

AN LCMV FILTER FOR SINGLE-CHANNEL NOISE CANCELLATION AND REDUCTION IN THE TIME DOMAIN

Jesper Rindom Jensen¹, Jacob Benesty^{1,2}, Mads Græsbøll Christensen¹, and Jingdong Chen³

¹Audio Analysis Lab, AD:MT, Aalborg University, Aalborg, Denmark, {jrj,mgc}@create.aau.dk

²INRS-EMT, University of Quebec, Montreal, Canada, benesty@emt.inrs.ca

³Northwestern Polytechnical University, Xi'an, China, jingdongchen@ieee.org

ABSTRACT

In this paper, we consider a recent class of optimal rectangular filtering matrices for single-channel speech enhancement. This class of filters exploits the fact that the dimension of the signal subspace is lower than that of the full space. Then, extra degrees of freedom in the filters, that are otherwise reserved for preserving the signal subspace, can be used for achieving an improved output signal-to-noise ratio (SNR). Interestingly, these filters unify the ideas of optimal filtering and subspace methods. We propose an optimal LCMV filter in this framework with minimum output power that passes the desired signal undistorted and cancels correlated noise. The cancellation was not facilitated by the filters derived so far in this framework. The results show that the proposed filter can achieve output SNRs similar to that of competing filter designs, while having a much higher output signal-to-interference ratio. This is showed for both synthetic and real speech signals.

Index Terms— Speech enhancement, interferer cancellation, LCMV, optimal filtering.

1. INTRODUCTION

The problem of speech enhancement has a rich and long history, but it remains a widely studied problem due to its occurrence in recent applications such as voice-over-IP, hearing aids, teleconferencing, mobile telephony, etc. Enhancement is vital in such systems as 1) noise has a detrimental impact on the perceived quality and intelligibility of speech signals and causes listener fatigue under extended exposure and 2) many speech processing systems or components are designed under the premise that only one, clean signal is present at the time. Even though more and more systems are now using multiple channels, i.e., microphone arrays, many systems today are still based on only a single channel, and this is also the context in which we will study the speech enhancement problem.

The speech enhancement problem can be posed as a filtering problem, wherein an estimate of the desired speech signal is obtained via filtering of the observed, noisy signal. An example of this is the classical Wiener filter. Such filtering approaches often require that either an estimate of the speech statistics or the noise statistics be found or known, and in the past decade, most efforts in improving speech enhancement algorithms has been devoted to the problem of estimating the noise statistics, with some examples being [1–3]. Recently, a number of important advances have, however, been made formulating different kinds of optimal filters. These include the adaptation of the linearly constrained minimum variance (LCMV) and the minimum variance distortionless response (MVDR) principles to speech enhancement [4, 5] in combination with the orthogonal [4] and harmonic decompositions [6], as well as the extension of these to non-causal filters [7].

An alternative approach to speech enhancement is the so-called subspace method [8], wherein bases of the signal and noise subspaces are obtained from the eigenvalue decomposition of the covariance matrix. Then enhancement is performed by modifying the eigenvalues corresponding to the signal and noise subspaces after which an estimate of the clean signal can be obtained. In the literature, the subspace methods are usually described as a competing approach to speech enhancement, although some interpretations of these approaches as filtering exist [9].

In this paper, we consider a new class of optimal filtering matrices that combine the notion of subspace-based enhancement with classical filtering approaches. As such, this approach unifies subspace and filtering methods in a common framework. More specifically, we propose an LCMV filter in this framework that enables cancellation of correlated noise, while at the same time suppressing the uncorrelated noise as much as possible. Cancellation of correlated noise was not possible with the filters derived in this framework so far [10].

The remainder of this paper is organized as follows. In Section 2, the signal model is introduced and the speech enhancement problem is stated, after which the linear filtering approach with a rectangular filtering matrix is introduced in Section 3 along with some useful performance measures. In Section 4, the LCMV filter is proposed, and its performance is studied in Section 5 on synthetic as well as real-life speech signals, and the results are discussed.

2. SIGNAL MODEL AND PROBLEM FORMULATION

The signal enhancement (or noise reduction) problem considered in this work is one of recovering the desired signal (or clean signal) $x(k)$, with k being the discrete-time index, from the noisy observation (sensor signal):

$$y(k) = x(k) + x_i(k) + v(k), \quad (1)$$

where $x_i(k)$ is a interfering signal correlated over time and $v(k)$ is an unwanted additive noise with no time correlation (i.e., white noise). Both noise sources, however, are assumed to be uncorrelated with $x(k)$. Moreover, all signals are considered to be real, zero mean, and stationary.

The signal model given in (1) can be put into a vector form by considering the L most recent successive time samples of the noisy signal, i.e.,

$$\mathbf{y}(k) = \mathbf{x}(k) + \mathbf{x}_i(k) + \mathbf{v}(k), \quad (2)$$

where $\mathbf{y}(k) = [y(k) \ y(k-1) \ \cdots \ y(k-L+1)]^T$ is a vector of length L , $(\cdot)^T$ denotes the transpose of a vector or a matrix, and $\mathbf{x}(k)$, $\mathbf{x}_i(k)$, and $\mathbf{v}(k)$ are defined in a similar way to $\mathbf{y}(k)$. As a result of the assumptions, the correlation matrix of size $L \times L$ of the noisy signal can be written as

$$\mathbf{R}_y = E[\mathbf{y}(k)\mathbf{y}^T(k)] = \mathbf{R}_x + \mathbf{R}_{x_i} + \mathbf{R}_v, \quad (3)$$

This work was supported in part by the Villum foundation.

where $E[\cdot]$ denotes the mathematical expectation, and $\mathbf{R}_x = E[\mathbf{x}(k)\mathbf{x}^T(k)]$, $\mathbf{R}_{x_i} = E[\mathbf{x}_i(k)\mathbf{x}_i^T(k)]$, and $\mathbf{R}_v = E[\mathbf{v}(k)\mathbf{v}^T(k)]$ are the correlation matrices of $\mathbf{x}(k)$, $\mathbf{x}_i(k)$, and $\mathbf{v}(k)$, respectively. The noise correlation matrix, \mathbf{R}_v , is assumed to be full rank, i.e., its rank is equal to L . In the rest of the paper, we assume that the rank of the desired signal correlation matrix, \mathbf{R}_x , is equal to P , that \mathbf{R}_{x_i} has rank Q , and that $P + Q < L$. This assumption is reasonable in several applications such as speech enhancement, where the speech signal can be modeled as the sum of a small number of sinusoids. In any case, for a given P and Q , we can always choose L much greater than $P + Q$. Then, the objective of signal enhancement (or noise reduction) is to estimate the desired signal vector, $\mathbf{x}(k)$, or any known linear transformation of it from $\mathbf{y}(k)$. This should be done in such a way that the noise is reduced as much as possible with little or no distortion of the desired signal.

Using the well-known eigenvalue decomposition, the desired signal correlation matrix can be diagonalized as

$$\mathbf{R}_x = \mathbf{Q}_x \mathbf{\Lambda}_x \mathbf{Q}_x^T, \quad (4)$$

where $\mathbf{Q}_x = [\mathbf{q}_{x,1} \ \mathbf{q}_{x,2} \ \cdots \ \mathbf{q}_{x,L}]$ is an orthogonal matrix, i.e., $\mathbf{Q}_x^T \mathbf{Q}_x = \mathbf{Q}_x \mathbf{Q}_x^T = \mathbf{I}_L$, with \mathbf{I}_L being the $L \times L$ identity matrix, and $\mathbf{\Lambda}_x = \text{diag}(\lambda_{x,1}, \lambda_{x,2}, \dots, \lambda_{x,L})$ is a diagonal matrix. The orthonormal vectors $\mathbf{q}_{x,1}, \mathbf{q}_{x,2}, \dots, \mathbf{q}_{x,L}$ are the eigenvectors corresponding, respectively, to the eigenvalues $\lambda_{x,1}, \lambda_{x,2}, \dots, \lambda_{x,L}$ of the matrix \mathbf{R}_x , where $\lambda_{x,1} \geq \lambda_{x,2} \geq \dots \geq \lambda_{x,P} > 0$ and $\lambda_{x,P+1} = \lambda_{x,P+2} = \dots = \lambda_{x,L} = 0$. Let $\mathbf{Q}_x = [\mathbf{T}_x \ \mathbf{Y}_x]$, where the $L \times P$ matrix \mathbf{T}_x contains the eigenvectors corresponding to the nonzero eigenvalues of \mathbf{R}_x and the $L \times (L - P)$ matrix \mathbf{Y}_x contains the eigenvectors corresponding to the null eigenvalues of \mathbf{R}_x . It can be verified that

$$\mathbf{I}_L = \mathbf{T}_x \mathbf{T}_x^T + \mathbf{Y}_x \mathbf{Y}_x^T. \quad (5)$$

Notice that $\mathbf{T}_x \mathbf{T}_x^T$ and $\mathbf{Y}_x \mathbf{Y}_x^T$ are two orthogonal projection matrices of rank P and $L - P$, respectively. Hence, $\mathbf{T}_x \mathbf{T}_x^T$ is the orthogonal projector onto the desired signal subspace where all the energy of the desired signal is concentrated and $\mathbf{Y}_x \mathbf{Y}_x^T$ is the orthogonal projector onto the null subspace. Using (5), we can write the desired signal vector as

$$\mathbf{x}(k) = \mathbf{Q}_x \mathbf{Q}_x^T \mathbf{x}(k) = \mathbf{T}_x \tilde{\mathbf{x}}(k), \quad (6)$$

where $\tilde{\mathbf{x}}(k) = \mathbf{T}_x^T \mathbf{x}(k)$ is the transformed desired signal vector of length P . Therefore, the signal model for noise reduction becomes

$$\mathbf{y}(k) = \mathbf{T}_x \tilde{\mathbf{x}}(k) + \mathbf{T}_{x_i} \tilde{\mathbf{x}}_i(k) + \mathbf{v}(k), \quad (7)$$

where $\tilde{\mathbf{x}}_i(k) = \mathbf{T}_{x_i}^T \mathbf{x}_i(k)$ is the transformed interfering signal vector of length Q , and \mathbf{T}_{x_i} contains the Q th most significant eigenvectors of \mathbf{R}_{x_i} . Fundamentally, from the L observations, we wish to estimate the P components of the transformed desired signal, i.e., $\tilde{\mathbf{x}}(k)$. Thanks to this transformation and the nullspace of \mathbf{R}_x , we are able to reduce the dimension of the desired signal vector that we want to estimate. Indeed, there is no need to use the subspace \mathbf{Y}_x since it contains no desired signal information. From (7), we give another form of the correlation matrix of $\mathbf{y}(k)$:

$$\mathbf{R}_y = \mathbf{T}_x \mathbf{R}_{\tilde{\mathbf{x}}} \mathbf{T}_x^T + \mathbf{R}_{x_i} + \mathbf{R}_v = \mathbf{T}_x \mathbf{\Lambda}_{\tilde{\mathbf{x}}} \mathbf{T}_x^T + \mathbf{R}_{x_i} + \mathbf{R}_v, \quad (8)$$

where

$$\begin{aligned} \mathbf{R}_{\tilde{\mathbf{x}}} &= E[\tilde{\mathbf{x}}(k)\tilde{\mathbf{x}}^T(k)] = \mathbf{T}_x^T \mathbf{R}_x \mathbf{T}_x = \mathbf{T}_x^T \mathbf{Q}_x \mathbf{\Lambda}_x \mathbf{Q}_x^T \mathbf{T}_x \\ &= \text{diag}(\lambda_{x,1}, \lambda_{x,2}, \dots, \lambda_{x,P}) = \mathbf{\Lambda}_{\tilde{\mathbf{x}}} \end{aligned} \quad (9)$$

and, obviously, $\mathbf{R}_x = \mathbf{T}_x \mathbf{R}_{\tilde{\mathbf{x}}} \mathbf{T}_x^T = \mathbf{T}_x \mathbf{\Lambda}_{\tilde{\mathbf{x}}} \mathbf{T}_x^T$.

3. LINEAR FILTERING

From the general linear filtering approach, we can estimate the desired signal vector, $\tilde{\mathbf{x}}(k)$, by applying a linear transformation to the observation signal vector, $\mathbf{y}(k)$, i.e.,

$$\begin{aligned} \tilde{\mathbf{z}}(k) &= \tilde{\mathbf{H}} \mathbf{y}(k) = \tilde{\mathbf{H}} [\mathbf{x}(k) + \mathbf{x}_i(k) + \mathbf{v}(k)] \\ &= \tilde{\mathbf{x}}_{\text{fd}}(k) + \tilde{\mathbf{x}}_{\text{ri}}(k) + \tilde{\mathbf{v}}_{\text{rn}}(k), \end{aligned} \quad (10)$$

where $\tilde{\mathbf{z}}(k)$ is supposed to be the estimate of $\tilde{\mathbf{x}}(k)$, $\tilde{\mathbf{H}} = [\tilde{\mathbf{h}}_1 \ \tilde{\mathbf{h}}_2 \ \cdots \ \tilde{\mathbf{h}}_P]^T$ is a rectangular filtering matrix of size $P \times L$, $\tilde{\mathbf{h}}_p = [\tilde{h}_{p,0} \ \tilde{h}_{p,1} \ \cdots \ \tilde{h}_{p,L-1}]^T$, $p = 1, 2, \dots, P$ are finite-impulse-response (FIR) filters of length L ,

$$\tilde{\mathbf{x}}_{\text{fd}}(k) = \tilde{\mathbf{H}} \mathbf{x}(k) = \tilde{\mathbf{H}} \mathbf{T}_x \tilde{\mathbf{x}}(k) \quad (11)$$

is the filtered transformed desired signal, while $\tilde{\mathbf{x}}_{\text{ri}}(k) = \tilde{\mathbf{H}} \mathbf{x}_i(k)$ and $\tilde{\mathbf{v}}_{\text{rn}}(k) = \tilde{\mathbf{H}} \mathbf{v}(k)$ are the residual interference and noise, respectively. As a result, the estimate of $\mathbf{x}(k)$ can be formed as

$$\mathbf{z}(k) = \mathbf{T}_x \tilde{\mathbf{z}}(k) = \mathbf{T}_x \tilde{\mathbf{H}} \mathbf{y}(k) = \mathbf{H} \mathbf{y}(k), \quad (12)$$

where

$$\mathbf{H} = \mathbf{T}_x \tilde{\mathbf{H}} = [\mathbf{h}_1 \ \mathbf{h}_2 \ \cdots \ \mathbf{h}_L]^T \quad (13)$$

is the filtering matrix of size $L \times L$ that leads to the estimation of $\mathbf{x}(k)$. The correlation matrix of $\tilde{\mathbf{z}}(k)$ is then

$$\mathbf{R}_{\tilde{\mathbf{z}}} = E[\tilde{\mathbf{z}}(k)\tilde{\mathbf{z}}^T(k)] = \mathbf{R}_{\tilde{\mathbf{x}}_{\text{fd}}} + \mathbf{R}_{\tilde{\mathbf{x}}_{\text{ri}}} + \mathbf{R}_{\tilde{\mathbf{v}}_{\text{rn}}}, \quad (14)$$

where $\mathbf{R}_{\tilde{\mathbf{x}}_{\text{fd}}} = \tilde{\mathbf{H}} \mathbf{R}_x \tilde{\mathbf{H}}^T = \tilde{\mathbf{H}} \mathbf{T}_x \mathbf{\Lambda}_x \mathbf{T}_x^T \tilde{\mathbf{H}}^T$, and $\mathbf{R}_{\tilde{\mathbf{v}}_{\text{rn}}} = \tilde{\mathbf{H}} \mathbf{R}_v \tilde{\mathbf{H}}^T$. We also observe that $\mathbf{R}_{\tilde{\mathbf{z}}} = \mathbf{T}_x \mathbf{R}_{\tilde{\mathbf{z}}} \mathbf{T}_x^T$ and $\text{tr}(\mathbf{R}_{\tilde{\mathbf{z}}}) = \text{tr}(\mathbf{R}_{\tilde{\mathbf{z}}})$, where $\text{tr}(\cdot)$ denotes the trace of a square matrix. The correlation matrix of $\tilde{\mathbf{z}}(k)$ or $\mathbf{z}(k)$ is helpful in defining meaningful performance measures.

We then define the most useful performance measures for time-domain signal enhancement in the single-channel case with a rectangular filtering matrix. The input SNR is a second-order measure that quantifies the level of noise present relative to the level of the desired signal. It is defined as

$$\text{iSNR} = \frac{\text{tr}(\mathbf{R}_x)}{\text{tr}(\mathbf{R}_{\text{in}})} = \frac{\sigma_x^2}{\sigma_{\text{in}}^2},$$

where $\sigma_x^2 = E[x^2(k)]$ and $\sigma_{\text{in}}^2 = E[x_i^2(k) + v^2(k)]$ are the variances of $x(k)$ and $[x_i(k) + v(k)]$, respectively.

The output SNR, obtained from (14), helps quantify the SNR after filtering. It is given by

$$\text{oSNR}(\tilde{\mathbf{H}}) = \frac{\text{tr}(\mathbf{R}_{\tilde{\mathbf{x}}_{\text{fd}}})}{\text{tr}(\mathbf{R}_{\tilde{\mathbf{x}}_{\text{ri}}} + \mathbf{R}_{\tilde{\mathbf{v}}_{\text{rn}}})} = \frac{\text{tr}(\tilde{\mathbf{H}} \mathbf{R}_x \tilde{\mathbf{H}}^T)}{\text{tr}(\tilde{\mathbf{H}} \mathbf{R}_{\text{in}} \tilde{\mathbf{H}}^T)}, \quad (15)$$

where $\mathbf{R}_{\text{in}} = \mathbf{R}_{x_i} + \mathbf{R}_v$. The objective is to find an appropriate $\tilde{\mathbf{H}}$ to make the output SNR greater than the input SNR. Consequently, the quality of the noisy signal will be enhanced. It can be shown that [4]

$$\text{oSNR}(\tilde{\mathbf{H}}) \leq \max_p \frac{\tilde{\mathbf{h}}_p^T \mathbf{R}_x \tilde{\mathbf{h}}_p}{\tilde{\mathbf{h}}_p^T \mathbf{R}_{\text{in}} \tilde{\mathbf{h}}_p} \leq \lambda_{\max}(\mathbf{R}_{\text{in}}^{-1} \mathbf{R}_x), \quad (16)$$

where $\lambda_{\max}(\mathbf{R}_{\text{in}}^{-1} \mathbf{R}_x)$ is the maximum eigenvalue of the matrix $\mathbf{R}_{\text{in}}^{-1} \mathbf{R}_x$. This shows how the output SNR is upper bounded. It is

easy to check that \mathbf{H} and $\tilde{\mathbf{H}}$, related through (13), yield the same output SNR, i.e.,

$$\text{oSNR}(\mathbf{H}) = \frac{\text{tr}(\mathbf{H}\mathbf{R}_x\mathbf{H}^T)}{\text{tr}(\mathbf{H}\mathbf{R}_{\text{in}}\mathbf{H}^T)} = \text{oSNR}(\tilde{\mathbf{H}}). \quad (17)$$

Fundamentally, there is no difference between $\tilde{\mathbf{H}}$ and \mathbf{H} . Both matrices lead to the same result as we should expect.

Another useful measure for enhancement methods is the signal-to-interference (SIR) ratio. The input SIR is defined as the ratio between the power of the desired signal and that of the interference, i.e.,

$$\text{iSIR} = \frac{\text{tr}(\mathbf{R}_x)}{\text{tr}(\mathbf{R}_{x_i})} = \frac{\sigma_x^2}{\sigma_{x_i}^2}, \quad (18)$$

where $\sigma_{x_i}^2$ is the variance $x_i(k)$. To quantify the SIR after noise reduction, we can consider the output SIR defined as

$$\text{oSIR}(\tilde{\mathbf{H}}) = \frac{\text{tr}(\mathbf{R}_{\tilde{x}_{\text{fd}}})}{\text{tr}(\mathbf{R}_{\tilde{x}_{\text{fi}}})} = \frac{\text{tr}(\tilde{\mathbf{H}}\mathbf{R}_x\tilde{\mathbf{H}}^T)}{\text{tr}(\tilde{\mathbf{H}}\mathbf{R}_{x_i}\tilde{\mathbf{H}}^T)} = \text{oSIR}(\mathbf{H}), \quad (19)$$

and if $\tilde{\mathbf{H}}\mathbf{T}_{x_i} = \mathbf{0}$, the interferer is removed completely by the filter, i.e., the output SIR is infinitely high.

We also have different measures regarding the distortion of the desired signal, with one example being the desired signal reduction factor defined as

$$\xi_{\text{sr}}(\tilde{\mathbf{H}}) = \frac{\text{tr}(\mathbf{R}_{\tilde{x}})}{\text{tr}(\mathbf{R}_{\tilde{x}_{\text{fd}}})} = \frac{\text{tr}(\Lambda_{\tilde{x}})}{\text{tr}(\tilde{\mathbf{H}}\mathbf{T}_x\Lambda_x\mathbf{T}_x^T\tilde{\mathbf{H}}^T)} = \xi_{\text{sr}}(\mathbf{H}). \quad (20)$$

Clearly, a rectangular filtering matrix that does not affect the desired signal requires the constraint $\tilde{\mathbf{H}}\mathbf{T}_x = \mathbf{I}_P$, where \mathbf{I}_P is the $P \times P$ identity matrix. Hence, $\xi_{\text{sr}}(\tilde{\mathbf{H}}) = 1$ in the absence of distortion and $\xi_{\text{sr}}(\tilde{\mathbf{H}}) > 1$ in the presence of distortion.

4. PROPOSED METHOD

Let us first introduce a combined noise term $\mathbf{w}(k) = \mathbf{x}_i(k) + \mathbf{v}(k)$ and decompose its correlation matrix as

$$\mathbf{R}_w = \mathbf{Q}_w\Lambda_w\mathbf{Q}_w^T, \quad (21)$$

where the orthogonal and diagonal matrices \mathbf{Q}_w and Λ_w are defined similarly to \mathbf{Q}_x and Λ_x , respectively. We assume that the positive eigenvalues of \mathbf{R}_w have the following structure: $\lambda_{w,1} \geq \lambda_{w,2} \geq \dots \geq \lambda_{w,Q} > \sigma_{\text{wn}}^2$ and $\lambda_{w,Q+1} = \lambda_{w,Q+2} = \dots = \lambda_{w,L} = \sigma_{\text{wn}}^2$, where $P + Q \leq L$. In this case, we can partition the unitary matrix as $\mathbf{Q}_w = [\mathbf{T}_{x_i} \quad \mathbf{T}_v]$, where the $L \times Q$ matrix \mathbf{T}_{x_i} contains the eigenvectors corresponding to the first Q eigenvalues of \mathbf{R}_w and the $L \times (L-Q)$ matrix \mathbf{T}_v contains the eigenvectors corresponding to the last $L-Q$ eigenvalues of \mathbf{R}_w . It is seen that $\mathbf{x}_i = \mathbf{T}_{x_i}\mathbf{T}_{x_i}^T\mathbf{w}$ corresponds to the correlated noise, $\mathbf{v} = \mathbf{T}_v\mathbf{T}_v^T\mathbf{w}$ corresponds to the uncorrelated noise, and $E(\mathbf{x}_i\mathbf{v}^T) = \mathbf{0}_{L \times L}$.

The linearly constrained minimum variance (LCMV) approach [11] consists of estimating $\mathbf{x}(k)$ without any distortion, completely removing the correlated noise, and attenuating the uncorrelated noise as much as possible. It follows that the constraints are

$$\tilde{\mathbf{H}}\mathbf{C}_{\text{xx}_i} = [\mathbf{I}_P \quad \mathbf{0}_{P \times Q}], \quad (22)$$

where $\mathbf{C}_{\text{xx}_i} = [\mathbf{T}_x \quad \mathbf{T}_v]$ is the constraint matrix of size $L \times (P+Q)$. Our optimization problem is now

$$\min_{\tilde{\mathbf{H}}} \text{tr}(\tilde{\mathbf{H}}\mathbf{R}_y\tilde{\mathbf{H}}^T) \quad \text{s. t.} \quad \tilde{\mathbf{H}}\mathbf{C}_{\text{xx}_i} = [\mathbf{I}_P \quad \mathbf{0}_{P \times Q}], \quad (23)$$

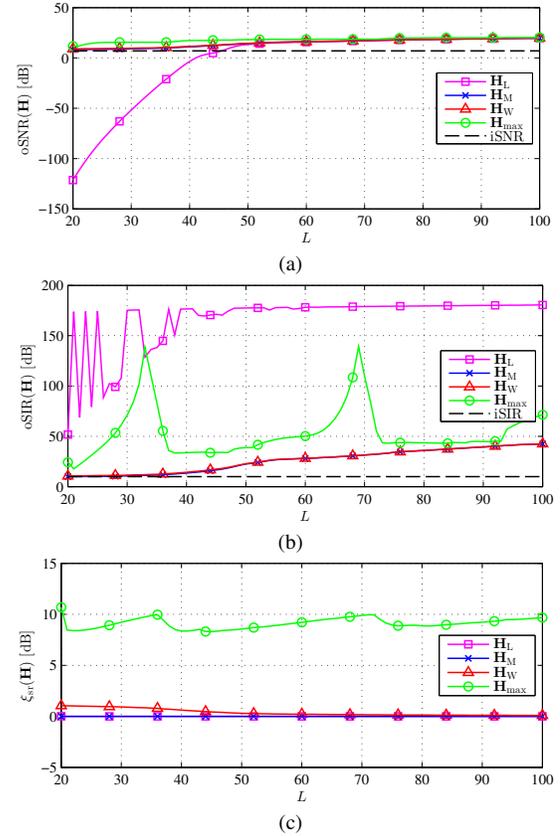


Figure 1: The (a) output SNRs, (b) output SIRs, and (c) signal reduction factors of the LCMV (\mathbf{H}_L), MVDR (\mathbf{H}_M), Wiener (\mathbf{H}_W), and maximum SNR (\mathbf{H}_{max}) filters when applied on a synthetic, noisy, periodic signal.

from which we find the LCMV filtering matrix as

$$\tilde{\mathbf{H}}_{\text{LCMV}} = [\mathbf{I}_P \quad \mathbf{0}_{P \times Q}] \left(\mathbf{C}_{\text{xx}_i}^T \mathbf{R}_y^{-1} \mathbf{C}_{\text{xx}_i} \right)^{-1} \mathbf{C}_{\text{xx}_i}^T \mathbf{R}_y^{-1}. \quad (24)$$

We immediately see from (24) that we must have $P + Q \leq L$, otherwise the matrix $\mathbf{C}_{\text{xx}_i}^T \mathbf{R}_y^{-1} \mathbf{C}_{\text{xx}_i}$ is not invertible. For $P + Q > L$, the LCMV does not exist. For $P + Q = L$, the LCMV simplifies to

$$\tilde{\mathbf{H}}_{\text{LCMV}} = [\mathbf{I}_P \quad \mathbf{0}_{P \times Q}] \mathbf{C}_{\text{xx}_i}^{-1}. \quad (25)$$

Finally, we see that the LCMV for the estimation of $\mathbf{x}(k)$ is

$$\mathbf{H}_{\text{LCMV}} = \mathbf{T}_x [\mathbf{I}_P \quad \mathbf{0}_{P \times Q}] \left(\mathbf{C}_{\text{xx}_i}^T \mathbf{R}_y^{-1} \mathbf{C}_{\text{xx}_i} \right)^{-1} \mathbf{C}_{\text{xx}_i}^T \mathbf{R}_y^{-1}. \quad (26)$$

5. RESULTS AND DISCUSSION

First, we evaluated the proposed filter on a synthetic, periodic signal, being mixed with an interfering periodic signal and white noise, which makes it possible to control the rank of the signal, interference and noise subspaces. In these experiments, the power ratios between the desired and interfering periodic signals and between the desired periodic signal and the white noise were 10 dB. Moreover, the desired periodic signal was constituted by 5 harmonics with a pitch of 0.175 rad/s, while the interfering periodic signal was constituted by 3 harmonics with a pitch of 0.31 rad/s. Then, covariance matrices needed for the filter design, i.e., \mathbf{R}_x , \mathbf{R}_{x_i} , and \mathbf{R}_v

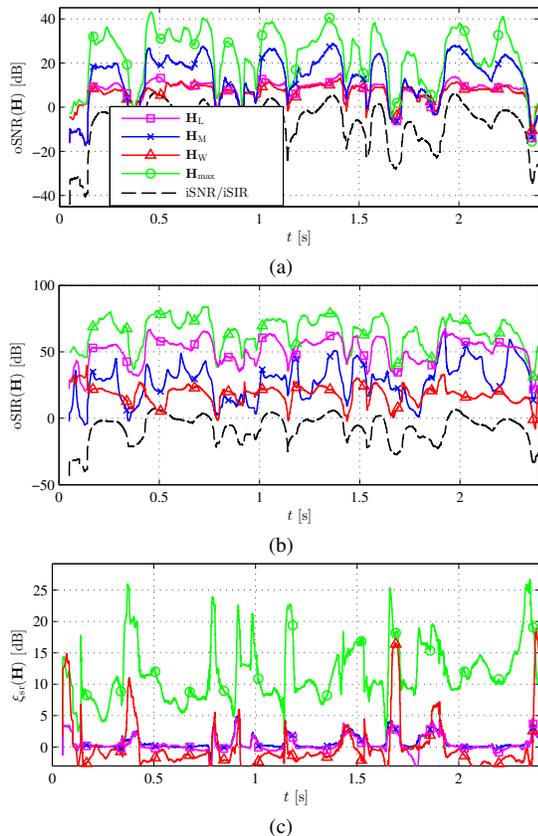


Figure 2: The estimated (a) output SNRs, (b) output SIRs, and (c) signal reduction factors of the LCMV (H_L), MVDR (H_M), Wiener (H_W), and maximum SNR (H_{max}) filters when applied on a real, noisy, speech signal.

were formed using the covariance matrix models for periodic signals and white noise, respectively. In this way, we get closed-form expressions for the performance measures of the filters, and do not need to estimate any signal statistics. Using this setup, we measured the output SNR, the output SIR, and the signal reduction factor for the proposed filter, and compared it with MVDR, Wiener, and maximum SNR filters [10]. The performance measures were evaluated as a function of the filter length, L , and they are shown in Fig. 1. As can be seen, the LCMV filter requires a longer filter length to achieve a high output SNR due to the extra constraints introduced. However, for $L > 50$, the LCMV can achieve much higher output SIR than all the other filters in the comparison, while having an output SNR similar to that of the MVDR and Wiener filters. As expected, the LCMV filter does not distort the desired signal in any case and thereby has a signal reduction factor of 0 dB.

Then, the LCMV filter was also evaluated on a real, speech signal. For this experiment, we used a 2.4 s long, female, speech excerpt from the Keele database [12]. We added white noise to the speech signal at an average ratio of 10 dB compared to the speech signal, and we added a periodic interferer at a speech-to-interference ratio of 0 dB. The interfering periodic signal was constituted by 3 unit amplitude harmonics with a pitch of 0.25 rad/s. Furthermore, the filter length was $L = 150$, the signal subspace rank was assumed to be $P = 20$, and the interfering signal subspace rank was $Q = 6$. To design the filters at each time instance,

we used outer product averaged estimates obtained directly from the past 400 samples of the clean speech, noise and interfering signals, respectively, as noise statistics estimation is out of the scope of this paper. Using this setup, the same filter as in the previous experiment were designed and applied to the noisy speech. Note that a whole vector of time-consecutive, desired signal sample estimates were obtained at each time instance, and these are thereby overlapping from one time instance to the following. The resulting enhanced signal was therefore obtained by averaging these over time. In Fig. 2, the resulting output SNRs, output SIRs, and signal reduction factors are depicted over time. As expected, the LCMV filter generally has a lower output SNR than the MVDR filter, but it also has a much higher output SIR.

6. REFERENCES

- [1] I. Cohen, "Noise spectrum estimation in adverse environments: improved minima controlled recursive averaging," *IEEE Trans. Speech Audio Process.*, vol. 11(5), pp. 466–475, 2003.
- [2] T. Gerkmann and R. C. Hendriks, "Unbiased MMSE-based noise power estimation with low complexity and low tracking delay," vol. 20(4), pp. 1383–1393, 2012.
- [3] R. C. Hendriks, R. Heusdens, J. Jensen, and U. Kjems, "Low complexity DFT-domain noise PSD tracking using high-resolution periodograms," *EURASIP J. on Advances in Signal Processing*, vol. 2009(1), pp. 15, 2009.
- [4] J. Benesty and J. Chen, *Optimal Time-Domain Noise Reduction Filters – A Theoretical Study*, Number VII. Springer, 1 edition, 2011.
- [5] M. G. Christensen and A. Jakobsson, "Optimal filter designs for separating and enhancing periodic signals," *IEEE Trans. Signal Process.*, vol. 58(12), pp. 5969–5983, 2010.
- [6] J. R. Jensen, J. Benesty, M. G. Christensen, and S. H. Jensen, "Enhancement of single-channel periodic signals in the time-domain," *IEEE Trans. Audio, Speech, Language Process.*, vol. 20(7), pp. 1948–1963, 2012.
- [7] J. R. Jensen, J. Benesty, M. G. Christensen, and S. H. Jensen, "Non-causal time-domain filters for single-channel noise reduction," *IEEE Trans. Audio, Speech, Language Process.*, vol. 20(5), pp. 1526–1541, 2012.
- [8] P. C. Hansen and S. H. Jensen, "Subspace-based noise reduction for speech signals via diagonal and triangular matrix decompositions: Survey and analysis," *EURASIP J. on Advances in Signal Processing*, vol. 2007(1), pp. 24, 2007.
- [9] P.C. Hansen and S.H. Jensen, "FIR filter representations of reduced-rank noise reduction," *IEEE Trans. Signal Process.*, vol. 46(6), pp. 1737–1741, 1998.
- [10] J. R. Jensen, J. Benesty, M. G. Christensen, and J. Chen, "A class of optimal rectangular filtering matrices for single-channel signal enhancement in the time-domain," *IEEE Trans. Audio, Speech, Language Process.*, pp. 1–11, 2013, submitted.
- [11] O. L. Frost, III, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60(8), pp. 926–935, 1972.
- [12] F. Plante, G. F. Meyer, and W. A. Ainsworth, "A pitch extraction reference database," in *Proc. Eurospeech*, 1995, pp. 837–840.