Modeling and Identification of LPTV Systems by Wavelets

Y. Dorfan A. Feuer * and B. Porat

Department of EE, Technion, Haifa 32000, Israel

Abstract

We propose a novel model for discrete Linear Periodic Time Varying (LPTV) systems using wavelets. The new model is compared with the 'raised model', which is commonly used for modeling LPTV systems. In fact, it turns out that the new model can be viewed as a generalization of the raised model.

The wavelets model will be shown to be particularly suitable for adaptive identification of LPTV systems. It offers a compromise between time- and frequency-based algorithms. Time resolution is needed for modeling reasons and minimizing processing delay. Frequency resolution enables faster convergence of adaptive algorithms in general and the Least Mean Square (LMS) algorithm used here, in particular. Simulations show that for a colored input using the new model results not only in faster convergence compared to the raised model based algorithm, but also produces a lower steady-state error. This, at no significant additional cost in numerical complexity.

1 Introduction

On-line identification of general Linear Time Variant (LTV) systems, although very important for a variety of applications, is still a relatively open issue. If the changes are sufficiently slow a Linear Time Invariant (LTI) model can be used and the changes are tracked by an adaptive algorithm [1], [2] and [3]. However, when these changes are fast, the time variations need to be modeled. If there are some underlying *constant* parameters which model the system, they can be estimated using adaptive or off-line algorithms. Furthermore, even if these parameters vary slowly an adaptive algorithm can track these variations. So,

* Corresponding author.
Email addresses: yuval.d@comsys.co.il (Y. Dorfan), feuer@ee.technion.ac.il (A. Feuer), boaz@ee.technion.ac.il (B. Porat).

for the identification of LTV systems, a priori knowledge about the variations in time is assumed. Usually one assumes that the changes are characterized by a finite set of functions [4], [5] and [6]. However, all these cases refer only to limited classes of LTV systems.

One important class of LTV systems includes the Linear Periodically Time Varying (LPTV) systems [7], [8]. For quite a number of applications periodicity in the time variation can be observed, see [9], [10] and [11]. A number of physical phenomena have periodic characteristics and using a LPTV model is very reasonable. Orbital motion, AC motors [12], rhythm of the heart, or periodic disturbances (vibrations) in helicopters [13] are just a few examples of such phenomena. Design of optimum periodic time varying filters for applications in diagnostics of combustion engines appears in [12] and periodic optimal control is discussed in [11]. Note that if a system is LTV, but its time variations consist of a fast mode which is periodic and a slow mode, an LPTV model can be used combined with an adaptive algorithm. The LPTV model then handles the fast changes and the adaptive algorithm the slow changes. A common way to model LPTV systems is via the raising method described in, e.g. [14].

In this paper we describe a new approach for modeling discrete LPTV systems with finite impulse responses using wavelets. We assume that the period is known a priori and, for simplicity, that it is a power of two (when the sampling time is under the user's control, he can choose it accordingly). In fact, we show that using wavelets can be viewed as a generalization of the raised model. Using wavelets in modeling LTV systems is discussed by Doroslovacki and Fan [4]. For the periodic case the authors assume that the period of the system and a set of functions that characterize the changes are known a priori.

After introducing the model we investigate its use for adaptive identification with the LMS algorithm. This algorithm is simple, has reasonable tracking abilities and low computational complexity. Its main disadvantage, as is well known, is its slow convergence. This disadvantage becomes more acute here since time scale for any adaptive algorithm is slowed down with LPTV systems, by a factor equal to their period.

The convergence rate disadvantage has motivated us to consider using wavelets in adaptive identification of linear systems. Wavelets have been used for adaptive identification of LTI systems. The convergence rate of the LMS algorithm is shown to be higher when wavelets are used [15], [16] and [17]. The motivation behind these algorithms is separation of the parameters, similar to Frequency Domain Least Mean Square (FDLMS) [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29] and [30]. The FDLMS uses the Fast Fourier Transform (FFT) for reasons of lower complexity and faster convergence rate.

Convergence rate is also a motivation for sub-band approaches [31], [32] and Gabor domain adaptive filtering [33]. Frequency resolution schemes enable separation and therefore faster convergence, while time resolution is needed for minimizing the delay and for accurate modeling of the unknown system's impulse response. Wavelets provide flexible trade-offs between time and frequency resolutions.

Subband approaches are used in applications such as echo cancellation [33] and equalizers [34]. Splitting a stationary input signal into different frequency bands and decimations, result in a lower spread of eigenvalues. A number of independent adaptive filters are updated at a reduced rate. The price is a larger steady-state error because of insufficient order estimation and aliasing. In some of those approaches an effort is made to minimize the aliasing error caused by non-ideal filters [35]. It is shown in [36] that decimated-based models used for LTI systems are in fact LPTV models. Moreover, these approaches are shown to be special cases of the new model we present here, which is more general for modeling both LPTV and LTI systems.

For many stationary and non-stationary inputs the wavelet transform is claimed to be very close to the Karhunen-Loēve Transform (KLT), which achieves exact diagonalization [37]. When the auto-correlation of the input signal is known a priori, the KLT can be used. Otherwise, it can be estimated. However, the complexity of such estimation is high [15]. Therefore, in [4], [16] and [37] wavelet transforms are used. This is a sub-optimal attempt to approach diagonalization, but suitable for real-time processing.

The next section briefly describes the raising technique. Section 3 presents a new wavelet model for LPTV systems which is analyzed in the 'raised domain'. Applying the new model to adaptive identification of LPTV systems is described in Section 4 and Section 5 summarizes some simulation experiment results. Conclusions are discussed in the last section.

2 Theoretical background

Let us start with a brief review of the *raised model* as one of the common ways to model LPTV systems. Generally, linear systems (LTV) can be modelled via their impulse response. Namely

$$y[n] = \sum_{\ell} h[n, \ell] x[\ell] \tag{1}$$

where $x[n], y[n]$ are the system input and output sequences respectively and $h[n, \ell]$ is the system response to an impulse sequence $\delta[n - \ell]$. When $h[n + N, \ell + N] = h[n, \ell]$ for every n, ℓ, the system is periodic (LPTV) with period

N and when $N = 1$ the system is time invariant (LTI). Using this property of the linear periodic system, (1) can be rewritten as

$$\begin{aligned} y[n] &= \sum_j \sum_{m=0}^{N-1} h[n, jN+m] x[jN+m] \\ &= \sum_j \sum_{m=0}^{N-1} h[n-jN, m] x[jN+m] \end{aligned} \tag{2}$$

or

$$\underline{Y}[k] = \sum_j H[k-j] \underline{X}[j] \tag{3}$$

where k is an N times slower time scale than n, and

$$\underline{X}[j] = \begin{bmatrix} x[jN] \\ x[jN+1] \\ \vdots \\ x[jN+N-1] \end{bmatrix}; \underline{Y}[k] = \begin{bmatrix} y[kN] \\ y[kN+1] \\ \vdots \\ y[kN+N-1] \end{bmatrix};$$

$$H[k-j] = \begin{bmatrix} h[kN,jN] & h[kN,jN+1] & \cdots & h[kN,jN+N-1] \\ h[kN+1,jN] & h[kN+1,jN+1] & \cdots & h[kN+1,jN+N-1] \\ \vdots & \vdots & \ddots & \vdots \\ h[kN+N-1,jN] & h[kN+N-1,jN+1] & \cdots & h[kN+N-1,jN+N-1] \end{bmatrix} \tag{4}$$

This is commonly referred to as the *raised model* of the LPTV system. Clearly, the raised form, (3), is a MIMO LTI system. Depending on the properties of the matrix sequence $\{H[k]\}$ one can talk about an FIR or IIR system with the corresponding transfer function matrix, $H(z) = \sum_k H[k]z^{-k}$. This sequence contains all the information of the LPTV system and identifying the system is identifying this sequence (or the corresponding transfer function matrix).

The raising concept appears in the 'filter bank' literature where it is referred to as 'polyphase transform' (see, e.g. [38]). Figure 1 illustrates the structure of the raised model.

The main observation we make in this model is that any LPTV system can be modeled as consisting of three parts: A generic (system independent) linear periodic transform[1] (the polyphase transform in the raised model case) and

[1] By a periodic transform T with period N we mean a transform with the property: $T\left(x\left[n+N\right]\right)\left[k\right] = T\left(x\left[n\right]\right)\left[k+1\right]$, where n and k are the times scales in the input and transform domains respectively.

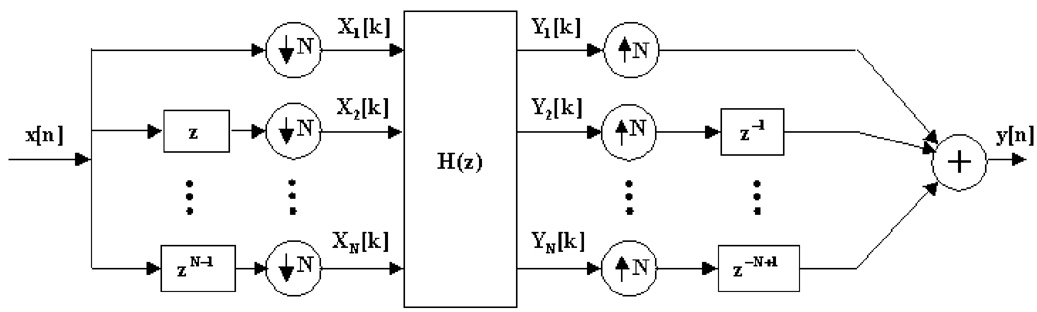

Fig. 1. The Raised Model (Polyphase transfomed).

its inverse and in between a LTI system which captures the particular system parameters. This is demonstrated in Figure 2. With this in mind we proceed to present our approach next.

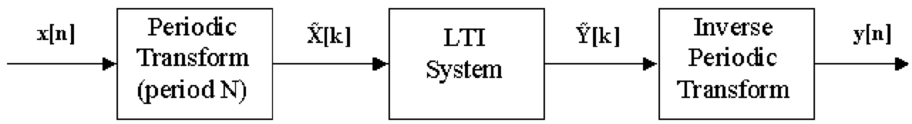

Fig. 2. Generalized LPTV model

3 A novel approach for modeling LPTV systems by wavelets

With the generic model of Figure 2 in mind, we recall that the Discrete Wavelet Transform (DWT) is a periodic transform, hence, can be used to model LPTV systems. The benefit of doing that will be discussed later on.

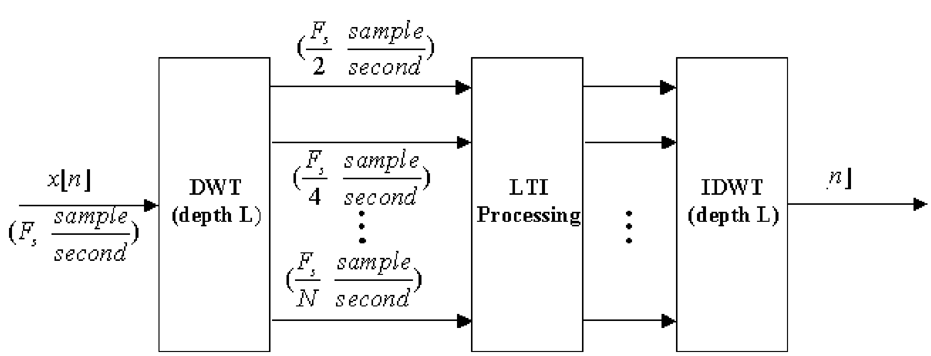

Fig. 3. A general model with DWT and IDWT.

We assume first that the period is such that $N = 2^L$. Then, the model takes the form presented in Figure 3. However, with this choice of transform, we face a problem with the linear processing block. Each input entry to this block has a different sampling rate. One possibility, commonly referred to as the subband approach (see [31] and [32]), is to have a separate LTI processing on each branch. This clearly limits the scope of this model. Motivated by the theory of Young about modeling LTV systems by the Continuous Wavelet Transform (CWT) [39] we prefer addressing this problem in a different way.

Our approach is to equalize the sampling rates of all branches and, as we will shortly show, it fits well into the structure of filter banks (which is used to implement the DWT). As is well known, dyadic filter banks are constructed using the blocks in Figure 4 for appropriately chosen filters H_1, H_0, G_1 and G_0.

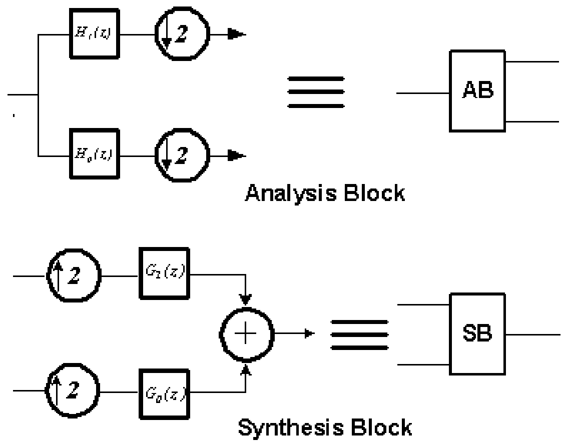

Fig. 4. Typical dyadic filter bank blocks.

When we choose $H_1 = 1$, $H_0 = z$, $G_1 = 1$ and $G_0 = z^{-1}$ AB and SB become the polyphase transform and its inverse, respectively (denoted then as PP and IPP). Combining these two sets of blocks we construct what we refer to as Wavelet Analysis Tree (WAT) and Wavelet Synthesis Tree (WST). These are presented in Figure 5. Note that, in this figure, if all the AB and SB blocks are also replaced by PP and IPP blocks respectively, we simply get the polyphase transform and its inverse (or, the raising process).

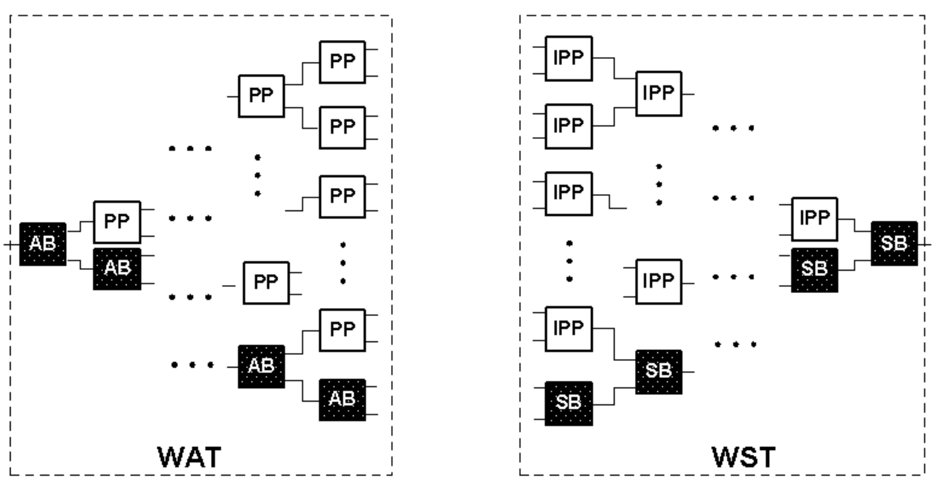

Fig. 5. The WAT and WST structures.

WAT has one input at rate F_s transformed into $N = 2^L$ outputs at rate F_s/N

each, while WST has $N = 2^L$ inputs at rate F_s/N each, transformed into one output at rate F_s. With WAT and WST replacing DWT and IDWT in Figure 3 and restricting ourselves to a MIMO FIR system for the LTI processing block we get the model we propose for LPTV system identification. This is presented in Figure 6 where we have denoted the output of WAT and the input to WST as $\underline{W}_x[k]$ and $\underline{W}_y[k]$ respectively.

Using the polyphase domain analysis as in [38] we can write

$$\underline{W}_x[k] = H_{DWT}[k] * \underline{X}[k]$$

and

$$\underline{Y}[k] = H_{IDWT}[k] * \underline{W}_y[k]$$

where $H_{DWT}[k]$ and $H_{IDWT}[k]$ are sequences of $N \times N$ matrices determined by the wavelets chosen and satisfy the perfect reconstruction condition, namely, $H_{IDWT}[k] * H_{DWT}[k] = \delta[k]I$ and $\underline{X}[k], \underline{Y}[k]$ are as in eqn.(4). Then the model in eqn. (3) can be rewritten as

$$H_{IDWT}[k] * \underline{W}_y[k] = H[k] * H_{IDWT}[k] * \underline{W}_x[k]$$

or

$$\underline{W}_y[k] = A\,[k] * \underline{W}_x[k]$$

where

$$H\,[k] = H_{IDWT}[k] * A[k] * H_{DWT}[k] \tag{5}$$

We restrict ourselves to $A\,[k]$ having a finite impulse response so that the model we consider is of the form:

$$\underline{W}_y[k] = \sum_{\ell=0}^{M-1} A_\ell \underline{W}_x[k-\ell] \tag{6}$$

where $A_\ell \in \Re^{N \times N}$ are the coefficient matrices and M is the memory length of the FIR. Since WAT and WST are generic, the LPTV system is parameterized through the $\{A_\ell\}$ matrices - MN^2 parameters which we wish to estimate.

Remark 1 *WAT and WST are a Perfect Reconstruction (PR) pair.*

Remark 2 *In Figure 5, replacing the AB and SB with PP and IPP the model reduces back to the raised model.*

Remark 3 *The structure offered by WAT and WST in Figure 5 can be generalized to any general tree-structured filter bank and wavelet packet (see e.g. [38]) by corresponding choices of the AB and SB trees.*

Remark 4 *The proposed model clearly induces a delay between its input and output (as does the raised model). This delay, denoted as d_0, is the result of*

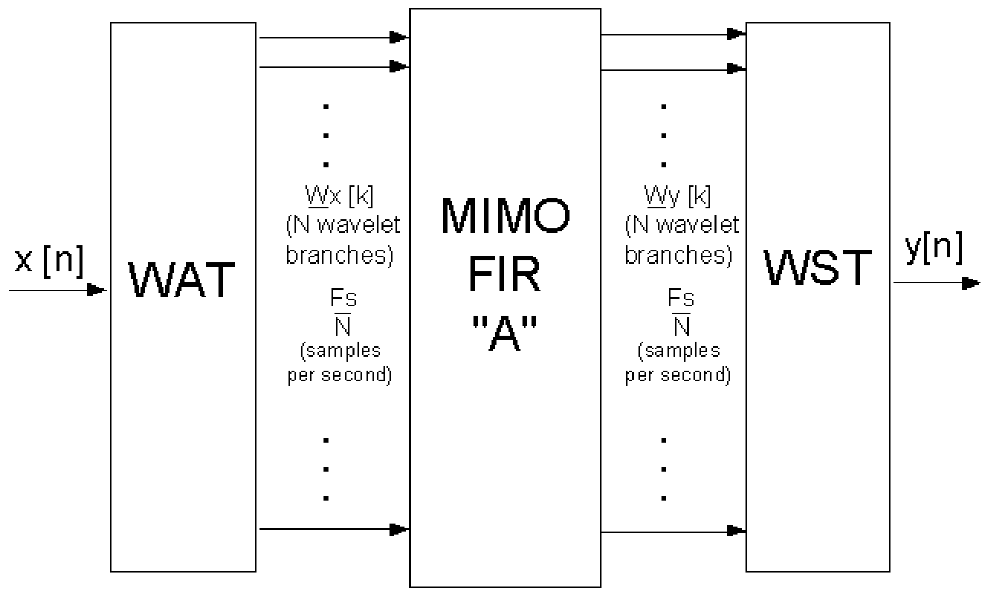

Fig. 6. The new wavelet model for LPTV systems.

the AB and PP blocks in the WAT which need to be delayed to be made causal. Assuming that each AB block needs to be delayed by d, the total delay is then

$$d_0 = d \cdot \sum_{i=0}^{L-1} 2^i = d \cdot (2^L - 1) = d \cdot (N-1) \tag{7}$$

Remark 5 *For the choice of Haar wavelets one gets $H_{DWT}[k] = \delta[k] H_H$ and $H_{IDWT}[k] = \delta[k] H_H^{-1}$ where H_H is a constant matrix. This means that the Haar wavelet choice is equivalent to the raised model up to a coordinate transformation.*

Remark 6 *Let the wavelet generating filter be of length ℓ. Then the resulting $H_{DWT}[k]$ and $H_{IDWT}[k]$ at depth L, are of length $\left\lceil \frac{(\ell-1)(N-1)}{N} \right\rceil + 1$ and the length of $H[k]$ is*

$$M_{total} = M + 2 \left\lceil \frac{(\ell-1)(N-1)}{N} \right\rceil \tag{8}$$

The new model (Figure 6) has the advantage of "almost diagonalization" of the input's auto-correlation matrix performed by the wavelet transform (because of its frequency localization property), a very important feature for adaptive filtering. The proposed model is, as far as we know, the first attempt to use the classical DWT of Mallat [40] for on-line modeling of LPTV systems. The structures of WAT and WST enable online processing with a low complexity and relatively small delay.

Choosing the wavelet function (or, equivalently, the filters H_1, H_0, G_1 and G_0) for the new model is an important issue. To minimize delay we wish to get as good a time localization as possible while to get an 'almost' diagonal auto-correlation matrix (which is a desired property for adaptive processing), frequency localization is desired. Wavelets provide a wide variety of choices and possible compromises between the two contradicting desired properties. They have localization qualities in both time and frequency, hence constitute a desired choice for our adaptive identification problem. Combined with our choice of FIR for the LTI part in the model, using wavelets has an additional benefit over the raised model. With the same number of parameters one can model a longer finite impulse response. This follows directly from eqn.(5) and is illustrated in Figure 7 where we note again (see Remark 5) that the model with the Haar wavelet is equivalent to the raised model.

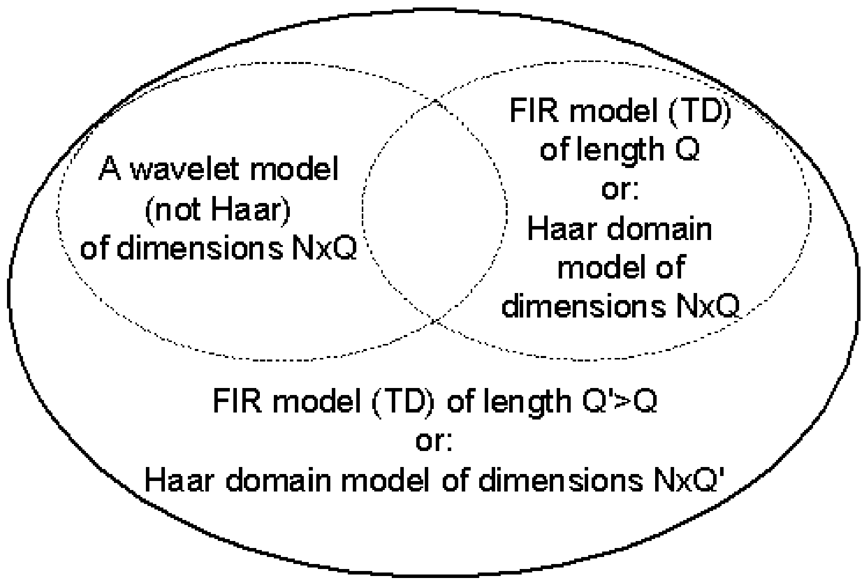

Fig. 7. Sub-domains of LPTV models.

In Figure 7, one sub-domain contains all LPTV FIR systems of length less than or equal to Q. It can be spanned by a raised model or by the Haar base (where $H[k] = H_H^{-1} A[k] H_H$), hence the length of $H[k]$ is the same as that of $A[k]$. The second sub-domain is spanned by another wavelet model (not Haar) of the same order. Since other wavelets are generated by higher order filters, their maximum memory length is larger than their dimension ($Q' = M_{total} \cdot N > Q = M \cdot N$, since here, the length of the resulting $H[k]$ will be larger than M, the length of $A[k]T$ - see eqn. (8)). There are systems which are part of both sub-domains and there are systems that can be modeled exactly only by one of them. If we want to model all systems of length Q' by a raised model, we need a larger number of parameters where the value of Q' depends on the wavelet filter and on N. This model spans the bigger sub-domain (circled by a thick solid line) in the figure. Hence, the choice of the wavelet clearly affects the resulting modeling error (or, in term of adaptive identification, the steady state performance). Other considerations for choosing the mother wavelet and the decomposition depth are processing delay time and complexity.

4 Applying the new model to adaptive identification of LPTV systems

The adaptive identification scheme we propose for our model is described in Figure 8. As we see, there are two types of errors one could consider for the adaptation algorithm - the error in the time domain, e, or the error in the transform domain $\underline{W}_e$. Note that we need to delay the system output y to compare it with its estimated value $\widehat{y}$.

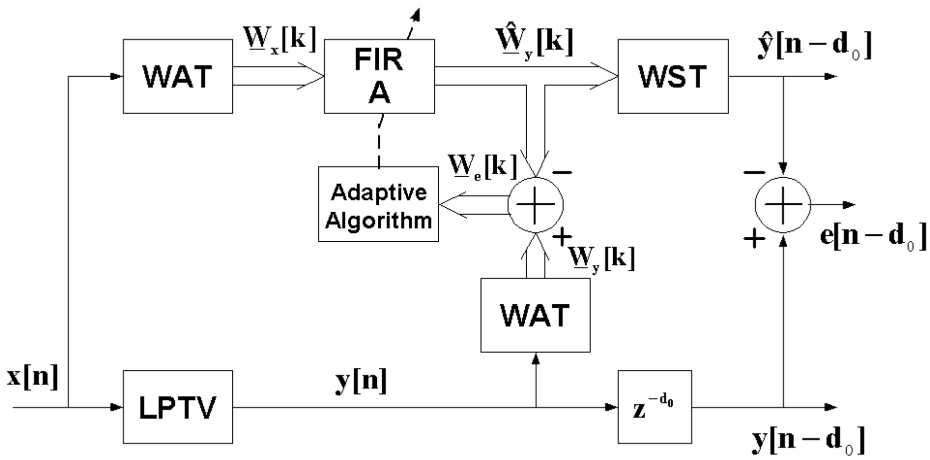

Fig. 8. The suggested adaptive scheme.

With the two errors one can define two corresponding measures of performance

$$J_{MS}\left[k\right] \triangleq \frac{1}{N} E\left\{\sum_{i=0}^{N-1} \left|e\left[kN-i\right]\right|^2\right\} \tag{9}$$

where $e[n] = y[n] - \widehat{y}\left[n\right]$ or

$$J_{W_{MS}}\left[k\right] \triangleq \frac{1}{N} E\left\{\sum_{i=1}^{N} \left|We_i\left[k\right]\right|^2\right\} \tag{10}$$

where $\underline{W}e\left[k\right] = \underline{W}y\left[k\right] - \underline{\widehat{W}}y\left[k\right]$ are the errors calculated in the wavelet domain (see Figure 8). Since we are assuming orthonormal filter bank the two criteria result in the same optimal choice for the FIR filter and, from Parseval's equality, the same optimal value. However, the corresponding adaptive algorithms will be different and we have chosen the latter (as seen in Figure 8).

Clearly, different choices of wavelets will result in different optimal solutions and, as discussed earlier, depending on the modeled LPTV system, one could be better than another.

While we have used the Mean Squared Error (MSE) as our performance criteria (resulting in the LMS as the adaptive algorithm), one could consider using Least Squares (LS), see eqn. (11), as the criteria (resulting in the RLS algorithm). This, as we well know, is a faster converging algorithm at the price of increased computation load.

$$J_{W_{LS}}[k] = \sum_{j=0}^{k} \underline{W}e[j]^T \underline{W}e[j] \tag{11}$$

Back to eqn. (10) and Figure 8 we observe that,

$$\min_{\substack{M-1 \\ \cup \\ \ell=0}A_\ell} (J_{W_{MS}}[k]) = \sum_{i=1}^{N} \min_{\substack{M-1 \\ \cup \\ \ell=0}A_\ell(i,:)} (J^i_{W_{MS}}[k])$$

where

$$J^i_{W_{MS}}[k] \triangleq E\left\{|We_i[k]|^2\right\} \ \forall i \in [1, N] \tag{12}$$

This is because every $We_i[k]$ is influenced by disjoint sets of parameters - the collection of all the ith rows of the matrix coefficients A_ℓ, of the MIMO FIR filter. The result is N parallel independent joint processing estimators based on the LMS algorithm. Note that the adaptive algorithm operates at a rate of $\frac{F_s}{N}$ samples per second.

We make the following claim:

Claim 1 *For any unknown LPTV system with period N and stationary input signal $x[n]$, the optimal solution, $[A_\ell]_{op}$ and the minimum value of $J_{W_{MS}}$ are given by:*

$$\begin{bmatrix} A_0 \ A_1 \ \cdots \ A_{M-1} \end{bmatrix}_{op} = \widetilde{P}_W \left(\widetilde{R}_{W_x}\right)^{-1} \left(\in \Re^{N \times NM}\right) \tag{13}$$

and

$$[J_{W_{MS}}]_{op} = tr\left[R_{W_y}[0]\right] - tr\left[\widetilde{P}_W \left(\widetilde{R}_{W_x}\right)^{-1} \widetilde{P}_W^T\right] \tag{14}$$

where

$$\widetilde{P}_W = \begin{bmatrix} P_W[0] \ P_W[1] \ \cdots \ P_W[M-1] \end{bmatrix} \left(\in \Re^{N \times NM}\right)$$

$$\widetilde{R}_{W_x} = \begin{bmatrix} R_{W_x}[0] & R_{W_x}[1] & \cdots & R_{W_x}[M-1] \\ R_{W_x}[1]^T & R_{W_x}[0] & \cdots & R_{W_x}[M-2] \\ \vdots & \vdots & \ddots & \vdots \\ R_{W_x}[M-1]^T & R_{W_x}[M-2]^T & \cdots & R_{W_x}[0] \end{bmatrix} \left(\in \Re^{NM \times NM}\right)$$

$$P_W[\ell] = E\left\{\underline{W}_y[k]\underline{W}_x[k-\ell]^T\right\} \left(\in \Re^{N \times N}\right)$$

$$R_{W_x}[\ell] = E\left\{\underline{W}_x[k]\underline{W}_x[k-\ell]^T\right\} \left(\in \Re^{N \times N}\right)$$

and

$$R_{W_y}[0] = E\left\{\underline{W}_y[k]\underline{W}_y[k]^T\right\} \left(\in \Re^{N \times N}\right)$$

Proof. *By straight forward substitutions and calculations.* ■

Remark 7 *The above result applies also to cyclo-stationary input signals with period N, since then, the regression vector $\underline{W}[k]$ is still stationary.*

Remark 8 *In view of Remark 5 it can readily be shown that if the LPTV system is such that for a particular choice of wavelet (see eqn. (3)) $H[k] = H_{IDWT}[k] * A^*[k] * H_{DWT}[k]$ and $A^*[k]$ is of length M, then $[A_\ell]_{op} = A^*[\ell]$.*

To write the N LMS equations explicitly we need to modify our notation. Let u_i denote the ith column of the $N \times N$ identity matrix and

$$C_i[k]^T \triangleq u_i^T \begin{bmatrix} A_0[k] & A_1[k] & \cdots & A_{M-1}[k] \end{bmatrix} \left(\in \Re^{1\times NM}\right) \tag{15}$$

Namely, $C_i[k]^T$ consists of all the $i-th$ rows of the matrices $A_\ell[k]$ and let

$$\widetilde{\underline{W}}_x[k] \triangleq \begin{bmatrix} \underline{W}_x[k] \\ \underline{W}_x[k-1] \\ \vdots \\ \underline{W}_x[k-M+1] \end{bmatrix} \left(\in \Re^{NM}\right) \tag{16}$$

Then

$$We_i[k] = Wy_i[k] - \underline{C}_i[k]^T \widetilde{\underline{W}}_x[k] \tag{17}$$

and, for each index $i \in [1, N]$, the corresponding LMS equation is

$$\underline{C}_i[k+1] = \underline{C}_i[k] + DWe_i[k]\widetilde{\underline{W}}_x[k] \tag{18}$$

To use the "almost diagonalization" feature of the wavelet transform wisely (which means, in terms of our notation, that $\widetilde{R}_{W_x}$ is almost diagonal), we introduce a diagonal matrix D, instead of a scalar step size, in the LMS equations (similar to the way it is done for adaptive identification of LTI systems [15]-[17]). Specifically,

$$D \triangleq diag(\mu_1^{Wav}, ..., \mu_{M\cdot N}^{Wav}) \tag{19}$$

where $\mu_1^{Wav}, ..., \mu_{M\cdot N}^{Wav}$ are the different step sizes and

$$\mu_i^{Wav} \triangleq \frac{\mu^{Wav}}{E\{\left(\widetilde{\underline{W}}_x[k]\right)_i^2\}} \tag{20}$$

with μ^{Wav} a common value.

Note that the regression vector is the same for all N parallel LMS algorithms, hence, so is the matrix D, which is chosen according to the statistical nature of the input signal.

When the input signal is stationary, the step sizes which relate to the same scale are equal. Then, there are only $L+1$ different values to be determined (a single value per scale). The convergence proof of the modified LMS can be found in [16]. Without this modification (namely, taking $\mu_1^{Wav} = ... = \mu_{M\cdot N}^{Wav}$) the advantage of the wavelet model for adaptive implementation is not utilized.

5 Experiments

In order to illustrate some of the ideas discussed here, Monte-Carlo simulations have been carried out and a sample of the results is presented. The experimental setup is described in Subsection 5.1 and the results in Subsection 5.2.

5.1 *Experimental setup*

We start by a general description of the experiments set up. Two distinct LPTV systems to be identified were chosen and three possible models for each.

5.1.1 *The unknown systems and the input signal*

Both LPTV systems used for our experiments are with a period $N = 8$ (i.e. $L = 3$). $h\,[n, \ell]$ for these systems is given in Figures 9 and 10. Note that these inpulse responses are presented in the ranges $0 \leq n \leq 48$ (or $0 \leq n \leq 64$) and $0 \leq \ell \leq 7$ and extend periodically as $h\,[n+8, \ell+8] = h\,[n, \ell]$. In both case the resulting $H[k]$ (of eqn. (3)) is FIR. For the first system we get $M_1 = 7$ and for the second $M_2 = 9$. In our experiments we want to illustrate two aspects of the advantages in using our wavelet model - faster convergence of the adaptive algorithm and potentially, improved modeling capability. So, for the first system, the model we use will have the same M and we compare convergence rates (here, we add some measurement noise, 30dB SNR, just so that we do not get zero steady state MSE). For the second system we get under modeling and the emphasis will be on comparison of the optimal solutions and the steady state errors.

In both cases the input $x[n]$, is a stationary Moving Average (MA) process of order 48. The frequency response of the input coloring filter (designed by Remez) is drawn in Figure 11.

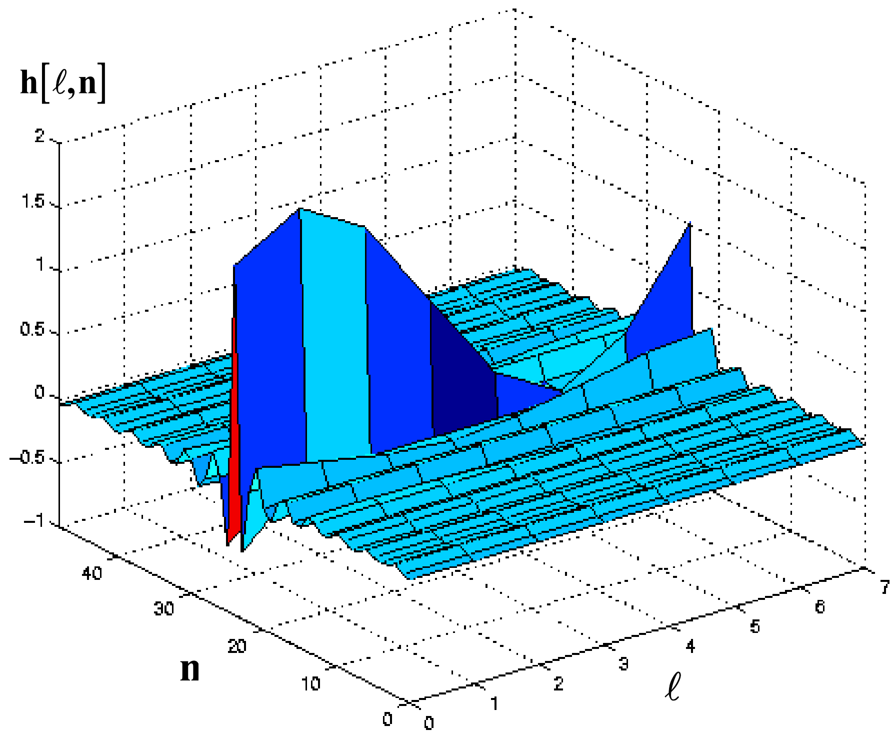

Fig. 9. A period of the impulse response for the first system.

5.1.2 The models and adaptive algorithms

We consider three possible models for each system, resulting in three adaptive identification schemes. In the first we assume that WAT and WST consist of only PP and IPP blocks (this is the 'raised model' [14]). The second is based on the Haar wavelet and the third on the 'db2' wavelet (see e.g. [38]). In all models, as said earlier, we choose $M = 7$. Hence, each identification algorithm consists of 8 parallel adaptations updating 56 parameters each at the rate 8 time slower than the rate of the input $x[n]$.

In the raised model all LMS elements are controlled by the same scalar step size, μ^{TD} (TD refers to time domain since in this case we do not use wavelets). The advantages of this model are its simple structure and low complexity. However, in the case of a stationary colored input signal its convergence rate might be too slow. The main reason for that is that the resulting regression vector (the input and its shifts) has correlated entires resulting in a large eigenvalue spread of its autocorrelation matrix and this is known to slow the LMS algorithm convergence [2]. We have already mentioned that convergence rate in LPTV system identification is more acute a problem since the number of iterations needed for convergence is multiplied by N, the period length of

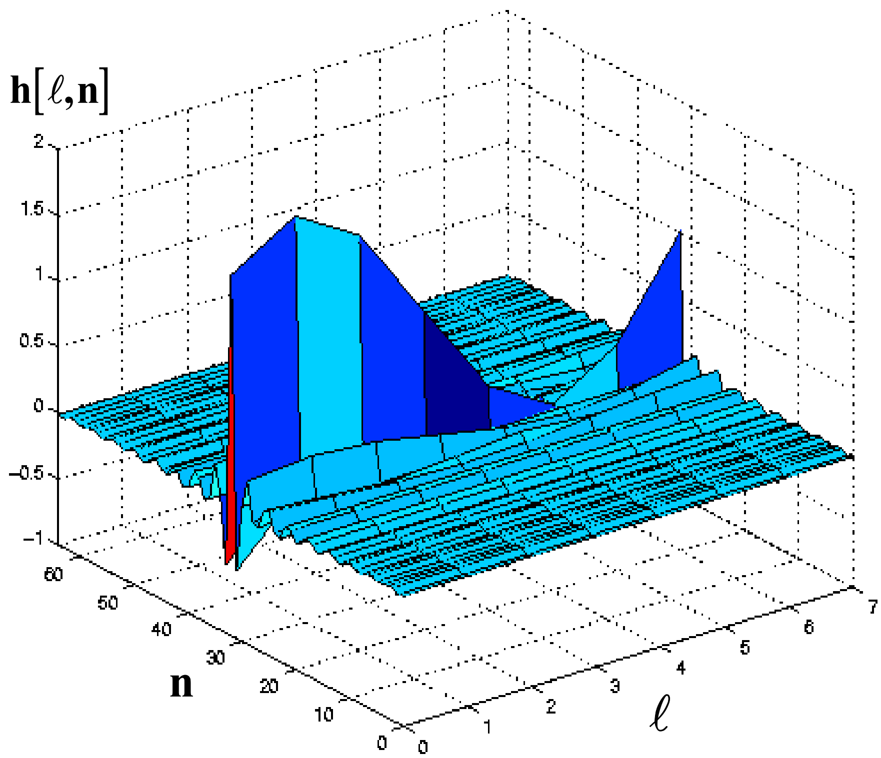

Fig. 10. A period of the impulse response for the second system.

the system.

Aiming to speed up convergence we use the new wavelet models and the corresponding algorithms. The model based on Haar wavelet, as we pointed out earlier, differs from the raised model only by a coordinate transformation (namely, $\underline{W}_x^{Haar}[k] = H\underline{X}[k]$ and $\underline{Y}[k] = H^{-1}\underline{W}_y^{Haar}[k]$ for some constant matrix H). Hence, its modeling capabilities are the same as the raised model. However, it does provide some improved resolution in frequency, hence we may expect to see some improvement in convergence speed over the raised model. This should be further enhanced with a longer wavelet filter ('db2') since the frequency resolution is improved, with an improved modeling capability too. We use decomposition depth $L = 3$ and MIMO FIR with memory length $M = 7$ with respect to the low rate, i.e. we have $7 \cdot 8 \cdot 8 = 448$ parameters in each model adaptively modified in each one of the algorithms according to eqns. (15)-(18).

For all LMS elements we use the same step size, denoted as μ^{Wav}. It is also denoted as μ^{Haar} and μ^{db2} for Haar and 'db2' cases. The values for μ_i^{wav} are chosen as suggested in Equation (20).

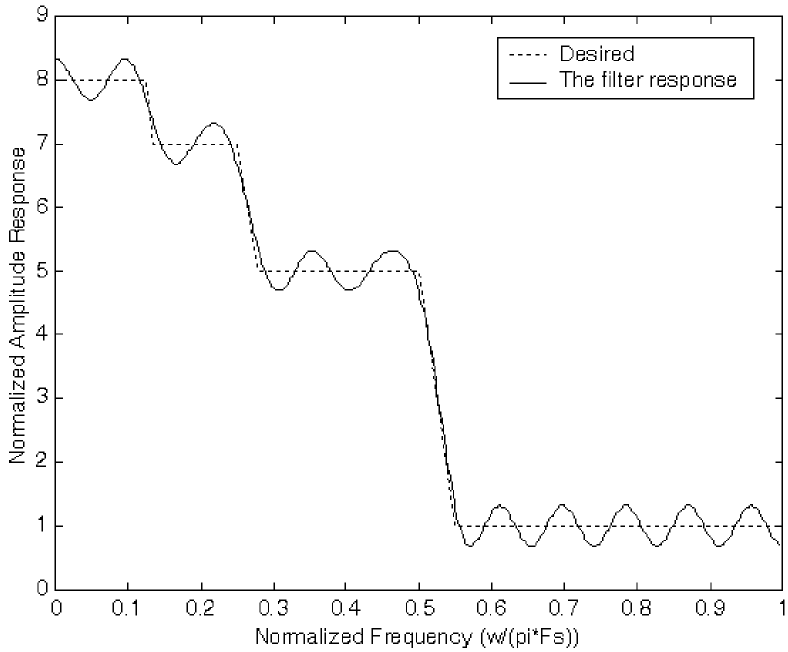

Fig. 11. The coloring filter designed by Remez, order 48.

The step sizes, μ^{TD}, μ^{Haar} and μ^{db2} are chosen in a way that the comparison between the algorithms is fair. The time domain step size, μ^{TD} is chosen to produce relative fast convergence. The wavelet domains step sizes, μ^{Haar} and μ^{db2} are chosen in a way that:

$$MSE_{ss}\left(TD\right) \geq MSE_{ss}\left(Haar\right) \geq MSE_{ss}\left('db2'\right) \tag{21}$$

where $MSE_{ss}\left(\cdot\right)$ is the Mean Square Error (MSE) in steady state for every one of the identification algorithms.

5.2 Results

We have carried out a total of six experiments–one per system per model. Each run consists of 80,000 data points (10,000 periods) and for each run we calculated the $MSE[k]$ according to

$$MSE\left[k\right] = \frac{1}{8}\sum_{i=1}^{8}\left|e\left[8k+i\right]\right|^2 \tag{22}$$

The number of Monte-Carlo experiments is 500 in each case and we present the ensemble average, $\overline{MSE[k]}$, results for each experiment.

The results are grouped according to the LPTV system identified. For each system we apply the three models and the corresponding algorithms. The results are then compared using two criterions. The first criterion is the steady-state MSE, MSE_{ss} and the second criterion is the convergence rate–the number of samples the algorithms takes to reach an MSE which is two times larger than MSE_{ss}.

5.2.1 *Sufficient order identification*

The results of applying the three models under consideration to the first unknown system are shown in Figure 12. The step size for the raised model algorithm, μ^{TD} , is 0.015 and the step sizes for the wavelets, μ^{Haar} and μ^{db2} are both 0.013.

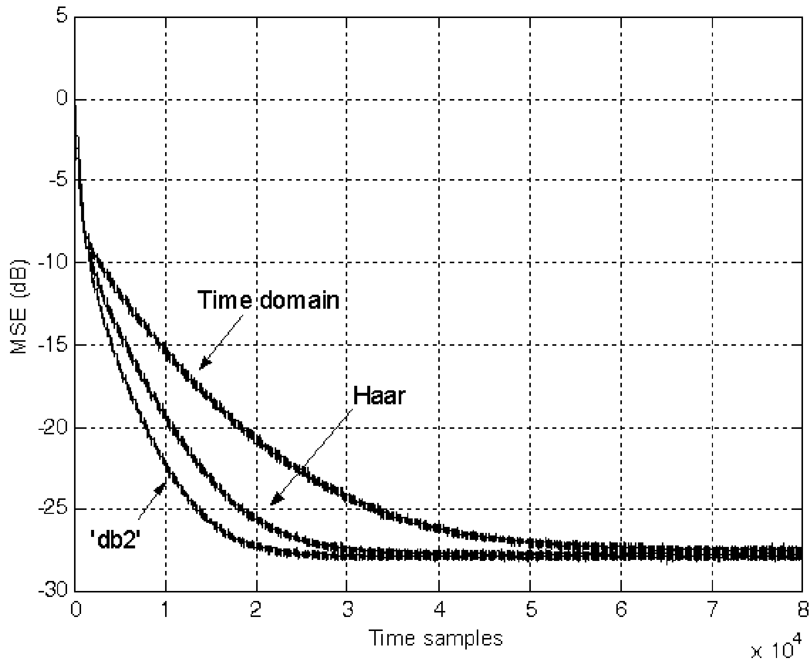

Fig. 12. Convergence curves for the sufficient order identification.

It can be observed that the raised model algorithm converges very slowly. The wavelet algorithms significantly speed up convergence, while steady-state error is also improved a little. The results confirm our earlier statements. The wavelet 'db2' yields better results than Haar thanks to its better frequency localization. We note that for 'db2' the result of lower steady-state error is not fundamental, but depends on the specific unknown system. Convergence might be even faster if we choose larger step size or longer wavelet filters, but the price might be a larger steady-state MSE and/or larger complexity and delay.

The results of this subsection are summarized in the following table:

	MSE_{ss}	rate (#samples till MSE=0.5MSE_{ss})
raised	-27.4680	29,736
Haar	-27.7699	17,512
'db2'	-27.9128	13,024

The values of $MSE_{ss}$are quite close, but the differences in convergence rates are significant. 'db2' converges about 2.3 times faster than the raise model algorithm.

5.2.2 Insufficient order identification

The results of applying the three models under consideration to the second unknown system are shown in Figure 13. The step size for the raise model algorithm, μ^{TD} is 0.015. The step size for the wavelet Haar, μ^{Haar} is 0.01 and for 'db2', μ^{db2} is 0.015.

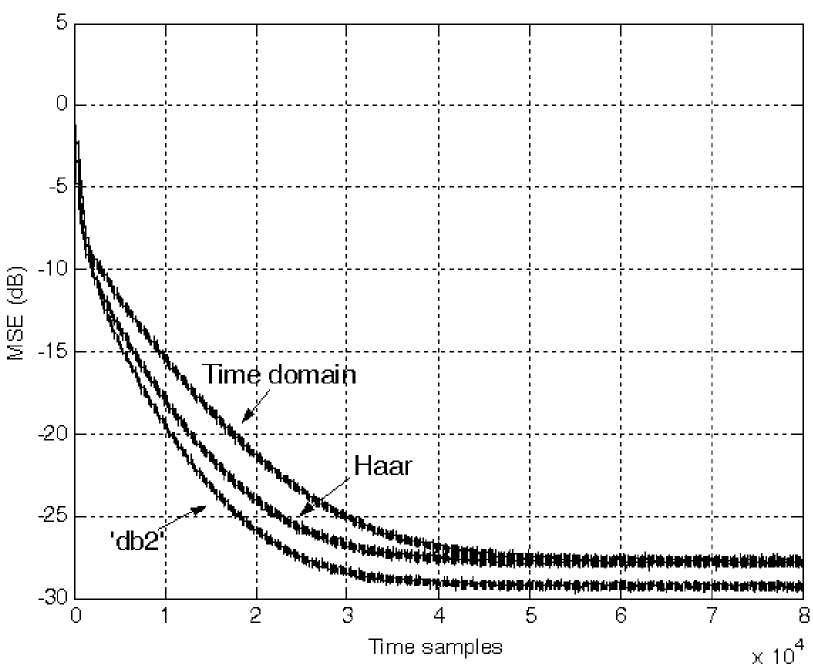

Fig. 13. Convergence curves for the insufficient order identification.

For this case we see that both the convergence rate and the steady state MSE are improved when the 'db2' model is used. The results of this subsection are summarized in the following table:

	$Jopt$	MSE_{ss}	rate (#samples till MSE=0.5MSE_{ss})
raised	-33.7	-27.782	28,168
Haar	-33.7	-27.8103	21,232
'db2'	-44.3	-29.263	20,400

We note here a significantly smaller optimal MSE, $Jopt$, in the 'db2' case. This clearly indicates that by tuning μ^{db2} one could improve steady state performance while still maintaining some of the convergence rate advantage over the raised model.

6 Conclusions

In this paper we first presented a generic model for LPTV systems by recognizing the commutative diagram in Figure 14, where $\widetilde{T}$ is a generic periodic transform (with the same period as the LPTV system). This opens a wide range of possible models for LPTV systems. On the other hand, when adaptive implementation is considered it is desired to preprocess the data to achieve (almost) diagonal data autocorrelation matrix. Recognizing that DWT has both of the above properties made it an excellent choice for our purposes.

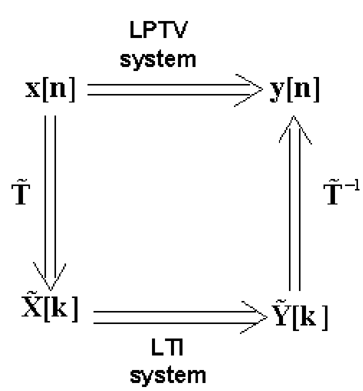

Fig. 14. A commutative diagram for LPTV systems.

Making use of the structure suggested by Mallat we also utilized the benefits of an efficient implementation of the DWT. Our modification of this basic structure coined as WAT and WST, significantly increases the utility of wavelets as a modeling vehicle for LPTV systems. While limiting the LTI part to an FIR system, as was demonstrated in our simulations, the use of wavelets has the potential of modeling a larger impulse response with a relatively small number of parameters.

In the context of adaptive implementation, the use of wavelets provides a wide spectrum of possibilities between the two extremes: Time domain (in our case, the raised model) and frequency domain. With each choice of wavelet we trade off frequency resolution (i.e. speed of convergence) with time resolution (i.e. time delay). Accordingly, the choice of a particular wavelet (or equivalently, the underlying filter), reflects a particular desired trade off combined with computation cost considerations.

References

[1] E. Eweda, "Comparison of RLS,LMS, and sign algorithms for tracking randomly time-varying channels," *IEEE trans. on signal processing*, vol. 42, pp. 2937–2944, November 1994.

[2] S. Haykin, *Adaptive Filter Theory.* Prentice-Hall, 1986.

[3] J. M. McCool and B. Widrow, "Principles and applications of adaptive filters: A tutorial review," *Proceedings of the 1980 IEEE International Symposium on Circuits and Systems*, pp. 1143–57, 1980.

[4] M. I. Doroslovacki and H. H. Fan, "Wavelet-based linear system modeling and adaptive filtering," *IEEE trans. on signal processing*, vol. 44, pp. 1156–1167, May 1996.

[5] M. K. Tsatsanis and G. B. Giannakis, "Time varying system identification and model validation using wavelets," *IEEE trans. on signal processing*, vol. 41, pp. 3512–3523, December 1993.

[6] E. Karlsson, "RLS polynomial lattice algorithms for modelling time-varying signals," *ICASSP 91: 1991 International Conference on Acoustics, Speech and Signal Processing*, vol. 5, pp. 3233–3236, 1991.

[7] M. Chakraborty and S. Prasad, "Multichannel time-varying ARMA model identification by least squares circular lattice structures," *ICASSP 91: 1991 International Conference on Acoustics, Speech and Signal Processing*, vol. 5, pp. 3229–3232, 1991.

[8] D. C. McLernon, "Parameter estimation of autoregressive processes with periodic coefficients," *1989 IEEE International Symposium on Circuits and Systems*, vol. 2, pp. 1350–1353, 1989.

[9] H. S. S. Ohno and H. Yoshida, "Adaptive blind equalization of multichannel FIR systems," *Proceedings of the 1999 IEEE International Symposium on Circuits and Systems VLSI*, vol. 3, pp. 66–69, 1999.

[10] S. S. K. Itamiya and K. Ando, "Stable model reference adaptive control for periodically LTV systems based on the maximal error evaluation," *Int. J. Control*, vol. 58, no. 1, pp. 51–64, 1993.

[11] J. L. Speyer and R. T. Evans, "A second variational theory for optimal periodic processes," *IEEE trans. on automatic control*, vol. AC-29, pp. 138–148, February 1984.

[12] C. T. D. Konig and J. F. Bohme, "Design of optimum periodic time varying filters for applications in combustion diagnosis of car engines," *1995 International Conference on Acoustics, Speech, and Signal Processing. Conference Proceedings*, vol. 3, pp. 1924–1927, 1995.

[13] C. R. Knospe and J. K. Havilland, "Pulse response method for vibration reduction in periodic dynamic systems," *Engineering notes*, vol. 15, no. 3, pp. 782–785, 1991.

[14] A. Feuer and G. C. Goodwin, *Sampling in Digital Signal Processing and Control*. Birkhauser, 1996.

[15] S. Hosur and A. H. Tewfik, "Wavelet transform domain adaptive FIR filtering," *IEEE trans. on signal processing*, vol. 45, pp. 617–630, March 1997.

[16] N. Erdol and F. Basbug, "Wavelet transform based adaptive filters: Analysis and new results," *IEEE trans. on signal processing*, vol. 44, pp. 2163–2171, September 1996.

[17] L. M. Dang, "Wavelet transform adaptive filtering," *SPIE*, vol. 2303, pp. 508–517, 1994.

[18] R. Merched and A. H. Sayed, "An embedding approach to frequency-domain and subband adaptive filtering," *IEEE trans. on signal processing*, vol. 48, pp. 2607–2619, September 2000.

[19] B. F. Boroujeny and S. Gazor, "Generalized sliding FFT and its application to implementation of block LMS adaptive filters," *IEEE trans. on signal processing*, vol. 42, pp. 532–538, March 1994.

[20] A. Feuer and R. Cristi, "On the steady state performance of frequency domain LMS algorithms," *IEEE Trans. on Signal Processing*, vol. 41, no. 1, pp. 419–423, 1993.

[21] P. M. G. G. Panda, A. M. Alvarez and C. F. N. Cowan, "A transform domain circular convolution for adaptive filtering," *IEEE trans. on acoustics, speech and signal processing*, vol. ASSP-35, pp. 1217–1220, August 1987.

[22] W. B. Mikhael and A. S. Spanias, "Comparison of several frequency-domain LMS algorithms," *IEEE trans. on circuits and systems*, vol. CAS-34, pp. 586–588, May 1987.

[23] J. C. Lee and C. K. Un, "Block realization of multirate adaptive digital filters," *IEEE trans. on acoustics, speech and signal processing*, vol. ASSP-34, pp. 105–117, February 1986.

[24] W. B. Mikhael and A. S. Spanias, "Performance enhancement of the frequency-domain LMS adaptive algorithm," *1986 IEEE international symposium on circuits and systems, New York*, vol. 1, pp. 349–352, May 1986.

[25] J. C. Lee and C. K. Un, "A reduced structure of the frequency-domain block LMS adaptive digital filter," *Proceedings of the IEEE*, vol. 72, pp. 1816–1818, December 1984.

[26] A. M. P. S. S. Narayan and M. J. Narasimha, "Transform domain LMS algorithm," *IEEE trans. on acoustics, speech and signal processing*, vol. ASSP-31, pp. 609–615, June 1983.

[27] D. Mansour and J. A. H. Gray, "Unconstrained frequency-domain adaptive filter," *IEEE trans. on acoustics, speech and signal processing*, vol. ASSP-30, pp. 726–734, October 1982.

[28] S. K. M. G. A. Clark and S. R. Parker, "Block implementation of adaptive digital filters," *IEEE trans. on circuits and systems*, vol. CAS-28, pp. 584–592, June 1981.

[29] F. A. Reed and P. L. Feintuch, "A comparison of LMS adaptive cancellers implemented in the frequency domain and the time domain," *IEEE trans. on circuits and systems*, vol. CAS-28, pp. 610–615, June 1981.

[30] J. M. M. Dentino and B. Widrow, "Adaptive filtering in the frequency domain," *Proceedings of the IEEE*, vol. 66, pp. 1658–9, December 1978.

[31] G. C. L. Hong and C. K. Chui, "A filter-bank-based kalman filtering technique for wavelet estimation and decomposition of random signals," *IEEE trans. on circuits and systems-II: analog and digital signal processing*, vol. 45, pp. 237–241, February 1998.

[32] M. R. Petraglia and S. K. Mitra, "Performance analysis of adaptive filter structures based on subband decomposition," *Proceedings of the Int. Symp. On Circ. and Systems*, pp. I. 60–3, June 1993.

[33] Y. Lu and J. M. Morris, "Gabor expansion for adaptive echo cancellation," *IEEE signal processing magazine*, pp. 68–80, March 1999.

[34] T. Fredrick and N. Erdol, "Subband equalizer performance for intersymbol-interference channels," *SPIE wavelet applications*, vol. 2242, pp. 69–77, 1994.

[35] A. Gilloire and M. Vetterli, "Adaptive filtering in subbands with critical sampling: Analysis, experiments, and application to acoustic echo cancellation," *IEEE trans. on signal processing*, vol. 40, pp. 1862–1875, August 1992.

[36] Y. Dorfan, "Modeling and identification of LPTV systems by wavelets," research thesis, The technion - IIT, Haifa, Israel, December 2000.

[37] N. Erdol and F. Basbug, "Use of the wavelet transform in optimal receiver design," *Applied Optics*, vol. 33, pp. 5303–7, August 1994.

[38] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding.* Prentice-Hall, 1995.

[39] R. K. Young, *Wavelet Theory and its Applications.* Kluwer academic publishers Massachusetts, 1993.

[40] S. Mallat, *A Wavelet Tour of Signal Processing.* Academic Press, 1999.