RESEARCH   **Open Access**

A general solution to the continuous-time estimation problem under widely linear processing

Ana María Martínez-Rodríguez, Jesús Navarro-Moreno, Rosa María Fernández-Alcalá[*] and Juan Carlos Ruiz-Molina

Abstract

A general problem of continuous-time linear mean-square estimation of a signal under widely linear processing is studied. The main characteristic of the estimator provided is the generality of its formulation which is applicable to a broad variety of situations, including finite or infinite intervals, different types of noises (additive and/or multiplicative, white or colored, noiseless observation data, etc.), capable of solving three estimation problems (smoothing, filtering or prediction), and estimating functionals of the signal of interest (derivatives, integrals, etc.). Its feasibility from a practical standpoint and a better performance with respect to the conventional estimator obtained from strictly linear processing is also illustrated.

Keywords: Continuous-time processing, Linear mean-square estimation problem, Widely linear processing
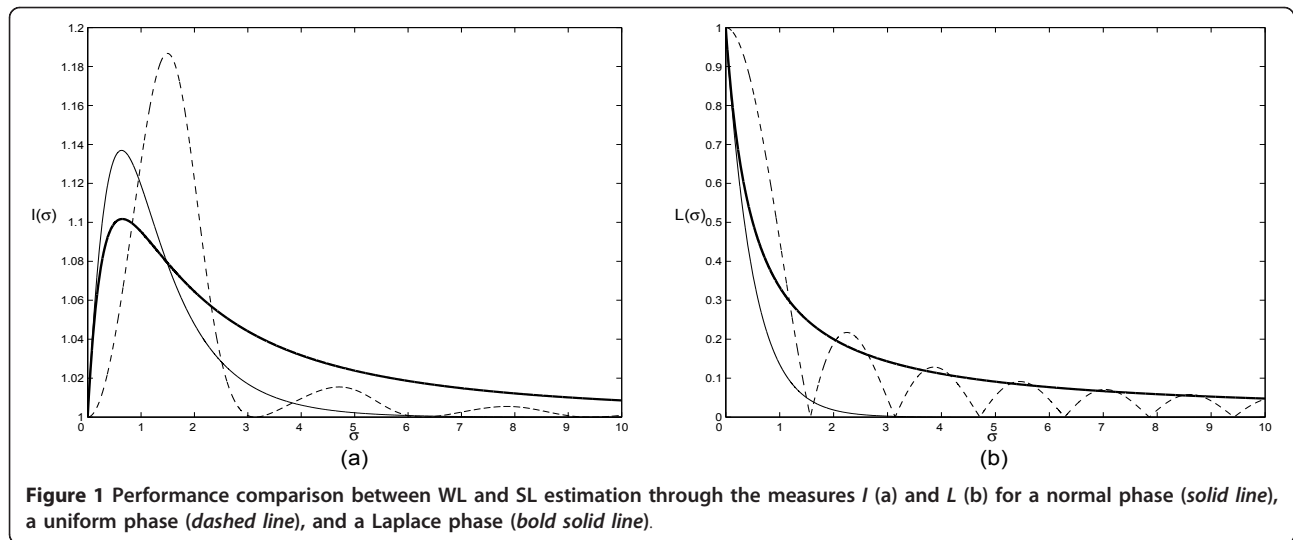
1 Introduction

In most engineering systems, the state variables represent some physical quantity that is inherently continuous in time (ground-motion parameters, atmospheric or oceanographic flow, and turbulence, etc.). Thus, the formulation of realistic models to represent a signal processing problem is one of the major challenges facing engineers and mathematicians today. Given that in many problems the incoming information is constituted by continuous-time series, the use of a continuous-time model will be a more realistic description of the underlying phenomena we are trying to model. For example, [1] gives techniques of continuous-time linear system identification, and [2] illustrates the use of stochastic differential equations for modeling dynamical phenomena (see also the references therein). Continuous-time processing is especially suitable when data are recorded continuously, as an approximation for discrete-time sampled systems when the sampling rate is high [3] and when data are sampled irregularly [4]. It is also necessary with applications that require high-frequency signal processing and/or very fast initial convergence rates. Analog realizations also result in a smaller integrated

circuit, lower power dissipation, and freedom from clocking and aliasing effects [5,6]. In such cases, the continuous-time solution becomes an adequate alternative to the discrete one since it allows real-time processing and alleviates the overload problem assuring more reliable overall operation of the system [7]. Moreover, the analytical tools developed in the continuous-time case might bring new insights to the analysis which are not possible in their discrete-time counterparts. In particular, [8] illustrates this fact in the problem of sorting continuous-time signals, [9] in the problem of nonfragile H_∞ filtering for a class of continuous-time fuzzy systems, and [10] in the study of the behavior of the continuous-time spectrogram.

The estimation problem is a topic of great interest in the statistical signal processing community. This problem has traditionally been solved by using a conventional or strictly linear (SL) processing. For instance, [11,12] deal with classical estimation problems (e.g., the Kalman-Bucy filter) under a real formalism, [13] tackles similar problems in the complex field, and [14] uses factorizable kernels for solving such problems. The main characteristic of the SL treatment is that it takes into account only the autocorrelation of the complex-valued observation process, ignoring its complementary function. That is, the only information considered for the

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

Figure 1 Performance comparison between WL and SL estimation through the measures *I* (a) and *L* (b) for a normal phase (*solid line*), a uniform phase (*dashed line*), and a Laplace phase (*bold solid line*).

building of the estimator is that supplied by the observation process, while the information provided by its conjugate is ignored. Cambanis [15] provided the more general solution to the problem of continuous-time linear mean-square (MS) estimation of a complex-valued signal on the basis of noisy complex-valued observations under a SL processing. In fact, Cambanis's approach is valid for any type of second-order signals and observation intervals, and it is not necessary to impose conditions such as stationarity, Gaussianity or continuity on the involved processes, nor restrictions of finite intervals.

Recently, it has been proved that the treatment of the linear MS estimation problem through widely linear (WL) processing, which takes into account both the observation process and its conjugate, leads to estimators with better performance than the SL ones in the sense that they show lower error variance. Specifically, and from a discrete-time perspective, the WL regression problem was tackled in [16], the prediction problem in a complex autoregressive modeling setting was addressed in [17,18] and later extended to autoregressive moving average prediction in [19]. Also, an augmented[1] affine projection algorithm based on the full second-order statistical information has been newly devised in [20]. Among the wide range of applications of WL processing is the analysis of communication systems [21], ICA models [22], quaternion domain [23], adaptive filters [24-26], etc.

The study of continuous-time estimation problems is also interesting because it provides precise information on some structural properties of the system under study [8,9]. For instance, an explicit expression of the MS error associated with the optimal estimator can be derived in this approach (e.g., see [12,13]). Notice that

this well-known result is independent of the number of available observations. In addition, the continuous-time solution becomes an excellent alternative to the discrete one when the number of available data is large. Discrete-time solutions involve the explicit calculation of matrix inverses whose dimensions depend on the number of observations (see, e.g., [16]). In practice, the process would be cumbersome or even prohibitive if this number were large (as occurs, e.g., in a major earthquake where the workload of the system increases suddenly).

The WL estimation problem under a continuous-time formulation was initially dealt with in [27,28] and [29]. More precisely, the particular problem of estimating a complex signal in additive complex white noise is solved in [27] or [28] through an improper version of the Karhunen-Loève expansion. A general result comparing the performance of WL and SL processing is also presented in which it is shown that the performance gain, measured by MS error, can be as large as 2. Finally, [29] provides an extension of the previous problem to the case in which the additive noise is made up of the sum of a colored component plus a white one. The handicaps of both solutions are: *i)* they are limited to MS continuous signals, *ii)* the signals must be defined on finite intervals, *iii)* the model for the observation process involves additive noise (white noise in the case of [27] and [28]), and *iv)* they are only devoted to solving a smoothing problem.

In this paper, we address a more general estimation problem than those solved in [27-29]. For that, we consider the general formulation of the estimation problem given in [15], and we solve it by using WL processing. The generality of this formulation allows the solution of a wide range of problems, including general second-order

processes, infinite observation intervals, additive and/or multiplicative noise, noiseless observations, estimation of functionals of the signal, etc. It also brings under a single framework three different kinds of estimation problems: prediction, filtering, and smoothing. Hence, all the above handicaps are avoided with the proposed solution. Specifically, we present two forms of the WL estimator depending on the nature, either proper or improper, of the observation process. Then, we state conditions to express such an estimator in closed form. Closed form expressions for the estimator are convenient from a computational point of view [11,12,15]. Three numerical examples show that the proposed solution is feasible and demonstrate the aforementioned generality. The first one compares the performance of the WL estimator in relation to the SL one by considering an observation process defined on an infinite interval and with multiplicative noise. The second concerns the problem of estimating a signal in nonwhite noise and illustrates its application with discrete data. Lastly, the third example considers the earthquake ground-motion representation problem and illustrates a possible real application.

The rest of this paper is organized as follows. In Section 2, we review the SL solution proposed in [15]. Section 3 presents the main results. We derive the new estimator and its associated MS error. Moreover, we prove the better performance of this in relation to the SL estimator, and we give conditions to obtain a closed form of the WL estimator. The results obtained in this section are first stated and then proved rigorously in an Appendix. This section also includes a brief description of how the technique can be implemented in practice. Finally, Section 4 contains three numerical examples illustrating the application of the suggested estimator, and a performance comparison between WL and SL estimation is carried out.

Throughout this paper, all the processes involved are complex, measurable and of second-order. Next, we introduce the basic notation. The real part of a complex number will be denoted by $\mathcal{R}\{\cdot\}$, the complex conjugate by $(\cdot)^*$, the conjugate transpose by $(\cdot)^{\mathrm{H}}$ and the orthogonality of two complex-valued random variables, say a and b, by $a \perp b$. Also, a.s. stands for almost surely and a.e. for almost everywhere.

2 Strictly Linear Estimation

A core problem in signal processing theory is the estimation of a signal from the information supplied by another signal. A very general formulation of this problem was provided by Cambanis in [15]. Specifically, let F and G be two functionals and $\{s(t), t \in S\}$ be a random signal, where S is any interval of the real line. Suppose that $s(t)$ is not observed directly and that we observe the process

$$x(t) = F(s(\tau), \tau \in S, t), \quad t \in T$$

where T is any interval of the real line. Based on the observations $\{x(t), t \in T\}$, the aim is to estimate a functional of $s(t)$

$$\xi(t) = G(s(\tau), \tau \in S, t), \quad t \in S'$$

S' being any interval of the real line.

As noted above, this formulation is very general and contains as particular cases a great number of classical estimation problems, such as estimation of signals in additive and/or multiplicative noise, estimation of signals observed through random channels, random channel identification, etc. [15]. It can also be adapted to treat filtering, prediction, and smoothing problems.

In order to proceed with the building of the Cambanis estimator, the second-order statistics of the processes involved are needed. Let $r_x(t, \tau)$ and $r_\xi(t, \tau)$ be the respective autocorrelation functions of $x(t)$ and $\xi(t)$. Let $c_x(t, \tau) = E[x(t)x(\tau)]$ denote the complementary autocorrelation function of $x(t)$. Moreover, we denote the cross-correlation functions of $\xi(t)$ with $x(t)$ and $x^*(t)$ by $\rho_1(t, \tau) = E[\xi(t)x^*(\tau)]$ and $\rho_2(t, \tau) = E[\xi(t)x(\tau)]$, respectively.

The weakness of the hypotheses imposed on the processes and the possibility of considering infinite intervals force us to construct measures other than Lebesgue measure. To avoid an excess of mathematical formalism, we do not follow the Cambanis exposition literally. Changing the measure is equivalent to searching for a function $F(t)$ such that

$$\int_T r_x(t, t)F(t)\mathrm{d}t < \infty \tag{1}$$

This function $F(t)$ can be selected by a trial-and-error method or by using the procedure given in [30], and in addition, it does not have to be unique. This freedom of choice is to be exploited appropriately in every particular case under consideration. For example, if $T = [T_i, T_f]$ and $x(t)$ is MS continuous, then we can select $F(t) = 1$. Some practical examples can be consulted in [31].

Condition (1) guarantees the existence of the eigenvalues and eigenfunctions, $\{\lambda_k\}$ and $\{\varphi_k(t)\}$, respectively, of $r_x(t, \tau)$. Next, we need an orthogonal basis of random variables built from the observation process and the Hilbert space spanned by it. The elements of such a basis take the form $\varepsilon_k = \int_T x(t)\phi_k^*(t)F(t)\mathrm{d}t$ a.s., and let $H(\varepsilon_k)$ be the Hilbert space spanned by the random variables $\{\varepsilon_k\}$. By using SL processing, the estimator $\hat{\xi}_{\mathrm{SL}}(t)$ proposed in [15] is calculated by projecting the process $\xi(t)$ onto $H(\varepsilon_k)$. As a consequence, $\hat{\xi}_{\mathrm{SL}}(t)$ is given by

$$\hat{\xi}_{\text{SL}}(t) = \sum_{k=1}^{\infty} b_k(t)\varepsilon_k, \quad t \in S'$$

with $b_k(t) = \frac{1}{\lambda_k} \int_T \rho_1(t, \tau)\phi_k(\tau)F(\tau)\mathrm{d}\tau$. Moreover, its associated MS error is

$$P_{\text{SL}}(t) = E[|\xi(t) - \hat{\xi}_{\text{SL}}(t)|^2] = r_\xi(t, t) - \sum_{k=1}^{\infty} \lambda_k b_k(t)b_k^*(t), \quad t \in S'.$$

3 Widely Linear Estimation

In general, complex-valued random processes are improper [24], and then the appropriate processing is the WL processing. In this section, we provide a new estimator, $\hat{\xi}_{\text{WL}}(t)$, by using WL processing and calculate its corresponding MS error, $P_{\text{WL}}(t) = E[|\xi(t) - \hat{\xi}_{\text{WL}}(t)|^2]$. To this end, we consider, together with the information supplied by the observation process, $x(t)$, the information provided by its conjugate, $x^*(t)$. Both processes are stacked in a vector giving rise to the augmented observation process, $\boldsymbol{x}(t) = [x(t), x^*(t)]'$, whose autocorrelation function is denoted by $\boldsymbol{r_x}(t, \tau) = E[\boldsymbol{x}(t)\boldsymbol{x}^{\text{H}}(\tau)]$. Notice that $\hat{\xi}_{\text{WL}}(t)$ receives the name of WL estimator because it depends linearly not only on $x(t)$ but also $x^*(t)$ in contrast with the conventional estimator.

In order to find an explicit form of the estimator and its error, we have to distinguish two possibilities in relation to the nature of $x(t)$: proper or improper. If $x(t)$ is proper, i.e., $cx(t, \tau) = 0$, then the expression for the estimator is

$$\hat{\xi}_{\text{WL}}(t) = \hat{\xi}_{\text{SL}}(t) + \sum_{k=1}^{\infty} \bar{b}_k(t)\varepsilon_k^*, \quad t \in S' \qquad (2)$$

where $\bar{b}_k(t) = \frac{1}{\lambda_k} \int_T \rho_2(t, \tau)\phi_k^*(\tau)F(\tau)\mathrm{d}\tau$, and with associated MS error

$$P_{\text{WL}}(t) = P_{\text{SL}}(t) - \sum_{k=1}^{\infty} \lambda_k \bar{b}_k(t)\bar{b}_k^*(t), \quad t \in S' \qquad (3)$$

Expressions (2) and (3) are derived in Theorem 1 in the "Appendix". These expressions extend to the SL ones since if $\rho_2(t, \tau) = 0$, then $\hat{\xi}_{\text{WL}}(t) = \hat{\xi}_{\text{SL}}(t)$ and $P_{\text{WL}}(t) = P_{\text{SL}}(t)$.

On the other hand, in the improper case ($c_x(t, \tau) \neq 0$), and unlike the proper case, it is not as quick to calculate an explicit and easily implemented expression of $\hat{\xi}_{\text{WL}}(t)$. The main difference between both cases is that now the members of the set $\{\varepsilon_k\} \cup \{\varepsilon_k^*\}$ are not orthogonal. In fact, we have

$$E[\varepsilon_k \varepsilon_l] = \int_T \int_T c_x(t, \tau)\phi_k^*(t)\phi_l^*(\tau)F(t)F(\tau)\mathrm{d}t\mathrm{d}\tau \neq 0, \quad k \neq l$$

Thus, the goal will be to calculate an orthogonal basis in the Hilbert space generated by $\{\varepsilon_k\}$ and $\{\varepsilon_k^*\}$, $H(\varepsilon_k, \varepsilon_k^*)$, which avoids this serious problem. This objective is attained in Lemma 1 in the "Appendix" by means of the eigenvalues, $\{\alpha_k\}$, and the corresponding eigenfunctions, $\phi_k(t)$, of $r_x(t, \tau)$. Following a similar reasoning to [28], it can be shown that the eigenfunctions $\phi_k(t)$ have the particular structure given by $\boldsymbol{\varphi}_k(t) = [f_k(t), f_k^*(t)]'$ and are orthonormal in the sense of (10). The elements of this new set are real random variables of the form

$$w_k = \int_T \boldsymbol{\varphi}_k^{\text{H}}(t)\boldsymbol{x}(t)F(t)\mathrm{d}t \quad \text{a.s.} = 2\mathcal{R}\left\{\int_T x(t)f_k^*(t)F(t)\mathrm{d}t\right\} \quad \text{a.s.} \qquad (4)$$

verifying that $E[w_n w_m] = \alpha_n \delta_{nm}$. By using this new set of variables, we can obtain the WL estimator explicitly

$$\hat{\xi}_{\text{WL}}(t) = \sum_{k=1}^{\infty} \psi_k(t)w_k, \quad t \in S' \qquad (5)$$

where $\psi_k(t) = \frac{1}{\alpha_k}(\int_T \rho_1(t, \tau)f_k(\tau)F(\tau)\mathrm{d}\tau + \int_T \rho_2(t, \tau)f_k^*(\tau)F(\tau)\mathrm{d}\tau)$, and its corresponding MS error is

$$P_{\text{WL}}(t) = r_\xi(t, t) - \sum_{k=1}^{\infty} \alpha_k \psi_k(t)\psi_k^*(t), \quad t \in S' \qquad (6)$$

Theorem 2 in the "Appendix" proves these assertions.

From a practical standpoint, it would be interesting to get a closed form for $\hat{\xi}_{\text{WL}}(t)$. For that, it is necessary to restrict the kind of processes considered so far. Theorem 3 in the "Appendix" gives conditions in order to express the estimator in the following way

$$\hat{\xi}_{\text{WL}}(t) = \int_T h_1(t, \tau)x(\tau)F(\tau)\mathrm{d}\tau + \int_T h_2(t, \tau)x^*(\tau)F(\tau)\mathrm{d}\tau \quad \text{a.s.} \qquad (7)$$

for some square integrable functions $h_1(t, \cdot)$ and $h_2(t, \cdot)$. Expression (7) is computationally more amenable than (2) or (5). The key question is whether the conditions of Theorem 3 are fulfilled. An example of the latter is the classical problem of estimating an improper complex-valued random signal in colored noise with an additive white part addressed in [29]. Specifically, the observation process considered is

$$x(t) = s(t) + n_c(t) + v(t), \qquad T_i \leq t \leq T_f < \infty$$

where $s(t)$ is an improper complex-valued MS continuous random signal, the colored noise component, n_c, is a complex-valued MS continuous stochastic process uncorrelated with $v(t)$, and $v(t)$ is a complex white noise uncorrelated with the signal $s(t)$. Note that the formulation of the estimation problem treated in [29] is much more restrictive than that studied in the present paper.

Finally, a remarkable advantage of the proposed estimator appears when $\xi(t)$ is a real process, and $x(t)$ is

still complex. In this case, $\hat{\xi}_{WL}(t)$ is real too. However, there is no reason for the SL estimator to be real, which is not convenient when we estimate a real functional. Moreover, if $x(t)$ is proper, then $\hat{\xi}_{WL}(t) = 2\mathcal{R}\{\hat{\xi}_{SL}(t)\}$ and its associated MS error is

$$P_{WL}(t) = r_\xi(t, t) - 2 \sum_{k=1}^{\infty} \lambda_k b_k(t) b_k^*(t), \quad t \in S'$$

which provides a decrease in the error that is twice as great as the SL estimator.

Notice also that the Hilbert space approach we have followed to derive the WL estimators allows us to give an alternative proof of the well-known fact that WL estimation outperforms SL estimation. The estimator $\hat{\xi}_{WL}(t)$ is really obtained by projecting the functional $\xi(t)$ onto the Hilbert space $H(\varepsilon_k, \varepsilon_k^*)$. Observe that $H(\varepsilon_k) \subseteq H(\varepsilon_k, \varepsilon_k^*)$ and then trivially by the projection theorem of the Hilbert spaces[2] [[12], Proposition VII. C.1], we have $P_{WL}(t) \le P_{SL}(t)$, for $t \in S'$, and hence, the WL estimator outperforms the SL one as regards its MS error.

3.1 Practical Implementation of the Estimator

We enumerate the necessary steps in implementing the estimation technique proposed for the estimator (5). Nevertheless, some comments are made on how the algorithm can be adapted to obtain (2). Moreover, the role played by (7) becomes clear at the end of the procedure. The steps are the following:

1) Determine the augmented statistics of the processes involved. In some practical applications, the second-order structure is initially known. In fact, it may be derived from experimental measurements or mathematical models. For instance, the information-bearing signal in the communications problem is purposely designed to have desired statistical properties [32]. Other examples can be consulted in [33,34].

2) Select a function $F(t)$ such that condition (1) holds. As noted above, this function $F(t)$ can be selected by a trial-and-error method or by using the procedure given in [30]. Notice that this function is not unique and, in general, there are many specifications possible.

3) Obtain the eigenvalues $\{\alpha_k\}$ and eigenfunctions $\{\phi_k(t)\}$ associated with $r_x(t, \tau)$. In general, determination of eigenvalues and eigenfunctions, except for a few cases, is a problem that is very involved, if not impossible. However, we can avoid the calculation of true eigenvalues and eigenfunctions by means of the Rayleigh-Ritz (RR) method, which is a procedure for numerically solving operator equations involving only elementary calculus and simple linear algebra (see [31,35] for a detailed study about the practical application of the RR method).

4) Truncate expressions (5) and (6) at n terms and substitute, if necessary, the true eigenvalues and eigenfunctions by the RR ones. This truncated version of the estimator, which is in fact a suboptimum estimator, can be calculated via the expression (7) with

$$h_1(t, \tau) = \sum_{k=1}^{n} \psi_k(t) f_k^*(\tau) \quad \text{and} \quad h_2(t, \tau) = \sum_{k=1}^{n} \psi_k(t) f_k(\tau)$$

and where both functions satisfy the conditions of Theorem 3.

Thus, we have replaced the computation of $2n$ integrals in the truncated version of (5) (or n integrals in the finite series obtained from (2)) by the computation of two integrals in (7), and hence, it entails a reduction in the error of approximation for a given precision.

Note that both the precision and the amount of computation required in applying this method depend heavily on the number n. An easy criterion[3] for determining an adequate level of truncation n without an unnecessary excess of computation can be the following: select n in such a way that $\sum_{k=1}^{n} \alpha_k$ represents at least 95% of the total variance of the process, $\sum_{k=1}^{\infty} \alpha_k = 2 \int_T r_x(t, t) F(t) dt$ (see the proof of Lemma 1 in the "Appendix").

5) Finally, from a discrete set of observations, $x_1, ..., x_N$, we can compute the integrals in (7) by means of

$$\int_T h_1(t, \tau) x(\tau) F(\tau) d\tau \approx \sum_{k=1}^{n} g_1(t, k) x_k$$

$$\int_T h_2(t, \tau) x^*(\tau) F(\tau) d\tau \approx \sum_{k=1}^{n} g_2(t, k) x_k^*$$

where the weights $g_1(t, k)$ and $g_2(t, k)$ are obtained via a suitable method that performs numerical integration with integrands constituted for discrete points. For example, using the Gill-Miller quadrature method [36] implemented by subroutine d01gaf from the NAG Toolbox for MAT-LAB or the trapezoidal rule (trapz function in MATLAB).

The only changes for implementing the estimator (2) are in steps 1 and 3, where we have to use $r_x(t, \tau)$ and their associated eigenvalues and eigenfunctions, $\{\lambda_k\}$ and $\{\varphi_k(t)\}$, instead.

4 Numerical Examples

Three examples illustrate the implementation of the proposed solution and show its capability to solve very general estimation problems. Example 1 shows a situation where true eigenvalues and eigenfunctions are available and aims at comparing the performance of WL processing in relation to SL processing. Example 2 applies the RR method to approximate the

eigenexpansion and also illustrates its implementation with discrete data. Finally, Example 3 considers an application in seismic signal processing in which the ground-motion velocity is estimated from seismic ground acceleration data.

4.1 Example 1

Assume that a real waveform $s(t)$ is transmitted over a channel that rotates it by some random phase θ and adds a noise $n(t)$. Unlike [28] and [29], we consider infinite observation intervals and a multiplicative quadratic noise in the observations. More precisely, $s(t)$ is defined on the real line, $S = \mathbb{R}$, with zero-mean and $r_s(t, \tau) = e^{-(t-\tau)^2}$. Thus, the observation process is given by

$$x(t) = e^{j\theta}s(t)n^2(t), \quad t \in T = \mathbb{R} \tag{8}$$

where $j = \sqrt{-1}$ and the noise $n(t)$ is a zero-mean Gaussian process with $r_n(t, \tau) = 3^{-1/2}p^{1/4}(t)p^{1/4}(\tau)$, where $p(t) = \sqrt{2/\pi}e^{-2t^2}$ (this type of process is studied in [34]). Three different probabilistic distributions for θ are taken: a uniform distribution on $(-\sigma, \sigma)$, a zero-mean normal with variance σ, and a Laplace distribution with zero-mean and variance σ. Several choices of σ will be used to show how the advantages of WL processing vary with the level of improperness of the observations. Finally, mutual independence of θ, $s(t)$ and $n(t)$ is assumed. The objective is to estimate $\dot{s}(t), t \in [0, 1]$, where $\dot{s}(t)$ denotes the MS derivative of $s(t)$.

We first notice that $\int_{-\infty}^{\infty} r_x(t, t)dt < \infty$, where $F(t) = 1$ has been selected by a trial-and-error method and thus, condition (1) is verified. This example is one of the particular cases where calculation of true eigenvalues and eigenfunctions is possible. In fact, $r_x(t, \tau)$ has eigenvalues $(1 + E[e^{2j\theta}]\bar{\lambda}_k$ and $(1 - E[e^{2j\theta}])\bar{\lambda}_k$ with respective associated eigenfunctions $[\phi_k(t)/\sqrt{2}, \phi_k(t)/\sqrt{2}]'$ and $[j\phi_k(t)/\sqrt{2}, -j\phi_k(t)/\sqrt{2}]'$, $k = 0, \dots,$ and where $\bar{\lambda}_k = \sqrt{\frac{2}{2+\sqrt{3}}}\left(\frac{1}{2+\sqrt{3}}\right)^k$, $\phi_k(t) = 2^k k! \frac{1}{3} 3^{3/4} e^{-(\sqrt{3}-1)t^2} H_k\left(\sqrt{2\sqrt{3}}t\right)$ and $H_k(t) = (-1)^k e^{t^2} \frac{\partial^k}{\partial t^k}e^{-t^2}$ are the Hermite polynomials. Moreover, we can check that the associated MS errors are the following:

$$P_{SL}(t) = 2 - E[e^{j\theta}]^2 \sum_{k=0}^{\infty} \frac{l_k^2(t)}{\bar{\lambda}_k} \quad \text{and} \quad P_{WL}(t) = 2 - \frac{2E[e^{j\theta}]^2}{1 + E[e^{2j\theta}]} \sum_{k=0}^{\infty} \frac{l_k^2(t)}{\bar{\lambda}_k}$$

with $l_k(t) = 3^{-1/2} \int_T \frac{\partial}{\partial t}r_s(t, \tau)p^{1/2}(\tau)\phi_k(\tau)d\tau$.

We use the measure

$$I = \frac{\int_0^1 P_{SL}(t)dt}{\int_0^1 P_{WL}(t)dt}$$

which is closely related to the performance measure considered in [29], to compare the performance of WL processing in relation to SL processing. For that, we have truncated the series in $P_{SL}(t)$ and $P_{WL}(t)$ at $n = 10$ terms (this approximate expansion explains 99.86% of the total variance of the process). The performance of both the SL and the WL estimators for $n = 10$ does not really vary substantially from the case of $n > 10$. Figure 1a depicts the measure I in function of σ for the three probabilistic distributions considered for θ. It turns out that the advantages of WL processing decrease in both cases as σ tends toward zero and as σ tends toward infinity. However, this occurs for different reasons. Another performance measure which helps in the interpretation is

$$L = \frac{|c_x(t, s)|}{|r_x(t, s)|}$$

which, for this example, takes the value $L = |E[e^{2j\theta}]|$. Figure 1b shows the index L as a function of σ for the three probabilistic distributions considered for θ. On the one hand, as σ tends toward zero, then the index L tends to one since in that limit the observation process becomes a real signal[4]. On the other hand, when σ increases, then L tends toward zero since $x(t)$ becomes a proper signal. The faster convergence to zero in the normal case and the slower one for the Laplace distribution are also observed.

4.2 Example 2

We study a generalization of the classical communication example addressed in [28] and [29]. Assume that a real waveform $s_1(t)$ is transmitted over a channel that rotates it by a standard normal phase θ_1 and adds a nonwhite noise $n(t)$. More precisely, $s_1(t)$ is defined on the interval [0, 1], with zero-mean and $r_{s_1}(t, \tau) = \min\{t, \tau\}$. Thus, the observation process is

$$x(t) = e^{j\theta_1}s_1(t) + n(t), \quad t \in [0, 1]$$

where the nonwhite noise $n(t)$ is obtained from a linear time-invariant system of the form $n(t) = e^{j\theta_2} \int_0^1 r_{s_1}(t, \tau)s_2(\tau)d\tau$, with θ_2 being a zero-mean normal random variable with variance 2 and $s_2(t)$ a standard Wiener process (these types of noises appear in [[37], p. 357]). Moreover, we assume that θ_1, θ_2, $s_1(t)$, and $s_2(t)$ are independent of each other. This example extends the cases studied in [28] and [29] since the considered noise here does not have a white component and thus, the previous solutions cannot be applied. The observations have been taken in the following time instants: $i/1000$, $i = 1, \dots, 1000$. The objective is to estimate $s(t) = e^{j\theta_1}s_1(t), t \in [0,1]$.

We first notice that $\int_0^1 r_x(t,t)\mathrm{d}t < \infty$, where $F(t) = 1$ has been selected since the processes involved are continuous and thus, condition (1) is verified. Now, to apply the RR method, we choose the Fourier basis of complex exponentials on $[0, 1]$, $\{\exp\{2\pi jk\}\}_{k=-\infty}^{\infty}$. Following the recommendations in step 5 of Section 3.1, we compute the integrals in (7) via the subroutines d01gaf and trapz (there were no significative differences between both methods).

Figure 2 depicts the MS error $P_{\mathrm{WL}}(t)$ together with the MS errors of the WL estimator obtained from the RR method with $n = 25$ and $n = 50$ terms in step 5 of the algorithm, which have been generated by Monte Carlo simulation (a total of 10,000 simulations were performed). We can see that the method may yield a sufficiently accurate solution with a short number n of terms while reducing the complexity of the problem significantly. Note that a truncated expansion at $n = 25$ terms explains 88.77% of the total variance of the process and the expansion with $n = 50$ terms 95.81%.

4.3 Example 3

The seismic ground acceleration can be represented by a uniformly modulated nonstationary process [33]. The modulated nonstationary process is obtained in the following way

$$s(t) = a(t)z(t)$$

where $a(t)$ is a time modulating function that could be a complex function, and $z(t)$ is a stationary process with zero-mean and known second-order moments. In general, the so-called exponential modulating function is adopted [38,39]. A common choice for $z(t)$ is the standard Ornstein-Uhlenbeck process with a particular version of the exponential modulating function given by $a(t) = \mathrm{e}^{-t}$ [[33], p. 38]. Thus, the seismic ground acceleration can be modeled as a stochastic signal $\{s(t), t \in S = \mathbb{R}^+\}$ with $r_s(t, \tau) = \mathrm{e}^{-(t+\tau)}\mathrm{e}^{-|t-\tau|}$. Consider the observation process

$$x(t) = \mathrm{e}^{j\theta}s(t), \quad t \in T = \mathbb{R}^+$$

where θ is a standard normal phase independent of $s(t)$. Now, the objective is to estimate the seismic ground velocity at instant $t \geq 2$, i.e., $\xi(t) = \int_0^1 s(\tau)\mathrm{d}\tau$, with $t \in S' = [2, \infty)$. A justification for considering infinite intervals on the basis of the stationarity property of $z(t)$ can be found in [40].

By using a trial-and-error method, we select $F(t) = \mathrm{e}^{-t}$ and then, (1) holds. For the case of infinite intervals, $T = \mathbb{R}^+$, the true eigenvalues and eigenfunctions of $r_x(t, \tau)$ are not known. We approximate them by means of the RR method. The RR eigenvalues and eigenfunctions of $r_x(t, \tau)$ are $(1 \pm \mathrm{e}^{-2})\bar{\lambda}_k$ and $[\tilde{\phi}_k(t)/\sqrt{2}, \tilde{\phi}_k(t)/\sqrt{2}]'$ and $[j\tilde{\phi}_k(t)/\sqrt{2}, -j\tilde{\phi}_k(t)/\sqrt{2}]'$, where $\tilde{\lambda}_k$ and $\tilde{\phi}_k(t)$ are the RR eigenvalues and eigenfunctions, respectively, of $r_x(t, \tau)$ obtained from the following trigonometric basis

$$\{1, \sqrt{2}\cos(2\pi\mathrm{e}^{-t}), \sqrt{2}\sin(2\pi\mathrm{e}^{-t}), \sqrt{2}\cos(4\pi\mathrm{e}^{-t}), \sqrt{2}\sin(4\pi\mathrm{e}^{-t}), \ldots\}$$
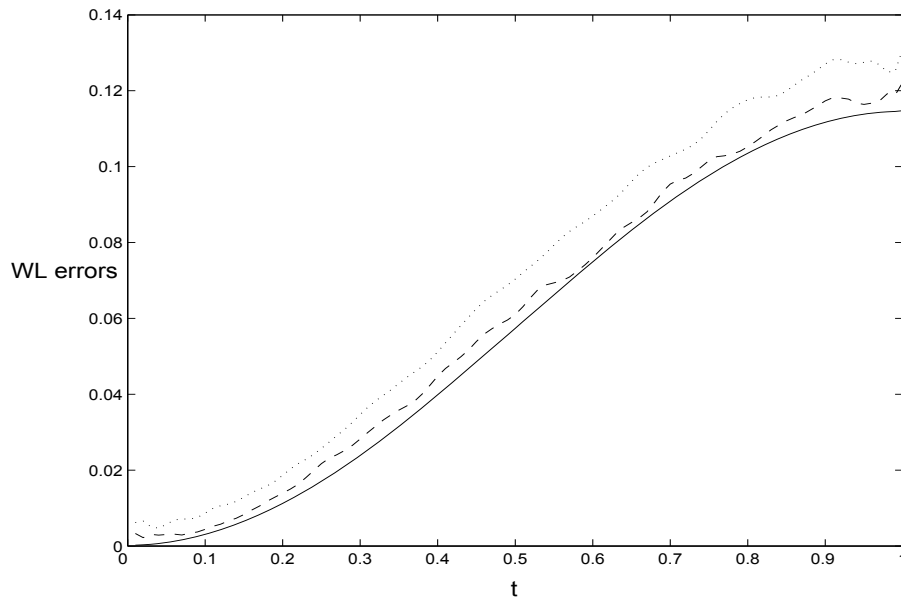
Figure 2 MS errors of the WL estimator (5) *(solid line)* and the estimator calculated in step 5 with *n* = 25 terms *(dotted line)* and with *n* = 50 terms *(dashed line)*.

In Figure 3, we compare the MS error of the SL estimator calculated with $n = 10$ terms with the MS errors of the WL estimator with $n = 2, 4$ and, 10 terms (which account for 57.60, 82.30 and 93.88% of the total variance of $x(t)$, respectively). We have limited the estimation interval to [2, 6] because of the observed stabilization of the MS errors for $t \geq 4$. Apart from the better performance of the WL estimator with respect to the SL estimator (as was to be expected), the rapid convergence of the RR estimators is also confirmed.

5 Concluding Remarks

A new WL estimator has been given for solving general continuous-time estimation problems. The formulation considered can be adapted in order to include as particular cases a great number of estimation problems of interest. The proposed estimator becomes a way that avoids explicit calculation of matrix inverses altogether and can be applied provided that the second-order characteristics of the processes involved are known. Such knowledge is usual in some practical problems in fields as diverse as seismic signal processing, signal detection, finite element analysis, etc. An alternative procedure is the stochastic gradient-based iterative solution called augmented complex least mean-square algorithm (see, e. g., [24]) in which the second-order statistics are estimated from data. However, if we wish to take advantage of the knowledge of the second-order characteristics and the number of observation data is very large, then the continuous-time solution is a recommended option.

Appendix

This "Appendix" is written following a rigorous mathematical formalism parallel to [15] or [30]. Condition (1) is indeed more restrictive than the one imposed in the works of Cambanis. Specifically, suppose μ a measure on $(T, \mathcal{B}(T))$ ($\mathcal{B}(T)$ is the σ-algebra of Lebesgue measurable subsets of T) which is equivalent to the Lebesgue measure and verifies

$$\int_T r_x(t, t) \mathrm{d}\mu(t) < \infty \tag{9}$$

The existence of μ satisfying (9) is proved in [30]. Cambanis also shows that (9) allows us to select a function $F(t)$ such that $\mathrm{d}\mu(t)/\mathrm{d}t = F(t)$ and (1) holds.
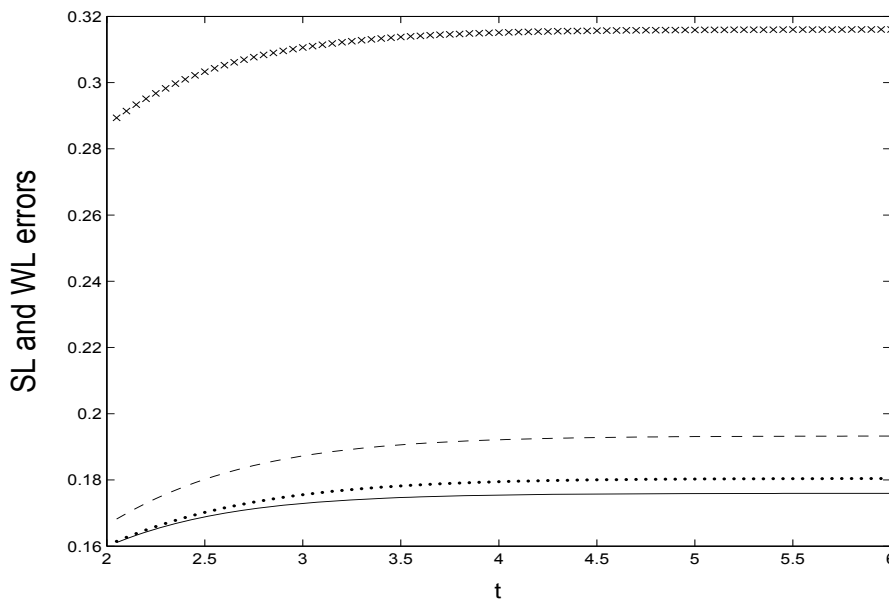
Theorem 1 *If $x(t)$ is proper, then*

$$\hat{\xi}_{\mathrm{WL}}(t) = \hat{\xi}_{\mathrm{SL}}(t) + \sum_{k=1}^{\infty} \bar{b}_k(t) \varepsilon_k^*, \qquad t \in S'$$

with $\bar{b}_k(t) = \frac{1}{\lambda_k} \int_T \rho_2(t, \tau) \phi_k^*(\tau) \mathrm{d}\mu(\tau)$. *Moreover, its associated MS error is*

$$P_{\mathrm{WL}}(t) = P_{\mathrm{SL}}(t) - \sum_{k=1}^{\infty} \lambda_k \bar{b}_k(t) \bar{b}_k^*(t), \qquad t \in S'$$

Proof: Firstly, notice that if $x(t)$ is proper, then the members of the set of random variables $\{\varepsilon_k\} \cup \{\varepsilon_k^*\}$ are orthogonal. Thus, the estimator $\hat{\xi}_{\mathrm{WL}}(t)$ is obtained by projecting the functional $\xi(t)$ onto the Hilbert space

Figure 3 MS errors for the SL estimator with $n = 10$ terms *(crossed line)* **and for the WL estimator with $n = 2$ terms** *(dashed line)*, **with $n = 4$ terms** *(dotted line)*, **and with $n = 10$ terms** *(solid line)*.

generated by $\{\varepsilon_k\}$ and $\{\varepsilon_k^*\}$, $H(\varepsilon_k, \varepsilon_k^*)$. Hence, the estimator can be expressed in the form $\hat{\xi}_{\text{WL}}(t) = \sum_{k=1}^{\infty} b_k(t)\varepsilon_k + \sum_{k=1}^{\infty} \bar{b}_k(t)\varepsilon_k^*$, where the coefficients $b_k(t)$ and $\bar{b}_k(t)$ are determined via the projection theorem of the Hilbert spaces. This result assures that $\xi(t) - \hat{\xi}_{\text{WL}}(t) \perp \{\varepsilon_k\} \cup \{\varepsilon_k^*\}$; that is, $E[\xi(t)\varepsilon_k^*] = E[\hat{\xi}_{\text{WL}}(t)\varepsilon_k^*]$ and $E[\xi(t)\varepsilon_k] = E[\hat{\xi}_{\text{WL}}(t)\varepsilon_k]$, for all k. Since $E[\hat{\xi}_{\text{WL}}(t)\varepsilon_k^*] = \lambda_k b_k(t)$, $E[\hat{\xi}_{\text{WL}}(t)\varepsilon_k^*] = \lambda_k b_k(t)$, $E[\xi(t)\varepsilon_k] = \int_T \rho_2(t, \tau)\phi_k^*(\tau)\mathrm{d}\mu(\tau)$, and $E[\hat{\xi}_{\text{WL}}(t)\varepsilon_k] = \lambda_k \bar{b}_k(t)$, then the first part of the result follows.

On the other hand, the corresponding MS error is

$$P_{\text{WL}}(t) = E[|\xi(t) - \hat{\xi}_{\text{WL}}(t)|^2] = r_\xi(t, t) - \sum_{k=1}^{\infty} \lambda_k b_k(t)b_k^*(t) - \sum_{k=1}^{\infty} \lambda_k \bar{b}_k(t)\bar{b}_k^*(t)$$

∎

We need the following Lemma before proving Theorem 2.

Lemma 1

$$H(w_k) = H(\varepsilon_k, \varepsilon_k^*)$$

Proof: From (9), we get that $r_x(t, \tau)$ is the kernel of an integral operator of $L_2(\mu \times \mu)$ into $L_2(\mu \times \mu)$, which is linear, self-adjoint, nonnegative-definite, and compact. Let $\{\alpha_k\}$ be their eigenvalues and $\{\phi_k(t)\}$ the corresponding eigenfunctions. The eigenfunctions $\varphi_k(t) = [f_k(t), f_k^*(t)]'$ are orthonormal in the following sense

$$\int_T \varphi_n^{\text{H}}(t)\varphi_m(t)\mathrm{d}\mu(t) = 2\mathcal{R}\left\{\int_T f_n^*(t)f_m(t)\mathrm{d}\mu(t)\right\} = \delta_{nm} \qquad (10)$$

Thus, the real random variables given by (4) are trivially orthogonal, i.e., $E[w_n w_m] = a_n \delta_{nm}$.

First, we prove that $H(w_k) \subseteq H(\varepsilon_k, \varepsilon_k^*)$. Let $H(\varepsilon_k^*)$ be the Hilbert space spanned by the random variables $\{\varepsilon_k^*\}$. From Theorem 6 of [30], we have $\int_T x(t)f_k^*(t)\mathrm{d}\mu(t)$ a.s. $\in H(\varepsilon_k)$ and $\int_T x^*(t)f_k(t)\mathrm{d}\mu(t)$ a.s. $\in H(\varepsilon_k^*)$ and hence it is trivial that $w_k \subseteq H(\varepsilon_k, \varepsilon_k^*)$.

Now, we demonstrate that $H(\varepsilon_k, \varepsilon_k^*) \subseteq H(w_k)$. For that, we begin to check that $\varepsilon_k \in H(w_k)$. By projecting $x(t)$ onto $H(w_k)$, we obtain that $x(t) = y(t) + v(t)$ with $y(t) = \sum_{k=1}^{\infty} f_k(t)w_k$ and $y(t)$ is perpendicular to $v(t)$. Thus, we have that $r_x(t, \tau) = r_y(t, \tau) + r_v(t, \tau)$ where $r_y(t, \tau) = E[y(t)y^*(\tau)]$ and $r_v(t, \tau) = E[v(t)v^*(\tau)]$. By the monotone convergence theorem and (10), we get that $\int_T r_x(t, t)\mathrm{d}\mu(t) = \frac{1}{2}\sum_{k=1}^{\infty} \alpha_k + \int_T r_v(t, t)\mathrm{d}\mu(t)$.

On the other hand, $\int_T r_x(t, t)\mathrm{d}\mu(t) = \frac{1}{2}\text{Tr}(r_x) = \frac{1}{2}\sum_{k=1}^{\infty} \alpha_k$, where $\text{Tr}(r_x)$ is the trace of the integral operator on $L_2(\mu \times \mu)$ with kernel $r_x(t, \tau)$.

Thus,

$$\int_T r_v(t, t)\mathrm{d}\mu(t) = 0 \qquad (11)$$

and hence

$$r_x(t, \tau) = r_y(t, \tau) \quad \text{a.e. } [\text{Leb} \times \text{Leb}] \text{ on } T \times T \qquad (12)$$

Now, we consider the integral $\eta_k = \int_T \gamma(t)\phi_k^*(t)\mathrm{d}\mu(t)$ a.s. From (12), we have

$$\int_T \int_T r_\gamma(t, \tau)\phi_k^*(t)\phi_k(\tau)\mathrm{d}\mu(t)\mathrm{d}\mu(\tau) = \lambda_k$$

and then $\eta_k \in H(w_k)$. Moreover, it follows that $E[|\varepsilon_k - \eta_k|^2] = 0$ and then $\varepsilon_k = \eta_k \in H(w_k)$.

Similarly, it can be proved that $\varepsilon_k^* \in H(w_k)$. ∎

Theorem 2 *If $x(t)$ is improper, then*

$$\hat{\xi}_{\text{WL}}(t) \sum_{k=1}^{\infty} \psi_k(t)w_k, \qquad t \in S'$$

where $\psi_k(t) = \frac{1}{\alpha_k}(\int_T \rho_1(t, \tau)f_k(\tau)\mathrm{d}\mu(\tau) + \int_T \rho_2(t, \tau)f_k^*(\tau)\mathrm{d}\mu(\tau))$. *Moreover, its corresponding MS error is*

$$P_{\text{WL}}(t) = r_\xi(t, t) - \sum_{k=1}^{\infty} \alpha_k \psi_k(t)\psi_k^*(t), \qquad t \in S'$$

Proof: Following a reasoning similar to that of proof of Theorem 1 and taking Lemma 1 into account, the result is immediate. ∎

In the next result, we provide conditions in order to hold (7).

Theorem 3 *The WL estimator can be expressed in the following closed form*

$$\hat{\xi}_{\text{WL}}(t) = \int_T h_1(t, \tau)x(\tau)\mathrm{d}\mu(\tau) + \int_T h_2(t, \tau)x^*(\tau)\mathrm{d}\mu(\tau) \quad a.s. \qquad (13)$$

for some $h_1(t, \cdot), h_2(t, \cdot) \in L_2(\mu)$ *if and only if for some* $h_1(t, \cdot), h_2(t, \cdot) \in L_2(\mu)$ *it is satisfied that*

$$\rho_1(t, \tau) = \int_T h_1(t, u)r_x(u, \tau)\mathrm{d}\mu(u) + \int_T h_2(t, u)c_x^*(u, \tau)\mathrm{d}\mu(u)$$

$$\rho_2(t, \tau) = \int_T h_1(t, u)c_x(u, \tau)\mathrm{d}\mu(u) + \int_T h_2(t, u)r_x^*(u, \tau)\mathrm{d}\mu(u) \qquad (14)$$

for $t \in S'$, a.e. $\tau \sim [\text{Leb}]$.
Proof: From (11), we have

$$x(t), x^*(t) \in H(w_k) \quad \text{for almost all } t \in T \text{ [Leb]} \qquad (15)$$

Suppose that $\hat{\xi}_{\text{WL}}(t)$ satisfies (13). It follows from $\xi(t) - \hat{\xi}_{\text{WL}}(t) \perp H(w_k)$ and (15) that $E[\xi(t)x^*(\tau)] = E[\hat{\xi}_{\text{WL}}(t)x^*(\tau)]$ and $E[\xi(t)x(\tau)] = E[\hat{\xi}_{\text{WL}}(t)x(\tau)]$, for almost all $\tau \in T$ [Leb], and thus we obtain (14).

Reciprocally, suppose that (14) holds. Define the process

$$\eta(t) = \int_T h_1(t,\tau)x(\tau)\mathrm{d}\mu(\tau) + \int_T h_2(t,\tau)x^*(\tau)\mathrm{d}\mu(\tau) \quad \text{a.s.}$$

Theorem 6 of [30] guarantees that $\eta(t) \in H(w_k)$. Moreover, from (14), we obtain that $\xi(t) - \eta(t) \perp x(\tau)$ and $\xi(t) - \eta(t) \perp x^*(t)$ for almost all $\tau \in T$ [Leb]. Hence, from the projection theorem of the Hilbert spaces $\hat{\xi}_{\mathrm{WL}}(t) = \eta(t)$ a.s. ■

6 Competing interests

The authors declare that they have no competing interests.

Note

[1] 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.

[2] This result is an extension of the more familiar orthogonality principle for finite-dimensional vector space (see, e.g., [12,13]).

[3] It should be remarked that this criterion only takes into account the information provided by $x(t)$ and the removed coefficients could be very informative about $\xi(t)$.

[4] Notice that the complex nature of $x(t)$ in (8) stems from the term $e^{j\theta}$. Hence, as $\sigma \to 0$, then the variance of θ vanishes and it becomes a degenerate random variable that only takes the value 0 with probability 1.

References
1. D Marelli, M Fu, A continuous-time linear system identification method for slowly sampled data. IEEE Trans Signal Process. **58**(5), 2521–2533 (2010)
2. L Murray, A Storkey, Particle smoothing in continuous time: a fast approach via density estimation. IEEE Trans Signal Process. **59**(3), 1017–1026 (2011)
3. L Guo, B Pasik-Duncan, Adaptive ontinuous-linear quadratic Gaussian control. IEEE Trans Autom Control **44**(9), 1653–1662 (1999). doi:10.1109/9.788532
4. M Mossberg, High-accuracy instrumental variable identification of continuous-time autoregressive processes from irregularly sampled noisy data. IEEE Trans Signal Process. **56**(8), 4087–4091 (2008)
5. W Edmonson, JC Palacios, CA Lai, H Latchman, A global optimization method for continuous-time adaptive recursive filters. IEEE Signal Process Lett. **6**(8), 199–201 (1999). doi:10.1109/97.774864
6. B Schell, Y Tsividis, Analysis and simulation of continuous-time digital signal processors. Signal Process. **89**(10), 2013–2026 (2009). doi:10.1016/j.sigpro.2009.04.005
7. AV Savkin, IR Petersen, SOR Moheimanic, Model validation and state estimation for uncertain continuous-time systems with missing discrete-continuous data. Comput Electr Eng. **25**, 29–43 (1999). doi:10.1016/S0045-7906(98)00024-X
8. P Ferreira, Sorting continuous-time signals and the analog median filter. IEEE Signal Process Lett. **7**(10), 281–283 (2000). doi:10.1109/97.870681
9. X Chang, G Yang, Nonfragile H_∞ filtering of continuous-time fuzzy systems. IEEE Trans Signal Process. **59**(4), 1528–1538 (2011)
10. JPS Bizarro, On the behavior of the continuous-time spectrogram for arbitrarily narrow windows. IEEE Trans Signal Process. **55**(5), 1793–1802 (2007)
11. HL Van Trees, *Detection, Estimation, and Modulation Theory. Part I* (Wiley, New York, 1968)
12. HV Poor, *An Introduction to Signal Detection and Estimation*, 2nd edn. (Springer, New York, 1994)
13. T Kailath, AH Sayed, B Hassibi, *Linear Estimation* (Prentice Hall, New Jersey, 2000)
14. RM Fernández-Alcalá, J Navarro-Moreno, JC Ruiz-Molina, A solution to the linear estimation problem with correlated signal and observation noise. Signal Process. **84**, 1973–1977 (2004). doi:10.1016/j.sigpro.2004.06.017
15. S Cambanis, A general approach to linear mean-square estimation problems. IEEE Trans Inf Theory **IT-19**(1), 110–114 (1973)
16. B Picinbono, P Chevalier, Widely linear estimation with complex data. IEEE Trans Signal Process. **43**(8), 2030–2033 (1995). doi:10.1109/78.403373
17. B Picinbono, P Bondon, Second-order statistics of complex signals. IEEE Trans Signal Process. **45**(2), 411–420 (1997). doi:10.1109/78.554305
18. P Rubin-Delanchy, AT Walden, Kinematics of complex-valued time series. IEEE Trans Signal Process. **56**(9), 4189–4198 (2008)
19. J Navarro-Moreno, ARMA prediction of widely linear systems by using the innovations algorithm. IEEE Trans Signal Process. **56**(7), 3061–3068 (2008)
20. Y Xia, CC Took, 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
21. H Gerstacker, R Schober, RA Lampe, Receivers with widely linear processing for frequency-selective channels. IEEE Trans Commun. **51**(9), 1512–1523 (2003). doi:10.1109/TCOMM.2003.816992
22. J Eriksson, V Koivunen, Complex random vectors and ICA models: identifiability, uniqueness, and separability. IEEE Trans Inf Theory. **52**(3), 1017–1029 (2006)
23. J Vía, D Ramírez, I Santamaría, Properness and widely linear processing of quaternion random vectors. IEEE Trans Inf Theory. **56**(7), 3502–3515 (2010)
24. DP Mandic, VSL Goh, *Complex Valued Nonlinear Adaptive Filters. Noncircularity, Widely Linear and Neural Models* (Wiley, New York, 2009)
25. T Adali, S Haykin, *Adaptive Signal Processing: Next Generation Solutions* (Wiley-IEEE Press, 2010)
26. CC Took, DP Mandic, A quaternion widely linear adaptive filter. IEEE Trans Signal Process. **58**(8), 4427–4431 (2010)
27. PJ Schreier, LL Scharf, *Statistical Signal Processing of Complex-Valued Data* (Cambridge University Press, Cambridge, 2010)
28. PJ Schreier, LL Scharf, TM Clifford, Detection and estimation of improper complex random signals. IEEE Trans Inf Theory **51**(1), 306–312 (2005). doi:10.1109/TIT.2004.839538
29. J Navarro-Moreno, MD Estudillo, RM Fernández-Alcalá, JC Ruiz-Molina, Estimation of improper complex random signals in colored noise by using the Hilbert space theory. IEEE Trans Inf Theory. **55**(6), 2859–2867 (2009)
30. S Cambanis, Representation of stochastic processes of second-order and linear operations. J Math Appl. (41), 603–620 (1973)
31. J Navarro-Moreno, JC Ruiz-Molina, RM Fernández-Alcalá, Approximate series representations of linear operations on second-order stochastic processes: application to simulation. IEEE Trans Inf Theory **52**(4), 1789–1794 (2006)
32. WA Gardner, LE Franks, An alternative approach to linear least squares estimation of continuous random processes, in *5th Annual Princeton Conferene Information Sciences and Systems* (1971)
33. RG Ghanem, PD Spanos, *Stochastic Finite Elements: A Spectral Approach* (Springer, New York, 1991)
34. CE Rasmussen, CKI Williams, *Gaussian Processes for Machine Learning*, (The MIT Press, 2006) http://www.gaussianprocess.org/gpml/
35. A Oya, J Navarro-Moreno, JC Ruiz-Molina, A numerical solution for multichannel detection. IEEE Trans Commun. **57**(6), 1734–1742 (2009)
36. PE Gill, GF Miller, An algorithm for the integration of unequally spaced data. Comput J, (15), 80–83 (1972)
37. JG Proakis, *Digital Communications* (McGraw-Hill, Newyork, 1989)

38. SC Liu, Evolutionary power spectral density of strong-motion earthquakes. Bull Seismol Soc Am. **60**(3), 891–900 (1970)

39. SS Wang, HP Hong, Quantiles of critical separation distance for non-stationary seismic excitations. Eng Struct. **28**, 985–991 (2006). doi:10.1016/j.engstruct.2005.11.003

40. A Zerva, V Zervas, Spatial variation of seismic ground motions: an overview. Appl Mech Rev. **55**(3), 271–297 (2002). doi:10.1115/1.1458013