Espinosa-Pulido et al. EURASIP Journal on Advances in Signal Processing 2012, **2012**:256
http://asp.eurasipjournals.com/content/2012/1/256

EURASIP Journal on Advances in Signal Processing
a SpringerOpen Journal

RESEARCH Open Access

Widely linear Markov signals

Juan A Espinosa-Pulido, Jesús Navarro-Moreno, Rosa M Fernández-Alcalá* and Juan C Ruiz-Molina

Abstract

The insufficiency to guarantee the existence of a state-space representation of the classical wide-sense Markov condition for improper complex-valued signals is shown and a generalization is suggested. New characterizations for wide-sense Markov signals which are based either on second-order properties or on state-space representations are studied in a widely linear setting. Moreover, the correlation structure of such signals is revealed and interesting results on modeling in both the forwards and backwards time directions are proved. As an application we give some recursive estimation algorithms obtained from the Kalman filter. The performance of the proposed results is illustrated in a numerical example in the areas of estimation and simulation.

Keywords: Modeling, Wide-sense Markov signals, Widely linear processing

1 Introduction

Markov signals are characterized by the condition that future development of these signals depends only on current states and not their history up to that time. In general, Markov processes are easier to model and analyze, and they do include interesting applications. Among others, estimation and detection are areas of signal processing where this kind of process has provided efficient solutions (see, e.g., [1,2]). Non-Markov processes in which the future state of a process depends on its whole history are generally harder to analyze mathematically [3]. In linear minimum-mean square error (MMSE) estimation theory, when the processes under consideration are not Gaussian, the classes of stochastic processes which are of practical importance are wide-sense Markov (WSM) processes. The concept of WSM signal is easier to check than the condition of (strictly) Markov since it involves only second-order characteristics [4]. In general, WSM processes (with the exception of Gaussian processes) are not Markov in the strict sense. The equivalence between the WSM condition and the state-space representation for the signal is really what makes WSM signals especially attractive in signal processing [1].

Widely linear (WL) processing is an emerging research area in the complex-valued signal analysis which gives significant performance gains with respect to strictly linear (SL) processing (excellent account of the topic and the literature can be found in [5,6]). It has proved to be a more useful approach than SL processing since complex-valued random signals are in general improper (i.e., they are correlated with their complex conjugates). Thus, the improper nature of most signals forces us to consider the so-called augmented statistics to entirely describe their second-order properties. Using augmented statistics means incorporating in the analysis the information supplied by the complex conjugate of the signal and examining properties of both the correlation and complementary correlation functions. SL processing operates ignoring this last function. Some areas of signal processing in which the treatment of improper signals by using a WL processing has proved to be beneficial are estimation [5-11], detection [12], modeling [8], and simulation [13].

A general characteristic of the articles devoted to studying WSM complex-valued signals is that they use a SL processing approach (see e.g., [1,14-16]). We will show by means of simple examples that the classical definition and the associated characterizations of WSM signals are incorrect for improper signals. The examples then motivate the extension of the concept of WSM signal to a WL setting and the study of new characterizations. Specifically, we introduce the concept of widely linear Markov (WLM) signals and we give different characterizations based either on second-order properties or on state-space representations from a WL processing point of view. The analysis is performed in both the forwards and backwards directions of time. We also provide a way to check

*Correspondence: rmfernan@ujaen.es
Department of Statistics and Operations Research, University of Jaén, Campus Las Lagunillas, 23071 Jaén, Spain

the WLM condition, similar to the well-known triangular property, based on augmented statistics and determine the correlation structure of WLM signals. The modeling part is the focus of this article. In this sense, WL forwards and backwards Markovian representations are suggested, the interrelation between them is studied and the connection with the WL autoregressive representations defined in [8] is established. These Markovian representations also become a starting point for the application of different recursive estimation algorithms. Thus, the application of the Kalman filter on the forwards and backwards representations yields different WL prediction, filtering and smoothing algorithms. The point, which is illustrated in an example, is that besides the well-known performance gain of the WL approach we also get more realistic results in simulation and modeling.

The article is organized as follows. In Section 2, we present some background material on complex-valued Markov signals, illustrate the incapacity of the usual WSM condition in order to characterize the state-space representation for improper signals and suggest the concept of WLM signal. Some preliminary characterizations are also given. Section 3 studies the correlation structure of WLM signals. In Section 4, we discuss the modeling problem for WLM signals and analyze the stationary case. The estimation problem is treated in Section 5. We apply our results in the fields of signal simulation and estimation by considering a numerical example in Section 6. A Section of conclusions ends the article. To preserve continuity in our presentation, all proofs are deferred to an Appendix 1.

2 Preliminaries

In this section, we give the main definitions, notation and auxiliary results. We also present two examples which motivate the necessity of the new concept introduced.

Bold capital letters will be used to refer to matrices and bold lower-case letters will be used to refer to vectors. The row j of any matrix $\boldsymbol{A}(\cdot)$ will be denoted by $\boldsymbol{A}_{[j]}(\cdot)$, the n-vector of zeros by $\mathbf{0}_n$ and the $n \times m$-matrix of zeros by $\mathbf{0}_{n\times m}$. Furthermore, the superscripts $*$, T, and H represent the complex conjugate, transpose, and complex transpose, respectively.

Let $\{x_t, t \in \mathbb{Z}\}$ be a zero-mean complex random signal with correlation function $r(t,s) = E[x_t x_s^*]$ and complementary correlation function $c(t,s) = E[x_t x_s]$. Most of the results in this article are valid for nonstationary signals. However, for some of them the stationary condition is necessary. The signal x_t is said to be of second-order wide-sense stationary (SOS) if the functions $r(t,s)$ and $c(t,s)$ depend on $t-s$. A zero-mean stochastic process w_t is called a doubly white noise if $E[w_t w_s^*] = e_1\delta(t-s)$ and $E[w_t w_s] = e_2\delta(t-s)$ with $|e_2| \le e_1$ (see [8] for a complete study of their characteristics). The linear MMSE estimator of x_t based on the set of observations $\{x_{t_1}, x_{t_2}, \ldots, x_{t_m}\}$ will be denoted by $\hat{x}(t|t_1, t_2, \ldots, t_m)$ and we will refer to it as the SL estimator.

The Markov condition on a signal $\{x_t, t \in \mathbb{Z}\}$ establishes the following identity for the conditional probability:

$$P(x_t \le x|x_{t_1}, x_{t_2}, \ldots, x_{t_m}) = P(x_t \le x|x_{t_1})$$

for all x and $t > t_1 > \cdots > t_m$. Doob [4] introduced a weaker concept based on the SL estimator which has received great attention in the literature (e.g., [1,14-16]). A signal x_t is called WSM if $\hat{x}(t|\tau \le s) = \hat{x}(t|s)$ for any $s < t$. Such signals have remarkable properties. For example, Beutler [14] showed that a signal x_t is WSM if, and only if, the function $\bar{k}(t,s) = r(t,s)r^{-1}(s,s)$ has the triangular property, i.e.,

$$\bar{k}(t,s) = \bar{k}(t,\tau)\bar{k}(\tau,s), \qquad t \ge \tau \ge s \tag{1}$$

Another characterization in terms of so-called Markovian state-space models can be found in [1]. They showed that a signal $\{x_t, t \ge 0\}$ is WSM if, and only if, it has a state-space model of the form

$$x_{t+1} = \bar{k}(t+1,t)x_t + u_t \tag{2}$$

where u_t is a white noise uncorrelated with x_0. Doob's definition was later generalized in [16] in the following sense: x_t is a WSM signal of order $n \ge 1$ if $\hat{x}(t|\tau \le s) = \hat{x}(t|s, s-1, \ldots, s-n+1)$ for any $s < t$. The authors also studied the second-order properties of such signals.

All these studies have a common characteristic: the information supplied by the complementary correlation function is ignored, i.e., the results are derived assuming implicitly that the signal is proper ($c(t,s) = 0$). As noted above, nowadays, the research activity in the field of the complex-valued signal is more and more focused on the better performing and less familiar WL processing. In this setting the SL MMSE estimator is replaced by the WL MMSE estimator, denoted by $\hat{x}^{WL}(t|t_1, t_2, \ldots, t_m)$, which uses the information of the augmented vector of observations $[x_{t_1}, x_{t_1}^*, x_{t_2}, x_{t_2}^*, \ldots, x_{t_m}, x_{t_m}^*]^T$. The immediate question that arises is whether the classical concept of WSM signals remains valid in the WL processing approach. The following two examples give us the answer.

Example 1. *Consider a signal $\{x_t, t \ge 0\}$ with correlation function $r(t,s) = \frac{1}{2}(e^{3|t-s|} + e^{|t-s|})$ and complementary correlation function $c(t,s) = \frac{1}{2}(e^{3|t-s|} - e^{|t-s|})$. It is easy to check that $r(t,s)$ does not satisfy the triangular property (1) and then, the signal cannot be modeled by a representation of the form (2). However, as we will show later, it is possible to find a state-space representation for such a signal given by (26). Thus, the classical WSM condition is clearly insufficient in the improper case to find a state-space representation for the signal involved.*

Example 2. *Assume that* $\{x_t, 1 \leq t \leq 100\}$ *is a signal with correlation and complementary correlation functions given by* $r(t,s) = (t/100 + 1)^{1/6}(s/100)^4$ *and* $c(t,s) = \mathrm{j}(s/100)^4$, *for* $s \leq t$, *respectively, with* $\mathrm{j} = \sqrt{-1}$. *Here, the triangular property (1) holds and then* x_t *has the representation*

$$x_{t+1} = \left(\frac{t+101}{t+100}\right)^{1/6} x_t + u_t \tag{3}$$

with x_t *uncorrelated with* u_t. *However, this model presents two important shortcomings in the WL processing framework: the noise* u_t *is correlated with* x_t^* *and the information supplied by* $c(t,s)$ *is ignored. Both problems can be avoided by considering a more competitive model for* x_t *obtained with the additional information of* x_t^*. *In fact, we can write an alternative state-space representation for* x_t *given by (27). An exhaustive study about the superiority of (27) against (3) is presented in Section 6.*

From these two simple examples we extract the following consequences: the classical definition of a WSM signal must be extended to deal with improper signals, this new concept must be characterized adequately to avoid the drawback shown in Example 1 and new results about modeling are necessary to exploit the information available in both x_t and x_t^* thus attaining better models for the signal as illustrated in Example 2. Next, we introduce such a definition in a WL processing setting.

Definition 1. *A complex-valued signal* $\{x_t, t \in \mathbb{Z}\}$ *is said to be WLM of order* $n \geq 1$, *briefly a WLM(n) signal, if the following condition holds*

$$\hat{x}^{WL}(t|\tau \leq s) = \hat{x}^{WL}(t|s, s-1, \ldots, s-n+1)$$

for any $s < t$.

Notice that this concept extends both the classical notion of WSM introduced by Doob in [4] and the later generalization given in [16].

In the rest of the section, we provide different characterizations of WLM(n) signals. For that, we need to introduce some additional notation. Denote the augmented forwards vector of order $n \geq 1$ of x_t as the $2n$-vector

$$\mathbf{x}_t = [x_t, x_t^*, x_{t-1}, x_{t-1}^*, \ldots, x_{t-n+1}, x_{t-n+1}^*]^T$$

and its correlation function by $\boldsymbol{R}(t,s) = E[\mathbf{x}_t\mathbf{x}_s^H]$. From now on, we assume that $\det\{\boldsymbol{R}_t\} \neq 0$ with $\boldsymbol{R}_t := \boldsymbol{R}(t,t)$. Moreover, we define the normalized correlation function as

$$\boldsymbol{K}(t,s) = \boldsymbol{R}(t,s)\boldsymbol{R}_s^{-1} \tag{4}$$

Similarly, we define the augmented backwards vector of order $n \geq 1$ of x_t as the $2n$-vector

$$\mathbf{x}_t^b = [x_{t+n-1}, x_{t+n-1}^*, x_{t+n-2}, x_{t+n-2}^*, \ldots, x_t, x_t^*]^T$$

The following results establish the relation between the signals x_t and their augmented forwards and backwards versions. We start first with the augmented forwards vector and we give a test similar to (1) for a signal being WLM(n).

Theorem 1. *The following statements are equivalent:*

1. $\{x_t, t \in \mathbb{Z}\}$ *is a WLM(n) signal.*
2. *For* $s < t$, *the WL MMSE estimator of* $\mathbf{x}_t$ *on the basis of the set* $\{\mathbf{x}_\tau, \mathbf{x}_\tau^*, \tau \leq s\}$ *is of the form*

$$\hat{\mathbf{x}}^{WL}(t|\tau \leq s) = \boldsymbol{K}(t,s)\mathbf{x}_s \tag{5}$$

3. *For* $t \geq \tau \geq s$,

$$\boldsymbol{K}(t,s) = \boldsymbol{K}(t,\tau)\boldsymbol{K}(\tau,s) \tag{6}$$

Now, we suggest a characterization based on the augmented backwards vector. This result also shows the independence from the time direction of the Markov condition.

Theorem 2. *The following statements are equivalent:*

1. $\{x_t, t \in \mathbb{Z}\}$ *is a WLM(n) signal.*
2. $\hat{x}^{WL}(t|\tau \geq s) = \hat{x}^{WL}(t|s, s+1, \ldots, s+n-1)$ *for any* $s > t$.
3. *For* $s > t$, *the WL MMSE estimator of* $\mathbf{x}_t^b$ *on the basis of the set* $\{\mathbf{x}_\tau^b, \mathbf{x}_\tau^{b^*}, \tau \geq s\}$ *is of the form*

$$\hat{\mathbf{x}}^{b^{WL}}(t|\tau \geq s) = \boldsymbol{K}(t+n-1, s+n-1)\mathbf{x}_s^b \tag{7}$$

3 Correlation structure of WLM(n) signals

In this section, the second-order properties of a WLM(n) signal $\{x_t, t \in \mathbb{Z}\}$ are analyzed. Specifically, we study the structure of the matrices $\boldsymbol{R}(t,s)$, $\boldsymbol{K}(t,s)$, $\boldsymbol{R}_t$, and $\boldsymbol{K}_t := \boldsymbol{K}(t+1,t)$.

Proposition 1. 1. *The following relations hold:*

$$\boldsymbol{K}_{[2(j+i)-1]}(t+j,t) = \Big[\underbrace{0,\ldots,0}_{2i-2}, 1, \underbrace{0,\ldots,0}_{2(n-i)+1}\Big], \quad j < n, \quad i = 1,\ldots,n-j \tag{8}$$

$$\boldsymbol{K}_{[2(j+i)]}(t+j,t) = \Big[\underbrace{0,\ldots,0}_{2i-1}, 1, \underbrace{0,\ldots,0}_{2(n-i)}\Big], \quad j < n, \quad i = 1,\ldots,n-j \tag{9}$$

$$\boldsymbol{K}_{[2+i]}(t+j+1,t) = \boldsymbol{K}_{[i]}(t+j,t), \quad j \geq 0, \quad i = 1,\ldots,2n-2 \tag{10}$$

$$\boldsymbol{K}_{[1]}(t+j+1,t)=\boldsymbol{K}_{[1]}(t+j+1,t+j)\boldsymbol{K}(t+j,t), \quad j\geq 0 \tag{11}$$

$$\boldsymbol{K}_{[2]}(t+j+1,t)=\boldsymbol{K}_{[2]}(t+j+1,t+j)\boldsymbol{K}(t+j,t), \quad j\geq 0 \tag{12}$$

2. *The matrix* $\boldsymbol{K}_t$ *is of the form*

$$\boldsymbol{K}_t=\begin{bmatrix} k_{1,t} & k_{2,t} & k_{3,t} & k_{4,t} & \cdots & k_{2n-3,t} & k_{2n-2,t} & k_{2n-1,t} & k_{2n,t}\\ k^*_{2,t} & k^*_{1,t} & k^*_{4,t} & k^*_{3,t} & \cdots & k^*_{2n-2,t} & k^*_{2n-3,t} & k^*_{2n,t} & k^*_{2n-1,t}\\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 \end{bmatrix} \tag{13}$$

where $k_{i,t}=k_i(t+1,t)$ *for* $i=1,\ldots,2n$ *and* $k_i(t+1,t)$ *is defined in (28).*

3. *The matrices* $\boldsymbol{R}(t,s)$ *and* $\boldsymbol{K}_t$ *satisfy the recursive equation*

$$\boldsymbol{R}(t+1,s)=\boldsymbol{K}_t\boldsymbol{R}(t,s), \qquad s\leq t \tag{14}$$

which has the solution

$$\boldsymbol{R}(t,s)=\boldsymbol{K}_{t-1}\cdots\boldsymbol{K}_s\boldsymbol{R}_s, \qquad s<t \tag{15}$$

Moreover,

$$\boldsymbol{R}_{t+1}=\boldsymbol{K}_t\boldsymbol{R}_t\boldsymbol{K}_t^H+\boldsymbol{Q}_t$$

where $\boldsymbol{Q}_t$ *is a* $2n\times 2n$*-matrix of the form*

$$\boldsymbol{Q}_t=\left[\begin{array}{c|c} \boldsymbol{A}_t & \boldsymbol{0}_{2\times 2n-2}\\ \hline \boldsymbol{0}_{2n-2\times 2} & \boldsymbol{0}_{2n-2\times 2n-2}\end{array}\right] \tag{16}$$

with

$$\boldsymbol{A}_t=\begin{bmatrix} a_{1,t} & a_{2,t}\\ a^*_{2,t} & a_{1,t}\end{bmatrix}$$

where $a_{1,t}$ *are real positive numbers and* $\boldsymbol{A}_t$ *is nonnegative definite.*

4 Modeling of WLM(n) signals

We aim to provide different ways of modeling for WLM(n) signals. The connection between stationary WLM(n) signals and the autoregressive representations defined in [8] is also established. First, we present a new characterization in which the equivalence between a WLM signal of order n and their forwards and backwards representations is given. Such representations show that a WLM(n) signal depends only on the n preceding or subsequent states and their conjugates.

Theorem 3. *A signal* $\{x_t, 0\leq t\leq m\}$ *is a WLM(*n*) if, and only if, it has the forwards and backwards representations*

$$x_{t+1}=\boldsymbol{k}_t^T\mathbf{x}_t+w_t, \qquad t\geq n-1 \tag{17}$$

$$x_t={\boldsymbol{k}^b}^T_{t+1}\mathbf{x}^b_{t+1}+w^b_{t+1}, \qquad t\leq m-n+1 \tag{18}$$

where $\boldsymbol{k}_t$, $\boldsymbol{k}^b_t$ *are* $2n$*-vectors, and* w_t, w^b_t *are doubly white noises such that*

$$E[\,w_t\mathbf{x}_{n-1}]=\mathbf{0}_{2n}, \qquad t\geq n-1 \tag{19}$$

$$E[\,w^b_t\mathbf{x}^b_{m-n+1}]=\mathbf{0}_{2n}, \qquad t\leq m-n+1$$

Now we state a parallel result to the classical one established for stationary WSM processes and autoregressive representations [16].

Corollary 1. *If* $\{x_t, 0\leq t\leq m\}$ *is a SOS WLM(*n*) signal, then* x_t *is the solution of the WL system defined in [8]*

$$x_{t+1}=\sum_{i=0}^{n-1}g_{1,i}x_{t-i}+\sum_{i=0}^{n-1}g_{2,i}x^*_{t-i}+w_t \tag{20}$$

where $g_{1,i}, g_{2,i}\in\mathbb{C}$, $i=1,\ldots,n-1$, *and* w_t *is a doubly white noise such that* $E[\,w_tw_t^*]=a_1$ *and* $E[\,w_tw_t]=a_2$.

We summarize the previous results in the following steps which provides forwards and backwards models for a WLM(n) signal:

Step 1: Define the $2n$-vector $\boldsymbol{k}_t$ such that $\boldsymbol{k}_t^T$ coincides with the first row of the matrix

$$\boldsymbol{K}_t:=\boldsymbol{R}(t+1,t)\boldsymbol{R}_t^{-1} \tag{21}$$

Similarly, we define the $2n$-vector $\boldsymbol{k}^b_{t+1}$ such that ${\boldsymbol{k}^b}^T_{t+1}$ is equal to the $2n-1$ row of the matrix

$$\boldsymbol{K}^b_{t+1}:=\boldsymbol{K}(t+n-1,t+n)=\boldsymbol{R}(t+n-1,t+n)\boldsymbol{R}^{-1}_{t+n} \tag{22}$$

Step 2: Consider the matrices

$$\boldsymbol{Q}_t=\boldsymbol{R}_{t+1}-\boldsymbol{K}_t\boldsymbol{R}_t\boldsymbol{K}_t^H \tag{23}$$

$$\boldsymbol{Q}^b_{t+1}=\boldsymbol{R}_{t+n-1}-\boldsymbol{K}^b_{t+1}\boldsymbol{R}_{t+n}{\boldsymbol{K}^b}^H_{t+1} \tag{24}$$

Step 3: The signal x_t can be represented by the following forwards and backwards models:

$$x_{t+1}=\boldsymbol{k}_t^T\mathbf{x}_t+w_t, \qquad t\geq n-1$$

$$x_t={\boldsymbol{k}^b}^T_{t+1}\mathbf{x}^b_{t+1}+w^b_{t+1}, \qquad t\leq m-n+1$$

where w_t is a doubly white noise uncorrelated with $\mathbf{x}_{n-1}$ for all $t\geq n-1$ and w^b_t is a doubly white noise uncorrelated with $\mathbf{x}_{m-n+1}$ for all $t\leq m-n+1$. Moreover, $E[\,w_tw_t^*]$ and $E[\,w_tw_t]$ are the

(1,1)-element and (1,2)-element of the matrix $\boldsymbol{Q}_t$, respectively. Similarly, $E[\,w_t^b w_t^{b^*}]$ and $E[\,w_t^b w_t^b]$ are the $(2n-1, 2n-1)$-element and $(2n-1, 2n)$-element of the matrix $\boldsymbol{Q}_t^b$, respectively.

In certain situations we have a forwards model of the form (17) for the signal x_t. It would be interesting to be able to obtain a backwards model directly from the forwards model. Next, we show a useful way to get our objective.

Proposition 2. *Given a forwards model of the form*

$$x_{t+1} = \boldsymbol{k}_t^T \mathbf{x}_t + w_t, \qquad n-1 \le t \le m \tag{25}$$

with w_t a doubly white noise uncorrelated with $\mathbf{x}_{n-1}$, then $\{x_t, 0 \le t \le m\}$ has the backwards representation

$$x_t = \boldsymbol{k^b}_{t+1}^T \mathbf{x}_{t+1}^b + w_{t+1}^b, \qquad 0 \le t \le m-n+1$$

where the $2n$-vector $\boldsymbol{k}_{t+1}^b$ satisfies that $\boldsymbol{k^b}_{t+1}^T$ is equal to the $2n-1$ row of the matrix $\boldsymbol{K}_{t+1}^b = \boldsymbol{R}_{t+n-1}\boldsymbol{K}_{t+n-1}^H\boldsymbol{R}_{t+n}^{-1}$ and w_t^b is a doubly white noise with the properties given in Step 3 above.

Example 1 (continued). *It is not difficult to check that x_t is a WLM(1) signal by using property (6). Hence, applying Steps 1–3 above, it has the state-space representation*

$$x_{t+1} = \frac{1}{2}(e^3+e)x_t + \frac{1}{2}(e^3-e)x_t^* + w_t \tag{26}$$

with w_t a doubly white noise uncorrelated with x_0 and x_0^*. Moreover, as x_t is also a SOS signal, this model is trivially its WL autoregressive representation.

Example 2 (continued). *From Theorem 1 and Steps 1–3, it follows that x_t is a WLM(1) signal and has the state-space representation*

$$x_{t+1} = \frac{10^{1/3}(t+101)^{1/6}(t+100)^{1/6}-10}{10^{1/3}(t+100)^{1/3}-10}x_t + \mathrm{j}\frac{10^{2/3}(-(t+101)^{1/6}+(t+100)^{1/6})}{10^{1/3}(t+100)^{1/3}-10}x_t^* + w_t \tag{27}$$

with w_t a doubly white noise uncorrelated with x_1 and x_1^*.

5 Estimation problem of WLM(n) signals

Once the modeling problem has been solved for WLM(n) signals, we address the MMSE estimation problem of such signals under a WL processing approach. The forwards and backwards representations given in Theorem 3 notably simplify the design of different recursive estimation algorithms. To this end, we use the Kalman recursions on the forwards representation to provide the solution for the prediction and filtering problems and on the backwards representation for the smoothing problem (see, e.g., [17,18]).

Suppose that we observe a WLM(n) signal $\{x_t, 0 \le t \le m\}$ via the process

$$y_t = h_t x_t + v_t, \qquad 0 \le t \le m$$

with v_t a doubly white noise such that $E[\,v_t v_t^*] = n_{1,t}$ and $E[\,v_t v_t] = n_{2,t}$ with $n_{1,t} > |n_{2,t}|$. Moreover, we assume that v_t is uncorrelated with x_s and x_s^* for all t,s.

Consider the 2-vector $\mathbf{y}_t = [\,y_t, y_t^*]^T$, the $2 \times 2n$ matrix

$$\boldsymbol{H}_t = \begin{bmatrix} h_t & 0 & 0 & \cdots & 0 \\ 0 & h_t^* & 0 & \cdots & 0 \end{bmatrix}$$

and the 2×2 matrix

$$\boldsymbol{N}_t = \begin{bmatrix} n_{1,t} & n_{2,t} \\ n_{2,t}^* & n_{1,t} \end{bmatrix}$$

5.1 Prediction and filtering cases

Denote the WL filtered estimator of x_t by $\hat{x}_t^{WL}$ and the one-step-ahead predictor of x_{t+1} by $\hat{x}_{t+1|t}^{WL}$, both obtained on the basis of the information provided by the set $\{y_0, y_0^*, \ldots, y_t, y_t^*\}$, and consider their associated errors $p_t = E[\,|x_t - \hat{x}_t^{WL}|^2]$ and $p_{t+1|t} = E[\,|x_{t+1} - \hat{x}_{t+1|t}^{WL}|^2]$. Also denote the estimate of $\mathbf{x}_{n-1}$ obtained from the information provided by $[\,y_{n-1}, y_{n-1}^*, \ldots, y_0, y_0^*]^T$ by $\hat{\mathbf{x}}_{n-1}$ and its associated error by $\boldsymbol{P}_{n-1}$. By combining the forwards representation (17) and the classical Kalman filter we present Algorithm 1 which provides these estimators in an efficient way.

Algorithm 1. WL filter and prediction

Require: $\mathbf{y}_t$, $\boldsymbol{H}_t$, $\boldsymbol{N}_t$, $\boldsymbol{K}_t$, $\boldsymbol{Q}_t$, $\boldsymbol{g} = [\,1, 0, \ldots, 0]^T$, $\hat{\mathbf{x}}_{n-1}$, and $\boldsymbol{P}_{n-1}$

Ensure: $\hat{x}_{t+1|t}^{WL}, \hat{x}_{t+1}^{WL}, p_{t+1|t}$, and p_{t+1}

for $t \ge n-1$ **do**
- $\hat{\mathbf{x}}_{t+1|t} \leftarrow \boldsymbol{K}_t\hat{\mathbf{x}}_t$
- $\boldsymbol{P}_{t+1|t} \leftarrow \boldsymbol{K}_t\boldsymbol{P}_t\boldsymbol{K}_t^H + \boldsymbol{Q}_t$
- $\boldsymbol{F}_{t+1} \leftarrow \boldsymbol{P}_{t+1|t}\boldsymbol{H}_{t+1}^H[\,\boldsymbol{H}_{t+1}\boldsymbol{P}_{t+1|t}\boldsymbol{H}_{t+1}^H + \boldsymbol{N}_{t+1}]^{-1}$
- $\hat{\mathbf{x}}_{t+1} \leftarrow \hat{\mathbf{x}}_{t+1|t} + \boldsymbol{F}_{t+1}[\,\mathbf{y}_{t+1} - \boldsymbol{H}_{t+1}\hat{\mathbf{x}}_{t+1|t}]$
- $\boldsymbol{P}_{t+1} \leftarrow \boldsymbol{P}_{t+1|t} - \boldsymbol{F}_{t+1}\boldsymbol{H}_{t+1}\boldsymbol{P}_{t+1|t}$
- $\hat{x}_{t+1|t}^{WL} \leftarrow \boldsymbol{g}^T\hat{\mathbf{x}}_{t+1|t}$
- $\hat{x}_{t+1}^{WL} \leftarrow \boldsymbol{g}^T\hat{\mathbf{x}}_{t+1}$
- $p_{t+1|t} \leftarrow \boldsymbol{g}^T\boldsymbol{P}_{t+1|t}\boldsymbol{g}$
- $p_{t+1} \leftarrow \boldsymbol{g}^T\boldsymbol{P}_{t+1}\boldsymbol{g}$

end for

5.2 Smoothing case

Next, we compute two WL smoothing estimators of x_t based on future data. The first smoother is obtained from the set of observations $\{y_t, y_t^*, y_{t+1}, y_{t+1}^*, \ldots, y_m, y_m^*\}$

and it will be denoted by $\hat{x}_t^{b^{WL}}$. The second one is derived from the information supplied by the set $\{y_{t+1}, y_{t+1}^*, \ldots, y_m, y_m^*\}$ and we will refer to it as $\hat{x}_{t|t+1}^{b^{WL}}$. The errors of both estimators are $p_t^b = E[|x_t - \hat{x}_t^{b^{WL}}|^2]$ and $p_{t|t+1}^b = E[|x_t - \hat{x}_{t|t+1}^{b^{WL}}|^2]$, respectively. The initial condition $\hat{\mathbf{x}}_{m-n+1}^b$ is the estimate of $\mathbf{x}_{m-n+1}^b$ obtained from the $2n+2$-vector $[y_{m-n+1}, y_{m-n+1}^*, \ldots, y_m, y_m^*]^T$ and $\boldsymbol{P}_{m-n+1}^b$ is its associated error. By applying the backwards Kalman recursions on the backwards model (18) we get Algorithm 2.

Algorithm 2. WL smoothing
Require: $\mathbf{y}_t$, $\boldsymbol{H}_t$, $\boldsymbol{N}_t$, $\boldsymbol{K}_{t+1}^b$, $\boldsymbol{Q}_{t+1}^b$, $\boldsymbol{l} = [0, \ldots, 0, 1, 0]^T$, $\hat{\mathbf{x}}_{m-n+1}^b$, and $\boldsymbol{P}_{m-n+1}^b$
Ensure: $\hat{x}_{t|t+1}^{b^{WL}}, \hat{x}_t^{b^{WL}}, p_{t|t+1}^b$, and p_t^b

for $t \leq m-n$ **do**
- $\hat{\mathbf{x}}_{t|t+1}^b \leftarrow \boldsymbol{K}_{t+1}^b \hat{\mathbf{x}}_{t+1}^b$
- $\boldsymbol{P}_{t|t+1}^b \leftarrow \boldsymbol{K}_{t+1}^b \boldsymbol{P}_{t+1}^b \boldsymbol{K}^{b^H}_{t+1} + \boldsymbol{Q}_{t+1}^b$
- $\boldsymbol{F}_t^b \leftarrow \boldsymbol{P}_{t|t+1}^b \boldsymbol{H}_t^H [\boldsymbol{H}_t \boldsymbol{P}_{t|t+1}^b \boldsymbol{H}_t^H + \boldsymbol{N}_t]^{-1}$
- $\hat{\mathbf{x}}_t^b \leftarrow \hat{\mathbf{x}}_{t|t+1}^b + \boldsymbol{F}_t^b [\mathbf{y}_t - \boldsymbol{H}_t \hat{\mathbf{x}}_{t|t+1}^b]$
- $\boldsymbol{P}_t^b \leftarrow \boldsymbol{P}_{t|t+1}^b - \boldsymbol{F}_t^b \boldsymbol{H}_t \boldsymbol{P}_{t|t+1}^b$
- $\hat{x}_{t|t+1}^{b^{WL}} \leftarrow \boldsymbol{l}^T \hat{\mathbf{x}}_{t|t+1}^b$
- $\hat{x}_t^{b^{WL}} \leftarrow \boldsymbol{l}^T \hat{\mathbf{x}}_t^b$
- $p_{t|t+1}^b \leftarrow \boldsymbol{l}^T \boldsymbol{P}_{t|t+1}^b \boldsymbol{l}$
- $p_t^b \leftarrow \boldsymbol{l}^T \boldsymbol{P}_t^b \boldsymbol{l}$

end for

6 Numerical example

This section is devoted to showing the advantages of representation (27) (model 2) in relation to (3) (model 1) in two fields of signal processing: simulation and estimation. Firstly, we use such models to simulate trajectories of x_t defined in Example 2. Specifically, 50,000 trajectories of both models have been generated via Montecarlo simulation. To assess the performance of the simulations we compare the true correlation and complementary correlation functions with the simulated ones. Figure 1a,b depicts the true correlation and complementary correlation functions of x_t, Figure 1c,d the simulated simulated ones corresponding to model 1 and Figure 1e,f the simulated ones for model 2. We can see that the simulated trajectories of model 1 pick up adequately the behavior of the correlation function. However, these trajectories are unable to show the basic characteristics of the complementary correlation function. This shortcoming does not appear with model 2 whose simulated trajectories yield accurate representations of the second-order moments of x_t. For more detail, the 2D sections of the true complementary function and the simulated ones with models 1 and 2 for $t = 60$ and $t = 90$, respectively, are shown in Figure 2a,b.

Finally, we compare the SL smoother obtained with model 1 and the WL smoother derived in Algorithm 2 for model 2. For the particular case in which $h_t = 1$ and $n_{1,t} = 1$, Figure 3a compares the error p_t^b obtained for $n_{2,t} = 0.25$ (dotted line) and $n_{2,t} = 0.8$ (solid line) with the counterpart SL error (dashed line). On the other hand, considering $n_{2t} = n_2$ and denoting the errors of the improper and proper smoothers for every value of n_2 by $p_t^b(n_2)$ and $\bar{p}_t^b(n_2)$, respectively, Figure 3b displays the mean of the difference between the SL and WL estimation errors, that is, $DE(n_2) = \frac{1}{100}\sum_{t=1}^{100}(\bar{p}_t^b(n_2) - p_t^b(n_2))$ with n_2 varying within the interval $[0, 1)$. As expected, both figures show that WL estimation outperforms SL estimation, that is, they illustrate the better performance of the improper smoother in relation to the proper one. From Figure 3b, we also come to the conclusion that this gain in performance decreases as n_2 reduces.

7 Conclusions

The limited utility of the classical WSM definition to characterize the existence of a state-space representation for improper random signals has been revealed. By means of two simple examples, we have shown that in some cases the triangular condition fails to hold for signals with a state-space representation or that there exist signals with autocorrelations satisfying the triangular property for which the associated state-space representations present drawbacks in relation to their WL counterparts. Thus, the definition of a WSM signal has been extended to deal with improper signals providing new characterizations for WLM signals based either on second-order properties or on state-space representations. Moreover, a way to check the WLM condition has been given and the correlation structure of WLM signals has been devised. Finally, WL forwards and backwards Markovian representations have been presented from which some applications are illustrated in the signal estimation and simulation fields.

Appendix 1

Proof of Theorem 1

To prove the implication 1) $\Rightarrow$ 2) observe that if x_t is a WLM(n) signal then for any $s < t$,

$$\hat{x}^{WL}(t|\tau \leq s) = k_1(t,s)x_s + k_2(t,s)x_s^* + \cdots + k_{2n-1}(t,s)x_{s-n+1} + k_{2n}(t,s)x_{s-n+1}^* \quad (28)$$

which implies that $\hat{\mathbf{x}}^{WL}(t|\tau \leq s)$ is of the form (5) with $\boldsymbol{K}(t,s)$ defined in (4). Moreover, the rows are of the form

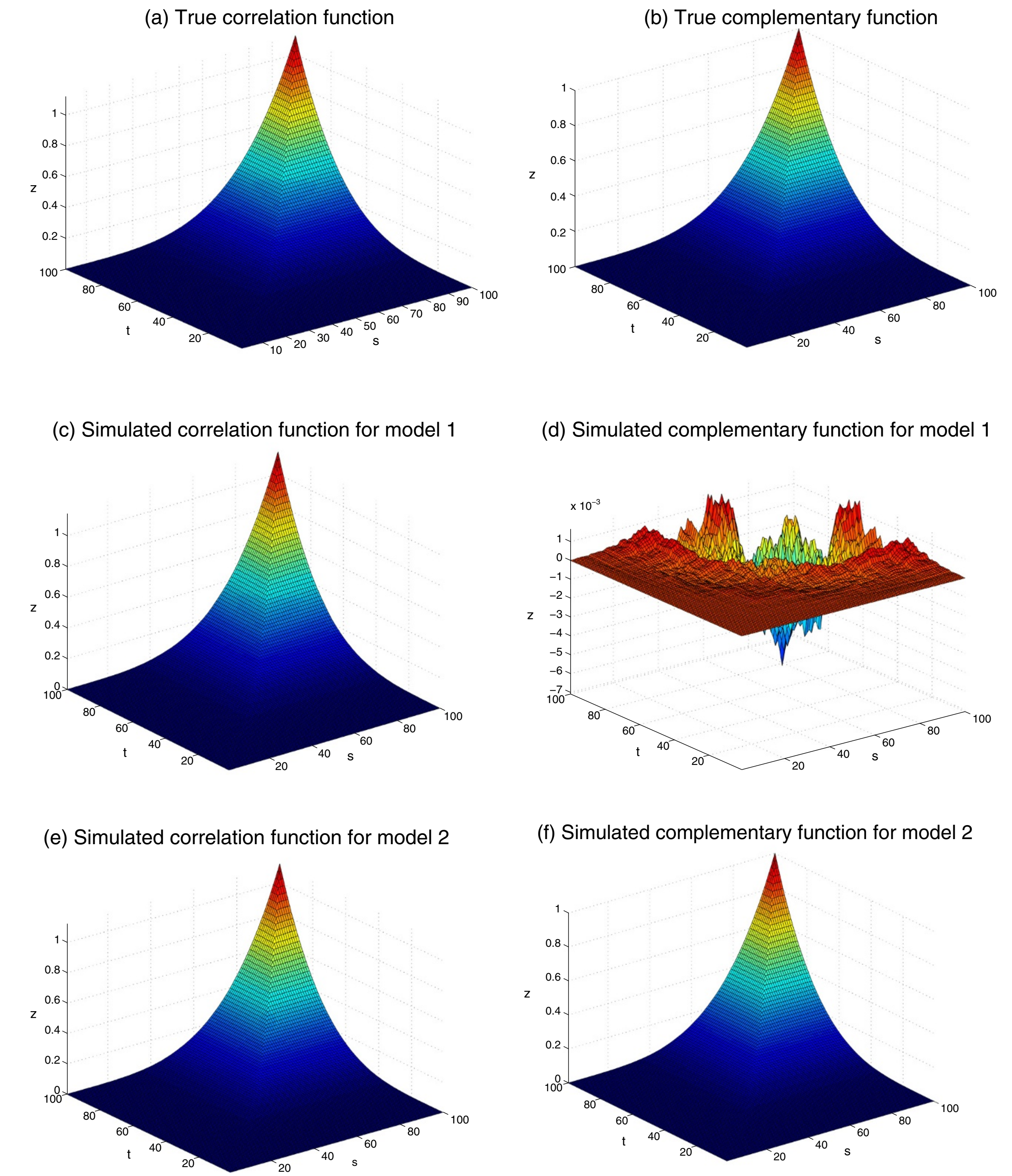

Figure 1 (a) True correlation function: (b) True complementary correlation function: (c) Simulated correlation function for model 1: (d) Simulated complementary correlation function for model 1: (e) Simulated correlation function for model 2: (f) Simulated complementary correlation function for model 2.

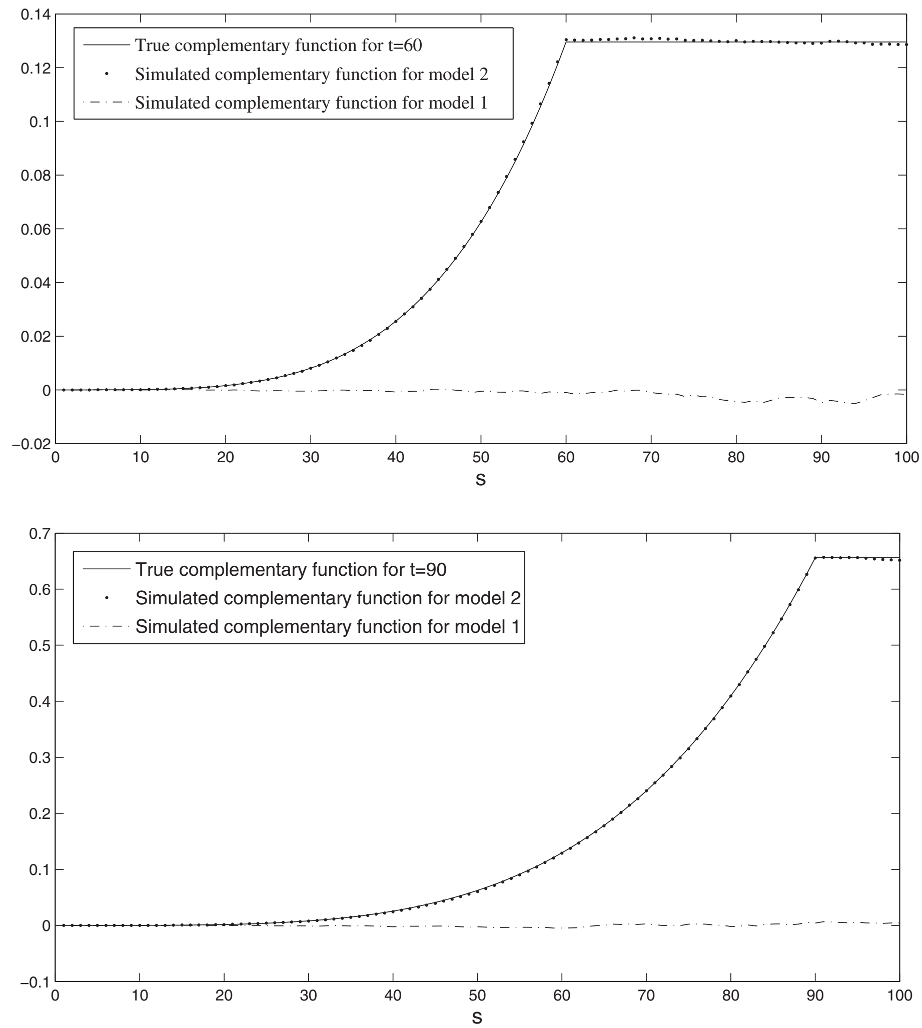

Figure 2 (a) 2D section of true and simulated complementary functions for $t = 60$: (b) 2D section of true and simulated complementary functions for $t = 90$.

$$\begin{aligned}\boldsymbol{K}_{[2i-1]}(t,s) &= [k_1(t-i+1,s), k_2(t-i+1,s), \ldots, k_{2n-1} \\ &\quad \times (t-i+1,s), k_{2n}(t-i+1,s)] \\ \boldsymbol{K}_{[2i]}(t,s) &= \left[k_2^*(t-i+1,s), k_1^*(t-i+1,s), \ldots, k_{2n}^* \right. \\ &\quad \left. \times (t-i+1,s), k_{2n-1}^*(t-i+1,s)\right]\end{aligned} \tag{29}$$

for $i = 1, \ldots, n$. The inverse implication, 2) $\Rightarrow$ 1), is checked similarly.

Finally, the proof of 2) $\Leftrightarrow$ 3) is similar to the one given in Theorem 1 of [16].

Proof of Theorem 2

The proof of 2) $\Leftrightarrow$ 3) is similar to that of Theorem 1 by taking into account that $E[\mathbf{x}_t^b \mathbf{x}_s^{bH}] = \boldsymbol{R}(t+n-1, s+n-1)$. Now, we prove 1) $\Leftrightarrow$ 3). Following a similar reasoning to that used in the proof of Theorem 1 in [16], we have that (7) is equivalent to the condition

$$\boldsymbol{K}(t,s) = \boldsymbol{K}(t,\tau)\boldsymbol{K}(\tau,s), \qquad t \le \tau \le s$$

and thus,

$$\boldsymbol{K}^H(s,t) = \boldsymbol{K}^H(\tau,t)\boldsymbol{K}^H(s,\tau) = (\boldsymbol{K}(s,\tau)\boldsymbol{K}(\tau,t))^H, \qquad t \ge \tau \ge s$$

from which, applying Theorem 1, it follows that x_t is a WLM(n) signal. In a similar way the implication 1) $\Rightarrow$ 3) is proven.

Proof of Proposition 1

Taking into account that $\hat{x}^{WL}(t+j-i|\tau \le t) = x_{t+j-i}$ for $j \le i \le n-1$ we obtain (8) and (9). Likewise, (13) follows from (29), (8), and (9).

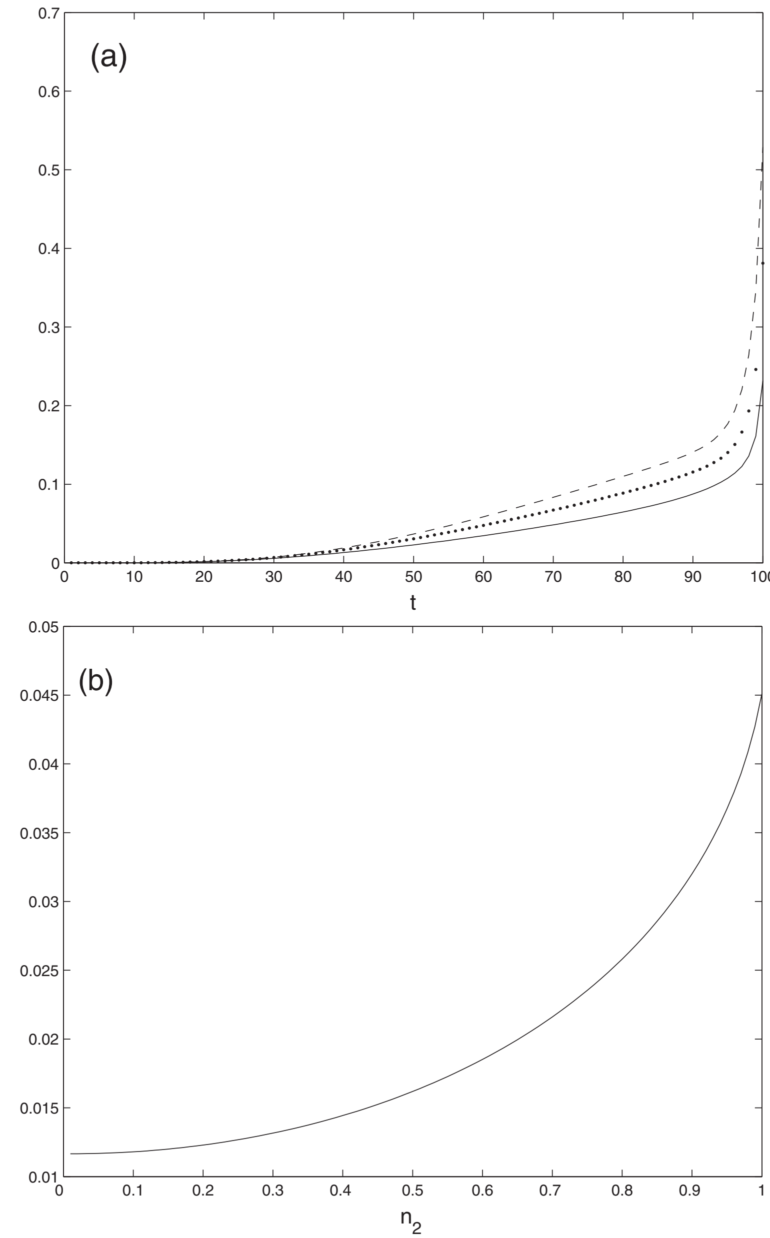

Figure 3 (a) Smoothing errors: WL smoothing errors p_t^b for $n_{2,t} = 0.25$ (dotted line) and $n_{2,t} = 0.8$ (solid line), and the SL smoothing error (dashed line): (b) Performance of the smoother estimate. Mean of the difference between the SL and WL estimation errors DE(n_2).

Now, from (6) we get

$$\boldsymbol{K}(t+j+1,t) = \boldsymbol{K}(t+j+1,t+j)\boldsymbol{K}(t+j,t), \qquad j \geq 0$$

and together with (13) we demonstrate (10), (11), and (12).

On the other hand, (14) and (15) can be proven following a similar reasoning to that of Theorem 2 in [16].

Finally, by using the Hilbert projection theorem and (5) we have

$$\mathbf{x}_{t+1} = \boldsymbol{K}_t\mathbf{x}_t + \boldsymbol{w}_t \tag{30}$$

where $\boldsymbol{w}_t = [\,w_t, w_t^*, 0, \ldots, 0]^T$ is the innovations process which, by construction, is uncorrelated with $\mathbf{x}_s$ for $t \geq s$. Thus,

$$\boldsymbol{R}_{t+1} = E\left[\mathbf{x}_{t+1}\mathbf{x}_{t+1}^H\right] = E\left[(\boldsymbol{K}_t\mathbf{x}_t + \boldsymbol{w}_t)(\boldsymbol{K}_t\mathbf{x}_t + \boldsymbol{w}_t)^H\right]$$
$$= \boldsymbol{K}_t\boldsymbol{R}_t\boldsymbol{K}_t^H + \boldsymbol{Q}_t$$

with $E[\,\boldsymbol{w}_t\boldsymbol{w}_t^H] = \boldsymbol{Q}_t$ given in (16).

Proof of Theorem 3

If x_t is a WLM(n) signal then, from (13) and (30), we have

$$x_{t+1} = k_{1,t}x_t + k_{2,t}x_t^* + \cdots + k_{2n-1,t}x_{t-n+1} + k_{2n,t}x_{t-n+1}^* + w_t \quad (31)$$

where w_t is the first component of $\boldsymbol{w}_t$. Hence, denoting $\boldsymbol{k}_t = \boldsymbol{K}_{[1]}^T(t+1,t) = [k_{1,t}, \ldots, k_{2n,t}]^T$ we obtain (17). On the other hand, from the Hilbert projection theorem and (7) we get

$$\mathbf{x}_t^b = \boldsymbol{K}(t+n-1, t+n)\mathbf{x}_{t+1}^b + \boldsymbol{w}_{t+1}^b \quad (32)$$

where $\boldsymbol{w}_t^b = [0, \ldots, 0, w_t^b, w_t^{b*}]^T$ is the backwards innovations process which, from construction, is uncorrelated with $\mathbf{x}_s$ for $t \leq s$. Hence, $x_t = \boldsymbol{K}_{[2n-1]}(t+n-1, t+n)\mathbf{x}_{t+1}^b + w_{t+1}^b$ with w_{t+1}^b the $2n-1$ component of $\boldsymbol{w}_{t+1}^b$. Thus, denoting ${\boldsymbol{k}^b}_{t+1}^T = \boldsymbol{K}_{[2n-1]}(t+n-1, t+n)$, (18) is obtained.

Conversely, suppose that x_t has the representation (17). Denote $\mathcal{H}$ the closed span generated by the set $\{x_\tau, x_\tau^*, \tau \leq t\}$. By using Proposition 2.3.2 of [19], to prove that $\hat{x}^{WL}(t|\tau \leq s) = \hat{x}^{WL}(t|s, s-1, \ldots, s-n+1)$ for any $s < t$ is equivalent to $\hat{x}^{WL}(t+1|\tau \leq t) = \hat{x}^{WL}(t+1|t, t-1, \ldots, t-n+1)$ for all t. Thus, projecting (17) onto $\mathcal{H}$ and taking Proposition 2.3.2 of [19] into account we have

$$\hat{x}^{WL}(t+1|\tau \leq t) = \boldsymbol{k}_t^T \boldsymbol{x}_t + \hat{w}^{WL}(t|\tau \leq t)$$

where $\hat{w}^{WL}(t|\tau \leq t)$ is the projection of w_t onto $\mathcal{H}$. The hypothesis (19) guarantees that w_t is uncorrelated with x_s and x_s^* for $t \geq s$. Hence, $\hat{w}^{WL}(t|\tau \leq t) = 0$ and x_t is a WLM(n) signal.

The proof for the backwards representation (18) is similar.

Proof of Corollary 1

Since x_t is a SOS signal then the matrices $\boldsymbol{R}(t+h, t)$, $h = 1, 2, \ldots$, are independent of t. Thus, from (4) we obtain $k_{i,t} = k_i$ for all i and t. Finally, taking (31) into account we have

$$x_{t+1} = \sum_{i=0}^{n-1} k_{2i+1}x_{t-i} + \sum_{i=0}^{n-1} k_{2i+2}x_{t-i}^* + w_t$$

which gives (20) defining $g_{1,i} = k_{2i+1}$ and $g_{2,i} = k_{2i+2}$.

Proof of Proposition 2

From (25) and Theorem 3 it follows that $\mathbf{x}_t^b$ has the representation (32). Then by using (22) we obtain

$$\begin{aligned}\boldsymbol{K}_{t+1}^b &= \boldsymbol{K}(t+n-1, t+n) = \boldsymbol{R}(t+n-1, t+n)\boldsymbol{R}_{t+n}^{-1} \\ &= \boldsymbol{R}^H(t+n, t+n-1)\boldsymbol{R}_{t+n}^{-1} = \boldsymbol{R}_{t+n-1}\boldsymbol{K}_{t+n-1}^H\boldsymbol{R}_{t+n}^{-1}\end{aligned}$$

and thus the result follows.

Abbreviations

MMSE, Minimum-mean square error; SL, Strictly linear; WL, Widely linear; WLM, Widely linear Markov; WSM, Wide-sense Markov.

Competing interests

The authors declare that they have no competing interests.

Received: 28 May 2012 Accepted: 29 October 2012
Published: 18 December 2012

References

1. T Kailath, AH Sayed, B Hassibi, *Linear Estimation*. (Prentince Hall, New Jersey, 2000)
2. HV Poor, Ch Chang, A reduced-complexity quadratic structure for the detection of stochastic signals. J. Acoust. Soc. Am. **78**(5), 1652–1657 (1985)
3. PW Jones, P Smith, *Stochastic Processes. An Introduction*. (Chapman &amp Hall/CRC, Boca Raton, 2010)
4. JL Doob, *Stochastic Processes*. (John Willey, New York, 1953)
5. DP Mandic, VSL Goh, *Complex Valued Nonlinear Adaptative Filters. Noncircularity, Widely Linear and Neural Models*. (John Willey, New York, 2009)
6. T Adali, S Haykin, *Adaptive Signal Processing: Next Generation Solutions*. (Wiley-IEEE Press, New York, 2010)
7. J Navarro-Moreno, MD Estudillo, RM Fernández-Alcalá, JC Ruiz-Molina, Estimation of improper complex-valued random signals in colored noise by using the hilbert space theory. IEEE Trans. Inf. Theory. **55**(6), 2859–2867 (2009)
8. B Picinbono, P Bondon, Second-order statistics of complex signals. IEEE Trans. Signal Process. **45**(2), 411–420 (1997)
9. VSL Goh, DP Mandic, An augmented extended kalman filter algorithm for complex-valued recurrent neural networks. Neural Comput. **19**, 1039–1055 (2007)
10. J Navarro-Moreno, ARMA prediction of widely linear systems by using the innovations algorithm. IEEE Trans. Signal Process. **56**(7), 3061–3068 (2008)
11. C Cheong Took, DP Mandic, A quaternion widely linear adaptive filter. IEEE Trans. Signal Process. **58**(8), 4427–4431 (2010)
12. S Buzzi, M Lops, S Sardellitti, Widely linear reception strategies for layered space-time wireless communications. IEEE Trans. Signal Process. **54**(6), 2252–2262 (2006)
13. P Rubin-Delanchy, AT Walden, Simulation of improper complex-valued sequences. IEEE Trans. Signal Process. **55**(11), 5517–5521 (2007)
14. F Beutler, Multivariate wide-sense Markov processes and prediction theory. Ann. Math. Stat. **34**(2), 424–438 (1963)
15. V Mandrekar, On multivariate wide-sense Markov processes. Nagoya Math. J. **3**, 7–19 (1968)
16. A Kasprzyk, W Szczotka, Covariance structure of wide-sense Markov processes of order $k \geq 1$. Appl. Math. **33**(2), 129–143 (2006)
17. DH Dini, DP Mandic, SJ Julier, A widely linear complex unscented Kalman filter. IEEE Signal Process. Lett. **18**(11), 623–626 (2011)
18. DH Dini, DP Mandic, Class of widely linear complex Kalman filters. IEEE Trans. Neural Netw. Learn. Syst. **23**(5), 775–786 (2012)
19. PJ Brockwell, RA Davis, *Time Series: Theory and Methods*, 2nd edn. (Springer-Verlag, New York, 1991)

doi:10.1186/1687-6180-2012-256
Cite this article as: Espinosa-Pulido *et al.*: **Widely linear Markov signals.** *EURASIP Journal on Advances in Signal Processing* 2012 **2012**:256.