

Non-Causal Time-Domain Filters for Single-Channel Noise Reduction

Jensen, Jesper Rindom; Benesty, Jacob; Christensen, Mads Græsbøll; Jensen, Søren Holdt

Published in:

I E E Transactions on Audio, Speech and Language Processing

DOI (link to publication from Publisher):

[10.1109/TASL.2012.2183872](https://doi.org/10.1109/TASL.2012.2183872)

Publication date:

2012

Document Version

Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):

Jensen, J. R., Benesty, J., Christensen, M. G., & Jensen, S. H. (2012). Non-Causal Time-Domain Filters for Single-Channel Noise Reduction. *I E E Transactions on Audio, Speech and Language Processing*, 20(5), 1526-1541. <https://doi.org/10.1109/TASL.2012.2183872>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Non-Causal Time-Domain Filters for Single-Channel Noise Reduction

Jesper Rindom Jensen, *Student Member, IEEE*, Jacob Benesty, Mads Græsbøll Christensen, *Senior Member, IEEE*, and Søren Holdt Jensen, *Senior Member, IEEE*

Abstract—In many existing time-domain filtering methods for noise reduction in, e.g., speech processing, the filters are causal. Such causal filters can be implemented directly in practice. However, it is possible to improve the performance of such noise reduction filtering methods in terms of both noise suppression and signal distortion by allowing the filters to be non-causal. Non-causal time-domain filters require knowledge of the future, and are therefore not directly implementable. If the observed signal is processed in blocks, however, the non-causal filters are implementable. In this paper, we propose such non-causal time-domain filters for noise reduction in speech applications. We also propose some performance measures that enable us to evaluate the performance of non-causal filters. Moreover, it is shown how some of the filters can be updated recursively. Using the recursive expressions, it is also shown that the output SNRs of the filters always increase as we increase the length of the filter when the desired signal is stationary. From both the theoretical and practical evaluations of the filters, it is clearly shown that the performance of time-domain filtering methods for noise reduction can be improved by introducing non-causality.

Index Terms—Linearly constrained minimum variance (LCMV) filter, maximum signal-to-noise ratio (SNR) filter, minimum variance distortionless response (MVDR) filter, noise reduction, non-causal filters, performance measures, time-domain filtering, Wiener filter.

I. INTRODUCTION

NOISE reduction is an important fundamental signal processing problem. In this paper, we consider generic noise reduction filters which are useful for enhancing any kind of desired signal. An example of a desired signal is speech which is commonly utilized in a multitude of applications such as telecommunications, teleconferencing, hearing-aids, and human-machine interfaces. In all these, the speech first has to be recorded using one or more microphones, and the speech will inevitably be corrupted by some degree of background noise.

Manuscript received August 02, 2011; revised November 13, 2011 and January 03, 2012; accepted January 04, 2012. Date of publication January 11, 2012; date of current version March 14, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Man-Wai Mak.

J. R. Jensen and S. H. Jensen are with the Department of Electronic Systems, Aalborg University, DK-9220 Aalborg, Denmark (e-mail: jrj@es.aau.dk; shj@es.aau.dk).

J. Benesty is with INRS-EMT, University of Quebec, Montreal, QC H5A 1K6, Canada (e-mail: benesty@emt.inrs.ca).

M. G. Christensen is with the Department of Architecture, Design, and Media Technology, Aalborg University, DK-9220 Aalborg, Denmark (e-mail: mgc@imi.aau.dk).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TASL.2012.2183872

The noise could be, for example, other interfering speakers, fan noise, car noise, etc. Since the noise will reduce the speech quality and intelligibility, it will most likely have a detrimental impact on speech applications. In hearing-aids, for example, decreased speech quality can cause listener fatigue. It is therefore highly important to develop noise reduction methods to reduce the impact of the noise in various signal processing applications. Over the years, numerous noise reduction methods have been proposed. For an overview of speech related noise reduction methods, see, e.g., [1], [2], and the references therein. In general, we can divide these speech related noise reduction methods into three groups, i.e., spectral-subtractive algorithms [3], statistical-model-based algorithms [4]–[7], and subspace algorithms [8]–[11]. The references, [3]–[5], and [8]–[10], refer to some of the pioneering work within these groups. Note that in the literature, noise reduction in speech applications is also termed speech enhancement.

Often, noise reduction methods rely on linear filtering. In such filtering methods, the noise reduction problem is formulated as a filter design problem. The goal of such filter design problems is to design a filter which attenuates the noise as much as possible while it only introduces an inconsiderable amount of distortion of the desired signal, e.g., speech. The filter can be derived directly in the time domain or in different transform domains. For example, it is possible to reduce the computational complexity by utilizing transform domain filters [12]. Two examples of transform domains are the Fourier [3], [9], [13], [14] and Karhunen–Loève [15], [16] domains. The filters can, though, be equivalently derived in all domains. In this paper, we consider time-domain filters only. Moreover, we restrict ourselves to the study of single-channel filters only.

Many existing time-domain filter designs for noise reduction are causal. In this paper, however, we propose novel non-causal filter designs, and we quantify the performance gain which can be obtained by exploiting non-causality. Note that we only consider the effects of introducing non-causality in the filter designs and not of introducing non-causality in the estimation of the signal and noise statistics since the statistics are assumed to be known exactly in most parts of the paper. The proposed filter designs are based on two different decompositions of the desired signal; three designs are based on an orthogonal decomposition [12], and one is based on a harmonic decomposition [17], [18]. The orthogonal decomposition based filters are suitable for enhancing any kind of desired signal since they are designed using the noise statistics, whereas the harmonic decomposition based filter is calculated from the statistics of the desired signal under the assumption that it is periodic. Periodicity or quasi-periodicity is a reasonable assumption for, e.g., short segments of voiced speech and musical instrument signals. The

two decomposition approaches both have advantages and disadvantages as discussed in [19]. For example, the orthogonal decomposition based filters can be used for enhancing any kind of desired signal, however, they are sensible to nonstationary noise since it is difficult to estimate the noise statistics when the desired signal is present. The harmonic decomposition filter, on the other hand, is robust against nonstationary noise since it is based on the statistics of the desired signal, but it will cause distortion of the desired signal when the periodicity assumption does not hold exactly. It was shown in [19] and [20] that the orthogonal and harmonic decomposition based filters are closely related, and that it is beneficial to use them jointly for speech enhancement.

In this paper, we generalize the mentioned decompositions such that they support the derivation of non-causal time-domain filters. Based on these generalized decompositions, we propose several performance measures suited for evaluation of non-causal filters. Moreover, we derive different non-causal orthogonal and harmonic decomposition based filters. Note that the causal filters proposed in [12], [17], [18], and [21] can be seen as special cases of the proposed designs. For the two particular cases where the filter is causal and anti-causal, respectively, we derive expressions for recursive updates of the orthogonal decomposition based filters and the maximum output signal-to-noise ratio of these. From these recursive expressions, it can be shown that the maximum output SNR always increases if we increase the filter order when the desired signal is stationary. We quantify the performance gain that can be achieved by introducing non-causality in the filter design. To this end, we assume that the desired signal is periodic and, therefore, has a harmonic structure. Using this assumption, we can obtain exact closed-form expressions for the performance measures which we can use to precisely quantify the theoretical gains that can be achieved by using non-causal filters. Finally, we apply the non-causal filters to noise reduction of noisy speech signals to show the practical benefits of introducing non-causality in the filter design.

The rest of the paper is organized as follows. In Section II, we introduce the signal model utilized in the paper, and we define the problem of designing non-causal time-domain filters for noise reduction. We then, in Section III, describe the concept of linear non-causal filtering for noise reduction for two different signal decompositions. Based on the different decompositions, we propose several performance measures for non-causal noise reduction filters in Section IV. In Section V, we propose new optimal non-causal noise reduction filters. We show, in Section VI, that some of the filters and their output signal-to-noise ratios can be updated recursively. In Section VII, we quantify the performance gain that can be obtained by introducing non-causality in the filter design. Finally, we conclude on the paper in Section IX.

II. SIGNAL MODEL AND PROBLEM STATEMENT

In this paper, we consider the benefits of introducing non-causality in optimal time-domain linear filters for noise reduction. The objective of such filters is to extract a zero-mean desired signal $x(n) \in \mathbb{R}$ from an observed signal $y(n) \in \mathbb{R}$ defined as

$$y(n) = x(n) + v(n) \quad (1)$$

where $v(n) \in \mathbb{R}$ is additive noise and n denotes the discrete-time index. The observed signal $y(n)$ could, for example, be a microphone recording and the desired signal $x(n)$ could be clean speech. In the rest of the paper, we assume that the noise $v(n)$ is a zero-mean random process which is uncorrelated with the desired signal.

In some parts of the paper, we assume that the desired signal is quasi-periodic. This is a reasonable assumption for voiced speech segments. By assuming this specific signal structure, we can obtain closed-form expressions for certain performance measures related to optimal filters which are applied on the observed signal. Ultimately, the closed-form performance measures enable easy quantification of the performance gain which can be obtained by introducing non-causality in noise reduction filters. This will become clear from the later sections. When the desired signal is quasi-periodic, we can express it in terms of a harmonic model. The signal model in (1) then becomes

$$y(n) = \sum_{l=1}^L A_l \cos(l\omega_0 n + \phi_l) + v(n) \quad (2)$$

where ω_0 is the fundamental frequency (aka. the pitch), L is the number of harmonics, A_l is the amplitude of the l th harmonic, and ϕ_l is the phase of the l th harmonic. In this paper, we consider the pitch, ω_0 , and the model order, L , as known parameters. Numerous methods for estimation of these parameters exist [17], [18], [22]–[28]. Using Euler's formula, we can also write (2) as

$$y(n) = \sum_{l=1}^L (a_l e^{j l \omega_0 n} + a_l^* e^{-j l \omega_0 n}) + v(n) \quad (3)$$

where $a_l = (A_l/2)e^{j\phi_l}$ is the complex amplitude of the l th harmonic, and $(\cdot)^*$ denotes the elementwise complex conjugate of a matrix/vector.

To make the notation simpler when deriving the optimal non-causal noise reduction filters, we stack consecutive samples of the observed signal $y(n)$ into a vector $\mathbf{y}(n_k) \in \mathbb{R}^{M \times 1}$ where $n_k = n + k$. The vector signal model is then given by

$$\mathbf{y}(n_k) = \mathbf{x}(n_k) + \mathbf{v}(n_k) \quad (4)$$

where

$$\mathbf{y}(n_k) = [y(n_k) \ y(n_k - 1) \cdots y(n_k - M + 1)]^T \quad (5)$$

with $(\cdot)^T$ denoting the matrix/vector transpose. Note that the definitions of the desired signal vector $\mathbf{x}(n_k)$ and the noise vector $\mathbf{v}(n_k)$ follow the definition of the observed signal vector $\mathbf{y}(n_k)$ in (5). Since the observed signal $x(n)$ and the noise $v(n)$ are uncorrelated by assumption, we can obtain a simple expression for the covariance matrix $\mathbf{R}_y \in \mathbb{R}^{M \times M}$ of $\mathbf{y}(n_k)$ as

$$\begin{aligned} \mathbf{R}_y &= \mathbb{E}[\mathbf{y}(n_k) \mathbf{y}^T(n_k)] \\ &= \mathbf{R}_x + \mathbf{R}_v \end{aligned} \quad (6)$$

where $\mathbb{E}[\cdot]$ is the mathematical expectation operator, $\mathbf{R}_x = \mathbb{E}[\mathbf{x}(n_k) \mathbf{x}^T(n_k)]$ is the covariance matrix of $\mathbf{x}(n_k)$, and $\mathbf{R}_v =$

$E[\mathbf{v}(n_k)\mathbf{v}^T(n_k)]$ is the covariance matrix of $\mathbf{v}(n_k)$. When $x(n)$ is quasi-periodic, we can also model \mathbf{R}_x as [29]

$$\begin{aligned}\mathbf{R}_x &\approx \mathbf{Z}_k(\omega_0)\mathbf{P}\mathbf{Z}_k^H(\omega_0) \\ &= \mathbf{Z}(\omega_0)\mathbf{P}\mathbf{Z}^H(\omega_0)\end{aligned}\quad (7)$$

with $(\cdot)^H$ denoting the complex conjugate transpose operator, and

$$\mathbf{P} = \text{diag}\{[|a_1|^2 \ |a_1^*|^2 \ \cdots \ |a_L|^2 \ |a_L^*|^2]\} \quad (8)$$

$$\mathbf{Z}_k(\omega_0) = \mathbf{Z}(\omega_0)\mathbf{S}(k) \quad (9)$$

$$\mathbf{Z}(\omega_0) = [\mathbf{z}(\omega_0) \ \mathbf{z}^*(\omega_0) \ \cdots \ \mathbf{z}(L\omega_0) \ \mathbf{z}^*(L\omega_0)] \quad (10)$$

$$\mathbf{z}(l\omega_0) = [1 \ e^{-jl\omega_0} \ \cdots \ e^{-jl\omega_0(M-1)}]^T \quad (11)$$

$$\mathbf{S}(k) = \text{diag}\{[e^{j\omega_0 k} \ e^{-j\omega_0 k} \ \cdots \ e^{jL\omega_0 k} \ e^{-jL\omega_0 k}]\} \quad (12)$$

where $\text{diag}\{\cdot\}$ denotes the construction of a diagonal matrix from a vector.

The objective in traditional noise reduction methods is to find a “good” estimate of $x(n)$ or $\mathbf{x}(n)$ from the observed signal vector $\mathbf{y}(n)$. Within the field of speech enhancement research there is, in general, consensus on that “good” means the noise should be reduced as much as possible while the desired signal remains undistorted or nearly undistorted in the noise reduction process. In this paper, we consider another approach on noise reduction where we instead estimate $x(n)$ from $\mathbf{y}(n_k)$ where $k \in [0; M-1]$. That is, we introduce non-causality in the estimation procedure which, eventually, can increase the amount of noise reduction.

III. NOISE REDUCTION USING NON-CAUSAL LINEAR FILTERS

Filtering methods constitute a commonly used group of methods for noise reduction tasks such as speech enhancement. In the majority of such filtering methods, a finite impulse response (FIR) filter is applied on the observed signal vector. If the filter is allowed to be non-causal, we can, in general, write the noise reduction filtering operation as

$$\begin{aligned}\hat{x}_k(n) &= \sum_{m=-k}^{M-1-k} h_{m,k} y(n-m) \\ &= \mathbf{h}_k^T \mathbf{y}(n_k)\end{aligned}\quad (13)$$

for $k \in [0; M-1]$ and where

$$\mathbf{h}_k = [h_{-k,k} \ h_{-k+1,k} \ \cdots \ h_{-k+M-1,k}]^T \quad (14)$$

and $\hat{x}_k(n)$ should be an estimate of $x(n)$. Traditionally, time-domain filters for noise reduction have been considered causal, i.e., they have been derived for $k = 0$ (see, e.g., [12] and the references therein). In this paper, however, we consider the general case where k can be any integer in the interval $[0; M-1]$. In practice, we can easily implement non-causal filters by doing block processing and allowing a small delay.

In the last couple of decades, several different causal filter designs have been proposed. The main difference between the designs is how the observed signal is decomposed. For the causal

filter design problem, we have, for example, the classical, the orthogonal, and the harmonic decompositions [12], [18]. In the following, we redefine the orthogonal and harmonic decompositions by introducing non-causality.

A. Orthogonal Decomposition

Recently, it was proposed to design causal time-domain noise reduction filter based on an orthogonal decomposition of the desired signal [21], [12]. By using this approach, it is clear that some components of the signal vector $\mathbf{x}(n)$ actually act as interference when we estimate the desired signal $x(n)$. Here, we generalize this decomposition by introducing non-causality to enable the estimation of $x(n)$ from $\mathbf{x}(n_k)$. If we apply the orthogonal decomposition with respect to $x(n)$ on the signal vector $\mathbf{x}(n_k)$, we get

$$\begin{aligned}\mathbf{x}(n_k) &= x(n)\boldsymbol{\rho}_{\mathbf{x},k} + \mathbf{x}_i(n_k) \\ &= \mathbf{x}_d(n_k) + \mathbf{x}_i(n_k)\end{aligned}\quad (15)$$

where

$$\boldsymbol{\rho}_{\mathbf{x},k} = \frac{E[\mathbf{x}(n_k)x(n)]}{E[x^2(n)]} \quad (16)$$

$$\begin{aligned}&= [\rho_x(k) \ \rho_x(k-1) \ \cdots \ \rho_x(k-M+1)]^T \\ \rho_x(m) &= \frac{E[x(n+m)x(n)]}{E[x^2(n)]}.\end{aligned}\quad (17)$$

Note that $\rho_x(m) = 1$ for $m = 0$. The elements in $\mathbf{x}_d(n_k)$ in (15) are the parts of the elements in $\mathbf{x}(n_k)$ which are proportional to the desired signal $x(n)$ and $\mathbf{x}_i(n_k)$ is the “interference” which is orthogonal to $\mathbf{x}_d(n_k)$. If we insert (15) into (13), we get

$$\begin{aligned}\hat{x}_k(n) &= \mathbf{h}_k^T \mathbf{x}_d(n_k) + \mathbf{h}_k^T \mathbf{x}_i(n_k) + \mathbf{h}_k^T \mathbf{v}(n_k) \\ &= x_{fd,k}(n) + x_{ri,k}(n) + v_{rn,k}(n)\end{aligned}\quad (18)$$

where $x_{fd,k}(n) = \mathbf{h}_k^T \mathbf{x}_d(n_k)$ is the filtered desired signal, $x_{ri,k}(n) = \mathbf{h}_k^T \mathbf{x}_i(n_k)$ is the residual interference, and $v_{rn,k}(n) = \mathbf{h}_k^T \mathbf{v}(n_k)$ is the residual noise. Since $\mathbf{x}_d(n_k)$, $\mathbf{x}_i(n_k)$ and $\mathbf{v}(n_k)$ are all orthogonal to each other, the variance of $\hat{x}_k(n)$ is given by

$$\sigma_{\hat{x}_k}^2 = \sigma_{x_{fd,k}}^2 + \sigma_{x_{ri,k}}^2 + \sigma_{v_{rn,k}}^2 \quad (19)$$

where

$$\sigma_{x_{fd,k}}^2 = \mathbf{h}_k^T \mathbf{R}_{\mathbf{x}_d,k} \mathbf{h}_k = \sigma_x^2 (\mathbf{h}_k^T \boldsymbol{\rho}_{\mathbf{x},k})^2 \quad (20)$$

$$\sigma_{x_{ri,k}}^2 = \mathbf{h}_k^T \mathbf{R}_{\mathbf{x}_i,k} \mathbf{h}_k \quad (21)$$

$$\sigma_{v_{rn,k}}^2 = \mathbf{h}_k^T \mathbf{R}_{\mathbf{v}} \mathbf{h}_k. \quad (22)$$

$\mathbf{R}_{\mathbf{x}_d,k} = \sigma_x^2 \boldsymbol{\rho}_{\mathbf{x},k} \boldsymbol{\rho}_{\mathbf{x},k}^T$ is the covariance matrix of $\mathbf{x}_d(n_k)$, $\sigma_x^2 = E[x^2(n)]$ is the variance of the desired signal, and $\mathbf{R}_{\mathbf{x}_i,k} = E[\mathbf{x}_i(n_k)\mathbf{x}_i^T(n_k)]$ is the covariance matrix of the interference $\mathbf{x}_i(n_k)$.

We can obtain the following error function for the orthogonal decomposition approach

$$e_k(n) = x_{fd,k}(n) + x_{ri,k}(n) + v_{rn,k}(n) - x(n). \quad (23)$$

Compared to the classical filtering approach, the orthogonal decomposition approach has an extra noise term, namely the residual interference $x_{ri,k}(n)$ [12]. Moreover, the desired signal $x_{fd,k}(n)$ is different from the desired signal in the classical filtering approach. The design goal is to minimize the effect of $x_{ri,k}(n)$ and $v_{rn,k}(n)$ while keeping the difference between $x_{fd,k}(n)$ and $x(n)$ small. These goals can obviously be fulfilled by minimizing the error function in the mean square error (MSE) sense, possibly under some constraints.

B. Harmonic Decomposition

The harmonic decomposition approach to noise reduction filter design is a special case of the classical approach to linear filtering. In the harmonic decomposition, it is assumed that the desired signal is periodic and modeled by the harmonic model in (3). The harmonic model has previously been applied in numerous pitch estimation methods [18]. Many real-life signals such as audio and voiced speech are quasi-periodic which makes the harmonic decomposition approach useful in practice. By using the harmonic model, we can write the signal vector $\mathbf{x}(n_k)$ as

$$\mathbf{x}(n_k) = \mathbf{Z}(\omega_0)\mathbf{a}(n_k) = \mathbf{x}'_d(n_k) \quad (24)$$

where

$$\mathbf{a}(n_k) = \begin{bmatrix} a_1 e^{j\omega_0(n+k)} & a_1^* e^{-j\omega_0(n+k)} & \dots \\ a_L e^{jL\omega_0(n+k)} & a_L^* e^{-jL\omega_0(n+k)} \end{bmatrix}^T. \quad (25)$$

In this approach, there is no interference since the harmonic model enables us to use all information embedded in $\mathbf{x}(n_k)$ in the estimation of $x(n)$. Moreover, the desired signal $x(n)$ is the $(k+1)$ th entry of the vector $\mathbf{Z}(\omega_0)\mathbf{a}(n_k)$, i.e., we can write it as

$$x(n) = \mathbf{z}_{r,k}(\omega_0)\mathbf{a}(n_k) \quad (26)$$

where $\mathbf{z}_{r,k}(\omega_0)$ is the $(k+1)$ th row of $\mathbf{Z}(\omega_0)$. An estimate of the desired signal $x(n)$ can be obtained by inserting (24) into (13). This yields

$$\begin{aligned} \hat{x}_k(n) &= \mathbf{h}_k^T \mathbf{x}'_d(n_k) + \mathbf{h}_k^T \mathbf{v}(n_k) \\ &= x'_{fd,k}(n) + v_{rn,k}(n). \end{aligned} \quad (27)$$

where $x'_{fd,k}(n) = \mathbf{h}_k^T \mathbf{x}'_d(n_k)$ is the filtered desired periodic signal. The orthogonality between the desired signal and the noise enables us to write the variance of $\hat{x}_k(n)$ as

$$\sigma_{\hat{x}_k}^2 = \sigma_{x'_{fd,k}}^2 + \sigma_{v_{rn,k}}^2. \quad (28)$$

Since the desired signal is assumed periodic in this approach, the variance of the filtered signal can also be written as

$$\sigma_{x'_{fd,k}}^2 = \mathbf{h}_k^T \mathbf{R}_{\mathbf{x}'_d,k} \mathbf{h}_k^T$$

$$\approx \mathbf{h}_k^T \mathbf{Z}(\omega_0) \mathbf{P} \mathbf{Z}^H(\omega_0) \mathbf{h}_k \quad (29)$$

where

$$\mathbf{R}_{\mathbf{x}'_d,k} = \mathbb{E} [\mathbf{x}'_d(n_k) \mathbf{x}'_d(n_k)^T] \approx \mathbf{Z}(\omega_0) \mathbf{P} \mathbf{Z}^H(\omega_0) \quad (30)$$

is the covariance matrix of $\mathbf{x}'_d(n_k)$. We can obtain the following error function for harmonic decomposition approach:

$$e_k(n) = x'_{fd,k}(n) + v_{rn,k}(n) - \mathbf{z}_{r,k}(\omega_0)\mathbf{a}(n_k). \quad (31)$$

A filter which minimizes the effect of the noise, $v_{rn,k}(n)$, and the difference between $x'_{fd,k}(n)$ and $\mathbf{z}_{r,k}(\omega_0)\mathbf{a}(n_k)$ can then be designed by minimizing (31), possibly under some constraints.

IV. PERFORMANCE MEASURES

Recently, several performance measures for noise reduction tasks were proposed in [12] and [21]. In this section, we generalize these performance measures to encompass non-causal filters. Note that while the measures are here derived for the orthogonal decomposition approach, they can easily be derived for the harmonic decomposition approach by replacing $\sigma_{x_{fd,k}}^2$ by $\sigma_{x'_{fd,k}}^2$ and $\sigma_{x_{ri,k}}^2$ by 0.

A. Noise Reduction

A common measure of noise reduction is the signal-to-noise ratio (SNR). Here, we consider two SNRs, i.e., the input SNR (iSNR) and the output SNR (oSNR). The iSNR is the SNR of the observed signal before filtering

$$\text{iSNR} = \frac{\sigma_x^2}{\sigma_v^2} \quad (32)$$

with $\sigma_v^2 = \mathbb{E}[v^2(n)]$ being the variance of the noise. The oSNR is defined as the SNR after filtering. When using the orthogonal decomposition, it is therefore given by

$$\text{oSNR}(\mathbf{h}_k) = \frac{\sigma_{x_{fd,k}}^2}{\sigma_{x_{ri,k}}^2 + \sigma_{v_{rn,k}}^2}. \quad (33)$$

Another measure is the noise reduction factor $\xi_{nr}(\mathbf{h}_k)$. This measure is defined as the ratio between the noise before and after noise reduction. That is, we can write the factor as

$$\xi_{nr}(\mathbf{h}_k) = \frac{\sigma_v^2}{\sigma_{x_{ri,k}}^2 + \sigma_{v_{rn,k}}^2}. \quad (34)$$

The noise reduction factor is expected to be larger than or equal to 1.

B. Signal Distortion

In many noise reduction methods, the desired signal is distorted in the process of noise reduction. One measure which quantifies this distortion is the desired signal reduction factor. This factor is defined as the ratio between the variances of the

desired signal before and after filtering, respectively. The measure can also be written as

$$\xi_{\text{dsr}}(\mathbf{h}_k) = \frac{\sigma_x^2}{\sigma_{x_{\text{fd},k}}^2}. \quad (35)$$

If there is no distortion, the desired signal reduction factor will be 1. Otherwise, it will be different from 1. According to (20), this implies that we must require that

$$\mathbf{h}_k^T \boldsymbol{\rho}_{\mathbf{x},k} = 1 \quad (36)$$

if a filter should be distortionless. This knowledge can, for example, be applied in the filter design by using it as a constraint.

When the desired signal is periodic, we can also consider the harmonic distortion incurred by the filter. The harmonic distortion measure was proposed in [19]. This measure is defined as the sum of the absolute differences between the powers of the sinusoids before and after noise reduction, i.e.,

$$\begin{aligned} \xi_{\text{hd}}(\mathbf{h}_k) &= 2 \sum_l^L |P_l - P_{\text{f},k,l}| \\ &= 2 \sum_l^L P_l |1 - \mathbf{h}_k^T \mathbf{z}(l\omega_0) \mathbf{z}^H(l\omega_0) \mathbf{h}_k| \end{aligned} \quad (37)$$

where $P_{\text{f},k,l}$ is the power of the l th harmonic after filtering using \mathbf{h}_k . The harmonic distortion measure will be zero when the harmonics are not distorted. Otherwise, it will be larger than zero. Note that the harmonics might be distorted even though (36) is fulfilled.

V. OPTIMAL NON-CAUSAL FILTERS FOR NOISE REDUCTION

In this section, we re-derive some optimal orthogonal and harmonic decomposition based noise reduction filters to obtain non-causal filters. The corresponding causal filters were derived in [12], [17], [18], and [21]. Note that all filters derived here except the harmonic decomposition based linearly constrained minimum variance (HDLCMV) filter are based on the orthogonal decomposition.

A. Maximum SNR

The maximum SNR filter, $\mathbf{h}_{\text{max},k}$, is a filter which maximizes the output SNR with respect to the estimation of $x(n)$. The output SNR is defined in (33). If we insert (20)–(22) into (33), we can also write the output SNR for an orthogonal decomposition based filter as

$$\text{oSNR}(\mathbf{h}_k) = \frac{\mathbf{h}_k^T \mathbf{R}_{\mathbf{x}_{\text{d},k}} \mathbf{h}_k}{\mathbf{h}_k^T \mathbf{R}_{\mathbf{in},k} \mathbf{h}_k} \quad (38)$$

where

$$\mathbf{R}_{\mathbf{in},k} = \mathbf{R}_{\mathbf{x}_{\text{i},k}} + \mathbf{R}_{\mathbf{v}} \quad (39)$$

is the covariance matrix of the interference-plus-noise. The expression in (38) can also be recognized as a generalized Rayleigh quotient [30]. This quotient is maximized when the filter, \mathbf{h}_k , equals the eigenvector $\mathbf{u}_{\text{max},k}$ corresponding to the

largest eigenvalue, $\lambda_{\text{max},k}$, of $\mathbf{R}_{\mathbf{in},k}^{-1} \mathbf{R}_{\mathbf{x}_{\text{d},k}}$. Clearly, $\mathbf{R}_{\mathbf{in},k}^{-1} \mathbf{R}_{\mathbf{x}_{\text{d},k}}$ is rank one, i.e.,

$$\begin{aligned} \lambda_{\text{max},k} &= \text{Tr} \left(\mathbf{R}_{\mathbf{in},k}^{-1} \mathbf{R}_{\mathbf{x}_{\text{d},k}} \right) \\ &= \sigma_x^2 \boldsymbol{\rho}_{\mathbf{x},k}^T \mathbf{R}_{\mathbf{in},k}^{-1} \boldsymbol{\rho}_{\mathbf{x},k} \\ &= \text{oSNR}(\mathbf{h}_{\text{max},k}) \end{aligned} \quad (40)$$

with $\text{Tr}(\cdot)$ denoting the trace operator. An important observation from the above expression is that, in general, $\lambda_{\text{max},p} \neq \lambda_{\text{max},q}$ for $p \neq q$. That is, the output SNR may be different for different k s which means that we may be able to improve the oSNR by introducing non-causality in the filter design.

From (40), we can readily see that $\mathbf{u}_{\text{max},k}$ and thus $\mathbf{h}_{\text{max},k}$ are given by

$$\mathbf{h}_{\text{max},k} = \alpha_k \mathbf{R}_{\mathbf{in},k}^{-1} \boldsymbol{\rho}_{\mathbf{x},k} \quad (41)$$

where $\alpha_k \neq 0$ is some arbitrary scaling factor. As it will become clear soon, the only difference between the orthogonal decomposition based filters described in this paper, is the scaling factor α_k .

B. Wiener

In the orthogonal decomposition-based Wiener (ODW) filter design, the filter is designed by minimizing the MSE. The MSE criterion, $J(\mathbf{h}_k)$, can be written as

$$\begin{aligned} J(\mathbf{h}_k) &= \text{E} [e_k^2(n)] \\ &= \sigma_x^2 (\mathbf{h}_k^T \boldsymbol{\rho}_{\mathbf{x},k} - 1)^2 + \mathbf{h}_k^T \mathbf{R}_{\mathbf{in},k} \mathbf{h}_k. \end{aligned} \quad (42)$$

The minimizer of $J(\mathbf{h}_k)$ is found by differentiating with respect to \mathbf{h}_k and equating with zero. If we do this, we get the following expression for the non-causal orthogonal decomposition based Wiener filter

$$\mathbf{h}_{\text{ODW},k} = \sigma_x^2 \mathbf{R}_{\mathbf{y}}^{-1} \boldsymbol{\rho}_{\mathbf{x},k}. \quad (43)$$

If we note that $\mathbf{R}_{\mathbf{y}}$ can also be written as

$$\mathbf{R}_{\mathbf{y}} = \sigma_x^2 \boldsymbol{\rho}_{\mathbf{x},k} \boldsymbol{\rho}_{\mathbf{x},k}^T + \mathbf{R}_{\mathbf{in},k} \quad (44)$$

and if we apply the matrix inversion lemma on $\mathbf{R}_{\mathbf{y}}^{-1}$, we can obtain another expression:

$$\mathbf{h}_{\text{ODW},k} = \frac{\sigma_x^2}{1 + \lambda_{\text{max},k}} \mathbf{R}_{\mathbf{in},k}^{-1} \boldsymbol{\rho}_{\mathbf{x},k} \quad (45)$$

for the Wiener filter. It appears from this expression that the ODW filter indeed maximizes the output SNR since it is just a scaled version of the maximum SNR filter where the scaling factor, $\alpha_{\text{W},k}$, is given by

$$\alpha_{\text{ODW},k} = \frac{\sigma_x^2}{1 + \lambda_{\text{max},k}}. \quad (46)$$

The output SNR of the non-causal ODW filter is therefore given by

$$\text{oSNR}(\mathbf{h}_{\text{ODW},k}) = \text{oSNR}(\mathbf{h}_{\text{max},k}). \quad (47)$$

C. Minimum Variance Distortionless Response

The minimum variance distortionless response (MVDR) filter (a.k.a. the Capon filter) was proposed by Capon in the context of spatial filtering [31], [32]. Here, the MVDR filter, or orthogonal decomposition-based MVDR (ODMVDR) filter as we term it, is used for temporal filtering, and it is designed on basis of the orthogonal decomposition. The ODMVDR filter is designed such that it minimizes the variances of both the residual interference, $x_{ri,k}(n)$, and the residual noise, $v_{rn,k}(n)$. Moreover, the ODMVDR filter is designed to be distortionless with respect to the desired signal. Such a filter design can be obtained by solving the following quadratic minimization problem:

$$\min_{\mathbf{h}_k} \mathbf{h}_k^T \mathbf{R}_{in,k} \mathbf{h}_k \quad \text{s.t.} \quad \mathbf{h}_k^T \boldsymbol{\rho}_{xx,k} = 1. \quad (48)$$

The well-known solution of this optimization problem is given by

$$\mathbf{h}_{\text{ODMVDR},k} = \frac{\mathbf{R}_{in,k}^{-1} \boldsymbol{\rho}_{xx,k}}{\boldsymbol{\rho}_{xx,k}^T \mathbf{R}_{in,k}^{-1} \boldsymbol{\rho}_{xx,k}}. \quad (49)$$

It turns out that the ODMVDR filter can be equivalently expressed as

$$\mathbf{h}_{\text{ODMVDR},k} = \frac{\mathbf{R}_y^{-1} \boldsymbol{\rho}_{xx,k}}{\boldsymbol{\rho}_{xx,k}^T \mathbf{R}_y^{-1} \boldsymbol{\rho}_{xx,k}}. \quad (50)$$

Another important expression for the ODMVDR filter is given by

$$\begin{aligned} \mathbf{h}_{\text{ODMVDR},k} &= \frac{\sigma_x^2}{\lambda_{\max,k}} \mathbf{R}_{in,k}^{-1} \boldsymbol{\rho}_{xx,k} \\ &= \frac{1 + \lambda_{\max,k}}{\lambda_{\max,k}} \mathbf{h}_{W,k} \end{aligned} \quad (51)$$

from which it is clear that the ODMVDR filter maximizes the output SNR. The scaling factor, $\alpha_{\text{ODMVDR},k}$, is given by

$$\alpha_{\text{ODMVDR},k} = \frac{\sigma_x^2}{\lambda_{\max,k}}. \quad (52)$$

That is, by using the $\mathbf{h}_{\text{ODMVDR},k}$, we can have maximum output SNR while not distorting the desired signal, $x(n)$. Since the ODMVDR filter is just a scaled version of the maximum SNR filter, its output SNR is given by

$$\text{oSNR}(\mathbf{h}_{\text{ODMVDR},k}) = \text{oSNR}(\mathbf{h}_{\max,k}). \quad (53)$$

D. Harmonic LCMV

The last filter, described in this section, is the HDLCMV filter. This filter design is inspired by the LCMV beamformer (a.k.a. the Frost beamformer) proposed by Frost in the context of spatial filtering [33]. Here, we derive a non-causal HDLCMV filter for temporal filtering. The HDLCMV filter is designed to extract periodic signals modeled by (2), i.e., it is suited for extraction of signals such as voiced speech and musical instruments. The causal version of the HDLCMV filter was proposed in [17].

Since the non-causal HDLCMV filter is based on the harmonic decomposition, it utilizes all the information in $\mathbf{x}(n_k)$

to estimate $x(n)$. In the harmonic decomposition, there is no interference term as opposed to in the orthogonal decomposition where we have $\mathbf{x}_i(n_k)$. Therefore, in the harmonic decomposition-based filter design, we only have to minimize the residual noise power, $\sigma_{v_{rn,k}}^2$, without distorting the signal too much. The HDLCMV filter, in particular, is designed to minimize $\sigma_{v_{rn,k}}^2$ without distorting the harmonics of the desired periodic signal, $x(n)$. Such a filter can be obtained by solving the following optimization problem:

$$\min_{\mathbf{h}_k} \mathbf{h}_k^T \mathbf{R}_v \mathbf{h}_k \quad \text{s.t.} \quad \mathbf{Z}_k^H \mathbf{h}_k = \mathbf{1} \Leftrightarrow \mathbf{Z}^H \mathbf{h}_k = \mathbf{z}_{r,k}^H \quad (54)$$

where $\mathbf{1} = [1 \dots 1]^T$. It can readily be verified that the constraint in (54) makes the filter distortionless with respect to both the desired signal reduction factor and the harmonic distortion measure, respectively, by applying the covariance matrix model in (35) and (37).

The well-known solution to the multiple constrained quadratic optimization problem in (54) is given by

$$\mathbf{h}_{\text{HDLCMV},k} = \mathbf{R}_v^{-1} \mathbf{Z} (\mathbf{Z}^H \mathbf{R}_v^{-1} \mathbf{Z})^{-1} \mathbf{z}_{r,k}^H. \quad (55)$$

In [19], it was shown that we can replace \mathbf{R}_v by \mathbf{R}_y in the above expression without changing the filter response. If we do this, we get the following equivalent expression for the HDLCMV filter:

$$\mathbf{h}_{\text{HDLCMV},k} = \mathbf{R}_y^{-1} \mathbf{Z} (\mathbf{Z}^H \mathbf{R}_y^{-1} \mathbf{Z})^{-1} \mathbf{z}_{r,k}^H. \quad (56)$$

This expression is interesting since we can find the optimal HDLCMV filter without knowing the noise statistics which is often a requirement in noise reduction methods. On the other hand, we need to know the pitch, ω_0 , and the number of harmonics, L , of the desired signal, $x(n)$. When the HDLCMV filter is applied to a noise corrupted periodic signal, the output SNR can be found by replacing $\sigma_{x_{fd,k}}^2$ by $\sigma_{x'_{fd,k}}^2$ and $\sigma_{x_{ri,k}}^2$ by 0 in (33). If we do this, we get the following expression:

$$\begin{aligned} \text{oSNR}(\mathbf{h}_{\text{HDLCMV},k}) &\approx \frac{\mathbf{h}_{\text{HDLCMV},k}^T \mathbf{Z} \mathbf{P} \mathbf{Z}^H \mathbf{h}_{\text{HDLCMV},k}}{\mathbf{h}_{\text{HDLCMV},k}^T \mathbf{R}_v \mathbf{h}_{\text{HDLCMV},k}} \\ &= \frac{\sigma_x^2}{\mathbf{z}_{r,k}^H (\mathbf{Z}^H \mathbf{R}_v^{-1} \mathbf{Z})^{-1} \mathbf{z}_{r,k}^H}. \end{aligned} \quad (57)$$

VI. RECURSIVE FILTER UPDATES AND THE MAXIMUM OUTPUT SNR

We now show how the ODW and ODMVDR filters presented in the previous section can be updated recursively. As a by product of this result, we also show how the maximum output SNR can be updated recursively which, eventually, proves that the maximum output SNR always increases when the filter order M is increased. Here, we only provide the recursion for $k = 0$, but it can be generalized for all k s. Note that the derived recursive expressions also holds for $k = M - 1$ due to prediction symmetry.

From (43) and (50), it is clear that the ODW and ODMVDR filters both depend on $\mathbf{R}_y^{-1} \boldsymbol{\rho}_{xx}$ for $k = 0$ where $\boldsymbol{\rho}_{xx} = \boldsymbol{\rho}_{xx,0}$. Therefore, to find recursive filter and output SNR expressions,

TABLE I
EFFICIENT AND RECURSIVE COMPUTATION OF $\mathbf{R}_y^{-1}\boldsymbol{\rho}_{xx}$

$$\begin{array}{l}
 E_0 = r(0) \\
 \text{for } m = 1, \dots, M \\
 \quad \gamma_{m-1} = p(m-1) - \mathbf{b}_{m-1}^T \mathbf{p}_{m-1} \\
 \quad \kappa_m = \frac{1}{E_{m-1}} \left[r(m) - \mathbf{b}_{m-1}^T \mathbf{J}_{m-1} \mathbf{r}_{b,m-1} \right] \\
 \quad \mathbf{g}_m = \begin{bmatrix} \mathbf{g}_{m-1} \\ 0 \end{bmatrix} - \frac{\gamma_{m-1}}{\kappa_{m-1}} \begin{bmatrix} \mathbf{b}_{m-1} \\ -1 \end{bmatrix} \\
 \quad \mathbf{b}_m = \begin{bmatrix} 0 \\ \mathbf{b}_{m-1} \end{bmatrix} - \kappa_m \mathbf{J}_m \begin{bmatrix} \mathbf{b}_{m-1} \\ -1 \end{bmatrix} \\
 \quad E_m = E_{m-1} (1 - \kappa_m^2) \\
 \text{end}
 \end{array}$$

we derive a recursive expression for $\mathbf{R}_y^{-1}\boldsymbol{\rho}_{xx}$. To simplify the derivations of the recursion, we introduce a slightly different notation. First, we define the length m observed signal vector as

$$\begin{aligned}
 \mathbf{y}_m(n) &= [y(n) y(n-1) \cdots y(n-m+1)]^T \\
 &= [\mathbf{y}_{m-1}^T(n) y(n-m+1)]^T.
 \end{aligned} \quad (58)$$

Using the above expression, we can write the covariance matrix of the observed signal as

$$\begin{aligned}
 \mathbf{R}_m &= \mathbb{E} [\mathbf{y}_m(n) \mathbf{y}_m^T(n)] \\
 &= \begin{bmatrix} \mathbf{R}_{m-1} & \mathbf{r}_{b,m-1} \\ \mathbf{r}_{b,m-1}^T & r(0) \end{bmatrix}
 \end{aligned} \quad (59)$$

where

$$\mathbf{r}_{b,m-1} = [r(m-1) r(m-2) \cdots r(1)]^T \quad (60)$$

$$r(i) = \mathbb{E} [y(n) y(n-i)], \quad i = 0, 1, \dots, m-1. \quad (61)$$

Using the new notation, we can write the Wiener–Hopf equations as

$$\mathbf{R}_m \mathbf{g}_m = \mathbf{p}_m \quad (62)$$

where

$$\begin{aligned}
 \mathbf{p}_m &= \boldsymbol{\rho}_{xx} \\
 &= [p(0) p(1) \cdots p(m-1)]^T \\
 &= [\mathbf{p}_{m-1}^T p(m-1)].
 \end{aligned} \quad (63)$$

We know from backward linear prediction theory that

$$\mathbf{R}_{m-1} \mathbf{b}_{m-1} = \mathbf{r}_{b,m-1} \quad (64)$$

where \mathbf{b}_{m-1} is the length $(m-1)$ optimal linear backward predictor. Moreover, we know that

$$\mathbf{R}_m \begin{bmatrix} -\mathbf{b}_{m-1} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ E_{m-1} \end{bmatrix} \quad (65)$$

with E_{m-1} being the prediction error energy defined as

$$E_{m-1} = r(0) - \mathbf{r}_{b,m-1}^T \mathbf{b}_{m-1}. \quad (66)$$

Consider now the following expression:

$$\begin{aligned}
 \mathbf{R}_m \begin{bmatrix} \mathbf{g}_{m-1} \\ 0 \end{bmatrix} &= \begin{bmatrix} \mathbf{R}_{m-1} & \mathbf{r}_{b,m-1} \\ \mathbf{r}_{b,m-1}^T & r(0) \end{bmatrix} \begin{bmatrix} \mathbf{g}_{m-1} \\ 0 \end{bmatrix} \\
 &= \begin{bmatrix} \mathbf{R}_{m-1} \mathbf{g}_{m-1} \\ \mathbf{r}_{b,m-1}^T \mathbf{g}_{m-1} \end{bmatrix}.
 \end{aligned} \quad (67)$$

If we use (64) on in the above expression we immediately see that

$$\mathbf{r}_{b,m-1} \mathbf{g}_{m-1} = \mathbf{b}_{m-1}^T \mathbf{p}_{m-1}. \quad (68)$$

Then we subtract (67) from (62) which yields

$$\mathbf{R}_m \left(\mathbf{g}_m - \begin{bmatrix} \mathbf{g}_{m-1} \\ 0 \end{bmatrix} \right) = \begin{bmatrix} \mathbf{0} \\ \gamma_{m-1} \end{bmatrix} \quad (69)$$

where

$$\gamma_{m-1} = p(m-1) - \mathbf{b}_{m-1}^T \mathbf{p}_{m-1}. \quad (70)$$

If we multiply both sides of (65) with γ_{m-1}/E_{m-1} and compare it to (69), we can obtain that

$$\mathbf{g}_m = \begin{bmatrix} \mathbf{g}_{m-1} \\ 0 \end{bmatrix} - \frac{\gamma_{m-1}}{E_{m-1}} \begin{bmatrix} \mathbf{b}_{m-1} \\ -1 \end{bmatrix}. \quad (71)$$

That is, if we use the above expression in connection with the Levinson–Durbin algorithm, we can calculate $\mathbf{R}_y^{-1}\boldsymbol{\rho}_{xx}$ recursively. The resulting algorithm is depicted in Table I. Note that in Table I, the matrix $\mathbf{J}_m \in \mathbb{R}^{m \times m}$ is defined as

$$\mathbf{J}_m = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \quad (72)$$

and κ_m can be interpreted as the reflection coefficient.

We can now use the algorithm in Table I to recursively calculate the orthogonal decomposition based Wiener and MVDR filters for $k=0$ using the definitions in (43) and (51), respectively. By doing this, we get the following recursive expressions for the filters:

$$\begin{aligned}
 \mathbf{h}_{\text{ODW},0,m} &= \sigma_x^2 \mathbf{g}_m \\
 &= \sigma_x^2 \left[\begin{bmatrix} \mathbf{g}_{m-1} \\ 0 \end{bmatrix} - \frac{\gamma_{m-1}}{\kappa_{m-1}} \begin{bmatrix} \mathbf{b}_{m-1} \\ -1 \end{bmatrix} \right]
 \end{aligned} \quad (73)$$

$$\begin{aligned}
 \mathbf{h}_{\text{ODMVDR},0,m} &= \frac{\mathbf{g}_m}{\mathbf{p}_m^T \mathbf{g}_m} \\
 &= \frac{\left[\begin{bmatrix} \mathbf{g}_{m-1} \\ 0 \end{bmatrix} - \frac{\gamma_{m-1}}{\kappa_{m-1}} \begin{bmatrix} \mathbf{b}_{m-1} \\ -1 \end{bmatrix} \right]}{\mathbf{p}_{m-1}^T \mathbf{h}_{m-1} + \frac{\gamma_{m-1}^2}{E_{m-1}}}
 \end{aligned} \quad (74)$$

where the third subscript on the filters denotes the filter order. By calculating the Wiener and MVDR filters using the recursive procedure in Table I, we can reduce the computational complexity significantly compared to when the filters are calculated directly using (43) and (50), respectively [34]. Similar recursive

filter expressions can be found for the non-causal filters where k is between 0 and $M - 1$.

Moreover, we can use the recursive algorithm developed in this section, to find a recursive expression for the maximum output SNR when using orthogonal decomposition based filters. Again, we only derive the recursive expression for $k = 0$ (and thereby also for $k = M - 1$), but it can be generalized to different k s. First, we have to rewrite the expression for the maximum output SNR. It can be seen that the covariance matrix, $\mathbf{R}_{\mathbf{x}_{i,k}}$, of the interference vector, $\mathbf{x}_i(n_k)$, is given by

$$\mathbf{R}_{\mathbf{x}_{i,k}} = \mathbf{R}_{\mathbf{x}} - \sigma_x^2 \boldsymbol{\rho}_{\mathbf{x},k} \boldsymbol{\rho}_{\mathbf{x},k}^T. \quad (75)$$

If we then insert (75) into (39) which is then inserted into (40), and if we use the matrix inversion lemma, we can show that

$$\begin{aligned} \text{oSNR}_{\max,k} &= \sigma_x^2 \boldsymbol{\rho}_{\mathbf{x},k}^T (\mathbf{R}_{\mathbf{y}} - \sigma_x^2 \boldsymbol{\rho}_{\mathbf{x},k} \boldsymbol{\rho}_{\mathbf{x},k}^T)^{-1} \boldsymbol{\rho}_{\mathbf{x},k} \\ &= \frac{1}{\left(\sigma_x^2 \boldsymbol{\rho}_{\mathbf{x},k}^T \mathbf{R}_{\mathbf{y}}^{-1} \boldsymbol{\rho}_{\mathbf{x},k} \right)^{-1} - 1}. \end{aligned} \quad (76)$$

We now consider the case where $k = 0$. In this case, we can use the recursive expressions in Table I to write

$$\begin{aligned} \boldsymbol{\rho}_{\mathbf{x}}^T \mathbf{R}_{\mathbf{y}}^{-1} \boldsymbol{\rho}_{\mathbf{x}} &= \mathbf{p}_m^T \mathbf{g}_m \\ &= \mathbf{p}_m^T \left(\begin{bmatrix} \mathbf{g}_{m-1} \\ 0 \end{bmatrix} - \frac{\gamma_{m-1}}{\kappa_{m-1}} \begin{bmatrix} \mathbf{b}_{m-1} \\ -1 \end{bmatrix} \right) \\ &= \mathbf{p}_{m-1}^T \mathbf{g}_{m-1} + \frac{\gamma_{m-1}^2}{E_{m-1}}. \end{aligned} \quad (77)$$

If we substitute (77) back into (76) we get

$$\text{oSNR}_{\max,0,m} = \frac{1}{\left[\sigma_x^2 \left(\mathbf{p}_{m-1}^T \mathbf{g}_{m-1} + \frac{\gamma_{m-1}^2}{E_{m-1}} \right) \right]^{-1} - 1}. \quad (78)$$

From the above expression, we can readily see that the output SNR will always increase as we increase m when the desired signal is stationary.

VII. STUDY OF OUTPUT SNR AND DISTORTION

In this section, we investigate the performance of all the non-causal filters proposed in this paper when the desired signal is periodic. The assumption of periodicity enables us to exactly quantify the gains which can be obtained by introducing non-causality in the filters since we can then model the requisite statistics with closed-form expressions. First, we conduct a study where we measure the performance of all the non-causal filters as a function of k . Then, we investigate the asymptotic behavior of the maximum output SNR for different k s.

A. Filter Performances for Small M

We now investigate the performance of the non-causal ODW, ODMVDR, and HDLCMV filters in terms of output SNR and harmonic distortion when the filters are applied on periodic signals. First, we derive closed-form expressions for the performance measures under the assumption of periodicity. When the desired signal is periodic, we know that

$$\boldsymbol{\rho}_{\mathbf{x},k} = \frac{\mathbf{R}_{\mathbf{x}} \mathbf{i}_k}{\mathbf{i}_k^T \mathbf{R}_{\mathbf{x}} \mathbf{i}_k}$$

$$\approx \mathbf{Z} \mathbf{P} \mathbf{z}_{r,k}^H \sigma_x^2 \quad (79)$$

where $\mathbf{i}_k \in \mathbb{R}^{M \times 1}$ is a vector of zeros except at the k th entry which is a 1. If we insert (79) into (40), we see that the output SNR for the ODW and ODMVDR filters is

$$\begin{aligned} \text{oSNR}(\mathbf{h}_{\text{ODW},k}) &= \text{oSNR}(\mathbf{h}_{\text{ODMVDR},k}) \\ &\approx \frac{\mathbf{z}_{r,k} \mathbf{P} \mathbf{Z}^H \mathbf{R}_{\mathbf{in},k}^{-1} \mathbf{Z} \mathbf{P} \mathbf{z}_{r,k}^H}{\sigma_x^2} \end{aligned} \quad (80)$$

when they are applied on periodic signals. The output SNR for the HDLCMV filter on periodic signals are given in (57). To find expressions for the harmonic distortion of the ODW and ODMVDR filters, we need expression for the filters for periodic signals. These can be obtained by inserting (79) into (43) and (50) which yields

$$\mathbf{h}_{\text{ODW},k} \approx \mathbf{R}_{\mathbf{y}}^{-1} \mathbf{Z} \mathbf{P} \mathbf{z}_{r,k}^H \quad (81)$$

$$\mathbf{h}_{\text{ODMVDR},k} \approx \frac{\mathbf{R}_{\mathbf{y}}^{-1} \mathbf{Z} \mathbf{P} \mathbf{z}_{r,k}^H}{\mathbf{z}_{r,k} \mathbf{P} \mathbf{Z}^H \mathbf{R}_{\mathbf{y}}^{-1} \mathbf{Z} \mathbf{P} \mathbf{z}_{r,k}^H}. \quad (82)$$

We can then obtain closed-form expression for the harmonic distortion of the ODW and ODMVDR filters by inserting (81) and (82) into (37):

$$\xi_{\text{hd}}(\mathbf{h}_{\text{ODW}}) \approx 2 \sum_{l=1}^L P_l \left| 1 - \left| \mathbf{z}_{r,k} \mathbf{P} \mathbf{Z}^H \mathbf{R}_{\mathbf{y}}^{-1} \mathbf{z}(l\omega_0) \right|^2 \right| \quad (83)$$

$$\xi_{\text{hd}}(\mathbf{h}_{\text{ODMVDR}}) \approx 2 \sum_{l=1}^L P_l \left| 1 - \frac{\sigma_x^4 \left| \mathbf{z}_{r,k} \mathbf{P} \mathbf{Z}^H \mathbf{R}_{\mathbf{y}}^{-1} \mathbf{z}(l\omega_0) \right|^2}{\left(\mathbf{z}_{r,k} \mathbf{P} \mathbf{Z}^H \mathbf{R}_{\mathbf{y}}^{-1} \mathbf{Z} \mathbf{P} \mathbf{z}_{r,k}^H \right)^2} \right|. \quad (84)$$

The harmonic distortion for the HDLCMV filter will always be zero due to its constraints.

In the following, we have evaluated the performances of the ODW, ODMVDR, and HDLCMV filters in different scenarios. We evaluated the performances when the filters were applied for enhancement of a periodic signal, $x(n)$, in noise, $v(n)$. The periodic signal was constituted by $L = 6$ harmonic sinusoids with a pitch of $\omega_0 = 0.1578$. The amplitudes of the harmonics were chosen to be

$$[A_1 \cdots A_6] = [1 \ 0.8 \ 0.5 \ 0.35 \ 0.2 \ 0.1]. \quad (85)$$

By using decreasing amplitudes, we can get insight into how the filters perform with respect to noise reduction of, for example, voiced speech. First, we evaluated the performances when the desired signal was corrupted by white Gaussian noise at an input SNR of 10 dB, and when the filter order was $M = 30$. In Fig. 1 the results are shown for different values of k . From these results, it is clear that the output SNR can be improved significantly by changing k compared to the traditional approach where $k = 0$. For the ODW and ODMVDR filters, the output SNR can be improved by ≈ 3 dB by choosing $k = 12$, and for the HDLCMV filter an improvement of ≈ 4 dB is obtainable by choosing $k = 14$. It is important to note that we do not necessarily introduce additional harmonic distortion by improving

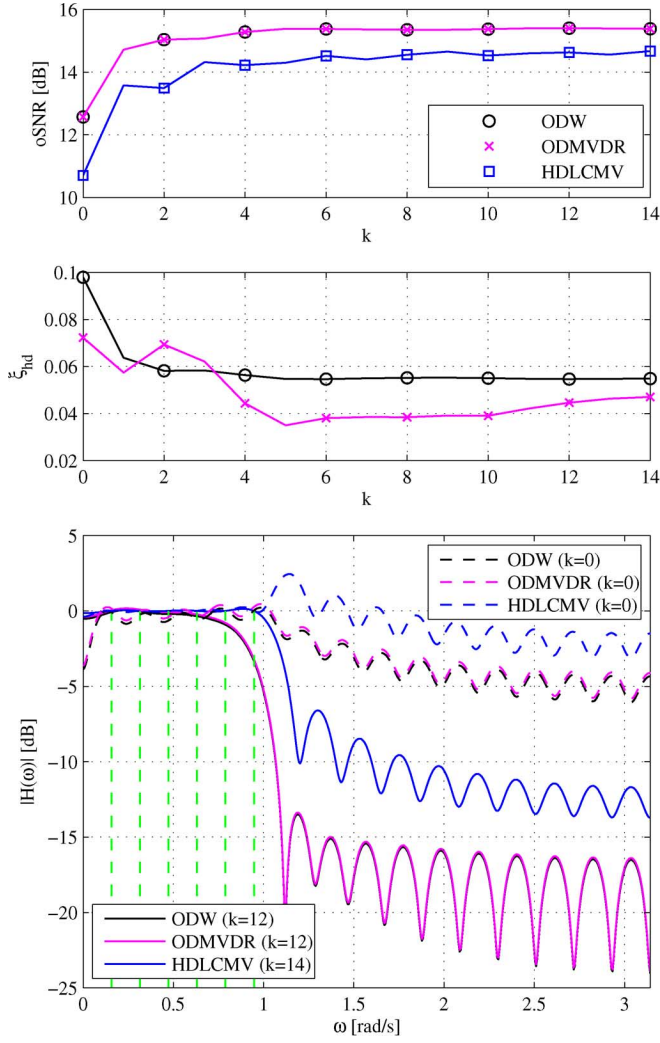


Fig. 1. Performance measures and filter responses of the ODW, ODMVDR, and HDLCMV filters for $M = 30$ when the noise is white Gaussian and the input SNR is 10 dB.

the output SNR by changing k . In this case, the harmonic distortion is also lowered compared to $k = 0$ for both the ODW and ODMVDR filters when the output SNR is maximized in k . Also, in Fig. 1, we have plotted the responses of the filters, for both $k = 0$ and for the k that maximizes the output SNR. It is clear from the filter responses, that the noise reduction can be improved significantly for $\omega > 1$ by choosing $k \neq 0$.

Then we conducted similar simulations, but with a filter order of $M = 50$. The results from these simulations are depicted in Fig. 2. Now, the gain that can be obtained by changing k is smaller. Compared to the case with $k = 0$, we can obtain a gain of ≈ 0.8 dB if we chose $k = 5$. Again, we can see that we can improve the output SNR and harmonic distortion simultaneously by changing k . We also plotted the filter responses. From these it is clear that we can obtain better noise reduction by choosing $k = 5$. This is especially so for high frequencies ($\omega > 1$).

We also conducted simulations where the noise was a sum of white Gaussian noise and sinusoidal noise. The sinusoidal noise source is used to investigate the impact of noise resembling voiced speech. In these simulations, the ratio between the

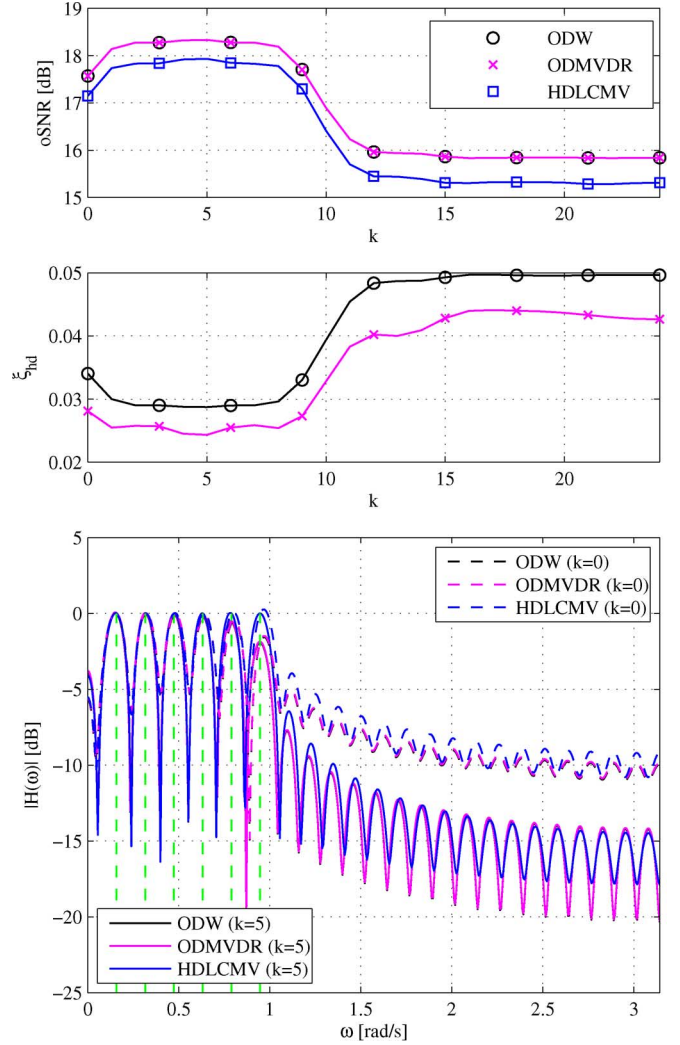


Fig. 2. Performance measures and filter responses of the ODW, ODMVDR, and HDLCMV filters for $M = 50$ when the noise is white Gaussian and the input SNR is 10 dB.

desired signal and the white Gaussian noise was 10 dB. The sinusoidal noise source was constituted by three harmonic sinusoids having a pitch of 0.1932. The amplitudes of the three harmonics of the sinusoidal noise source were [1 0.9 0.3]. The input SNR is therefore ≈ -0.09 dB in these simulations. Again, we conducted some simulations where the performances of the filters were evaluated for different k s. The results for a filter order of $M = 30$ are shown in Fig. 3. In this case, no improvement can be obtained for the ODW and ODMVDR filters compared to $k = 0$. For the HDLCMV filter a small improvement of ≈ 0.1 dB can be obtained by choosing $k = 6$ instead of $k = 0$. While the output SNR for the ODMVDR filter cannot be improved by changing k , its harmonic distortion can still be reduced. If we take a look on the filter responses, we can see that the HDLCMV filter for $k = 6$ provides significantly more noise reduction for $\omega > 1$ compared to when $k = 0$.

The simulations with sinusoidal noise were also conducted for a filter order of $M = 100$. The results from this simulation are depicted in Fig. 4. In these simulations, we see that the output SNR can be improved by ≈ 0.5 dB by changing k from 0 to 10 for the ODW and ODMVDR filters, and from 0 to 9 for

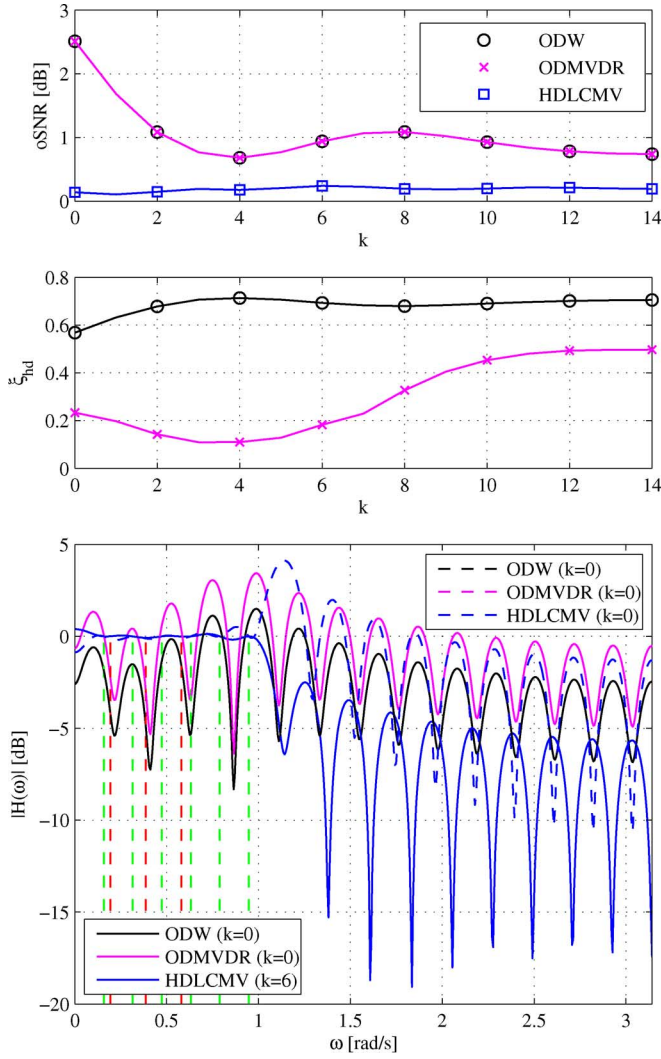


Fig. 3. Performance measures and filter responses of the ODW, ODMVDR, and HDLCMV filters for $M = 30$ when the noise is a sum of sinusoidal noise and white Gaussian. The ratio between the desired signal and the white noise is 10 dB and the input SNR is ≈ -0.09 dB.

the HDLCMV filter. We can see that the harmonic distortion of the ODW and ODMVDR filters are also improved by changing k . Again, it is clear from the frequency responses of the filters that we can obtain significantly more noise reduction for high frequencies ($\omega > 1$) by optimizing the output SNR over k .

From the results in Figs. 1–4, we can conclude that the k that maximizes the output SNR is dependent on the filter length, the noise, the fundamental frequency and the number of harmonics. To the extent of our knowledge, there is no simple expression for this optimal k and, in practice, it therefore has to be estimated by maximizing over the estimated output SNRs for all k s in $[0; \lfloor (M-1)/2 \rfloor]$ at every time instance.

B. Filter Performances for Large M

We now consider the performances of the filters when we let M approach infinity. Recall that the maximum output SNR for the orthogonal decomposition based filters can be written as

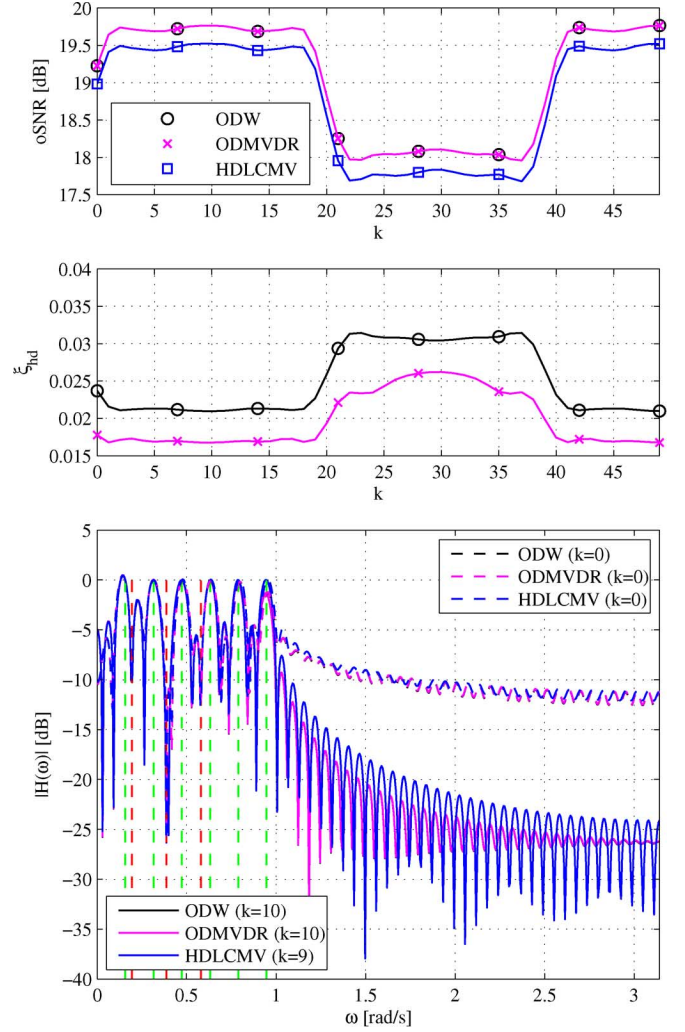


Fig. 4. Performance measures and filter responses of the ODW, ODMVDR, and HDLCMV filters for $M = 100$ when the noise is a sum of sinusoidal noise and white Gaussian. The ratio between the desired signal and the white noise is 10 dB and the input SNR is ≈ -0.09 dB.

$$\text{oSNR}_{\max,k} = \frac{1}{\left(\sigma_x^2 \rho_{xx,k}^T \mathbf{R}_y^{-1} \rho_{xx,k}\right)^{-1} - 1}. \quad (86)$$

Inserting (79) in (86) and applying (79) in the left-hand side of the denominator in (86) yields

$$\begin{aligned} \sigma_x^2 \rho_{xx,k}^T \mathbf{R}_y^{-1} \rho_{xx,k} &= \frac{\mathbf{z}_{r,k} \mathbf{P} \mathbf{Z}^H (\mathbf{Z} \mathbf{P} \mathbf{Z}^H + \mathbf{R}_v)^{-1} \mathbf{Z} \mathbf{P} \mathbf{z}_{r,k}^H}{\sigma_x^2} \\ &= \frac{\mathbf{z}_{r,k} \mathbf{P} \mathbf{C} \mathbf{P} \mathbf{z}_{r,k}^H}{\sigma_x^2} \end{aligned} \quad (87)$$

where

$$\mathbf{C} = \mathbf{Z}^H (\mathbf{Z} \mathbf{P} \mathbf{Z}^H + \mathbf{R}_v)^{-1} \mathbf{Z}. \quad (88)$$

When the noise is a summation of white Gaussian noise and sinusoidal interferers, it can be shown that [19]

$$\lim_{M \rightarrow \infty} \mathbf{C} = \mathbf{P}^{-1}. \quad (89)$$

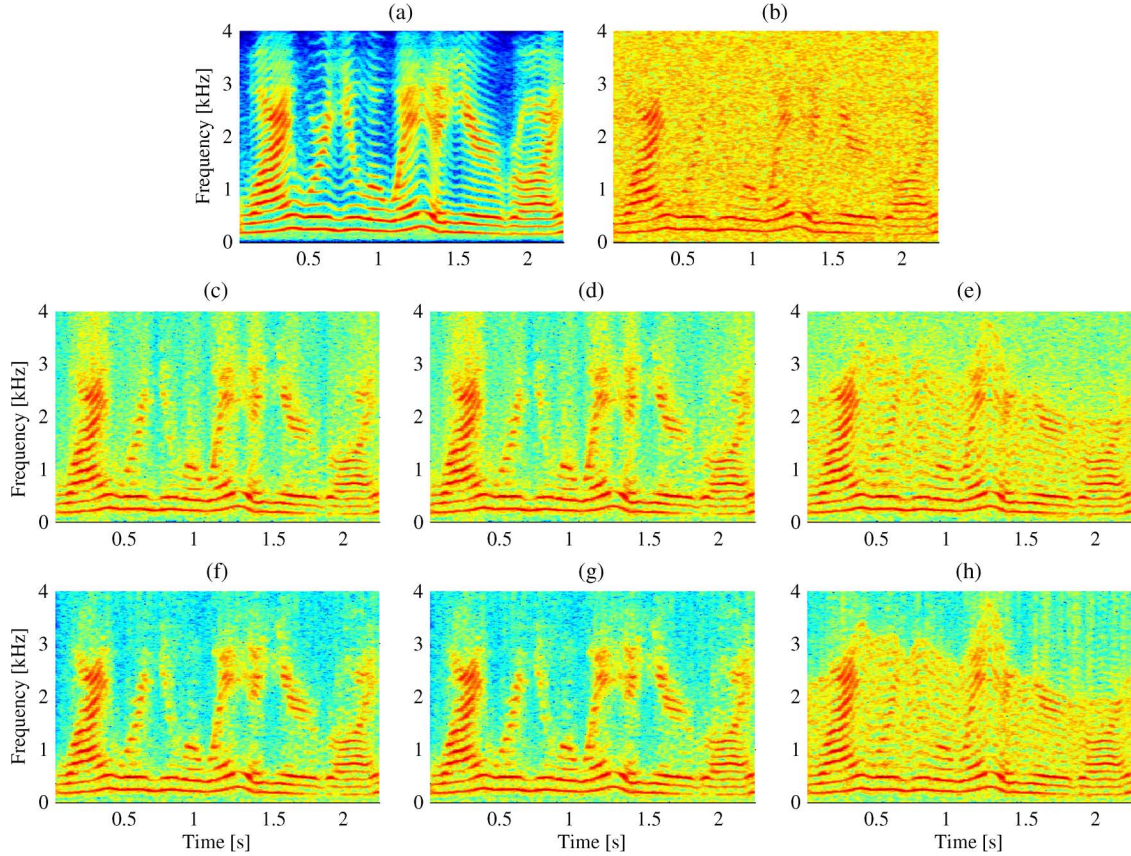


Fig. 5 Spectrograms of (a) a clean female speaker signal, (b) a female speaker signal in white noise at an iSNR of 5 dB, (c) an enhanced signal obtained using a causal ODW filter, (d) an enhanced signal obtained using a causal ODMVDR filter, (e) an enhanced signal obtained using a causal HDLCMV filter, (f) an enhanced signal obtained using a non-causal OD Wiener filtering scheme, (g) an enhanced signal obtained using a non-causal ODMVDR filtering scheme, and (h) an enhanced signal obtained using a non-causal HDLCMV filtering scheme. The enhancement filters were all of length $M = 60$

If we combine (86), (87), and (89) we can see that

$$\lim_{M \rightarrow \infty} \text{oSNR}_{\max,k} = \infty. \quad (90)$$

That is, when M becomes very large and the noise is a sum of white Gaussian noise and sinusoidal noise, the maximum output SNR of the orthogonal decomposition filter will approach ∞ for all values of k . The same will be the case for the HDLCMV filter since it equals the ODMVDR filter for large M [19].

In the following, we consider the asymptotic behavior of the harmonic distortion of the filters. Again, we assume that the desired signal is periodic. We know that the HDLCMV filter always has no harmonic distortion due to its constraints, so this filter is not considered in this investigation. The expression for the ODW and ODMVDR filters when the desired signal is periodic are given in (81) and (82). If we then let M approach infinity we can see that

$$\begin{aligned} \lim_{M \rightarrow \infty} \mathbf{h}_{\text{ODMVDR},k} &= \lim_{M \rightarrow \infty} \mathbf{h}_{\text{HDLCMV},k} = \mathbf{R}_y^{-1} \mathbf{Z} \mathbf{P} \mathbf{z}_{r,k}^H \\ &= \mathbf{h}_{\text{ODW},k}. \end{aligned} \quad (91)$$

It can now be seen that the harmonic distortion of the ODW and ODMVDR filters approaches zero when M is increased.

This can be seen by inserting (91) into (37), and by letting M approach infinity, which yields

$$\begin{aligned} \lim_{M \rightarrow \infty} \xi_{\text{hd}}(\mathbf{h}_{\text{ODW},k}) &= \lim_{M \rightarrow \infty} \xi_{\text{hd}}(\mathbf{h}_{\text{ODMVDR},k}) \\ &= \xi_{\text{hd}}(\mathbf{h}_{\text{HDLCMV},k}) = 0. \end{aligned} \quad (92)$$

In summary, all filters show the same asymptotic performances for all k s both with respect to noise reduction and distortion. This motivates using the orthogonal and harmonic decomposition based filters jointly as considered in [19], [20] for $k = 0$ since they have complementary advantages and disadvantages. However, this will not be treated in this paper.

VIII. EXAMPLE: NOISE REDUCTION OF SPEECH

In the following, we demonstrate the applicability of the proposed non-causal filters on real-life signals. In particular, we consider noise reduction of speech recordings. First, we considered a 2.2 seconds long speech segment sampled at 8 kHz. The segment contains a female speaker uttering the sentence “Why where you away a year Roy?”. In Fig. 5(a), the spectrogram of the clean speech signal is plotted. As it can be seen from this spectrogram, the speech signal used in the first experiment on real-life speech contains voiced speech only. This was chosen to allow for the evaluation of the HDLCMV filter which is only applicable on (quasi-)periodic signals. We added white Gaussian

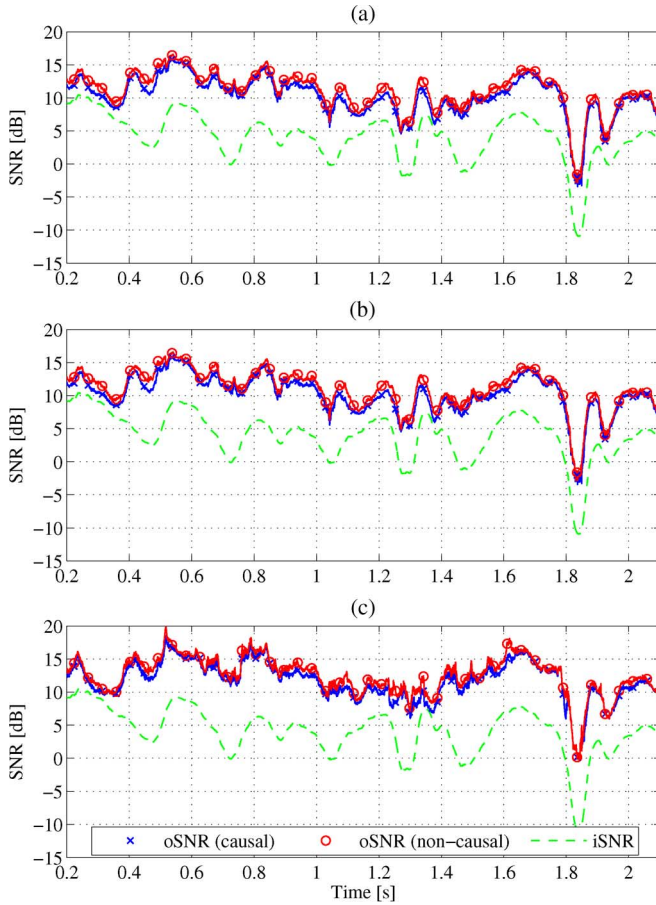


Fig. 6. Estimated output SNRs over time for causal and non-causal (a) ODW, (b) ODMVDR, and (c) HDLCMV filtering schemes.

noise to the speech signal at an average input SNR of 5 dB, and the spectrogram of the noisy signal is depicted in Fig. 5(b). The noisy signal was then enhanced using different causal and non-causal filtering schemes; we used causal and non-causal ODW, ODMVDR, and HDLCMV filtering schemes involving filters of length $M = 60$. In the filtering schemes, we used (43), (51) and (56) for different k ; for the causal filtering schemes, k was set to 0 whereas, for the non-causal filtering schemes, k was chosen such that the estimated output SNRs of the filters were maximized at every time instance. Note that the applied filters require that the noise and/or signal statistics are known or estimated in practice which justifies that we require knowledge about the output SNRs. In all filtering schemes, we recalculated the filters and their output SNRs at every time instance, n , using the estimated observed signal, desired signal and noise statistics ($\hat{\mathbf{R}}_y$, $\hat{\mathbf{R}}_x$, and $\hat{\mathbf{R}}_v$). The statistics were estimated from the previous $N = 400$ samples of the observed signal, desired signal and noise, respectively. We focus on comparing the performance of causal and non-causal filters in this section, so we assume that the noise signal is always available. In practice, we can estimate the noise statistics during silences by using a voice activity detector (VAD) if the noise is stationary, or we can estimate the noise statistics even in periods with voice activity using, e.g., [14], [35]. The ODW and ODMVDR filters were calculated using $\hat{\mathbf{R}}_y$ and $\hat{\mathbf{R}}_v$, whereas the HDLCMV filter

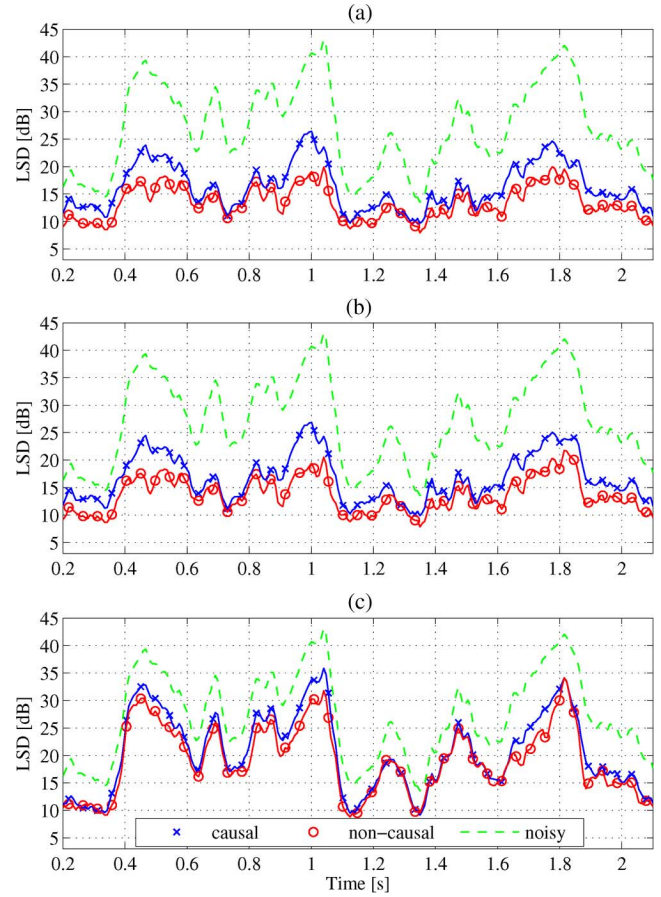


Fig. 7. Estimated log-spectral distances over time for causal and non-causal (a) ODW, (b) ODMVDR, and (c) HDLCMV filtering schemes.

was calculated using $\hat{\mathbf{R}}_y$, the pitch estimated at every time instance, and a fixed harmonic model order of $L = 13$. We estimated the pitch using the orthogonality based subspace method in [17], [18] which is freely available online.¹ The model order, on the other hand, was chosen on basis of an inspection of the spectrogram in Fig. 5(a). Furthermore, in the calculations of the HDLCMV filter, we regularized the covariance matrix of the observed signal as in [36]

$$\hat{\mathbf{R}}_{y,\text{reg}} = (1 - \gamma)\hat{\mathbf{R}}_y + \gamma \frac{\text{Tr}\{\hat{\mathbf{R}}_y\}}{M} \mathbf{I}. \quad (93)$$

The regularization is necessary due to estimation errors on the signal statistics and mismatch between the assumed harmonic model and the speech signal. We experienced that $\gamma = 0.7$ gives consistently good results in terms of oSNR and perceptual scores.

The spectrograms of the resulting enhanced signals obtained using the described simulation setup are depicted in Fig. 5(c)–(h). It is clearly indicated by these spectrograms that the non-causal filtering schemes reduce the noise more than their causal counterparts. At the same time, the non-causal filtering schemes do not introduce additional distortion of the desired signal compared to the causal filters. To support the rather subjective observations on the noise reduction performances, we also estimated the output SNRs for both the

¹<http://www.morganclaypool.com/page/multi-pitch>.

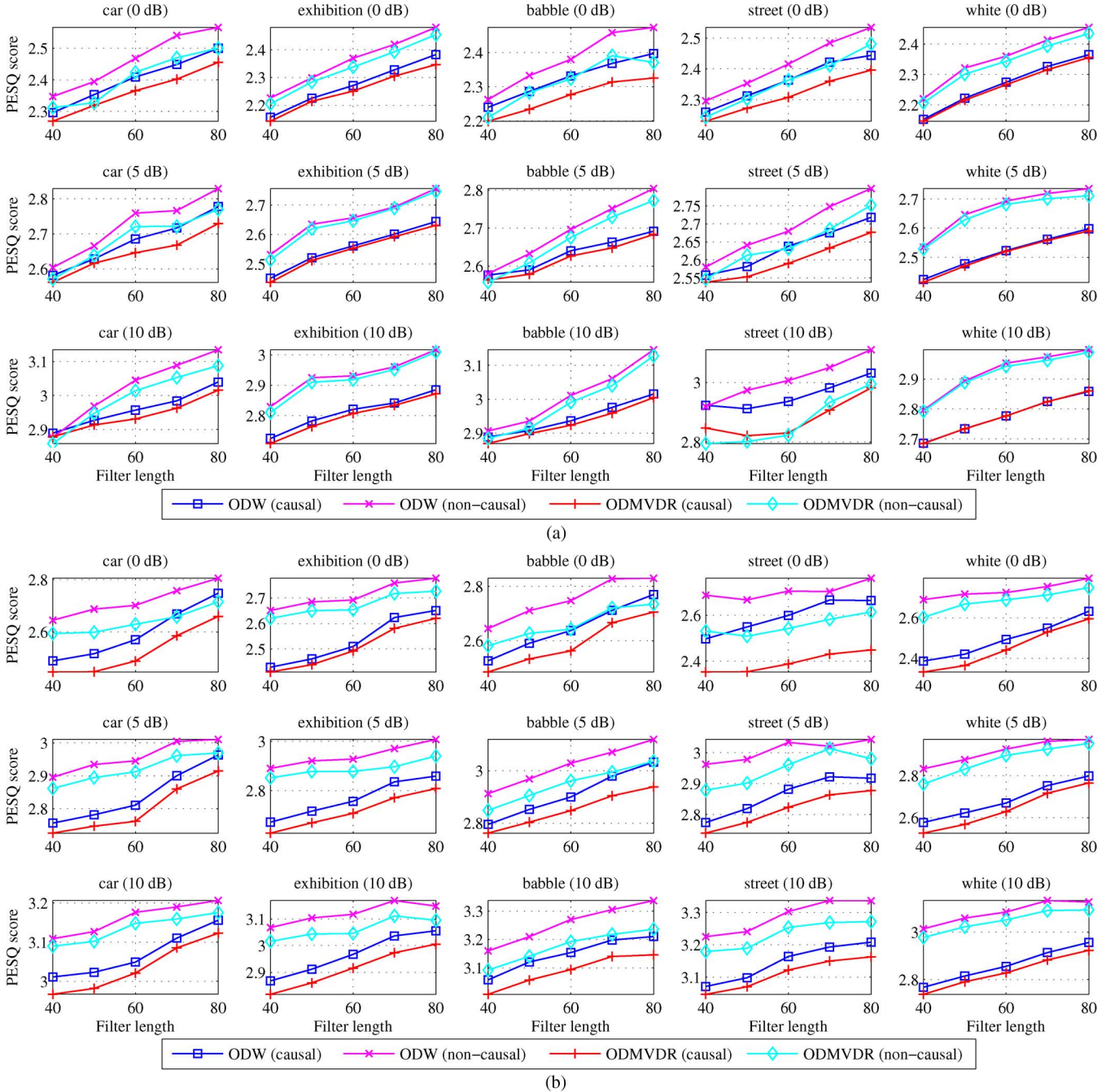


Fig. 8. PESQ scores for noisy (a) female and (b) male speech signals enhanced using the causal ODW filter, the non-causal maximum SNR ODW filtering scheme, the causal ODMVDR filter, and the non-causal maximum SNR ODMVDR filtering scheme. The scores were measured in different noise scenarios for different filter lengths.

non-causal filtering schemes and the causal filters at each time instance. The estimated output SNRs are depicted in Fig. 6. As expected, the non-causal filtering schemes has higher output SNRs at every time instance compared to the causal filters. This is expected, since we maximize the output SNR at every time instance in the non-causal filtering schemes. It seems from the results in Fig. 6 that the HDLCMV filter outperforms both the ODW and ODMVDR filters in terms of output SNR. This cannot be concluded, however, since the output SNRs for the orthogonal and harmonic decomposition are defined differently. In practice, the orthogonal decomposition based filters actually show superior noise reduction performances compared

to the HDLCMV filter according to our listening experience. We also measured the log-spectral distance (LSD) between the clean signal and the enhanced signals over time, and the results are depicted in Fig. 7. First of all, we can see from these results that the enhanced signals obtained using the non-causal filtering schemes have lower LSDs compared to the enhanced signals obtained using the causal filters at almost every time instance. That is, these results indicate that the non-causal filtering schemes have better distortion properties compared to the causal filters. Moreover, we can see from the results in Fig. 7 that both the ODW and ODMVDR filters outperform the HDLCMV filter in terms of LSDs. This supports our previous

claim on that, in practice, the ODW and ODMVDR filters introduce less distortion of the desired signal compared to the HDLCMV filter.

The results from the previous simulations indicate that we can achieve better noise reduction and distortion performances by using non-causal filtering schemes instead of non-causal filters. However, these results do not necessarily reflect the achievable perceptual improvement in performance. Therefore, we conducted another real-life experiment on speech where we considered enhancement of female (sp12.wav) and male (sp02.wav) speech signals in noise. The speech signals are part of the NOIZEUS speech corpus [37]. Note that since the utilized speech signals contain segments of unvoiced speech, we only evaluate the ODW and ODMVDR filters in this experiment. Using the ODW and ODMVDR non-causal filtering schemes and causal filters considered in the previous experiment, we conducted simulations where we enhanced the female and male speech signals in different noise scenarios, for different input SNRs, and for different filter lengths. The necessary statistics for this experiment were estimated as in the previous simulations. In each simulation, we measured the “Perceptual Evaluation of Speech Quality” (PESQ) scores [38] of the different enhanced signals compared to the clean speech signals using a freely available online toolbox.² The PESQ score is an objective measure which reflects the perceptual quality of a speech signal. That is, we can use PESQ scores to evaluate the practical applicability of the non-causal filtering schemes versus causal filters. The PESQ scores resulting from the simulations involving the female and male speech signals are shown in Fig. 8. First, we observe from these results that the ODW non-causal filtering schemes and causal filters outperform the ODMVDR non-causal filtering schemes and causal filters, respectively. Moreover, we observe that, in the simulations with female speech, the non-causal filtering schemes outperform their causal counterparts in almost all scenarios. Only for some noise types (street and car), for a 10-dB iSNR, and for small filter lengths, the causal filters get similar or a slightly better PESQ scores compared to the non-causal filtering schemes. Finally, we observe that the non-causal OD Wiener and ODMVDR filtering schemes outperforms their causal versions in all of the considered scenarios with male speech. Furthermore, by listening to the enhanced signals, it is our experience that the non-causal filtering schemes indeed outperforms the corresponding causal filter versions in terms of noise reduction in most scenarios. The enhanced signals used in our informal listening test can be found at the demo website³ for the paper.

IX. CONCLUSION

In this paper, we proposed novel non-causal time-domain filters for noise reduction in, e.g., speech applications. The proposed filters are based on the orthogonal and harmonic signal decompositions. To enable the design of non-causal filters from these decompositions, we generalized the decompositions. We

also proposed performance measures for evaluating non-causal time-domain filters based on the generalized decompositions. On a side note, we showed how the non-causal orthogonal decomposition based filters can be updated recursively when the filter order is increased. This was shown for the two particular cases where the filters are either causal or anti-causal. A by-product of these recursive updates is that we can also show how the output SNR is updated recursively which, eventually, proves that the output SNR is always increased when we increase the filter length and the desired signal is stationary. We also conducted theoretical evaluations of the filters. In these evaluations, we assumed that the desired signal is periodic and thereby has a harmonic structure. By making this assumption, it is possible to obtain exact closed-form expressions for the performance measures of the filters. The theoretical evaluations showed that we can indeed improve both the output SNR and the harmonic distortion of the filters simultaneously by allowing the filters to be non-causal. Moreover, we applied the non-causal filters for noise reduction of noisy real-life speech signals. These simulations showed that the non-causal filters can achieve more noise reduction compared to the causal filters in practice in terms of output SNR, log-spectral distance and PESQ scores.

REFERENCES

- [1] J. Benesty, S. Makino, and J. Chen, *Speech Enhancement*, ser. Signals and Communication Technology. New York: Springer, 2005.
- [2] P. Loizou, *Speech Enhancement: Theory and Practice*. Boca Raton, FL: CRC, 2007.
- [3] S. Boll, “Suppression of acoustic noise in speech using spectral subtraction,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-27, no. 2, pp. 113–120, Apr. 1979.
- [4] R. McAulay and M. Malpass, “Speech enhancement using a soft-decision noise suppression filter,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-28, no. 2, pp. 137–145, Apr. 1980.
- [5] Y. Ephraim and D. Malah, “Speech enhancement using a minimum mean-square error log-spectral amplitude estimator,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, no. 2, pp. 443–445, Apr. 1985.
- [6] K. V. Sørensen and S. V. Andersen, “Rayleigh mixture model-based hidden markov modeling and estimation of noise in noisy speech signals,” *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 3, pp. 901–917, Mar. 2007.
- [7] S. Srinivasan, J. Samuelsson, and W. B. Kleijn, “Codebook-based bayesian speech enhancement for nonstationary environments,” *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 2, pp. 441–452, Feb. 2007.
- [8] M. Dendrinos, S. Bakamidis, and G. Carayannis, “Speech enhancement from noise: A regenerative approach,” *Speech Commun.*, vol. 10, no. 1, pp. 45–57, 1991.
- [9] Y. Ephraim and H. L. Van Trees, “A signal subspace approach for speech enhancement,” *IEEE Trans. Speech Audio Process.*, vol. 3, no. 4, pp. 251–266, Jul. 1995.
- [10] S. H. Jensen, P. C. Hansen, S. D. Hansen, and J. A. Sørensen, “Reduction of broad-band noise in speech by truncated QSVD,” *IEEE Trans. Speech Audio Process.*, vol. 3, no. 6, pp. 439–448, Nov. 1995.
- [11] 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. Adv. Signal Process.*, vol. 2007, p. 24, 2007.
- [12] J. Benesty and J. Chen, *Optimal Time-Domain Noise Reduction Filters—A Theoretical Study*, ser. SpringerBriefs in Electrical and Computer Engineering, 1st ed. New York: Springer, 2011, vol. VII.
- [13] J. S. Lim, Ed., *Speech Enhancement*. Englewood Cliffs, NJ: Prentice-Hall, 1983.
- [14] R. Martin, “Noise power spectral density estimation based on optimal smoothing and minimum statistics,” *IEEE Trans. Speech Audio Process.*, vol. 9, no. 5, pp. 504–512, Jul. 2001.

²<http://www.utdallas.edu/~loizou/speech/software.htm>.

³http://kom.aau.dk/~jrj/Demo/non_causal_filt/demo.html

- [15] J. Chen, J. Benesty, and Y. Huang, "Study of the noise-reduction problem in the Karhunen-Loève expansion domain," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 17, no. 4, pp. 787–802, May 2009.
- [16] J. Benesty, J. Chen, and Y. Huang, "Speech enhancement in the Karhunen-Loève expansion domain," *Synth. Lectures Speech Audio Process.*, vol. 7, no. 1, pp. 1–112, 2011.
- [17] M. G. Christensen, P. Stoica, A. Jakobsson, and S. H. Jensen, "Multi-pitch estimation," *Elsevier Signal Process.*, vol. 88, no. 4, pp. 972–983, 2008.
- [18] M. G. Christensen and A. Jakobsson, "Multi-pitch estimation," *Synth. Lectures Speech Audio Process.*, vol. 5, no. 1, pp. 1–160, 2009.
- [19] J. R. Jensen, J. Benesty, M. G. Christensen, and S. H. Jensen, "Enhancement of single-channel periodic signals in the time-domain," 2011, unpublished manuscript.
- [20] J. R. Jensen, J. Benesty, M. G. Christensen, and S. H. Jensen, "Joint filtering scheme for nonstationary noise reduction," 2012, unpublished manuscript.
- [21] J. Chen, J. Benesty, Y. Huang, and T. Gaensler, "On single-channel noise reduction in the time domain," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2011, pp. 277–280.
- [22] H. Li, P. Stoica, and J. Li, "Computationally efficient parameter estimation for harmonic sinusoidal signals," *Elsevier Signal Process.*, vol. 80, no. 9, pp. 1937–1944, 2000.
- [23] K. W. Chan and H. C. So, "Accurate frequency estimation for real harmonic sinusoids," *IEEE Signal Process. Lett.*, vol. 11, no. 7, pp. 609–612, Jul. 2004.
- [24] A. de Cheveigné and H. Kawahara, "YIN, a fundamental frequency estimator for speech and music," *J. Acoust. Soc. Amer.*, vol. 111, no. 4, pp. 1917–1930, 2002.
- [25] V. Emiya, B. David, and R. Badeau, "A parametric method for pitch estimation of piano tones," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2007, vol. 1, pp. 249–252.
- [26] S. Godsill and M. Davy, "Bayesian harmonic models for musical pitch estimation and analysis," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 13–17, 2002, vol. 2, pp. 1769–1772.
- [27] P. Stoica and Y. Selen, "Model-order selection: A review of information criterion rules," *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 36–47, Jul. 2004.
- [28] M. G. Christensen, A. Jakobsson, and S. H. Jensen, "Joint high-resolution fundamental frequency and order estimation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 5, pp. 1635–1644, Jul. 2007.
- [29] P. Stoica and R. Moses, *Spectral Analysis of Signals*. Upper Saddle River, NJ: Pearson Education, Inc., 2005.
- [30] J. N. Franklin, *Matrix Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1968.
- [31] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969.
- [32] J. Capon, "Maximum-Likelihood Spectral Estimation," in *Nonlinear Methods of Spectral Analysis*. New York: Springer-Verlag, 1983.
- [33] O. L. Frost, III, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, no. 8, pp. 926–935, Aug. 1972.
- [34] C. Paleologu, J. Benesty, and S. Ciochina, "Sparse adaptive filters for echo cancellation," *Synth. Lectures Speech Audio Process.*, vol. 6, no. 1, pp. 1–124, 2010.
- [35] D. Ealey, H. Kelleher, and D. Pearce, "Harmonic tunnelling: Tracking non-stationary noises during speech," in *Proc. Eurospeech*, 2001.
- [36] F. van der Heijden, R. P. W. Duin, D. de Ridder, and D. M. J. Tax, *Classification, Parameter Estimation and State Estimation—An Engineering Approach Using MATLAB*. New York: Wiley, 2004.
- [37] Y. Hu and P. C. Loizou, "Subjective comparison and evaluation of speech enhancement algorithms," *Speech Commun.*, vol. 49, no. 7–8, pp. 588–601, 2007.
- [38] *Perceptual Evaluation of Speech Quality (PESQ): An Objective Method for End-to-End Speech Quality Assessment of Narrow-Band Telephone Networks and Speech Codecs*, TU-T Rec. P.862, 02/2001.



Jesper Rindom Jensen was born in Ringkøbing, Denmark, in August 1984. He received the B.Sc. degree from Aalborg University, Aalborg, Denmark, in 2007. In the fall 2007, he was enrolled in the elite candidate program in wireless communications at Aalborg University, and received the M.Sc. degree (*cum laude*) for completing the elite candidate education at Aalborg University in 2009. He is currently pursuing the Ph.D. degree in the Department of Electronic Systems, Aalborg University.

He has been a Visiting Researcher at University of Quebec, INRS-EMT, Montreal, QC, Canada. He has published several conference papers and a paper in the IEEE SIGNAL PROCESSING LETTERS on the topics of parameter estimation and coding of single- and multichannel audio, and speech signals. Among others, his research interests are digital signal processing theory, parametric analysis, modeling and coding of audio and speech signals, and microphone array processing.



Jacob Benesty was born in 1963. He received the M.S. degree in microwaves from Pierre and Marie Curie University, Paris, France, in 1987, and the Ph.D. degree in control and signal processing from Orsay University, France, in April 1991.

During the Ph.D. degree (from November 1989 to April 1991), he worked on adaptive filters and fast algorithms at the Centre National d'Etudes des Télécommunications (CNET), Paris. From January 1994 to July 1995, he worked at Telecom Paris University on multichannel adaptive filters and acoustic echo cancellation. From October 1995 to May 2003, he was first a Consultant and then a Member of the Technical Staff at Bell Laboratories, Murray Hill, NJ. In May 2003, he joined the University of Quebec, INRS-EMT, Montreal, QC, Canada, as a Professor. His research interests are in signal processing, acoustic signal processing, and multimedia communications. He is the inventor of many important technologies. In particular, he was the Lead Researcher at Bell Labs who conceived and designed the world-first real-time hands-free full-duplex stereo-phonetic teleconferencing system. Also, he and Gaensler conceived and designed the world-first PC-based multiparty hands-free full-duplex stereo conferencing system over IP networks. He is the editor of the book series *Springer Topics in Signal Processing* (Springer, 2009).

Dr. Benesty was the cochair of the 1999 International Workshop on Acoustic Echo and Noise Control and the general cochair of the 2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. He was a member of the IEEE Signal Processing Society Technical Committee on Audio and Electroacoustics and a member of the editorial board of the *EURASIP Journal on Applied Signal Processing*. He is the recipient, with Morgan and Sondhi, of the IEEE Signal Processing Society 2001 Best Paper Award. He is the recipient, with Chen, Huang, and Doclo, of the IEEE Signal Processing Society 2008 Best Paper Award. He is also the coauthor of a paper for which Huang received the IEEE Signal Processing Society 2002 Young Author Best Paper Award. In 2010, he received the "Gheorghe Cartianu Award" from the Romanian Academy. In 2011, he received the Best Paper Award from IEEE WASPAA for a paper that he published with Chen. He has coauthored and coedited/coauthored many books in the area of acoustic signal processing. He is also the lead editor-in-chief of the reference *Springer Handbook of Speech Processing* (Berlin: Springer-Verlag, 2007).



Mads Græsbøll Christensen (S'00–M'05–SM'11) was born in Copenhagen, Denmark, in March 1977. He received the M.Sc. and Ph.D. degrees from Aalborg University, Aalborg, Denmark, in 2002 and 2005, respectively.

He was formerly with the Department of Electronic Systems, Aalborg University, and is currently an Associate Professor in the Department of Architecture, Design and Media Technology. He has been a Visiting Researcher at Philips Research Labs, Ecole Nationale Supérieure des Télécommunications (ENST), University of California, Santa Barbara (UCSB), and Columbia University. He has published about 100 papers in peer-reviewed conference proceedings and journals and is coauthor (with A. Jakobsson) of the book *Multi-Pitch Estimation* (Morgan & Claypool, 2009). His research interests

include digital signal processing theory and methods with application to speech and audio, in particular parametric analysis, modeling, and coding.

Dr. Christensen has received several awards, namely an IEEE International Conference on Acoustics, Speech, and Signal Processing Student Paper Contest Award, the Spar Nord Foundation's Research Prize for his Ph.D. dissertation, and a Danish Independent Research Councils Young Researcher's Award. He is an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS.



Søren Holdt Jensen (S'87–M'88–SM'00) received the M.Sc. degree in electrical engineering from Aalborg University, Aalborg, Denmark, in 1988, and the Ph.D. degree in signal processing from the Technical University of Denmark, Lyngby, Denmark, in 1995.

Before joining the Department of Electronic Systems of Aalborg University, he was with the Telecommunications Laboratory of Telecom Denmark, Ltd, Copenhagen, Denmark; the Electronics Institute of the Technical University of Denmark; the Scientific Computing Group of Danish Computing Center for

Research and Education (UNI•C), Lyngby; the Electrical Engineering Department of Katholieke Universiteit Leuven, Leuven, Belgium; and the Center for PersonKommunikation (CPK) of Aalborg University. He is Full Professor and is currently heading a research team working in the area of numerical algorithms and signal processing for speech and audio processing, image and video processing, multimedia technologies, and digital communications.

Prof. Jensen was an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and *Elsevier Signal Processing*, and is currently Member of the Editorial Board of *EURASIP Journal on Advances in Signal Processing*. He is a recipient of an European Community Marie Curie Fellowship, former Chairman of the IEEE Denmark Section, and Founder and Chairman of the IEEE Denmark Section's Signal Processing Chapter. He is member of the Danish Academy of Technical Sciences and was in January 2011 appointed as member of the Danish Council for Independent Research-Technology and Production Sciences by the Danish Minister for Science, Technology and Innovation.