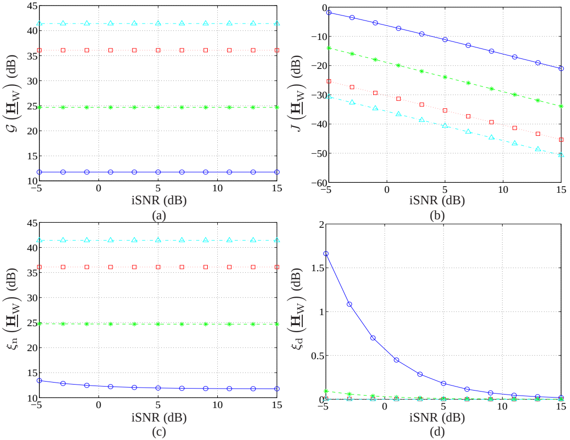



Figure 4: (a) The gain in SNR, (b) the MMSE, (c) the noise reduction factor, and (d) the desired-signal reduction factor of the Wiener filtering matrix for different numbers of sensors, $M$: $M = 1$ (solid line with circles), $M = 2$ (dashed line with asterisks), $M = 5$ (dotted line with squares), and $M = 10$ (dash-dot line with triangles).

Figure 5 shows a realization of the noise corrupted signal received at the first sensor, $y_1(t)$, and filtered signals for $\mathrm{iSNR} = -5$ dB and different numbers of sensors.

The filtered signal, $z(t)$, is obtained by taking at each $t$ the first element of $\mathbf{z}(t) = \underline{\mathbf{H}}_{\mathrm{W}}\ \underline{\mathbf{y}}(t)$.

Obviously, as the number of sensors increases, the Wiener filtering matrix better enhances the harmonic signal.

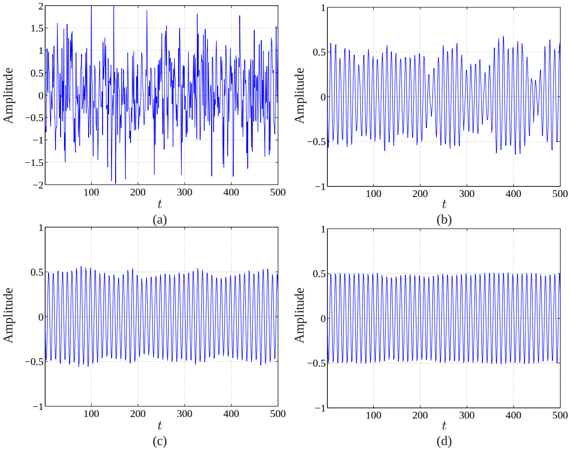



Figure 5: Example of noise corrupted and filtered sinusoidal signals for different numbers of sensors, $M$: (a) noise corrupted signal received at the first sensor, $y_1(t)$ ($\mathrm{iSNR} = -5$ dB), and filtered signals for (b) $M = 1$ [$\mathrm{oSNR}\left(\underline{\mathbf{H}}_{\mathrm{W}}\right) = 6.76$ dB], (c) $M = 2$ [$\mathrm{oSNR}\left(\underline{\mathbf{H}}_{\mathrm{W}}\right) = 19.68$ dB], and (d) $M = 5$ [$\mathrm{oSNR}\left(\underline{\mathbf{H}}_{\mathrm{W}}\right) = 31.09$ dB].

# Variable Span Wiener Filtering

From the formulation given in (36), we propose a variable span (VS) Wiener filtering matrix [5], [6]:

$$\underline{\mathbf{H}}_{\mathrm{W},Q} = \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{x}}} \sum_{q=1}^{Q} \frac{\underline{\mathbf{t}}_q \underline{\mathbf{t}}_q^T}{1 + \underline{\lambda}_q}, \tag{43}$$

where $1 \leq Q \leq ML$. We see that $\underline{\mathbf{H}}_{\mathrm{W},ML} = \underline{\mathbf{H}}_{\mathrm{W}}$ and for $Q = 1$, we obtain the maximum SNR filtering matrix with minimum MSE:

$$\underline{\mathbf{H}}_{\max,1} = \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{x}}} \frac{\underline{\mathbf{t}}_1 \underline{\mathbf{t}}_1^T}{1 + \underline{\lambda}_1}, \tag{44}$$

since

$$\mathrm{oSNR}\left(\underline{\mathbf{H}}_{\max,1}\right) = \underline{\lambda}_1. \tag{45}$$

# Example 2

Returning to Example 1, we now assume a desired signal, $x(t)$, with the autocorrelation sequence:

$$E\left[x(t)x(t')\right] = \alpha^{\left|t-t'\right|},\ -1 < \alpha < 1.$$

The desired signal needs to be recovered from the noisy observation, $\underline{\mathbf{y}}(t) = \underline{\mathbf{x}}(t) + \underline{\mathbf{v}}(t)$.

Since the desired source is at the broadside direction, the correlation matrix of $\underline{\mathbf{x}}(t)$ is

$$\mathbf{R}_{\underline{\mathbf{x}}} = \mathbf{1}_M \otimes \mathbf{R}_{\mathbf{x}_1},$$

where $\left[\mathbf{R}_{\mathbf{x}_1}\right]_{i,j} = \alpha^{|i-j|}$.

The input SNR is

$$\mathrm{iSNR} = 10 \log \frac{1}{\sigma_u^2 + \sigma_w^2} \quad (\mathrm{dB}).$$

The optimal filter $\underline{\mathbf{H}}_{\mathrm{W},Q}$ is obtained from (43).

To demonstrate the performance of the VS Wiener filtering matrix, we choose $\alpha = 0.8$, $L = 10$, and $M = 5$.

Figure 6 shows plots of the gain in SNR, $\mathcal{G}\left(\underline{\mathbf{H}}_{\mathrm{W},Q}\right)$, the MSE, $J\left(\underline{\mathbf{H}}_{\mathrm{W},Q}\right)$, the noise reduction factor, $\xi_{\mathrm{n}}\left(\underline{\mathbf{H}}_{\mathrm{W},Q}\right)$, and the desired-signal reduction factor, $\xi_{\mathrm{d}}\left(\underline{\mathbf{H}}_{\mathrm{W},Q}\right)$, as a function of the input SNR for several values of $Q$.

For a given input SNR, the higher is the value of $Q$, the lower are the MSE and the desired-signal reduction factor, but at the expense of lower gain in SNR and lower noise reduction factor.

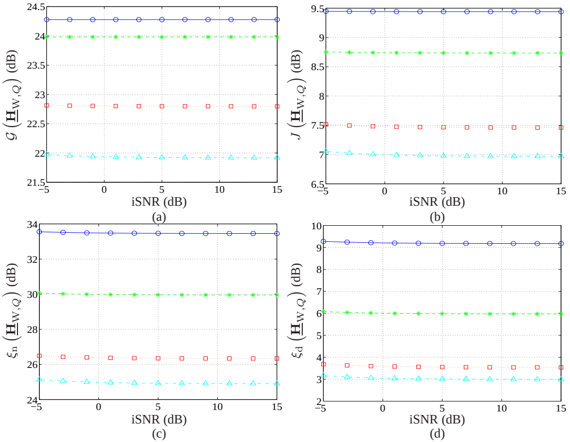



Figure 6: (a) The gain in SNR, (b) the MSE, (c) the noise reduction factor, and (d) the desired-signal reduction factor of the VS Wiener filtering matrix for several values of $Q$: $Q = 1$ (solid line with circles), $Q = 2$ (dashed line with asterisks), $Q = 5$ (dotted line with squares), and $Q = 9$ (dash-dot line with triangles).

# MVDR Filtering Matrix

We can also try to minimize the distortion-based MSE. Taking the gradient of $J_{\mathrm{d}}\left(\underline{\mathbf{A}}\right)$ with respect to $\underline{\mathbf{A}}$ and equating the result to zero, we get

$$\underline{\mathbf{A}}\ \underline{\mathbf{\Lambda}} = \underline{\mathbf{I}}_{\mathrm{i}}\mathbf{R}_{\underline{\mathbf{x}}}\underline{\mathbf{T}}. \tag{46}$$

Since $\underline{\mathbf{\Lambda}}$ is not invertible, we can take its pseudo-inverse. Then, a solution to (46) is

$$\underline{\mathbf{A}}_{\mathrm{MVDR}} = \underline{\mathbf{I}}_{\mathrm{i}}\mathbf{R}_{\underline{\mathbf{x}}}\underline{\mathbf{T}}\ \underline{\mathbf{\Lambda}}'^{-1}, \tag{47}$$

where

$$\underline{\mathbf{\Lambda}}'^{-1} = \mathrm{diag}\left(\underline{\lambda}_1^{-1}, \underline{\lambda}_2^{-1}, \ldots, \underline{\lambda}_P^{-1}, 0, \ldots, 0\right). \tag{48}$$

Therefore, the MVDR filtering matrix is

$$\begin{aligned} \underline{\mathbf{H}}_{\mathrm{MVDR}} &= \underline{\mathbf{A}}_{\mathrm{MVDR}} \underline{\mathbf{T}}^T \\ &= \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{x}}} \sum_{p=1}^{P} \frac{\underline{\mathbf{t}}_p \underline{\mathbf{t}}_p^T}{\underline{\lambda}_p} \\ &= \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{v}}} \sum_{p=1}^{P} \underline{\mathbf{t}}_p \underline{\mathbf{t}}_p^T . \end{aligned} \tag{49}$$

Now, let us show that (49) is the MVDR filtering matrix:

With $\underline{\mathbf{H}}_{\mathrm{MVDR}}$, the filtered desired-signal vector is

$$\begin{aligned} \mathbf{x}_{\mathrm{fd}}(t) &= \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{v}}} \sum_{p=1}^{P} \underline{\mathbf{t}}_p \underline{\mathbf{t}}_p^T \underline{\mathbf{x}}(t) \\ &= \underline{\mathbf{I}}_{\mathrm{i}} \left( \mathbf{I}_{ML} - \mathbf{R}_{\underline{\mathbf{v}}} \sum_{i=P+1}^{ML} \underline{\mathbf{t}}_i \underline{\mathbf{t}}_i^T \right) \underline{\mathbf{x}}(t) \\ &= \mathbf{x}_1(t) - \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{v}}} \sum_{i=P+1}^{ML} \underline{\mathbf{t}}_i \underline{\mathbf{t}}_i^T \underline{\mathbf{x}}(t) = \mathbf{x}_1(t), \end{aligned} \tag{50}$$

where we have used (10) and (11). Then, it is clear that

$$\upsilon_{\mathrm{d}} \left( \underline{\mathbf{H}}_{\mathrm{MVDR}} \right) = 0, \tag{51}$$

proving that, indeed, $\underline{\mathbf{H}}_{\mathrm{MVDR}}$ is the MVDR filtering matrix.

### Property

*With the MVDR filtering matrix given in (49), the output SNR is always greater than or equal to the input SNR, i.e.,* $\mathrm{oSNR}\left(\underline{\mathbf{H}}_{\mathrm{MVDR}}\right) \geq \mathrm{iSNR}$.

# Controlled Distortion Filtering Matrix

From the MVDR filtering matrix, we can propose the controlled distortion (CD) filtering matrix:

$$\underline{\mathbf{H}}_{\mathrm{CD},P'} = \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{x}}} \sum_{p'=1}^{P'} \frac{\underline{\mathbf{t}}_{p'} \underline{\mathbf{t}}_{p'}^T}{\underline{\lambda}_{p'}}, \tag{52}$$

where $1 \leq P' \leq P$. We observe that $\underline{\mathbf{H}}_{\mathrm{CD},P} = \underline{\mathbf{H}}_{\mathrm{MVDR}}$ and for $P' = 1$, we obtain the maximum SNR filtering matrix with minimum distortion:

$$\underline{\mathbf{H}}_{\max,0} = \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{x}}} \frac{\underline{\mathbf{t}}_1 \underline{\mathbf{t}}_1^T}{\underline{\lambda}_1}, \tag{53}$$

since

$$\mathrm{oSNR}\left(\underline{\mathbf{H}}_{\max,0}\right) = \underline{\lambda}_1. \tag{54}$$

# Example 3

Returning to Example 2, we now employ the CD filtering matrix, $\underline{\mathbf{H}}_{\mathrm{CD},P'}$, given in (52).

Figure 7 shows plots of the gain in SNR, $\mathcal{G}\left(\underline{\mathbf{H}}_{\mathrm{CD},P'}\right)$, the MSE, $J\left(\underline{\mathbf{H}}_{\mathrm{CD},P'}\right)$, the noise reduction factor, $\xi_{\mathrm{n}}\left(\underline{\mathbf{H}}_{\mathrm{CD},P'}\right)$, and the desired-signal reduction factor, $\xi_{\mathrm{d}}\left(\underline{\mathbf{H}}_{\mathrm{CD},P'}\right)$, as a function of the input SNR for several values of $P'$.

For a given input SNR, the higher is the value of $P'$, the lower are the MSE and the desired-signal reduction factor, but at the expense of lower gain in SNR and lower noise reduction factor.

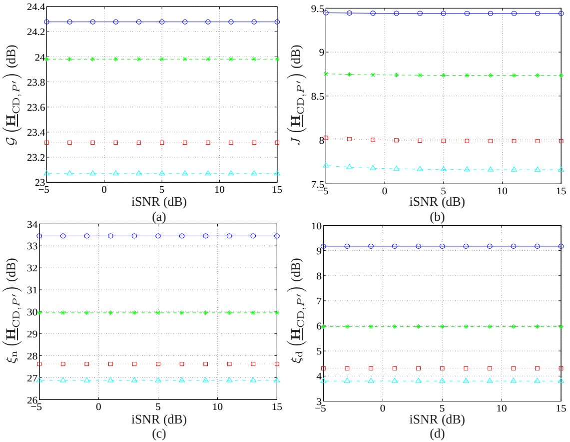



Figure 7: (a) The gain in SNR, (b) the MSE, (c) the noise reduction factor, and (d) the desired-signal reduction factor of the CD filtering matrix as a function of the input SNR for several values of $P'$: $P' = 1$ (solid line with circles), $P' = 2$ (dashed line with asterisks), $P' = 3$ (dotted line with squares), and $P' = 4$ (dash-dot line with triangles).

# Tradeoff Filtering Matrix

Another practical approach that can compromise between noise reduction and desired-signal distortion is the tradeoff filtering matrix obtained by

$$\min_{\underline{\mathbf{A}}} J_{\mathrm{d}}\left(\underline{\mathbf{A}}\right) \quad \text{subject to} \quad J_{\mathrm{n}}\left(\underline{\mathbf{A}}\right) = \aleph \mathrm{tr}\left(\mathbf{R}_{\mathbf{v}_1}\right), \tag{55}$$

where $0 < \aleph < 1$ to ensure that filtering achieves some degree of noise reduction.

We find that the optimal filtering matrix is

$$\underline{\mathbf{H}}_{\mathrm{T},\mu} = \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{x}}} \sum_{i=1}^{ML} \frac{\underline{\mathbf{t}}_i \underline{\mathbf{t}}_i^T}{\mu + \underline{\lambda}_i}, \tag{56}$$

where $\mu \geq 0$ is a Lagrange multiplier.

For $\mu = 1$, we get the Wiener filtering matrix.

### Property

*With the tradeoff filtering matrix given in (56), the output SNR is always greater than or equal to the input SNR, i.e.,* $\mathrm{oSNR}\left(\underline{\mathbf{H}}_{\mathrm{T},\mu}\right) \geq \mathrm{iSNR},\ \forall \mu \geq 0.$

# Example 4

Returning to Example 2, we now employ the tradeoff filtering matrix, $\underline{\mathbf{H}}_{\mathrm{T},\mu}$, given in (56).

Figure 8 shows plots of the gain in SNR, $\mathcal{G}\left(\underline{\mathbf{H}}_{\mathrm{T},\mu}\right)$, the MSE, $J\left(\underline{\mathbf{H}}_{\mathrm{T},\mu}\right)$, the noise reduction factor, $\xi_{\mathrm{n}}\left(\underline{\mathbf{H}}_{\mathrm{T},\mu}\right)$, and the desired-signal reduction factor, $\xi_{\mathrm{d}}\left(\underline{\mathbf{H}}_{\mathrm{T},\mu}\right)$, as a function of the input SNR for several values of $\mu$.

For a given input SNR, the higher is the value of $\mu$, the higher are the gain in SNR and the noise reduction factor, but at the expense of higher desired-signal reduction factor.

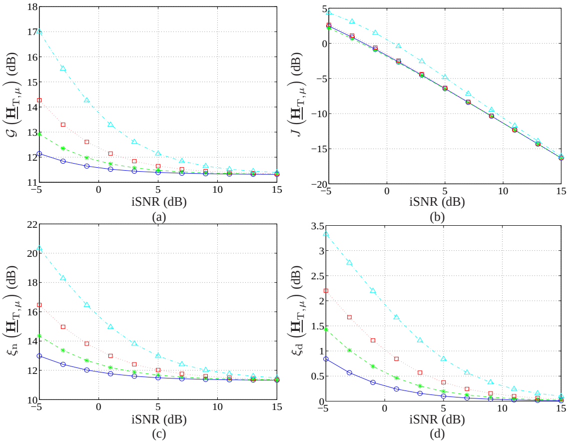



Figure 8: (a) The gain in SNR, (b) the MSE, (c) the noise reduction factor, and (d) the desired-signal reduction factor of the tradeoff filtering matrix as a function of the input SNR for several values of $\mu$: $\mu = 0.5$ (solid line with circles), $\mu = 1$ (dashed line with asterisks), $\mu = 2$ (dotted line with squares), and $\mu = 5$ (dash-dot line with triangles).

# General Subspace Filtering Matrix

From what we have seen so far, we can propose a very general subspace (GS) noise reduction filtering matrix [7]:

$$\underline{\mathbf{H}}_{\mu,Q} = \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{x}}} \sum_{q=1}^{Q} \frac{\underline{\mathbf{t}}_q \underline{\mathbf{t}}_q^T}{\mu + \underline{\lambda}_q}, \tag{57}$$

where $1 \leq Q \leq ML$.

This form encompasses most known optimal filtering matrices.

Indeed, it is clear that

$$\begin{array}{lll} \underline{\mathbf{H}}_{1,ML} = \underline{\mathbf{H}}_{\mathrm{W}} & \underline{\mathbf{H}}_{1,Q} = \underline{\mathbf{H}}_{\mathrm{W},Q} & \underline{\mathbf{H}}_{1,1} = \underline{\mathbf{H}}_{\mathrm{max},1} \\ \underline{\mathbf{H}}_{0,P} = \underline{\mathbf{H}}_{\mathrm{MVDR}} & \underline{\mathbf{H}}_{0,P'} = \underline{\mathbf{H}}_{\mathrm{CD},P'} & \underline{\mathbf{H}}_{0,1} = \underline{\mathbf{H}}_{\mathrm{max},0} \\ \underline{\mathbf{H}}_{\mu,ML} = \underline{\mathbf{H}}_{\mathrm{T},\mu} & & \end{array}$$

Table 1: Optimal linear filtering matrices for multichannel signal enhancement in the time domain.

| | |
|---|---|
| Wiener: | $\underline{\mathbf{H}}_{\mathrm{W}} = \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{x}}} \sum_{i=1}^{ML} \frac{\underline{\mathbf{t}}_i \underline{\mathbf{t}}_i^T}{1 + \underline{\lambda}_i}$ |
| VS Wiener: | $\underline{\mathbf{H}}_{\mathrm{W},Q} = \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{x}}} \sum_{q=1}^{Q} \frac{\underline{\mathbf{t}}_q \underline{\mathbf{t}}_q^T}{1 + \underline{\lambda}_q}, \ 1 \leq Q \leq ML$ |
| MVDR: | $\underline{\mathbf{H}}_{\mathrm{MVDR}} = \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{x}}} \sum_{p=1}^{P} \frac{\underline{\mathbf{t}}_p \underline{\mathbf{t}}_p^T}{\underline{\lambda}_p}$ |
| CD: | $\underline{\mathbf{H}}_{\mathrm{CD},P'} = \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{x}}} \sum_{p'=1}^{P'} \frac{\underline{\mathbf{t}}_{p'} \underline{\mathbf{t}}_{p'}^T}{\underline{\lambda}_{p'}}, \ 1 \leq P' \leq P$ |
| Maximum SNR: | $\underline{\mathbf{H}}_{\max,\mu} = \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{x}}} \frac{\underline{\mathbf{t}}_1 \underline{\mathbf{t}}_1^T}{\mu + \underline{\lambda}_1}, \ \mu \geq 0$ |
| Tradeoff: | $\underline{\mathbf{H}}_{\mathrm{T},\mu} = \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{x}}} \sum_{i=1}^{ML} \frac{\underline{\mathbf{t}}_i \underline{\mathbf{t}}_i^T}{\mu + \underline{\lambda}_i}, \ \mu \geq 0$ |
| GS: | $\underline{\mathbf{H}}_{\mu,Q} = \underline{\mathbf{I}}_{\mathrm{i}} \mathbf{R}_{\underline{\mathbf{x}}} \sum_{q=1}^{Q} \frac{\underline{\mathbf{t}}_q \underline{\mathbf{t}}_q^T}{\mu + \underline{\lambda}_q}, \ \mu \geq 0, \ 1 \leq Q \leq ML$ |

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

[2] M. Brandstein and D. B. Ward, Eds., *Microphone Arrays: Signal Processing Techniques and Applications*. Berlin, Germany: Springer-Verlag, 2001.

[3] J. Benesty and J. Chen, *Optimal Time-domain Noise Reduction Filters–A Theoretical Study*. Springer Briefs in Electrical and Computer Engineering, 2011.

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

[5] J. Benesty, M. G. Christensen, and J. R. Jensen, *Signal Enhancement with Variable Span Linear Filters*. Berlin, Germany: Springer-Verlag, 2016.

[6] J. R. Jensen, J. Benesty, and M. G. Christensen, "Noise reduction with optimal variable span linear filters," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 24, pp. 631–644, Apr. 2016.

[7] J. Benesty, J. R. Jensen, M. G. Christensen, and J. Chen, *Speech Enhancement–A Signal Subspace Perspective*. Academic Press, 2014.

[8] P. C. Hansen and S. H. Jensen, "Prewhitening for rank-deficient noise in subspace methods for noise reduction," *IEEE Trans. Signal Process.*, vol. 53, pp. 3718–3726, Oct. 2005.

[9] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Third Edition. Baltimore, Maryland: The Johns Hopkins University Press, 1996.

[10] S. M. Nørholm, J. Benesty, J. R. Jensen, and M. G. Christensen, “Single-channel noise reduction using unified joint diagonalization and optimal filtering,” *EURASIP J. Advances Signal Process.*, 2014, 2014:37 (11 pages).