REPRESENTATION AND IDENTIFICATION OF SYSTEMS IN THE WAVELET TRANSFORM DOMAIN

Yekutiel Avargel and Israel Cohen *
Department of Electrical Engineering, Technion - Israel Institute of Technology, Technion City, Haifa 32000, Israel
{kutiav@tx,icohen@ee}.technion.ac.il

ABSTRACT
In this paper, we introduce an explicit representation of linear time-invariant system in the discrete-time wavelet transform (DTWT) domain. It is shown that crossband filters between subbands are required for perfect representation of the system. These filters depend on the DTWT parameters and on the system impulse response, and are shown to be time-varying. An approximate representation based on band-to-band filters without crossband filters is employed for system identification in the wavelet domain. We show that for longer and stronger input signals, longer band-to-band filters may be estimated. Experimental results validate the theoretical analysis and demonstrate the proposed system identification approach.

KEY WORDS
System identification, modeling, estimation, wavelet transform.

1 Introduction

Time-frequency domain is often more advantageous than time domain for linear time-invariant (LTI) system identification, mainly due to the lower computational complexity and faster convergence rate [1]. However, time-frequency techniques generally produce aliasing effects, which necessitate crossband filters between the subbands [1, 2]. The influence of these crossband filters on a system identifier implemented in the short-time Fourier transform (STFT) domain has been recently investigated [2], and explicit expressions for the STFT representation of LTI systems have been derived.

In contrast to the fixed time-frequency resolution of the STFT, the wavelet transform provides good localization both in frequency and time domains, and, as such, has attracted significant research in system identification and subband filtering [3, 4, 5]. In [3], the nonuniform filter banks interpretation of the discrete-time wavelet transform (DTWT) is used to perform linear filtering by directly convolving the subband signals and combining the results. In another scheme [4], it was shown that the DTWT of the system output signal can be computed by a weighted combination of the DTWT of shifted versions of the input signal. The use of the undecimated DTWT, which is linear and shift invariant, was introduced in [5] to overcome the lack of shift invariance and to implement time-domain convolution. However, none of the existing approaches provides an explicit representation of the system in the DTWT domain. A typical system identification scheme in the DTWT domain is illustrated in Fig. 1, where the block $\hat{\mathbf{A}}$ represents the DTWT model of the system.

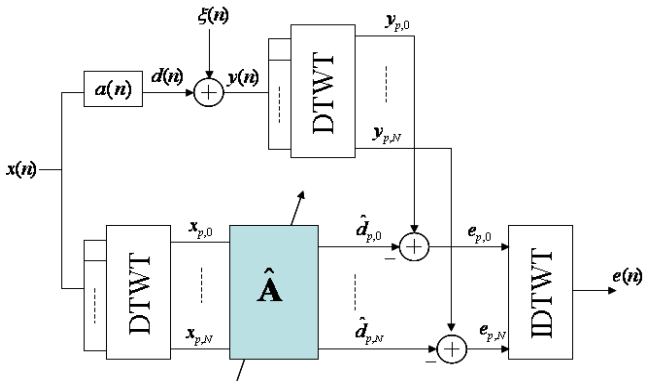

Figure 1. System identification scheme in the DTWT domain. The unknown system $a(n)$ is modeled by the block $\hat{\mathbf{A}}$ in the DTWT domain.

In this paper, we represent LTI systems in the DTWT domain and show that crossband filters between subbands are necessary for perfect representation. We derive relations between the crossband filters in the DTWT domain and the impulse response in the time domain. In contrast to the time-invariance property of the crossband filters in the STFT domain [2], the crossband filters in the DTWT domain are shown to be time-varying, due to nonuniform decimation factor over frequency-bands. Nonetheless, the band-to-band filters (*i.e.*, the filters that relate identical frequency-bands of input and output signals) remain time invariant. Furthermore, we show that under certain conditions, system representation in the DTWT domain can be approximated with only band-to-band filters. We show that as the signal-to-noise ratio (SNR) increases, or as more input data is available, longer band-to-band filters may be estimated to achieve the minimal mean-square error (MSE). Experimental results are provided to support the theoretical analysis.

The paper is organized as follows. In Section 2, we briefly review the DTWT. In Section 3, we derive explicit expressions for the representation of LTI systems in

*This research was supported by the Israel Science Foundation (grant no. 1085/05).

the DTWT domain. In Section 4, we consider an offline system identification in the DTWT domain using a least squares (LS) optimization criterion. Finally, in Section 5, we present simulation results to validate the theoretical analysis.

2 The discrete wavelet transform

In this section, we introduce the DTWT and relate it to nonuniform filter banks (for further details, see *e.g.*, [6] and the references therein).

Let $x(n) \in \ell^2(\mathbb{Z})$ denote a discrete-time signal, and let $x_{p,k}$ be the N-level wavelet coefficients at frequency-band k ($0 \leq k \leq N$) and at frame index p. The DTWT is commonly interpreted as a tree structured filter bank. Specifically, the N-level wavelet decomposition of $x(n)$ uses a low-pass filter $h(n)$ and a high-pass filter $g(n)$ to split the original space in two. One of the resulting half spaces is then divided in two, etc., such that the signal is decomposed into $N+1$ adjacent octave bands[1]. Similarly, the inverse DTWT (IDTWT), *i.e.*, reconstruction of $x(n)$ from its DTWT representation $x_{p,k}$, has also a tree structure with synthesis low-pass filter $\bar{h}(n)$ and high-pass filter $\bar{g}(n)$. In order to perfectly recovered $x(n)$ from $x_{p,k}$, the analysis and synthesis filters must satisfy perfect reconstruction constraints [6].

The DTWT is closely related to nonuniform filter banks, and these relations have been studied extensively (*e.g.*, [6]). In particular, we consider a decomposition of the signal $x(n)$ by using the nonuniform filter bank as illustrated in Fig. 2(a). By nonuniform we mean that the analysis filters have nonuniform bandwidths and that they are followed by an unequal decimation factor 2^{k+1}. Let $H(z)$ be the z-transform of the low-pass filter $h(n)$, and let $G(z)$, $\bar{H}(z)$ and $\bar{G}(z)$ be defined similarly. Then, using the "Nobel identities" [7], it is easy to verify that the analysis filters $H_k(z)$ are given by

$$H_k(z) = \begin{cases} G(z) & ; \quad k = 0 \\ G\left(z^{2^k}\right) \prod_{i=0}^{k-1} H\left(z^{2^i}\right); & k = 1, ..., N-1 \\ \prod_{i=0}^{k-1} H\left(z^{2^i}\right) & ; \quad k = N \end{cases} \tag{1}$$

Similarly, the inverse wavelet transform can be represented in terms of a synthesis (nonuniform) filter bank, as shown in Fig. 2(b). The synthesis filters $F_k(z)$ are given by

$$F_k(z) = \begin{cases} \bar{G}(z) & ; \quad k = 0 \\ \bar{G}\left(z^{2^k}\right) \prod_{i=0}^{k-1} \bar{H}\left(z^{2^i}\right); & k = 0, 1, ..., N-1 \\ \prod_{i=0}^{k-1} \bar{H}\left(z^{2^i}\right) & ; \quad k = N \end{cases} \tag{2}$$

Considering the nonuniform filter bank representation of the DTWT, the wavelet coefficients $x_{p,k}$ at each frequency-band k, can be expressed as

[1] Note that low values of k correspond to high frequency range.

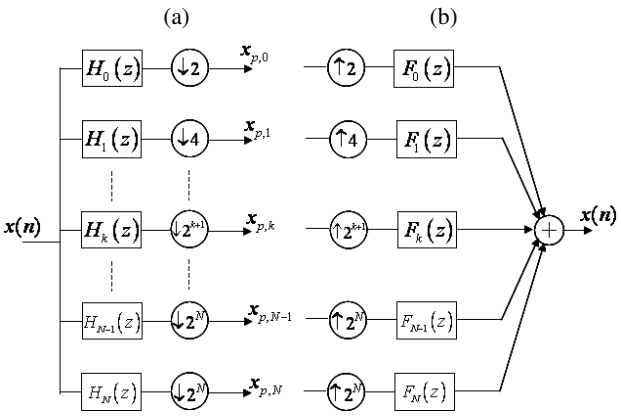

Figure 2. (a) Analysis and (b) synthesis nonuniform filter bank interpretation of the DTWT.

$$x_{p,k} = \begin{cases} \sum_m x(m)\, h_k\left(2^{k+1}p - m\right); & k = 0, ..., N-1 \\ \sum_m x(m)\, h_k\left(2^N p - m\right) & ; \quad k = N \end{cases} \tag{3}$$

where $h_k(n)$ is the inverse z-transform of $H_k(z)$. Similarly, the reconstruction of $x(n)$ from its wavelet coefficients $x_{p,k}$ can be written as

$$\begin{aligned} x(n) &= \sum_{k=0}^{N-1} \sum_p x_{p,k}\, f_k\left(n - 2^{k+1}p\right) \\ &\quad + \sum_p x_{N,p}\, f_N\left(n - 2^N p\right), \end{aligned} \tag{4}$$

where $f_k(n)$ is the inverse z-transform of $F_k(z)$. Let us define $\tilde{\psi}_{p,k}(n)$ and $\psi_{p,k}(n)$, as

$$\tilde{\psi}_{p,k}(n) = \begin{cases} \tilde{h}_k\left(n - 2^{k+1}p\right) & ; \quad k = 0, 1, ..., N-1 \\ \tilde{h}_k\left(n - 2^N p\right) & ; \quad k = N \end{cases} \tag{5}$$

and

$$\psi_{p,k}(n) = \begin{cases} f_k\left(n - 2^{k+1}p\right) & ; \quad k = 0, 1, ..., N-1 \\ f_k\left(n - 2^N p\right) & ; \quad k = N \end{cases} \tag{6}$$

where $\tilde{h}_k(n) \triangleq h_k(-n)$. Using (5) and (6), the DTWT and IDTWT of $x(n)$ can be written, respectively, as

$$x_{p,k} = \sum_m x(m) \tilde{\psi}^*_{p,k}(m) \tag{7}$$

and

$$x(n) = \sum_p \sum_{k=0}^{N} x_{p,k}\, \psi_{p,k}(n), \tag{8}$$

where * denotes complex conjugation. Here $\psi_{p,k}(n)$ are the wavelet basis functions, and the weights $x_{p,k}$ are the wavelet coefficients of $x(n)$ with respect to the above basis. Expressions (7)-(8) represent the DTWT and IDTWT of a discrete signal $x(n)$ in terms of basis functions, and

will be used in the following sections for deriving an explicit representation of an LTI system in the DTWT domain. It is worth noting that when orthonormal basis functions are considered, the analysis and synthesis filters satisfy $f_k(n) = h_k^*(-n)$ [7].

3 Representation of LTI systems in the DTWT domain

In this section, we derive explicit expressions for the representation of LTI systems in the DTWT domain, and show that crossband filters between subbands are essential for perfect modeling of the system.

Let $a(n)$ denote a length L_a impulse response of an LTI system, whose input $x(n)$ and output $d(n)$ are related by

$$d(n) = \sum_{i=0}^{L_a-1} a(i)x(n-i)\,. \tag{9}$$

Using (7) and (9), the DTWT of $d(n)$ can be written as

$$d_{p,k} = \sum_{m,\ell} a(\ell)x(m-\ell)\tilde{\psi}_{p,k}^*(m)\,. \tag{10}$$

Substituting (8) for $x(n)$ into (10), we obtain

$$d_{p,k} = \sum_{k'=0}^{N}\sum_{p'} x_{p-p',k'}\, a_{p',k,k'}(p)\ , \tag{11}$$

where

$$a_{p',k,k'}(p) = \sum_{m,\ell} \psi_{p-p',k'}(m-l)\tilde{\psi}_{p,k}^*(m)a(\ell) \tag{12}$$

may be interpreted as a response to an impulse $\delta_{p',k-k'}$ in the time-frequency domain (the impulse response is translation varying in both time and frequency axes). An explicit relation between the time-frequency domain impulse response $a_{p',k,k'}(p)$ and the time-domain impulse response $a(n)$ is achieved by substituting (5) and (6) into (12), resulting in

$$\begin{aligned} a_{p',k,k'}(p) &= \sum_{m,\ell} f_{k'}\left(m-\ell-2^{\min(k'+1,N)}(p-p')\right) \\ &\quad \times \tilde{h}_k\left(m-2^{\min(k+1,N)}p\right)a(\ell) \\ &= \{a(n)*\phi_{k,k'}(n)\}|_{n=\lambda_{k,k'}(p,p')} \\ &\triangleq \bar{a}_{n,k,k'}|_{n=\lambda_{k,k'}(p,p')} \end{aligned} \tag{13}$$

where $*$ denotes convolution with respect to the time index n,

$$\phi_{k,k'}(n) \triangleq \sum_m \tilde{h}_k(m)\, f_{k'}(n+m) \tag{14}$$

and $\lambda_{k,k'}(p,p') = \left(2^{\min(k+1,N)} - 2^{\min(k'+1,N)}\right)p + 2^{\min(k'+1,N)}p'$. The $\min(\cdot)$ operator is attributable to the equal decimation factor used at the last two frequency-bands ($k = N-1, N$). Equation (11) indicates that the temporal signal $d_{p,k}$, for a given frequency-band index k, is related via the time-varying filters $a_{p',k,k'}(p)$ to all the frequency-bands k' ($k' = 0, 1, \ldots, N$) of the input signal $x_{p,k'}$. We refer to $a_{p',k,k'}(p)$ for $k = k'$ as a band-to-band filter and for $k \neq k'$ as a crossband filter. The crossband filters are used for canceling the aliasing effects caused by the subsampling. It is worth noting that in contrast with the STFT representation of LTI systems [2], for which the crossband filters are time invariant, in the DTWT domain these filters are time-varying. The time variation of the filters are represented by the dependence of the system response $a_{p',k,k'}(p)$ on the frame index p. This dependence, however, vanishes when $k = k'$, which indicates the time invariance of the band-to-band filters $a_{p',k,k}$. The time variations of the crossband filters are a consequence of utilizing an unequal decimation factor at each frequency-band.

The significance of the crossband filters can be well illustrated by applying the discrete-time Fourier transform (DTFT) to the undecimated crossband filter $\bar{a}_{n,k,k'}$ [defined in (13)] with respect to the time index n:

$$\bar{A}_{k,k'}(\theta) = \sum_n \bar{a}_{n,k,k'}\, e^{-jn\theta} = A(\theta)H_k(\theta)\,F_{k'}(\theta)\ , \tag{15}$$

where $A(\theta)$, $H_k(\theta)$ and $F_{k'}(\theta)$ are the DTFT of $a(n)$, $h_k(n)$ and $f_{k'}(n)$, respectively. Equation (15) implies that the number of crossband filters required for the representation of an impulse response is mainly determined by the analysis and synthesis filters, while the length of the crossband filters (with respect to the time index n) is related to the length of the impulse response. Had both $h(n)$ and $\bar{h}(n)$ been ideal halfband low-pass filters and had $g(n)$ and $\bar{g}(n)$ been ideal halfband high-pass filters, a perfect DTWT representation of the system $a(n)$ could be achieved by using just the band-to-band filter $a_{p',k,k}$, since in this case the product of $H_k(\theta)$ and $F_{k'}(\theta)$ is identically zero for $k \neq k'$. However, the low-pass and high-pass filters are practically not ideal and therefore, $\bar{A}_{k,k'}(\theta)$ and $\bar{a}_{n,k,k'}$ are not zero for $k \neq k'$. Figure 3 illustrates the magnitude response of a 6-band filter bank corresponding to a 5-level wavelet decomposition, using a Daubechies orthonormal wavelet of length 64. It can be seen that a substantial overlap exists between the analysis filters due to the compact support of the low-pass filter $h(n)$. It is worth noting that since we employ orthonormal wavelet bases [such that $f_k(n) = h_k^*(-n)$], only the overlap between the analysis filters $h_k(n)$ is of interest. Figure 4 illustrates the energy of the crossband filters, defined in dB by

$$E_{k,k'} = 10\log_{10}\sum_n |\bar{a}_{n,k,k'}|^2\ , \tag{16}$$

at the third frequency-band ($k = 3$), and for 5-level Daubechies wavelet with prototype low-pass filter lengths $L = 4, 16$ and 64. We use a synthetic room impulse response $a(n)$ of length $L_a = 1000$ based on a statistical

reverberation model, which exhibits a reverberation time of $T_{60} = 50\ ms$ (for further simulation details see Section 5). It can be seen that the energy of a crossband filter from frequency-band k' to frequency-band k decreases as $|k - k'|$ increases, since the overlap between adjacent analysis filters becomes smaller. Clearly, this overlap is determined by the compact support of the time-domain low-pass wavelet function $h(n)$. As L, the length of $h(n)$, increases, a smaller overlap is obtained and lower crossband filters energy is achieved, as shown in Fig. 4. As a result, for large L values, relatively few crossband filters need to be considered in order to capture most of the energy of the DTWT representation of $a(n)$. We observe from Fig. 4 that for $L = 64$, for instance, most of the energy of $\bar{a}_{n,3,k'}$ is concentrated in only three filters ($k' = 2, 3$ and 4). In the following sections, for the sake of simplicity, we assume that the analysis and synthesis filters are selective enough so that adjacent filters have insignificant overlap with each other, and therefore no crossband filters should be considered. Denoting by L_{a_k} the length of the band-to-band filter at the kth frequency-band, it is easy to verify from (13) that

$$L_{a_k} = \left\lceil \frac{L_a + L_{h_k} + L_{f_k} - 2}{2^{k+1}} \right\rceil , \qquad (17)$$

where L_{h_k} and L_{f_k} are the length of the analysis filter $h_k(n)$ and the synthesis filter $f_k(n)$, respectively, at the kth frequency-band. Using (1) and (2), we obtain after some manipulations

$$L_{h_k} = L_{f_k} = 2^k (2L - 1) - (L - 1) \qquad (18)$$

which can be substituted into (17) to obtain

$$L_{a_k} = \left\lceil \frac{L_a - 2L}{2^{\min(k+1,N)}} \right\rceil + 2L - 1 , \qquad (19)$$

where L is the length of the low-pass and high-pass filters [*i.e.*, $h(n)$, $g(n)$, $\tilde{h}(n)$ and $\tilde{g}(n)$]. Equation (19) indicates that the length of the band-to-band filter at each frequency-band decreases as k increases[2], which is in contrast with the fixed-length filters in STFT-based identification schemes [2]. Note that in many applications, such as acoustic echo cancellation, the length of the system impulse response is much larger than that of the analysis/synthesis filters, such that (19) can be approximated as

$$L_{a_k} \approx \left\lceil \frac{L_a}{2^{\min(k+1,N)}} \right\rceil . \qquad (20)$$

4 System identification in the DTWT domain

In this section, we consider an offline system identification in the DTWT domain using the LS criterion for the estimation of the band-to-band filter in each frequency-band.

[2]Note that the length of the band-to-band filter in the last frequency-band $k = N$ is equal to that of $k = N - 1$ [see (13)].

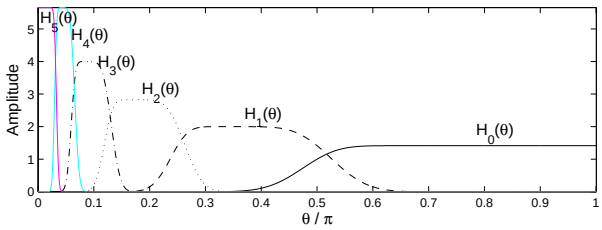

Figure 3. Magnitude responses of analysis filters in a 6-band nonuniform filter bank using a prototype Daubechies low-pass filter of length 64.

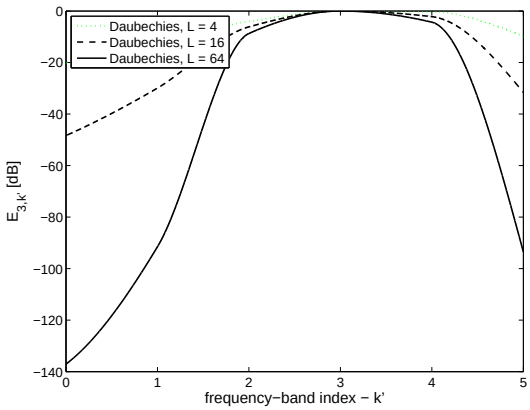

Figure 4. Energy of the crossband filters $\bar{a}_{n,3,k'}$ for a synthetic room impulse response $a(n)$.

Consider the DTWT-based system identification scheme as illustrated in Fig. 1. The system output signal $y(n)$ is given by

$$y(n) = d(n) + \xi(n) = a(n) * x(n) + \xi(n) , \qquad (21)$$

where $a(n)$ is the impulse response of the unknown LTI system, and $\xi(n)$ is the corrupting noise signal. From (21) and (11), the DTWT of $y(n)$ may be written as

$$y_{p,k} = d_{p,k} + \xi_{p,k} = \sum_{k'=0}^{N} \sum_{p'} x_{p-p',k'}\, a_{p',k,k'}\,(p) + \xi_{p,k} . \qquad (22)$$

Let P_k denote the number of samples in the time-trajectory of $y_{p,k}$. The subscript k in P_k indicates the unequal length of $y_{p,k}$ in each frequency-band, due to the frequency-dependent decimation factor. Then, (22) can be written in a vector form as

$$\mathbf{y}_k = \mathbf{d}_k + \boldsymbol{\xi}_k , \qquad (23)$$

where

$$\mathbf{y}_k = \left[\begin{array}{cccc} y_{0,k} & y_{1,k} & y_{2,k} & \cdots \quad y_{P_k-1,k} \end{array} \right]^T \qquad (24)$$

represents the DTWT coefficients of the output signal in the kth frequency-band, and the vectors $\mathbf{d}_k$ and $\boldsymbol{\xi}_k$ are defined similarly.

Let $\hat{a}_{p',k,k}$ be an estimate of the (time-invariant) band-to-band filter $a_{p',k,k}$, and let $\hat{d}_{p,k}$ be the resulting estimate

of $d_{p,k}$, *i.e.*,

$$\hat{d}_{p,k} = \sum_{p'=0}^{L_{a_k}-1} \hat{a}_{p',k,k}\, x_{p-p',k}\,. \tag{25}$$

We disregard the crossband filters in the identification process, relying on the assumption that the overlap between $H_k(\theta)$ and $F_{k'}(\theta)$ for $k \neq k'$ is small enough. However, when the overlap is relatively large, ignoring the crossband filters yields a model mismatch which may degrade the system estimate accuracy and result in an insufficient MSE performance. This point will be further demonstrated in Section 5. Let $\hat{\mathbf{a}}_k = \left[\begin{array}{cccc} \hat{a}_{0,k,k} & \hat{a}_{1,k,k} & \cdots & \hat{a}_{L_{a_k}-1,k,k} \end{array}\right]^T$ denote the LS estimate of the band-to-band filter at frequency-band k:

$$\begin{aligned} \hat{\mathbf{a}}_k &= \arg\min_{\mathbf{a}_k} \|\mathbf{y}_k - \mathbf{X}_k\mathbf{a}_k\|^2 \\ &= \left(\mathbf{X}_k^H\mathbf{X}_k\right)^{-1}\mathbf{X}_k^H\mathbf{y}_k\,, \end{aligned} \tag{26}$$

where $\mathbf{y}_k$ is defined in (24), $\mathbf{X}_k$ represents an $P_k \times L_{a_k}$ Toeplitz matrix with $x_{m-\ell,k}$ being its (m,ℓ)th term, and $\mathbf{X}_k^H\mathbf{X}_k$ is assumed to be not singular. An estimate of the desired signal in the DTWT domain, using only the band-to-band filter, is then given by

$$\hat{\mathbf{d}}_k = \mathbf{X}_k\hat{\mathbf{a}}_k = \mathbf{X}_k\left(\mathbf{X}_k^H\mathbf{X}_k\right)^{-1}\mathbf{X}_k^H\mathbf{y}_k\,. \tag{27}$$

The model defined in (25) for the system identification contains $N+1$ filters, each of length $L_{a_k} = \left\lceil L_a/2^{\min(k+1,N)}\right\rceil$, $k = 0,...,N$, resulting in L_a coefficients that should be estimated for identifying the impulse response $a(n)$ in the DTWT domain. It is well known, however, that the optimal model order, *i.e.*, the number of model coefficients that should be estimated to attain the minimum MSE (MMSE), is affected by the level of noise in the data and the length of the observable data [8]. Here the model order is determined by the length of the band-to-band filters L_{a_k}. Consequently, as the SNR increases or as more data is employable, the optimal model order increases, and correspondingly longer band-to-band filters can be estimated. Note that the time-domain impulse response length L_a determines the length of the band-to-band filters in each frequency-band [see (20)]. Therefore, denoting by $\hat{L}_a$ the length of $a(n)$ that is practically employed for the identification process, the resulting MSE is defined by

$$\epsilon(\hat{L}_a) = \frac{E\left\{\left(d(n) - \hat{d}_{\hat{L}_a}(n)\right)^2\right\}}{E\left\{d^2(n)\right\}}\,, \tag{28}$$

where $\hat{d}_{\hat{L}_a}(n)$ is the inverse DTWT of the estimated desired signal $\hat{d}_{p,k}$ using band-to-band filters of lengths $\hat{L}_{a_k} = \left\lceil \hat{L}_a/2^{\min(k+1,N)}\right\rceil$. The optimal model order is therefore given by

$$\hat{L}_{a,opt} = \arg\min_{\hat{L}_a} \epsilon(\hat{L}_a)\,. \tag{29}$$

The influence of the power and length of the input signal on the optimal model order is investigated in the next section.

5 Experimental results

In this section, we present experimental results that verify the theoretical analysis. We use a synthetic room impulse response $a(n)$ based on a statistical reverberation model, which generates a room impulse response as a realization of a nonstationary stochastic process $a(n) = u(n)\beta(n)e^{-\alpha n}$, where $u(n)$ is a step function, $\beta(n)$ is a zero-mean white Gaussian noise and α is related to the reverberation time T_{60} (the time for the reverberant sound energy to drop by 60 dB from its original value). In the following simulations, the sampling rate is 16 kHz, the length of the impulse response is set to 62.5 ms ($L_a = 1000$), α corresponds to $T_{60} = 50$ ms and $\beta(n)$ is unit-variance zero-mean white Gaussian noise. We employ a 5-level Daubechies wavelet ($N = 5$) of length $L = 64$. The input signal $x(n)$ and the additive noise signal $\xi(n)$ are uncorrelated zero-mean white Gaussian processes with variances σ_x^2 and σ_ξ^2, respectively, and the SNR is defined by σ_x^2/σ_ξ^2.

Figure 5 shows the MSE curves $\epsilon(\hat{L}_a)$ [see (28)], for several $\hat{L}_a$ values, as a function of the input SNR obtained by an input signal of length 0.5 sec [Fig. 5(a)] and a longer signal of length 2 sec [Fig. 5(b)]. It can be seen that as the SNR increases, a lower MSE value can be obtained by utilizing longer band-to-band filters (larger $\hat{L}_a$). We observe that assuming the true system order ($\hat{L}_a = L_a = 1000$) not necessarily improves the system identifier performance. Figure 5(a) shows that when the SNR is lower than -30 dB, assuming a length of $\hat{L}_a = 100$ samples ($= 0.1L_a$) yields the minimal MSE, and enables a decrease of 7 dB in the MSE value relative to that achieved by assuming $\hat{L}_a = 1000$ (true system length). When considering SNR values higher than -30 dB, the inclusion of 300 samples in the model ($\hat{L}_a = 300$) is preferable. Moreover, a comparison of Figs. 5(a) and (b) indicates that when the signal length increases (while the SNR remains constant), longer band-to-band filters should be considered in order to attain the MMSE. The relatively high MSE value obtained in this experiment is attributable to the significance overlap exists between adjacent filters (see Fig 3), which necessitates the estimation of crossband filters. Note that surprisingly, a lower MSE is achieved for the shorter signal [Fig. 5(a)] at high SNR values. This result, however, is somehow misleading since the proposed model is not accurate and a model mismatch is introduced by ignoring the crossband filters. If the model was accurate and all crossband filters were estimated, a lower MSE would have been achieved by increasing the signal length. As was explained in Section 3, ignoring the crossband filters is justified by assuming a long low-pass filter, such that the overlap between adjacent frequency-bands is negligible. To validate this assumption, we repeat the previous experiment for several low-pass filter lengths. Figure 6 shows the resulting MSE curves as a function of $\hat{L}_a$ for analysis Daubechies low-pass filter of lengths $L = 4, 8, 16$ and 32, obtained for a 25 dB SNR and a 2 sec length input signal. Indeed, a lower MSE value is achieved with increasing L. Figure 6

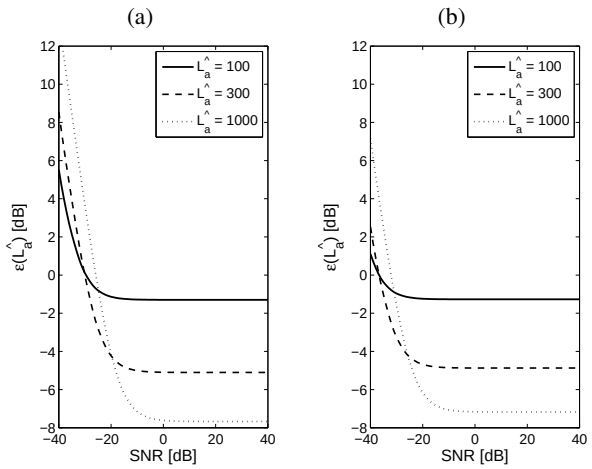

Figure 5. MSE curves as a function of the input SNR for white Gaussian signals. (a) Signal length is 0.5 sec. (b) Signal length is 2 sec.

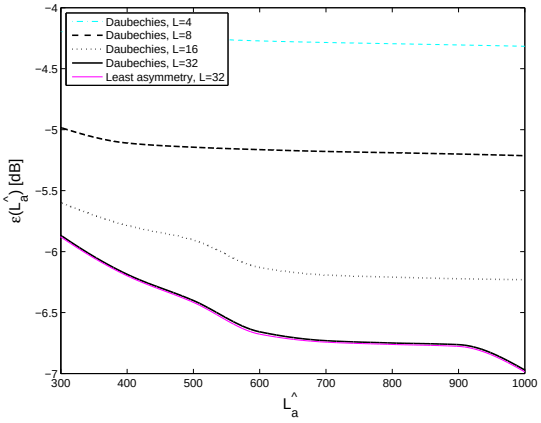

Figure 6. MSE curves as a function of $\hat{L_a}$ for several low-pass filter lengths (L).

also compares the Daubechies wavelet, which is associated with *minimum-phase* filters, to the least asymmetry wavelet associated with near *linear-phase* filters (both of length 32). No improvement is visible by using the least asymmetry filter, which indicates that the linearity of the phase is not critical for efficiently representing an LTI system in the DTWT domain. The representation is mainly influenced by the filter's frequency response amplitude rather than its phase.

6 Conclusion

We have presented LTI systems in the DTWT domain, and showed that time-varying crossband filters are required for a perfect representation. We showed that not only do the crossband filters vary in time but also their length changes with frequency. When using an approximate representation without crossband filters, the system identification performance is greatly affected by the assumed lengths of band-to-band filters, which are related to the SNR and length of input signal. As the SNR or the signal length increases, longer band-to-band filters may be estimated. Further improvement is obtainable by incorporating crossband filters into the identification process. However, the time variation of crossband filters has to be carefully considered when estimating these filters.

7 Acknowledgments

The authors thank Prof. Shalom Raz for teaching a graduate course on "Time-frequency methods and their applications" in the Technion, which inspired and encouraged this work.

References

[1] A. Gilloire and M. Vetterli, "Adaptive filtering in subbands with critical sampling: Analysis, experiments, and application to acoustic echo cancellation," *IEEE Trans. Signal Processing*, vol. 40, no. 8, pp. 1862–1875, Aug. 1992.

[2] Y. Avargel and I. Cohen, "System identification in the short-time Fourier transform domain with crossband filtering," *IEEE Trans. Audio, Speech, Language Process.*, vol. 15, no. 4, pp. 1305–1319, May 2007.

[3] P. P. Vaidyanathan, "Orthonormal and biorthonormal filter banks as convolvers and convolutional coding gain," *IEEE Trans. Signal Processing*, vol. 41, no. 6, pp. 2110–2130, Jun. 1993.

[4] M. Sandler, "Linear time-invariant systems in the wavelet domain," *IEE Seminar: Time-scale and Time-Frequency Analysis and Applications*, pp. 2/1–2/6, Feb. 2000.

[5] H. Guo and C. S. Burrus, "Convolution using the undecimated discrete wavelet transform," in *Proc. Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, May. 1996, pp. 1291–1294.

[6] M. Vetterli and C. Herley, "Wavelets and filter banks: theory and design," *IEEE Trans. Signal Processing*, vol. 40, no. 9, pp. 2207–2232, Sep. 1992.

[7] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[8] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Upper Saddle River, New Jersey: Prentice-Hall, 1999.