RESEARCH **Open Access**

Prediction on widely factorizable signals

Rosa M Fernández-Alcalá[*], Jesús Navarro-Moreno and Juan C Ruiz-Molina

Abstract

This article extends the solutions to the prediction problem for factorizable real random signals to the class of improper complex-valued random signals. For that, the concept of widely factorizable signals is introduced and several real examples of signals having a widely factorizable correlation function are presented. A widely linear processing is considered in the design of both linear and nonlinear prediction algorithms which are computationally feasible from the practical standpoint. These algorithms are valid for stationary as well as nonstationary signals and they can be applied from only the knowledge of the second-order properties of the augmented vectors involved, it not being necessary to know if the signal satisfies a state-space model.

Keywords: complex-valued improper signals, prediction problems, widely factorizable signals, widely linear processing.

1 Introduction

In recent research, estimation theory becomes very relevant within the complex random signals field. In fact, although traditionally the treatment of this type of problem has consisted in mere extensions of the vectorial real-valued estimation algorithms to the complex plane, a complex formalism has a special value in the description of some physical systems in such diverse fields as communications, oceanography, meteorology and optics among others (see, for example, [1] and the references therein).

The classic processing with such signals, called strictly linear (SL), assumes that the signals involved are proper.[a] However, this assertion is not always justified. The improper nature of some signals requires that the pseudocorrelation function must be taken into account in describing and characterizing their second order properties completely [2-4].

With this motivation, [2] introduces a new methodology called widely linear (WL) whose more notable characteristic is that it utilizes not only the observed signal but also its conjugate to obtain estimators with better behavior than the conventional ones in the sense of reducing the mean square estimation error. Moreover, this kind of processing has become very usual in the last decade for designing linear and nonlinear estimation algorithms from a discrete-time [1,5-9] as well as a continuous-time perspective of the problem [10]. Specifically, focussing our attention on the discrete case, the recent books of Mandic and Goh [1] and Adali and Haykin [11] about WL adaptive systems can be considered as two reference texts in this area which provide a unified treatment of linear and nonlinear complex-valued adaptive filters.

On the other hand, knowledge of second-order statistics is a key assumption in solving estimation problems (see, for example, [12]). In practice this knowledge is available because second-order statistics of the problem have been measured experimentally or they are understandable enough from the physical mechanism [13]. In this framework, WL estimation algorithms for computing all types (filtering, prediction, and smoothing) of estimates have been devised in [8] for second-order stationary (SOS) signals, i.e., those signals with constant mean function and both correlation and pseudocorrelation functions only dependent on the difference of time instants. Although the WL estimation problem considered is very general, its applicability is limited to SOS signals.

In the real field, an alternative estimation methodology based on correlation information has been developed to solve very general linear and nonlinear estimation problems for the class of factorizable signals (see, for example, [14,15]). This type of signal is characterized by having a factorizable kernel which is a very general condition valid for stationary and nonstationary signals. This

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

formulation does not require postulating a dynamical model for the signal and thus, it is useful whenever the physical mechanism generating the signal of interest is not known or if it impossible to determine (for instance, if it does not satisfy a state-space model).

In this article, our objective is to extend this last methodology to the complex plane. For that, we introduce a new class of signals, called widely factorizable, whose correlation function of the augmented vector formed by the signal and its conjugate is a factorizable kernel. Then, a WL processing is employed in the development of both linear and nonlinear prediction algorithms for widely factorizable signals observed in the presence of noise. With this aim, in Section 2 widely factorizable signals are defined and also the basic notation and concepts about complex-valued signals are summarized. Next, the linear augmented complex prediction problem is addressed in Section 3 where a recursive algorithm is provided for computing the optimal WL prediction estimate as well as its associated error. This section also includes three numerical examples which show the enhancement of the proposed WL predictor in relation to the SL solution. Finally, by following a similar reasoning to the extended Kalman filter (EKF), the nonlinear augmented complex prediction problem is solved in Section 4 where a numerical example is also developed in order to compare the proposed WL algorithm with two conventional nonlinear augmented complex techniques, the WL EKF and the WL unscented Kalman filter (UKF) suggested in [1]. Furthermore, these solutions are also compared with the SL EKF and SL UKF.

2 Preliminaries

This section covers some of the more basic notions associated with the complex-valued random signals field and introduces a new class of signals, called widely factorizable. Moreover, the notation and hypotheses held throughout the article are also established in this section.

First of all, note that all vectors will be denoted by bold small letters and bold capital letters will be used for matrices. Also, row k of any matrix $\mathbf{A}(\cdot)$ will be denoted by $\mathbf{a}_{[k]}(\cdot)$. Furthermore, $\mathbf{0}_p$ and $\mathbf{0}_{p \times q}$ represent the p-vector and the $p \times q$-matrix, respectively, whose elements are all zeros.

Unless indicated to the contrary, throughout this article we consider an improper complex-valued signal $\{s(t_i), t_i \in T\}$, $T = \{t_1, t_2, \ldots\}$, with zero-mean and correlation function $r_s(t_i, t_j) = E[s(t_i)s^*(t_j)]$, where the superscript '*' represents the complex conjugate.

The signal $\{s(t_i), t_i \in T\}$ is said to be factorizable if there exist two l-vectors $\boldsymbol{\alpha}(t_i)$ and $\boldsymbol{\beta}(t_i)$ such that its correlation function $r_s(t_i, t_j)$ can be expressed in the form

$$r_s(t_i, t_j) = \begin{cases} \boldsymbol{\alpha}^{\mathrm{T}}(t_i)\boldsymbol{\beta}^*(t_j), & t_i \geq t_j \\ \boldsymbol{\beta}^{\mathrm{T}}(t_i)\boldsymbol{\alpha}^*(t_j), & t_i \leq t_j \end{cases} \tag{1}$$

where the superscript 'T' denotes the transpose. Note that this type of signal is very general and includes both stationary and nonstationary signals. However, a new class of signals is possible by imposing the condition of factorizable kernel on the correlation function $\mathbf{R}_s(t_i, t_j) = E[\mathbf{s}(t_i)\mathbf{s}^{\mathrm{H}}(t_j)]$, with the superscript 'H' denoting the conjugate transpose, of the augmented vector $\mathbf{s}(t_i) = [s(t_i), s^*(t_i)]^{\mathrm{T}}$. This type of signal, called *widely factorizable*, is introduced next.

Definition 2.1 *A signal $\{s(t_i), t_i \in T\}$ is said to be widely factorizable if and only if there exist two $2 \times m$-matrices $\mathbf{A}(t_i)$ and $\mathbf{B}(t_i)$ such that the correlation function $\mathbf{R}_s(t_i, t_j)$ of the augmented vector $\mathbf{s}(t_i)$ can be expressed in the form*

$$\mathbf{R}_s\left(t_i, t_j\right) = \begin{cases} \mathbf{A}\left(t_i\right)\mathbf{B}^H\left(t_j\right), & t_i \geq t_j \\ \mathbf{B}\left(t_i\right)\mathbf{A}^H\left(t_j\right), & t_i \leq t_j \end{cases} \tag{2}$$

where the superscript 'H' denotes the conjugate transpose.

Note that condition (2) implies (1), however condition (1) does not assure that the correlation function of the augmented vector satisfies (2). As a consequence, we have that all widely factorizable signals are also factorizable but the converse does not hold.

Some illustrative examples of widely factorizable signals are the following:

i) The rotation of a real factorizable zero-mean signal $x(t_i)$ by an independent random phase θ, $s(t_i) = e^{\theta\,\mathrm{j}}x(t_i)$, with $\mathrm{j} = \sqrt{-1}$. Assume

$$r_x\left(t_i, t_j\right) = \begin{cases} \boldsymbol{a}^{\mathrm{T}}\left(t_i\right)\boldsymbol{b}\left(t_j\right), & t_i \geq t_j \\ \boldsymbol{b}^{\mathrm{T}}\left(t_i\right)\boldsymbol{a}\left(t_j\right), & t_i \leq t_j \end{cases} \tag{3}$$

with $\boldsymbol{a}(t_i)$ and $\boldsymbol{b}(t_i)$ two real l-vectors. Thus,

$$\mathbf{A}(t_i) = \begin{pmatrix} \mathbf{0}_l^{\mathrm{T}} & \boldsymbol{a}^{\mathrm{T}}(t_i) \\ \boldsymbol{a}^{\mathrm{T}}(t_i) & \mathbf{0}_l^{\mathrm{T}} \end{pmatrix}, \quad \mathbf{B}^H(t_i) = \begin{pmatrix} \varphi_\theta\,(-2)\,\boldsymbol{b}\,(t_i) & \boldsymbol{b}\,(t_i) \\ \boldsymbol{b}\,(t_i) & \varphi_\theta\,(2)\,\boldsymbol{b}\,(t_i) \end{pmatrix} \tag{4}$$

where $\phi_\theta(\cdot)$ is the characteristic function of θ.

ii) The seismic ground acceleration can be represented by a uniformly modulated nonstationary process given by $s(t_i) = d(t_i)x(t_i)$, where $x(t_i)$ is a stationary process with zero mean and known second-order statistics, and $d(t_i)$ is the time modulating function. Both $x(t_i)$ and $d(t_i)$ can be either real or complex-valued [16]. Thus, if $d(t_i)$ is complex and $x(t_i)$ is real and factorizable with correlation function given by (3) then, $s(t_i)$ is widely factorizable being

$$\mathbf{A}(t_i) = \begin{pmatrix} d\,(t_i)\,\boldsymbol{a}^{\mathrm{T}}\,(t_i) \\ d^*\,(t_i)\,\boldsymbol{a}^{\mathrm{T}}\,(t_i) \end{pmatrix}, \quad \mathbf{B}^H(t_i) = \begin{pmatrix} \boldsymbol{b}\,(t_i)\,d^*\,(t_i) & \boldsymbol{b}\,(t_i)\,d\,(t_i) \end{pmatrix}$$

Common choices are $d(t_i) = e^{j2\pi\omega_0 t_i}$ with ω_0 a constant frequency and $x(t_i)$ the Ornstein-Uhlenbeck process.

On the other hand, if both $d(t_i)$ and $x(t_i)$ are complex and $x(t_i)$ is also widely factorizable with

$$\mathbf{R_x}(t_i, t_j) \begin{cases} \mathbf{E}(t_i)\, \mathbf{F}^H(t_j), & t_j \geq t_j \\ \mathbf{F}(t_i)\, \mathbf{E}^H(t_j), & t_i \leq t_j \end{cases}$$

then, $s(t_i)$ is widely factorizable with

$$\mathbf{A}(t_i) = \begin{pmatrix} d(t_i) & 0 \\ 0 & d^*(t_i) \end{pmatrix} \mathbf{E}(t_i), \qquad \mathbf{B}^H(t_i) = \mathbf{F}^H(t_i) \begin{pmatrix} d^*(t_i) & 0 \\ 0 & d(t_i) \end{pmatrix}$$

iii) In electromagnetic theory, the time-varying position of the electric field vector can be represented as the following signal $s(t_i) = Ae^{jw_0 t_i} + Be^{-jw_0 t_i}$ where the expressions for the random coefficients A and B can be found in [[17], p. 7]. Hence, $s(t_i)$ is widely factorizable with

$$\mathbf{A}(t_i) = \begin{pmatrix} r_A e(t_i) + r_{AB}^* e^*(t_i)\, \gamma_A^* e^* + \gamma_{AB}^* e(t_i) \\ r_B e^*(t_i) + r_{AB} e(t_i)\, \gamma_B^* e(t_i) + \gamma_{AB}^* e^*(t_i) \\ \gamma_B e^*(t_i) + \gamma_{AB} e(t_i)\, r_B e(t_i) + r_{AB}^* e^*(t_i) \\ \gamma_{AB} e^*(t_i) + \gamma_A e(t_i)\, r_A e^*(t_i) + r_{AB} e(t_i) \end{pmatrix}^T \quad (5)$$

$$\mathbf{B}^H(t_i) = \begin{pmatrix} e(t_i)\, e^*(t_i) & 0 & 0 \\ 0 & 0 & e(t_i)\, e^*(t_i) \end{pmatrix}^H$$

where $e(t_i) = e^{jw_0 t_i}$, $r_A = E[AA^*]$, $r_{AB} = E[AB^*]$, $\gamma_A = E[AA]$, $\gamma_B = E[BB]$, and $\gamma_{AB} = E[AB]$.

iv) A signal widely used in many areas of signal processing is a linear frequency modulation or chirp with random phase. The chirp process can be expressed as $s(t_i) = e^{j\pi(2\alpha t_i + \beta t_i^2 + 2\theta)}$ where θ is the random phase, α determines the starting instantaneous frequency of the chirp and β is the chirp rate. Then, the chirp process is widely factorizable with

$$\mathbf{A}(t_i) = \begin{pmatrix} c(t_i) & 0 \\ 0 & c^*(t_i) \end{pmatrix}$$

$$\mathbf{B}^H(t_i) = \begin{pmatrix} (1 - |\varphi_\theta(2\pi)|^2) c^*(t_i) & (\varphi_\theta(4\pi) - \varphi_\theta^2(2\pi)) c(t_i) \\ (\varphi_\theta(4\pi) - \varphi_\theta^2(2\pi))^* c^*(t_i) & (1 - |\varphi_\theta(2\pi)|^2) c(t_i) \end{pmatrix}$$

with $c(t_i) = e^{j\pi(2\alpha t_i + \beta t_i^2)}$ and $\phi_\theta(\cdot)$ the characteristic function of θ.

v) An application of the complex Ornstein-Uhlenbeck process is the description of the motion of the instantaneous axis of the Earth's rotation [18]. This motion has an 1 year period and if it is removed, there remains the so-called Chandler Wobble, which has a period of about 435 days (14 months). Kolmogorov proposed the complex stochastic process $s(t_i) = x(t_i) + jy(t_i) = \sigma e^{j2\pi t_i} + \xi(t_i)$ to describe the Chandler Wobble, i.e., the motion of the pole, where $x(t_i)$ and $y(t_i)$ are the coordinates of the deviation of the instantaneous pole from the North Pole. In that model the first term is a periodical component, and the second term $\xi(t_i)$ is a complex Ornstein-Uhlenbeck process. It is not difficult to check that the signal $s(t_i)$ is proper and then,

$$\mathbf{A}(t_i) = \begin{pmatrix} \frac{1}{\lambda} e^{(\omega j - \lambda)t_i} & 0 \\ 0 & \frac{1}{\lambda} e^{-(\omega j \cdot \lambda)t_i} \end{pmatrix}, \qquad \mathbf{B}^H(t_i) = \begin{pmatrix} e^{(-\omega j + \lambda)t_i} & 0 \\ 0 & e^{(\omega j + \lambda)t_i} \end{pmatrix}$$

where $\lambda > 0$ is the drift parameter and $\omega \in \mathbb{R}$ is the period.

On the other hand, we provide a simple example of a factorizable signal which is not widely factorizable: a zero-mean complex-valued signal $\{s(t_i),\ t_i = i/100,\ i = 1, ..., 100\}$ with $r_s(t_i, t_j) = 2e^{-|t_i - t_j|}$ and $\rho_s(t_i, t_j) = j0.5 e^{t_i t_j}$.

In the following, we also assume that the signal of interest $s(t_i)$ is widely factorizable in the sense of Definition 2.1.

Finally, $\mathbf{R_{sy}}(t_i, t_j) = E[\mathbf{s}(t_i)\mathbf{y}^H(t_j)]$ denotes the cross-correlation function between any two augmented signals $\mathbf{s}(t_i)$ and $\mathbf{y}(t_i)$, and $\mathbf{r_{sy}}(t_i, t_j) = E[s(t_i)\mathbf{y}^H(t_j)]$ represents the cross-correlation function between $s(t_i)$ and the augmented vector $\mathbf{y}(t_i)$.

3 Linear augmented complex prediction

Assume that the signal $s(t_i)$ established in Section 2 is observed through the following linear equation:

$$y(t_i) = g(t_i)\, s(t_i) + v(t_i), \qquad t_1 \leq t_i \leq t_n \quad (6)$$

where $g(t_i)$ is a deterministic complex-valued function and $v(t_i)$ is a doubly white noise[b] correlated with $s(t_i)$. Moreover, the cross-correlation function and the augmented signal $\mathbf{s}(t_i)$ and the augmented noise $\mathbf{v}(t_i)$ is of the form

$$\mathbf{R_{sv}}(t_i, t_j) = \begin{cases} \mathbf{C}(t_i)\, \mathbf{D}^H(t_j), & t_i \geq t_j \\ \mathbf{E}(t_i)\, \mathbf{F}^H(t_j), & t_i \leq t_j \end{cases} \quad (7)$$

where $\mathbf{C}(t_i)$, $\mathbf{D}(t_i)$, $\mathbf{E}(t_i)$, and $\mathbf{F}(t_i)$ are matrices of dimensions $2 \times l$, $2 \times l$, $2 \times l'$, and $2 \times l'$, respectively.

We consider the problem of obtaining the optimal (in the sense of minimizing the WL mean square error) estimator[c] of the signal $s(t_k)$ as a function of the information given by the observations $\{y(t_1), ..., y(t_n), y^*(t_1), ..., y^*(t_n)\}$, for $t_k \geq t_n$. It is known that such an estimator can be expressed as a linear function of the set of augmented observations $\{\mathbf{y}(t_1), ..., \mathbf{y}(t_n)\}$ as follows [2]

$$\hat{s}(t_k|t_n) = \sum_{j=1}^{n} \mathbf{h}^T(t_k, t_j, t_n)\, \mathbf{y}(t_j), \qquad t_k \geq t_n \quad (8)$$

where the 2D vector $\mathbf{h}(t_k, t_j, t_n)$, called the impulse response function, satisfies the equation

$$\mathbf{r}_{sy}\left(t_k, t_j\right) = \sum_{i=1}^{n} \mathbf{h}^{T}\left(t_k, t_i, t_n\right)\mathbf{R}\left(t_i, t_j\right) + \mathbf{h}^{T}\left(t_k, t_j, t_n\right)\sum, \qquad t_1 \le t_j \le t_n, \ t_k \ge t_n \qquad (9)$$

where $\mathbf{R}(t_i, t_j) = \mathbf{G}(t_i)\,\mathbf{R}_s(t_i, t_j)\mathbf{G}^{H}(t_j) + \mathbf{G}(t_i)\mathbf{R}_{sv}(t_i, t_j) + \mathbf{R}_{sv}^{H}(t_j, t_i)\mathbf{G}^{H}(t_j)$, with $\mathbf{G}(t_i)$ denoting the 2×2-diagonal matrix $\mathbf{G}(t_i) = \mathrm{diag}(g(t_i), g^*(t_i))$, and $E[\mathbf{v}(t_i)\mathbf{v}^{H}(t_i)] = \Sigma$.

From (2) and (7), it is easy to check that $\mathbf{r}_{sy}(t_k, t_j)$ and $\mathbf{R}(t_i, t_j)$ can be written as follows:

$$\mathbf{r}_{sy}\left(t_k, t_j\right) = \begin{cases} \boldsymbol{\psi}_{[1]}(t_k)\,\boldsymbol{\Gamma}^{H}\left(t_j\right), & t_k \ge t_j \\ \boldsymbol{\pi}_{[1]}(t_k)\,\boldsymbol{\Phi}^{H}\left(t_j\right), & t_k \le t_j \end{cases} \qquad (10)$$

$$\mathbf{R}\left(t_i, t_j\right) = \begin{cases} \boldsymbol{\Phi}\left(t_i\right)\boldsymbol{\Gamma}^{H}\left(t_j\right), & t_i \ge t_j \\ \boldsymbol{\Gamma}\left(t_i\right)\boldsymbol{\Phi}^{H}\left(t_j\right), & t_i \le t_j \end{cases} \qquad (11)$$

where $\boldsymbol{\psi}_{[1]}(t_k)$ is the first row of the $2 \times q$-matrix $\boldsymbol{\Psi}(t_i) = [\mathbf{A}(t_i), \mathbf{C}(t_i), \mathbf{0}_{2 \times l'}]$, $\boldsymbol{\pi}_{[1]}(t_k)$ is the first row of the $2 \times q$-matrix $\boldsymbol{\Pi}(t_i) = [\mathbf{B}(t_i), \mathbf{0}_{2 \times l}, \mathbf{E}(t_i)]$, $\boldsymbol{\Phi}(t_i) = [\mathbf{G}(t_i)\mathbf{A}(t_i), \mathbf{G}(t_i)\mathbf{C}(t_i), \mathbf{F}(t_i)]$, and $\boldsymbol{\Gamma}(t_i) = [\mathbf{G}(t_i)\mathbf{B}(t_i), \mathbf{D}(t_i), \mathbf{G}(t_i)\mathbf{E}(t_i)]$ are also matrices of dimensions $2 \times q$, with $q = m + l + l'$.

Although the problem is completely determined from the computation of the impulse response function by solving Equation (9), our aim here is to provide a recursive algorithm for its computation. Next, the recursive formulas for computing the estimator (8) and its associated error $p(t_k|t_n) = E[|s(t_k) - \hat{s}(t_k|t_n)|^2]$ are devised.

Theorem 3.1 *The optimal WL estimate $\hat{s}(t_k|t_n)$ defined in (8) can be recursively computed as follows:*

$$\hat{s}(t_k|t_n) = \boldsymbol{\psi}_{[1]}(t_k)\boldsymbol{\epsilon}(t_n), \qquad t_k \ge t_n \qquad (12)$$

where the q-vector $\boldsymbol{\epsilon}(t_n)$ is recursively computed from the expression

$$\boldsymbol{\epsilon}(t_n) = \boldsymbol{\epsilon}(t_{n-1}) + \mathbf{J}(t_n, t_n)\left[\mathbf{y}(t_n) - \boldsymbol{\Phi}(t_n)\boldsymbol{\epsilon}(t_{n-1})\right]$$
$$\boldsymbol{\epsilon}(t_0) = \mathbf{0}_q$$

with the $q \times 2$-matrix $\mathbf{J}(t_n, t_n)$ given by the equation

$$\mathbf{J}(t_n, t_n) = [\boldsymbol{\Gamma}^{H}(t_n) - \mathbf{Q}(t_{n-1})\boldsymbol{\Phi}^{H}(t_n)]\boldsymbol{\Omega}^{-1}(t_n) \qquad (14)$$

with the 2×2-matrix $\boldsymbol{\Omega}(t_n) = \Sigma + [\boldsymbol{\Gamma}(t_n) - \boldsymbol{\Phi}(t_n)\mathbf{Q}(t_{n-1})]\boldsymbol{\Phi}^{H}(t_n)$ and the $q \times q$-matrix $\mathbf{Q}(t_n)$ satisfying the recursive equation

$$\mathbf{Q}(t_n) = \mathbf{Q}(t_{n-1}) + \mathbf{J}(t_n, t_n)\left[\boldsymbol{\Gamma}(t_n) - \boldsymbol{\Phi}(t_n)\mathbf{Q}(t_{n-1})\right]$$
$$\mathbf{Q}(t_0) = \mathbf{0}_{q \times q} \qquad (15)$$

Moreover, the associated error is given by the expression

$$p(t_k|t_n) = r_s(t_k, t_k) - \boldsymbol{\psi}_{[1]}(t_k)\mathbf{Q}(t_n)\boldsymbol{\psi}_{[1]}^{H}(t_k), \qquad t_k \ge t_n \qquad (16)$$

Proof. From (10) and (11), Equation (9) can be rewritten as

$$\mathbf{h}^{T}(t_k, t_j, t_n)\sum = \boldsymbol{\psi}_{[1]}(t_k)\,\boldsymbol{\Gamma}^{H}(t_j) - \sum_{i=1}^{n} \mathbf{h}^{T}(t_k, t_i, t_n)\,\mathbf{R}(t_i, t_j)$$

Then, if we introduce a function $\mathbf{J}(t_j, t_n)$ satisfying the equation

$$\mathbf{J}(t_j, t_n)\sum = \boldsymbol{\Gamma}^{H}(t_j) - \sum_{i=1}^{n} \mathbf{J}(t_i, t_n)\,\mathbf{R}(t_i, t_j) \qquad (17)$$

we obtain that

$$\mathbf{h}^{T}(t_k, t_j, t_n) = \boldsymbol{\psi}_{[1]}(t_k)\mathbf{J}(t_j, t_n) \qquad (18)$$

and then, substituting (18) in (8), and defining the function

$$\boldsymbol{\epsilon}(t_n) = \sum_{i=1}^{n} \mathbf{J}(t_i, t_n)\mathbf{y}(t_i) \qquad (19)$$
$$\boldsymbol{\epsilon}(t_0) = \mathbf{0}_q$$

the Equation (12) for the optimal estimator is devised.

Now, subtracting the Equation (17) for t_n and t_{n-1} and taking (11) into account, we can write

$$[\mathbf{J}(t_j, t_n) - \mathbf{J}(t_j, t_{n-1})]\sum = -\mathbf{J}(t_n, t_n)\,\boldsymbol{\Phi}(t_n)\,\boldsymbol{\Gamma}^{H}(t_j) - \sum_{i=1}^{n-1}[\mathbf{J}(t_i, t_n) - \mathbf{J}(t_i, t_{n-1})]\,\mathbf{R}(t_i, t_j)$$

Thus, from (17), we have the relation

$$\mathbf{J}(t_j, t_n) - \mathbf{J}(t_j, t_{n-1}) = -\mathbf{J}(t_n, t_n)\boldsymbol{\Phi}(t_n)\mathbf{J}(t_j, t_{n-1}) \qquad (20)$$

As a consequence, subtracting the Equation (19) for t_n and t_{n-1} and using (20) in the resulting equation, the recursive expression (13) is obtained.

Next, we proceed to derive expression (14) for $\mathbf{J}(t_n, t_n)$. By taking $t_j = t_n$ in (17) and using (11), we have

$$\mathbf{J}(t_n, t_n)\sum = \boldsymbol{\Gamma}^{H}(t_n) - \sum_{i=1}^{n} \mathbf{J}(t_i, t_n)\,\boldsymbol{\Gamma}(t_i)\,\boldsymbol{\Phi}^{H}(t_n) = [\boldsymbol{\Gamma}^{H}(t_n) - \mathbf{Q}(t_n)\,\boldsymbol{\Phi}^{H}(t_n)] \qquad (21)$$

where we have introduced the $q \times q$-matrix

$$\mathbf{Q}(t_n) = \sum_{i=1}^{n} \mathbf{J}(t_i, t_n)\,\boldsymbol{\Gamma}(t_i) \qquad (22)$$
$$\mathbf{Q}(t_0) = \mathbf{0}_{q \times q}$$

Moreover, if we subtract $\mathbf{Q}(t_{n-1})$ from $\mathbf{Q}(t_n)$, and use (20) and (22) in the resulting expression, the recursive Equation (15) for $\mathbf{Q}(t_n)$ is derived. Finally, using (15) in (21), it is easy to check that $\mathbf{J}(t_n, t_n)$ satisfies the expression (14).

Finally, in order to derive expression (16) for the error $p(t_k|t_n)$ associated with the above estimate, we remark that, from the orthogonal projection lemma, this function can be expressed as

$$p(t_k|t_n) = r_s(t_k, t_k) - E[\hat{s}(t_k|t_n)\hat{s}^*(t_k|t_n)]$$

Then, substituting (12) in the above equation and using (17) and (19), we check that

$$p(t_k|t_n) = r_s(t_k, t_k) - \boldsymbol{\psi}_{[1]}(t_k) E[\boldsymbol{\epsilon}(t_n)\boldsymbol{\epsilon}^H(t_n)] \boldsymbol{\psi}_{[1]}^H(t_k) = r_s(t_k, t_k) - \boldsymbol{\psi}_{[1]}(t_k) \sum_{i=1}^{n} J(t_i, t_n) \boldsymbol{\Gamma}(t_i) \boldsymbol{\psi}_{[1]}^H(t_k)$$

As a consequence, from (22), (16) is obtained.

Remark 1 *When* $\{s(t_i), t_i \in T\}$ *is a factorizable real-valued signal with correlation function of the form* (1), *and the observations of the signal verify the complex-valued linear Equation* (6) *with*

$$r_{sv}(t_i, t_j) = \begin{cases} \mathbf{c}^T(t_i)\mathbf{D}^H(t_j), & t_i \geq t_j \\ \mathbf{e}^T(t_i)\mathbf{F}^H(t_j), & t_i \leq t_j \end{cases}$$

where $\mathbf{c}(t_i)$ *and* $\mathbf{e}(t_i)$ *are vectors of dimensions* l *and* l', *respectively, and* $\mathbf{D}(t_i)$ *and* $\mathbf{F}(t_i)$ *are matrices of respective dimensions* $2 \times l$ *and* $2 \times l'$, *we obtain that Algorithm 3.1 holds, replacing the involved* $2 \times q$-*matrices* $\boldsymbol{\Psi}(t_i)$, $\boldsymbol{\Pi}(t_i)$ *by the vectors* $\boldsymbol{\psi}^T(t_i) = [\boldsymbol{\alpha}^T(t_i), \mathbf{c}^T(t_i), \mathbf{0}_l^T]$ *and* $\boldsymbol{\pi}^T(t_i) = [\boldsymbol{\beta}^T(t_i), \mathbf{0}_l^T, \mathbf{e}^T(t_i)]$, *and taking the matrices* $\boldsymbol{\Phi}(t_i) = [\mathbf{g}(t_i)\boldsymbol{\alpha}^T(t_i), \mathbf{g}(t_i)\mathbf{c}^T(t_i), \mathbf{F}(t_i)]$ *and* $\boldsymbol{\Gamma}(t_i) = [\mathbf{g}(t_i)\boldsymbol{\beta}^T(t_i), \mathbf{D}(t_i), \mathbf{g}(t_i)\mathbf{e}^T(t_i)]$, *with* $\mathbf{g}(t_i) = [g(t_i), g^*(t_i)]^T$.

Remark 2 *The efficiency of Algorithm 3.1 is closely related to the dimensions* l, l', *and* m *of the matrices involved in the factorizations* (2) *and* (7). *Indeed, the computational complexity of this algorithm is of order* q, *with* $q = m + l + l'$, *and thus, it involves a further complication in implementation and an increased computational burden as* q *grows. Since the factorization of the covariance is not unique then, the key question is in choosing the one which minimizes the dimension* q. *There exist simple cases, as illustrated previously, where the factorization is easily obtained. Nevertheless, in those more complex cases where this factorization is not trivial, one can use several methods available in the literature to get a factorization with minimum dimension (see, e.g.,* [19]).

3.1 Numerical examples

The advantages of the proposed Algorithm with respect to the SL solution are illustrated here through three numerical examples. The first one involves real correlation matrices and analyzes the effectiveness of the WL processing with respect to the SL one in terms of the impropriety degree of the observations. In the second example, complex correlation matrices are considered and the resulting WL and SL estimation errors are graphically compared. Finally, the third example shows a real application to seismic signal processing.

3.1.1 Example

Let $\{x(t_i), t_1 \leq t_i \leq t_{100}\}$, with $t_i = i/100$, $i = 1, ..., 100$, be an Ornstein-Uhlenbeck process with correlation function

$$r_x(t_i, t_j) = \exp(-|t_i - t_j|), \quad t_1 \leq t_i, t_j \leq t_{100}$$

which is transmitted over a channel that rotates it by a standard normal phase θ and adds a doubly white Gaussian noise $v(t_i)$ correlated with the signal with

$$\mathbf{R}_{xv}(t_i, t_j) = \begin{pmatrix} \frac{1}{25}t_i^2 t_j & \frac{1}{25}t_i^2 t_j \\ \frac{1}{25}t_i^2 t_j & \frac{1}{25}t_i^2 t_j \end{pmatrix}$$

Thus, the signal of interest is $s(t_i) = e^{\theta j}x(t_i)$ and the observations $y(t_i)$ are of the form (6) with $g(t_i) = 1$. Moreover, we assume that θ is independent of $x(t_i)$ and $v(t_i)$.

Note that the correlation matrix of the augmented signal $\mathbf{s}(t_i)$ can be expressed in the form (2), where $\mathbf{A}$ and $\mathbf{B}$ are as in (4) with $l = 1$, $a(t_i) = e^{-t_i}$, $b(t_i) = e^{t_i}$, and $\phi_\theta(-2) = \phi_\theta(2) = e^{-2}$. Moreover, the cross-correlation matrix between $\mathbf{s}(t_i)$ and $\mathbf{v}(t_i)$ is of the form (7) with $\mathbf{D}(t_i) = \mathbf{F}(t_i) = [t_i/25, t_i/25]^T$ and $\mathbf{C}(t_i) = \mathbf{E}(t_i) = [e^{-1/2}t_i^2, e^{-1/2}t_i^2]^T$.

In this example, we consider the problem of computing the fixed-lead predictor $\hat{s}(t_{k+10}|t_k)$. As a measure for comparing the performance of the WL and SL fixed-lead predictors we use the one defined in [6], the mean square of the difference between both errors, for

$$\sum = \begin{pmatrix} 2 & \tau \\ \tau & 2 \end{pmatrix}$$

with τ varying within the interval $[1, 2]$:

$$\frac{1}{100} \sum_{k=1}^{100} (\tilde{p}_\tau(t_{k+10}|t_k) - p_\tau(t_{k+10}|t_k))^2 \quad (24)$$

with $p_\tau(t_{k+10}|t_k)$ and $\tilde{p}_\tau(t_{k+10}|t_k)$ denoting the WL and SL fixed-lead prediction errors, respectively, for every value τ. The results obtained are displayed in Figure 1

Figure 1 Performance of fixed-lead predictors. Quadratic mean of the difference between the WL and SL fixed-lead prediction errors.
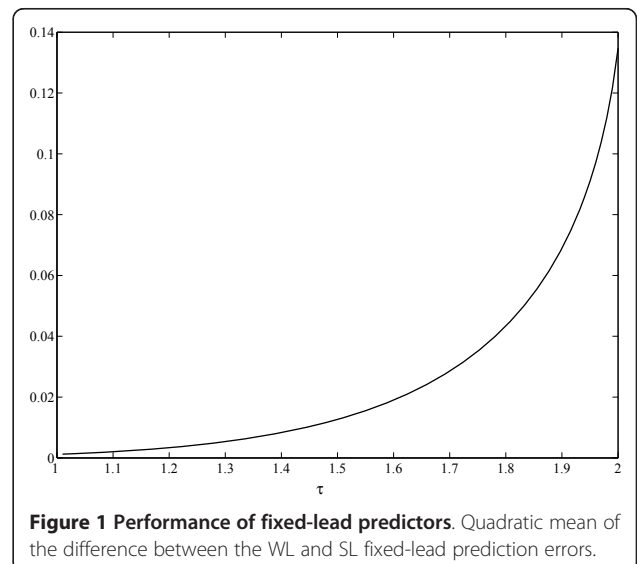
which shows that, not only the WL fixed-lead predictor presents a better behavior than the SL fixed-lead predictor but also the difference between both errors (in the mean square sense) increases with τ, and hence, the WL technique becomes more effective.
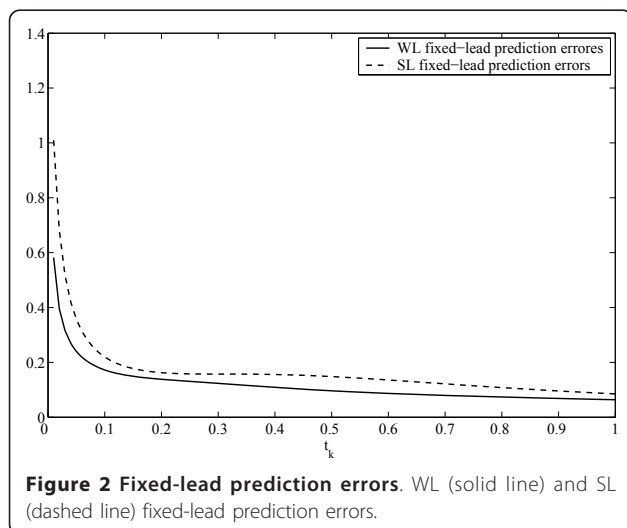
3.1.2 Example

Let $\{s(t_i), t_1 \leq t_i \leq t_{100}\}$, with $t_i = i/100$, $i = 1, ..., 100$, be a signal of the form
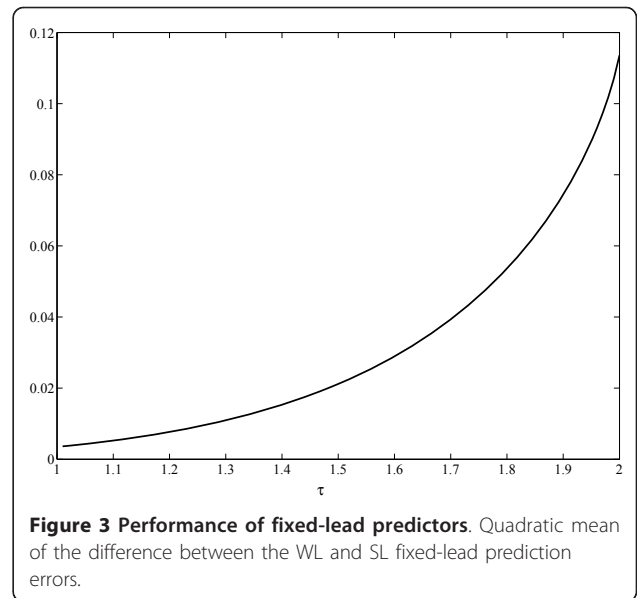
$$s(t_i) = Ae^{jt_i} + Be^{-jt_i}$$

where A and B are complex random variables. In this example, the signal is assumed to be widely stationary, that is $E[AA] = E[BB] = E[AB^*] = 0$ (see [[17], p. 24]), and also we consider that $E[AA^*] = E[BB^*] = 1$ and $E[AB] = -0.8$. Thus, by using (5), the correlation matrix of the signal $s(t_i)$ can be expressed in the form (2). Moreover, we consider that the observations $y(t_i)$ are of the form (6) with $g(t_i) = 1$ and where the noise $v(t_i)$ is uncorrelated with the signal and its augmented variance matrix is

$$\sum = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

On the basis of the set of observations $\{y(t_1), y(t_2), ..., y(t_{100})\}$, we consider the problem of computing the fixed-lead predictor $\hat{s}(t_{k+10}|t_k)$. Then, Algorithm 3.1 is used to obtain the WL fixed-lead prediction error $p_\tau(t_{k+10}|t_k)$ which is compared with the SL fixed-lead prediction error $\bar{p}_\tau(t_{k+10}|t_k)$ in Figure 2. As could be expected, this figure shows that the WL fixed-lead predictor presents a better behavior than the SL fixed-lead predictor. Finally, Figure 3 depicts the performance measure given by (24) with $\tau \in [1, 2]$ and Σ as in (23). Again, the improved precision attained with the WL fixed-lead predictor with

Figure 3 Performance of fixed-lead predictors. Quadratic mean of the difference between the WL and SL fixed-lead prediction errors.

respect to the SL fixed-lead predictor is observed as τ increases.
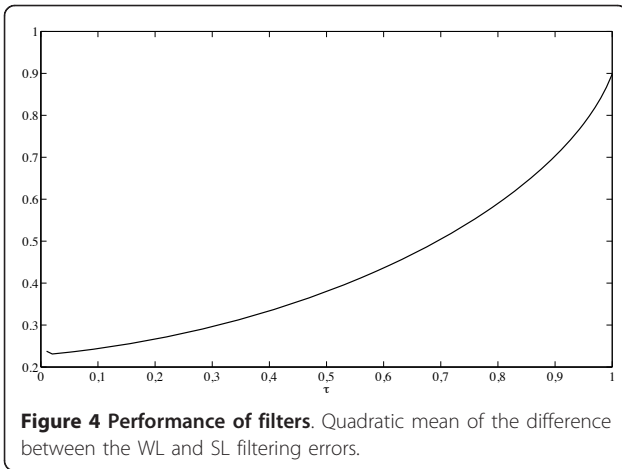
3.1.3 Example

As indicated in Section 2, uniformly modulated nonstationary processes are often used to model seismic records, especially acceleration records. The modulated nonstationary process is given by $s(t_i) = d(t_i)x(t_i)$, where $x(t_i)$ is a stationary process with zero mean and known second-order statistics, and $d(t_i)$ is the time modulating function. A stochastic earthquake model commonly used for $x(t_i)$ is the Kanai-Tajimi process (see, e.g., [20]). It is well-known that the Kanai-Tajimi earthquake model is covariance equivalent with the subset of the ARMA(2,1) model corresponding to a unit value of the spring-dashpot input ratio [21]. For firm ground conditions, at moderate epicentral distance, Kanai and Tajimi have suggested specific values for the parameters in the equation of motion in continuous time whose corresponding discrete ARMA(2,1) model is [21]

$$x(t_i) - 1.604x(t_{i-1}) + 0.686x(t_{i-2}) = e(t_i) - 0.767e(t_{i-1})$$

with $\sigma_e^2 = 39.08$ and $\Delta t = t_i - t_{i-1} = 0.02$ s. This ARMA model has a factorizable correlation function given by (3). Moreover, we have chosen the modulating function $d(t_i) = e^{jt_i}$ and the augmented error covariance matrix of the form

$$\sum = \begin{pmatrix} 1 & \tau \\ \tau & 1 \end{pmatrix}$$

with $\tau \in [0, 1)$. Here we have studied the filtering problem. Figure 4 illustrates the enhancement of the WL filter in relation to the SL one by using the measure

Figure 2 Fixed-lead prediction errors. WL (solid line) and SL (dashed line) fixed-lead prediction errors.

Figure 4 Performance of filters. Quadratic mean of the difference between the WL and SL filtering errors.

$$\frac{1}{100} \sum_{k=1}^{100} \left(\tilde{p}_\tau(t_k|t_k) - p_\tau(t_k|t_k)\right)^2$$

Similar to the previous examples, we observe the superiority of the WL estimate as τ increases.

4 Nonlinear augmented complex prediction

Given the same conditions on the signal established in Section 2, we suppose that the ob-servation process can be given by a nonlinear relation of the form

$$y(t_i) = z(s(t_i), t_i) + v(t_i), \qquad t_1 \le t_i \le t_n \qquad (25)$$

where $z(\cdot)$ is a complex-valued nonlinear function and the signal $s(t_i)$ is uncorrelated with the noise $v(t_i)$. As in the linear case, the aim here is to estimate the signal $s(t_k)$ on the basis of the observation set $\{y(t_1), ..., y(t_n), y^*(t_1), ..., y^*(t_n)\}$, with $t_k \ge t_n$.

The solution to this problem is addressed by following a similar philosophy to the EKF [22]. Specifically, following the basic idea of the EKF the nonlinear function $z(s(t_n), t_n)$ is linearized at each time instant by a first-order Taylor series expanded about the estimated signal $\hat{s}(t_n|t_{n-1})$

$$z(s(t_n), t_n) \approx z(\hat{s}(t_n|t_{n-1}), t_n) + \frac{\partial z(s, t_n)}{\partial s}|_{s=\hat{s}(t_n|t_{n-1})}(s(t_n) - \hat{s}(t_n|t_{n-1}))$$

Consequently, we can proceed to approximate the nonlinear observation Equation (25) as shown by

$$\bar{y}(t_n) \approx g(t_n)s(t_n) + v(t_n) \qquad (26)$$

where

$$\bar{y}(t_n) = y(t_n) - z(\hat{s}(t_n|t_{n-1}), t_n) + g(t_n)\hat{s}(t_n|t_{n-1})$$

and

$$g(t_n) = \frac{\partial z(s, t_n)}{\partial s}\bigg|_{s=\hat{s}(t_n|t_{n-1})} \qquad (27)$$

and thus, $s(t_k)$ can be estimated in terms of the set of observations $\{\bar{y}(t_1), \ldots, \bar{y}(t_n), \bar{y}^*(t_1), \ldots, \bar{y}^*(t_n)\}$ from the relation (26). For that, the formulas given in Algorithm 3.1 for the WL predictor can be used. Note that in the case of the signal and observation noise being uncorrelated, Algorithm 3.1 holds with $\boldsymbol{\Psi}(t_i) = \mathbf{A}(t_i)$, $\boldsymbol{\Pi}(t_i) = \mathbf{B}(t_i)$, $\boldsymbol{\Phi}(t_i) = \mathbf{G}(t_i)\mathbf{A}(t_i)$ and $\boldsymbol{\Gamma}(t_i) = \mathbf{G}(t_i)\mathbf{B}(t_i)$. Then, as in the EKF, in the resulting formulas we also use the linearized observation function (27) in place of the previous function $g(t_n)$ and the term $\mathbf{G}(t_n)\mathbf{A}(t_n)\epsilon(t_{n-1})$ is replaced by the vector

$$\mathbf{z}\left(\hat{s}(t_n|t_{n-1}), t_n\right) = \left[z\left(\hat{s}(t_n|t_{n-1}), t_n\right), z^*\left(\hat{s}(t_n|t_{n-1}), t_n\right)\right]^{\mathrm{T}}$$

Next the formulas of the proposed Algorithm are summarized.

Theorem 4.1 *A nonlinear augmented complex predictor of the signal $s(t_k)$ based on the set of nonlinear observations $\{y(t_1), ..., y(t_n), y^*(t_1), ..., y^*(t_n)\}$ of the form (25) can be determined through the equation*

$$\hat{s}(t_k|t_n) = \mathbf{a}_{[1]}(t_k)\boldsymbol{\epsilon}(t_n)$$

where the m-vector $\epsilon(t_n)$ is recursively computed from the expression

$$\boldsymbol{\epsilon}(t_n) = \boldsymbol{\epsilon}(t_{n-1}) + \mathbf{J}(t_n, t_n)[\mathbf{y}(t_n) - \mathbf{z}(\hat{s}(t_n|t_{n-1}), t_n)]$$
$$\boldsymbol{\epsilon}(t_0) = \mathbf{0}_m$$

with

$$\mathbf{J}(t_n, t_n) = [\mathbf{B}^H(t_n) - \mathbf{Q}(t_{n-1})\mathbf{A}^H(t_n)]\mathbf{G}^H(t_n)\boldsymbol{\Omega}^{-1}(t_n)$$

where $\boldsymbol{\Omega}(t_n) = \boldsymbol{\Sigma} + \mathbf{G}(t_n)[\mathbf{B}(t_n) - \mathbf{A}(t_n)\mathbf{Q}(t_{n-1})]\mathbf{A}^H(t_n)\mathbf{G}^H(t_n)$, $\mathbf{G}(t_n) = \mathrm{diag}(g(t_n), g^*(t_n))$, with $g(t_n) = \partial z(s, t_n)/\partial s|_{s=\hat{s}(t_n|t_{n-1})}, and \; \mathbf{Q}(t_n)$ satisfies the recursive equation

$$\mathbf{Q}(t_n) = \mathbf{Q}(t_{n-1}) + \mathbf{J}(t_n, t_n)\mathbf{G}(t_n)[\mathbf{B}(t_n) - \mathbf{A}(t_n)\mathbf{Q}(t_{n-1})]$$
$$\mathbf{Q}(t_0) = \mathbf{0}_{m \times m}$$

Remark 3 *Similarly to Remark 1, if the signal $\{s(t_i), t_i \in T\}$ is a real-valued signal with factorizable kernel of the form (1) which is observed through a complex-valued nonlinear equation of the form (25), Algorithm 4.1 holds with $\mathbf{a}_{[1]}(t_i) = \boldsymbol{\alpha}^T(t_i)$, $\mathbf{g}(t_i) = [g(t_i), g^*(t_i)]^T$ and replacing the $2 \times q$-matrices $\mathbf{A}(t_i)$ and $\mathbf{B}(t_i)$ by the vectors $\boldsymbol{\alpha}(t_i)$ and $\boldsymbol{\beta}(t_i)$, respectively.*

4.1 Numerical example

In this Example, the estimation of a real signal on the basis of a set of complex-valued observations is considered. Specifically, let $\{s(t_i), t_1 \le t_i \le t_{100}\}$ be a real Wiener process with unit variance parameter and the observation equation

$$\gamma(t_i) = e^{s(t_i)j} + v(t_i), \quad t_i = i/100, \quad i = 1, \dots, 100$$

where $v(t_i) = e^{\theta j} u(t_i)$, with θ a standard normal phase and $u(t_i)$ a white Gaussian noise with unit spectral height and uncorrelated with the signal $s(t_i)$.
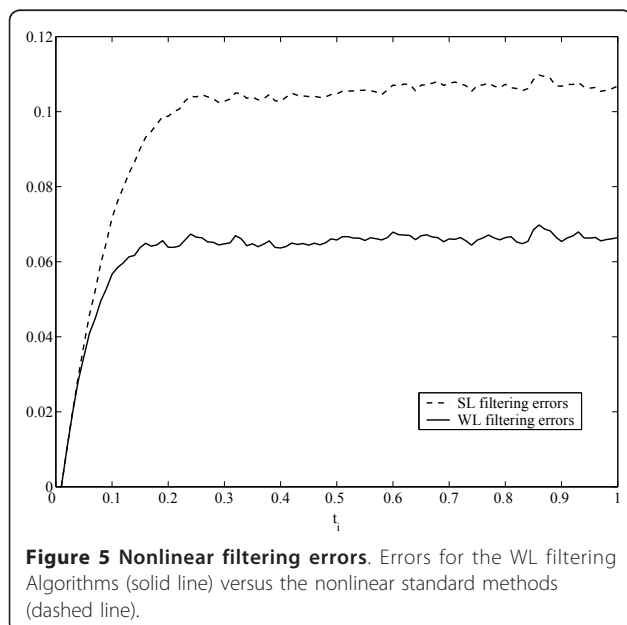
In this case, the correlation function of the signal $s(t_i)$ can be expressed in the form (1) with $\alpha(t_i) = 1$ and $\beta(t_i) = t_i$.

With the aim of examining the good behavior of the WL solution proposed in Algorithm 4.1, the estimation error of the WL filter is compared with the errors associated with SL and WL conventional Algorithms. Specifically, the standard EKF and UKF and the WL EKF and WL UKF proposed in [1] have been implemented. For that, we use the fact that the signal $s(t_i)$ obeys the state equation

$$s(t_i) = s(t_{i-1}) + w(t_i), \quad t_i = i/100, \quad i = 1, \dots, 100$$

with $w(t_i)$ a centered Gaussian signal with variance parameter 10^{-2}.

On the other hand, for computing the estimation errors, Monte Carlo simulations have been performed. Figure 5 shows the results obtained with 5000 sample paths, confirming the better behavior of a WL processing in the nonlinear estimation problem. In fact, the dashed line represents the errors associated with the SL EKF and SL UKF (the differences between them are negligible) and the solid line depicts the errors associated with the WL UKF, WL EKF and the filter given in Algorithm 4.1 (again the differences between them are negligible). Obviously, the similar behavior shown here by the three WL filters has not to be repeated in other examples. As occurs in the standard estimation techniques, there is not a best nonlinear estimator either. In each application one has to pick the appropriate nonlinear estimation method. Really, in every particular case one has to choose the estimator which is found to best trade off various properties such as estimation accuracy, ease of implementation, numerical robustness, and computational burden [22]. Note that unlike UKF and EKF, Algorithm 4.1 does not require a state space model but only the knowledge of the second-order statistics of the processes involved.

Endnotes

[a]A zero-mean complex-valued signal $\{s(t_i), t_i \in T\}$, $T = \{t_1, t_2, \dots, \}$ is said to be proper if the pseudocorrelation function, $\rho_s(t_i, t_j) = E[s(t_i)s(t_j)]$, is null for all $t_i, t_j \in T$. Otherwise, it is called improper. [b]$v(t_i)$ is said to be a doubly white noise if $E[v(t_i)v^*(t_j)] = \sigma_1 \delta_{ij}$ and $E[v(t_i)v(t_j)] = \sigma_2 \delta_{ij}$, with $|\sigma_2| < \sigma_1$ and δ_{ij} stands for the Kronecker delta function [3]. [c]A simple application of the Hilbert space projection theorem shows that WL estimation outperforms SL estimation for general complex-valued signals. Specifically, denote $\overline{\mathrm{sp}}\{\gamma(t_1), \dots, \gamma(t_n)\}$ and $\overline{\mathrm{sp}}\{\gamma(t_1), \dots, \gamma(t_n), \gamma^*(t_1), \dots, \gamma^*(t_n)\}$ the closed spans of the following sets $\{y(t_1), \dots, y(t_n)\}$ and $\{y(t_1), \dots, y(t_n), y^*(t_1), \dots, y(t_n)\}$, respectively. Let $\tilde{s}(t_k|t_n)$ and $\hat{s}(t_k|t_n)$ be the projections of $s(t_k)$, $t_k \geq t_n$, onto the spaces $\overline{\mathrm{sp}}\{\gamma(t_1), \dots, \gamma(t_n)\}$ and $\overline{\mathrm{sp}}\{\gamma(t_1), \dots, \gamma(t_n), \gamma^*(t_1), \dots, \gamma^*(t_n)\}$ respectively. Thus, $\tilde{s}(t_k|t_n)$ is the SL estimate of $s(t_k)$ and $\hat{s}(t_k|t_n)$ is its WL estimate. Since $\overline{\mathrm{sp}}\{\gamma(t_1), \dots, \gamma(t_n)\} \subseteq \overline{\mathrm{sp}}\{\gamma(t_1), \dots, \gamma(t_n), \gamma^*(t_1), \dots, \gamma^*(t_n)\}$, then, by the projection theorem, it follows that the mean square error of $\hat{s}(t_k|t_n)$ is smaller than that of $\hat{s}(t_k|t_n)$.

Figure 5 Nonlinear filtering errors. Errors for the WL filtering Algorithms (solid line) versus the nonlinear standard methods (dashed line).

Abbreviations

EKF: extended Kalman filter; SL: strictly linear; SOS: second-order stationary; UKF; un-scented Kalman filter; WL: widely linear.

Competing interests

The authors declare that they have no competing interests.

References

1. DP Mandic, VSL Goh, *Complex Valued Nonlinear Adaptative Filters. Noncircularity, Widely Linear and Neural Models*, Wiley, New York, (2009)
2. B Picinbono, P Chevalier, Widely linear estimation with complex data. IEEE Trans Signal Process. **43**(8), 2030–2033 (1995). doi:10.1109/78.403373
3. B Picinbono, P Bondon, Second-order statistics of complex signals. IEEE Trans Signal Process. **45**(2), 411–420 (1997). doi:10.1109/78.554305
4. C CheongTook, DP Mandic, Augmented second-order statistics of quaternion random signals. Signal Process. **91**(2), 214–224 (2011). doi:10.1016/j.sigpro.2010.06.024

5. VSL Goh, DP Mandic, An augmented extended kalman filter algorithm for complex-valued recurrent neural networks. Neural Comput. **19**, 1039–1055 (2007). doi:10.1162/neco.2007.19.4.1039
6. J Navarro-Moreno, ARMA prediction of widely linear systems by using the innovations algorithm. IEEE Trans Signal Process. **56**(7), 3061–3068 (2008)
7. P Rubin-Delanchy, AT Walden, Kinematics of complex-valued time series. IEEE Trans Signal Process. **56**(9), 4189–4198 (2008)
8. J Navarro-Moreno, J Moreno-Kayser, RM Fernández-Alcalá, JC Ruiz-Molina, Widely linear estimation algorithms for second-order stationary signals. IEEE Trans Signal Process. **57**(12), 4930–4935 (2009)
9. Y Xia, C CheongTook, DP Mandic, An augmented affine projection algorithm for the filtering of noncircular complex signals. Signal Process. **90**(6), 1788–1799 (2010). doi:10.1016/j.sigpro.2009.11.026
10. AM Martínez-Rodríguez, J Navarro-Moreno, RM Fernández-Alcalá, JC Ruiz-Molina, A general solution to the continuous-time estimation problem under widely linear processing. EURASIP J Adv Signal Process. (accepted, 2011)
11. T Adali, S Haykin, *Adaptive Signal Processing: Next Generation Solutions* Wiley, New York, (2010)
12. S Cambanis, A general approach to linear mean-square estimation problems. IEEE Trans Inf Theory. **IT-19**, 110–114 (1973)
13. HV Poor, *An Introduction to Signal Detection and Estimation*, 2nd edn. (Springer-Verlag, New York, 1998)
14. M Sugisaka, The design of on-line least-squares estimators given covariance specifications via an imbedding method. Appl Math Comput. **13**, 55–85 (1983). doi:10.1016/0096-3003(83)90030-9
15. RM Fernández-Alcalá, J Navarro-Moreno, JC Ruiz-Molina, On the smoothing estimation problem for the intensity of a DSMPP. Methodol Comput Appl Probab 1–12 (2010), http://dx.doi.org/10.1007/s11009-010-9165-z
16. A Moustafa, I Takewaki, Use of probabilistic and deterministic measures to identify unfavourable earthquake records. J Zhejiang Univ Sci A. **10**, 619–634 (2009), http://dx.doi.org/10.1631/jzus.A0930001
17. PJ Schreier, LL Scharf, *Statistical Signal Processing of Complex-Valued Data* (Cambridge University Press, New York, 2010)
18. M Arató, S Baran, M Ispány, Functionals of complex Ornstein-Uhlenbeck processes. Comput Math Appl. **37**, 1–13 (1999)
19. A Baggeroer, State variable analysis procedures, in *Detection, Estimation, and Modulation Theory*, Wiley, New York, 286–327 (1971)
20. S Wang, H Hong, Quantiles of critical separation distance for nonstationary seismic excitations. Eng Struct. **28**, 985–991 (2006). doi:10.1016/j.engstruct.2005.11.003
21. JP Conte, KS Pister, SA Mahin, Nonstationary ARMA modeling of seismic motions. Soil Dyn Earthquake Eng. **11**, 411–426 (1992). doi:10.1016/0267-7261(92)90005-X
22. S Haykin, *Kalman Filtering and Neural Networks* (John Wiley & Sons, Inc., New York, 2001)