We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Adaptive-FRESH Filtering

Omar A. Yeste Ojeda and Jesús Grajal
Universidad Politécnica de Madrid
Spain

1. Introduction

Adaptive filters are self-designing systems for information extraction which rely for their operation on a recursive algorithm (Haykin, 2001). They find application in environments where the optimum filter cannot be applied because of lack of knowledge about the signal characteristics and/or the data to be processed. This a priori knowledge required for the design of the optimum filter is commonly based on stochastic signal models, which traditionally are stationary. However, the parameters of many signals found in communication, radar, telemetry and many other fields can be represented as periodic functions of time (Gardner, 1991). When this occurs, stationary signal models cannot exploit the signal periodicities, while cyclostationary models become more suitable since they represent more reliably the statistical signal properties.[1]

Firstly, let us review some of the key points of cyclostationary signals, while introducing some definitions that will be used throughout this chapter. We have said that most of signals used in many fields such as communication, radar, or telemetry exhibit statistical parameters which vary periodically with time. The periodicities of these parameters are related to the parameters of the signal modulation, such as the carrier frequency or the chip rate among others (Gardner, 1986; 1994). Let the second-order[2] auto-correlation function of a zero-mean stochastic signal be defined as:

$$R_{xx}(t, \lambda) \triangleq \mathrm{E}\left\{x(t)\, x^*(\lambda)\right\} \tag{1}$$

A signal is said to be (second-order) cyclostationary if, and only if, its (second-order) auto-correlation function is a periodic function of time, namely with period T. Therefore, it can be expanded in a Fourier series (Giannakis, 1998):

$$R_{xx}(t, \lambda) = \sum_{p=-\infty}^{\infty} R_{xx}^{\alpha_p}(t - \lambda)\, e^{j2\pi\alpha_p\lambda} \tag{2}$$

where $\alpha_p = p/T$ are called the *cycle frequencies* of $x(t)$, and the set of all the cycle frequencies is referred to as the *cyclic spectrum*. The Fourier coefficients of the expansion in (2) are named

---

[1] Cyclostationary signal models are a more general class of stochastic processes which comprise the stationary ones. Therefore, they always represent the statistical properties of the signal at least as well as the stationary ones.

[2] Since the optimality criterion used in this chapter is based on the Mean Squared Error (MSE), only the second-order statistical moments are of interest to us. The first-order statistical moment (i.e. the mean) is zero by assumption. Throughout this chapter, the second-order cyclostationarity is exploited. For brevity, hereinafter the cyclostationarity and correlation functions are assumed to be of second order, even without explicit mention.

cyclic auto-correlation functions and are computed as:

$$R_{xx}^{\alpha_p}(\tau) \triangleq \frac{1}{T} \int_0^T R_{xx}(t+\tau, t) \ e^{-j\frac{2\pi}{T}pt} \ dt \tag{3}$$

In practice, the signal periodicities are often incommensurable with each other, which yields that the auto-correlation function in (1) is not periodic, but an almost-periodic function of time. In this general case (in the sense that periodic functions are a particular case of almost-periodic ones), the signal is said to be almost-cyclostationary, and its auto-correlation function allows its expansion as a generalized Fourier series:

$$R_{xx}(t, \lambda) = \sum_{\alpha_p \in A_{xx}} R_{xx}^{\alpha_p}(t - \lambda) \ e^{j2\pi\alpha_p\lambda} \tag{4}$$

where the set A_{xx} stands for the cyclic spectrum of $x(t)$, and is generally composed of the sum and difference of integer multiples of the signal periodicities (Gardner, 1987; Gardner et al., 1987; Napolitano & Spooner, 2001). Additionally, the definition of the cyclic auto-correlation functions must be changed accordingly:

$$R_{xx}^{\alpha}(\tau) \triangleq \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} R_{xx}(t+\tau, t) \ e^{-j2\pi\alpha t} \ dt \tag{5}$$

One of the most important properties of almost-cyclostationary signals concerns the existence of correlation between their spectral components. The periodicity of the auto-correlation function turns into spectral correlation when it is transformed to the frequency domain. As a result, almost-cyclostationarity and spectral correlation are related in such a way that a signal exhibits almost-cyclostationary properties if, and only if, it exhibits spectral correlation too. Let $X_{\Delta f}(t, f)$ be the spectral component of $x(t)$, around time t and frequency f, and with spectral bandwidth Δf:

$$X_{\Delta f}(t, f) = \int_{t-\frac{1}{2\Delta f}}^{t+\frac{1}{2\Delta f}} x(u) \ e^{-j2\pi fu} \ du \tag{6}$$

The spectral correlation function of the signal $x(t)$ is defined as:

$$S_{xx}^{\alpha}(f) \triangleq \lim_{\Delta f \to 0} \Delta f \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \mathrm{E} \left\{ X_{\Delta f}(t, f) \ X_{\Delta f}^*(t, f-\alpha) \right\} \ dt \tag{7}$$

which represents the time-averaged correlation between two spectral components centered at frequencies f and $f - \alpha$, as their bandwidth tends to zero. It can be demonstrated that the spectral correlation function matches de Fourier transform of the cyclic correlation functions (Gardner, 1986), that is:

$$S_{xx}^{\alpha}(f) = \int_{-\infty}^{\infty} R_{xx}^{\alpha}(\tau) \ e^{-j2\pi f\tau} \ d\tau \tag{8}$$

where the inherent relationship between almost-cyclostationarity and spectral correlation is fully revealed. Intuitively, Eq. (8) means that the spectral components of almost-cyclostationary signals are correlated with other spectral components which are spectrally separated at the periodicities of their correlation function, i.e. their cyclic spectrum. For this reason the cycle frequency is also known as spectral separation (Gardner, 1991). Note that, at cycle frequency zero, the cyclic auto-correlation function in (5) represents the stationary

(or time-averaged) part of the nonstationary auto-correlation function in (3). Therefore, it is straightforward from (8) that the spectral correlation function matches the Power Spectral Density (PSD) at $\alpha = 0$, and also represents the auto-correlation of the signal spectral components, which is indicated by (7).

The key result from the above is that, since almost-cyclostationary signals exhibit spectral correlation, a single spectral component can be restored not only from itself, but also from other components which are spectrally separated from it as indicated by the cyclic spectrum of the signal. It is clear that a simple Linear Time-Invariant (LTI) filter cannot achieve this spectral restoration, but intuition says that a kind of filter which incorporates frequency shifts within its structure could. Contrary to optimal linear filtering of stationary signals, which results in time-invariant filters, the optimal linear filters of almost-cyclostationary signals are Linear Almost-Periodically Time-Variant (LAPTV) filters (also known as *poly-periodic* filters (Chevalier & Maurice, 1997)). This optimality can be understood in the sense of Signal-to-Noise Ratio (SNR) maximisation or in the sense of Minimum Mean Squared Error (MMSE, the one used in this chapter), where the optimal LAPTV filter may differ depending on the criterion used. Therefore, LAPTV filters find application in many signal processing areas such as signal estimation, interference rejection, channel equalization, STAP (Space-Time Adaptive Processing), or watermarking, among others (see (Gardner, 1994) and references therein, or more recently in (Adlard et al., 1999; Chen & Liang, 2010; Chevalier & Blin, 2007; Chevalier & Maurice, 1997; Chevalier & Pipon, 2006; Gameiro, 2000; Gelli & Verde, 2000; Gonçalves & Gameiro, 2002; Hu et al., 2007; Li & Ouyang, 2009; Martin et al., 2007; Mirbagheri et al., 2006; Ngan et al., 2004; Petrus & Reed, 1995; Whitehead & Takawira, 2004; 2005; Wong & Chambers, 1996; Yeste-Ojeda & Grajal, 2008; Zhang et al., 2006; 1999)).

This chapter is devoted to the study of LAPTV filters for adaptive filtering. In the next sections, the fundamentals of LAPTV filters are briefly described. With the aim of incorporating adaptive strategies, the theoretical development is focused on the FRESH (FREquency SHift) implementation of LAPTV filters. FRESH filters are composed of a set of frequency shifters followed by an LTI filter, which notably simplifies the analysis and design of adaptive algorithms. After reviewing the theoretical background of adaptive FRESH filters, an important property of adaptive FRESH filters is analyzed: Their capability to operate in the presence of errors in the LAPTV periodicities. This property is important because small errors in these periodicities, which are quite common in practice due to non-ideal effects, can make the use of LAPTV filters unfeasible.

Finally, an application example of adaptive FRESH filters is used at the end of this chapter to illustrate their benefit in real applications. In that example, an adaptive FRESH filter constitutes an interference rejection subsystem which forms part of a signal interception system. The goal is to use adaptive FRESH filtering for removing the unwanted signals so that a subsequent subsystem can detect any other signal present in the environment. Therefore, the interference rejection and signal detection problems can be dealt with independently, allowing the use of high sensitivity detectors with poor interference rejection properties.

2. Optimum linear estimators for almost cyclostationary processes

This section is devoted to establishing the optimality of LAPTV filters for filtering almost-cyclostationary signals. The theory of LAPTV filters can be seen as a generalization of the classical Wiener theory for optimal LTI filters, where the signals involved are no longer stationary, but almost-cyclostationary. Therefore, optimal LTI filters become a

particularization of LAPTV ones, as stationary signals can be seen as a particularization of almost-cyclostationary ones.

The Wiener theory defines the optimal (under MMSE criterion) LTI filter for the estimation of a desired signal, $d(t)$, given the input (or observed) signal $x(t)$, when both $d(t)$ and $x(t)$ are jointly stationary. In this case, their auto- and cross-correlation functions do not depend on the time, but only on the lag, and the estimation error results constant too. Otherwise, the estimation error becomes a function of time. The Wiener filter is still the optimal LTI filter if, and only if, $d(t)$ and $x(t)$ are jointly stationarizable processes (those which can be made jointly stationary by random time shifting) (Gardner, 1978). In this case, the Wiener filter is optimal in the sense of Minimum Time-Averaged MSE (MTAMSE). For instance, jointly almost-cyclostationary processes (those whose auto- and cross-correlation functions are almost-periodic functions of time) are always stationarizable. Nonetheless, if $x(t)$ and $d(t)$ are jointly almost-cyclostationary, it is possible to find an optimal filter which minimizes the MSE at all instants, which becomes a Linear Almost-Periodically Time-Variant (LAPTV) filter (Gardner, 1994). This result arises from the orthogonality principle of optimal linear estimators (Gardner, 1986), which is developed next.

Let $\widehat{d}(t)$ be the estimate of $d(t)$ from $x(t)$, obtained through the linear filter $h(t, \lambda)$:

$$\widehat{d}(t) = \int_{-\infty}^{\infty} h(t, u) \, x(u) \, du \tag{9}$$

The orthogonality principle establishes that, if $h(t, \lambda)$ is optimum, then input signal $x(t)$ and the estimation error $(\varepsilon(t) = d(t) - \widehat{d}(t))$ are orthogonal with each other: [3]

$$\mathrm{E}\left\{\varepsilon(t) \, x^*(\lambda)\right\} = 0 \, , \quad \forall \lambda, t \in \mathbb{R} \tag{10}$$

$$R_{dx}(t, \lambda) = \int_{-\infty}^{\infty} h_\Gamma(t, u) \, R_{xx}(u, \lambda) \, du \, , \quad \forall \lambda, t \in \mathbb{R} \tag{11}$$

where

$$R_{uv}(t, \lambda) \triangleq \mathrm{E}\left\{u(t) \, v^*(\lambda)\right\} \tag{12}$$

stands for the cross-correlation function between $u(t)$ and $v(t)$, and $h_\Gamma(t, \lambda)$ stands for the optimal LAPTV filter (where the meaning of the subindex Γ will be clarified in the next paragraphs). Since $d(t)$ and $x(t)$ are jointly almost-cyclostationary by assumption, their auto- and cross-correlation functions are almost-periodic functions of time, and therefore they can be expanded as a generalized Fourier series (Corduneanu, 1968; Gardner, 1986):

$$R_{dx}(t, \lambda) = \sum_{\alpha_k \in \mathrm{A}_{dx}} R_{dx}^{\alpha_k}(t - \lambda) \, e^{j2\pi\alpha_k\lambda} \tag{13}$$

$$R_{xx}(t, \lambda) = \sum_{\beta_p \in \mathrm{B}_{xx}} R_{xx}^{\beta_p}(t - \lambda) \, e^{j2\pi\beta_p\lambda} \tag{14}$$

where A_{dx} and B_{xx} are countable sets consisting of the cycle frequencies of $R_{dx}(t, \lambda)$ and $R_{xx}(t, \lambda)$, respectively. In addition, the cyclic cross- and auto-correlation functions $R_{dx}^\alpha(\tau)$

---

[3] Up to this point, the signals considered are real valued and therefore the complex-conjugation operator can be ignored. However, it is incorporated in the formulation for compatibility with complex signals, which will be considered in following sections.

and $R_{xx}^{\alpha}(\tau)$ are computed as generalized Fourier coefficients (Gardner, 1986):

$$R_{dx}^{\alpha}(\tau) \triangleq \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} R_{dx}(t+\tau, t) \, e^{-j2\pi\alpha t} \, dt \tag{15}$$

$$R_{xx}^{\alpha}(\tau) \triangleq \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} R_{xx}(t+\tau, t) \, e^{-j2\pi\alpha t} \, dt \tag{16}$$

Substitution of (13) and (14) in (11) yields the condition:

$$\sum_{\alpha_k \in A_{dx}} R_{dx}^{\alpha_k}(t-\lambda) \, e^{j2\pi\alpha_k\lambda} = \sum_{\beta_p \in B_{xx}} \int_{-\infty}^{\infty} h_{\Gamma}(t,u) \, R_{xx}^{\beta_p}(u-\lambda) \, e^{j2\pi\beta_p\lambda} \, du \tag{17}$$

This condition can be satisfied for all $t, \lambda \in \mathbb{R}$ if $h_{\Gamma}(t, \lambda)$ is also an almost-periodic function of time, and therefore, can be expanded as a generalized Fourier series (Gardner, 1993; Gardner & Franks, 1975):

$$h_{\Gamma}(t, \lambda) \triangleq \sum_{\gamma_q \in \Gamma} h_{\Gamma}^{\gamma_q}(t-\lambda) \, e^{j2\pi\gamma_q\lambda} \tag{18}$$

where Γ is the minimum set containing A_{dx} and B_{xx} which is closed in the addition and subtraction operations (Franks, 1994). The Fourier coefficients in (18), $h_{\Gamma}^{\gamma_q}(\tau)$, can be computed analogously to Eq. (15) and (16). Then, the condition in (17) is developed by using the definition in (18), taking the Fourier transform, and augmenting the sets A_{dx} and B_{xx} to the set Γ, which they belong to, yielding the following condition:

$$\sum_{\alpha_k \in \Gamma} S_{dx}^{\alpha_k}(f) \, e^{j2\pi\alpha_k\lambda} = \sum_{\beta_p, \gamma_q \in \Gamma} H_{\Gamma}^{\gamma_q}(f) \, S_{xx}^{\beta_p}(f-\gamma_q) \, e^{j2\pi(\beta_p+\gamma_q)\lambda} \tag{19}$$

which must be satisfied for all $f, \lambda \in \mathbb{R}$, and where the Fourier transform of the cyclic cross- and auto-correlation functions,

$$S_{uv}^{\alpha}(f) = \int_{-\infty}^{\infty} R_{uv}^{\alpha}(\tau) \, e^{-j2\pi f\tau} \, d\tau \tag{20}$$

stands for the spectral cross- and auto-correlation function, respectively (Gardner, 1986). Finally, we use the fact that two almost-periodic functions are equal if and only if their generalized Fourier coefficients match (Corduneanu, 1968), which allows to reformulate (19) as the design formula of optimal LAPTV filters:

$$S_{dx}^{\alpha_k}(f) = \sum_{\gamma_q \in \Gamma} H_{\Gamma}^{\gamma_q}(f) \, S_{xx}^{\alpha_k-\gamma_q}(f-\gamma_q), \quad \forall \alpha_k \in \Gamma, \, \forall f \in \mathbb{R} \tag{21}$$

Note that the sets of cycle frequencies A_{dx} and B_{xx} are, in general, subsets of Γ. Consequently, the condition in (21) makes sense under the consideration that $S_{dx}^{\alpha}(f) = 0$ if $\alpha \notin A_{dx}$, and $S_{xx}^{\alpha}(f) = 0$ if $\alpha \notin B_{xx}$, which is coherent with the definitions in (15) and (16). Furthermore, (21) is also coherent with the classical Wiener theory. When $d(t)$ and $x(t)$ are jointly stationary, then the sets of cycle frequencies A_{dx}, B_{xx}, and Γ consist only of cycle frequency zero, which yields that the optimal linear estimator is LTI and follows the well known expression of Wiener filter:

$$S_{dx}^{0}(f) = H_{\Gamma}^{0}(f) \, S_{xx}^{0}(f) \tag{22}$$

Let us use a simple graphical example to provide an overview of the implications of the design formula in (21). Consider the case where the signal to be estimated, $s(t)$, is corrupted
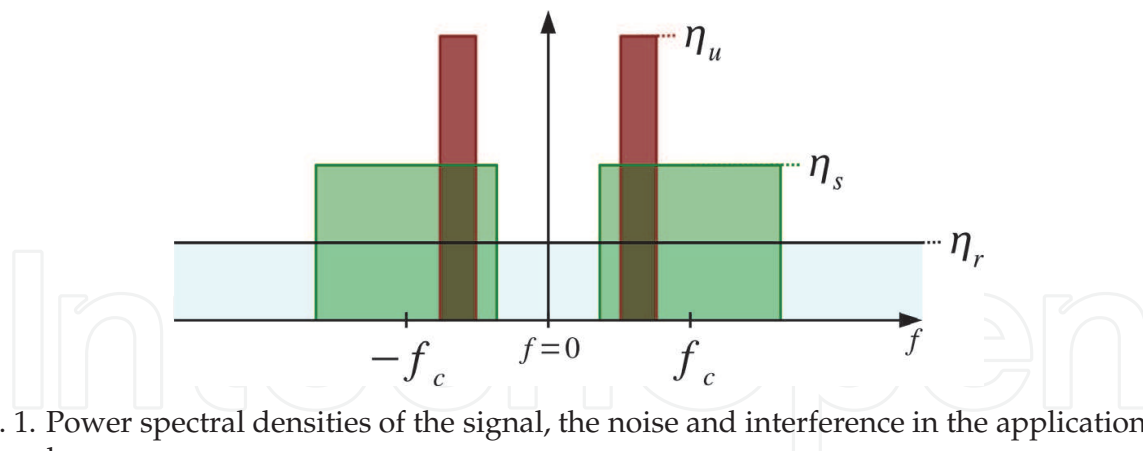
Fig. 1. Power spectral densities of the signal, the noise and interference in the application example.

by additive stationary white noise, $r(t)$, along with an interfering signal, $u(t)$, to form the observed signal, that is:

$$x(t) = s(t) + r(t) + u(t) \tag{23}$$

$$d(t) = s(t) \tag{24}$$

Assuming that $s(t)$, $r(t)$, and $u(t)$ are statistically independent processes, the design formula becomes:

$$S_{ss}^{\alpha_k}(f) = \sum_{\gamma_q \in \Gamma} H_{\Gamma}^{\gamma_q}(f) \left[S_{ss}^{\alpha_k - \gamma_q}(f - \gamma_q) + S_{rr}^{\alpha_k - \gamma_q}(f - \gamma_q) + S_{uu}^{\alpha_k - \gamma_q}(f - \gamma_q) \right] \tag{25}$$

In the following, let us consider the PSDs plotted in Fig. 1 for the signal, the noise and the interference, all of which are flat in their spectral bands, with PSD levels η_s, η_r and η_u, respectively. Let us further simplify the example by assuming that the signal is received with a high Signal-to-Noise Ratio (SNR), but low Signal-to-Interference Ratio (SIR), so that $\eta_u \gg \eta_s \gg \eta_r$. The Wiener filter can be obtained directly from Fig. 1, and becomes:

$$H_w(f) = \begin{cases} \frac{\eta_s}{\eta_s + \eta_r + \eta_u} \approx 0 & f \in B_u \\ \frac{\eta_s}{\eta_s + \eta_r} \approx 1 & f \notin B_u, \ f \in B_s \\ 0 & f \notin B_s \end{cases} \tag{26}$$

where B_u and B_s represent the frequency intervals comprised in the spectral bands of the interference and the signal, respectively. Thus, the Wiener filter is not capable of restoring the signal spectral components which are highly corrupted by the interference, since they are almost cancelled at its output.

On the contrary, an LAPTV filter could restore the spectral components cancelled by the Wiener filter depending on the spectral auto-correlation function of the signal and the availability at the input of correlated spectral components of the signal which are not perturbed by the interference. In our example, consider that the signal $s(t)$ is Amplitude Modulated (AM). Then, the signal exhibits spectral correlation at cycle frequencies $\alpha = \pm 2f_c$, in addition to cycle frequency zero, which means that the signal spectral components at

positive frequencies are correlated with those components at negative frequencies (Gardner, 1987).[4] The spectral correlation function of the AM signal is represented in Fig. 2.
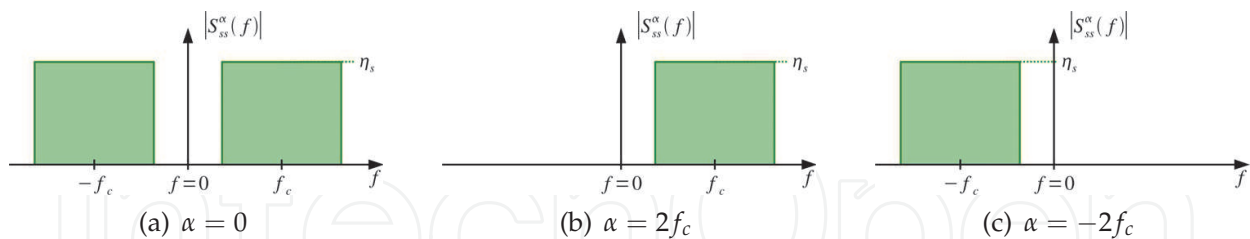

Fig. 2. Spectral correlation function of the signal with AM modulation.

The design formula in (25) states that the Fourier coefficients of the optimal LAPTV filter represent the coefficients of a linear combination in which frequency shifted version of $S_{xx}^{\alpha}(f)$ are combined in order to obtain $S_{dx}^{\alpha}(f)$. For simplicity, let us suppose that the set of cycle frequencies Γ only consists of the signal cyclic spectrum, that is $\Gamma = \{-2f_c, 0, 2f_c\}$, so that the design formula must be solved only for these values of α_k. Suppose also that the cycle frequencies $\pm 2f_c$ are exclusive of $s(t)$, so that $S_{xx}^{\pm 2f_c}(f) = S_{ss}^{\pm 2f_c}(f)$. This is coherent with the assumption that noise is stationary and with Fig. 1, where the bandwidth of the interference is narrower than $2f_c$, and therefore none of its spectral components is separated $2f_c$ in frequency. Fig. 3 graphically represents the conditions imposed by the design formula (25) on the Fourier coefficients of the optimal LAPTV filter, where each column stands for the different equations as α_k takes different values from Γ. The plots in the first three rows stand for the amplitude of the frequency shifted versions of the spectral auto-correlation function of the signal, the noise and the interference, while the plots in the last row represent the spectral cross-correlation between the input and the desired signals. The problem to be solved is to find the filter Fourier coefficients, $\{H_{\Gamma}^{-2f_c}(f), H_{\Gamma}^{0}(f), H_{\Gamma}^{2f_c}(f)\}$, which multiplied by the spectral correlation functions represented in the first three rows, and added together, yield the spectral cross-correlation function in the last row.

Firstly, let us pay attention to the spectral band of the interference at positive frequencies. From Fig. 3, the following equation system apply:

$$\begin{cases} S_{xx}^{0}(f)\, H_{\Gamma}^{0}(f) + S_{xx}^{-2f_c}(f - 2f_c)\, H_{\Gamma}^{2f_c}(f) = S_{ss}^{0}(f) \\[2mm] S_{xx}^{2f_c}(f)\, H_{\Gamma}^{0}(f) + S_{xx}^{0}(f - 2f_c)\, H_{\Gamma}^{2f_c}(f) = S_{ss}^{2f_c}(f) \qquad , f \in B_u^{+} \\[2mm] S_{xx}^{0}(f + 2f_c)\, H_{\Gamma}^{-2f_c}(f) = 0 \end{cases} \qquad (27)$$

[4] The definition of the spectral correlation function used herein (see Eq. (7)) differs in the meaning of frequency from the definition used by other authors, as in (Gardner, 1987). Therein, the frequency stands for the mean frequency of the two spectral components whose correlation is computed. Both definitions are related with each other by a simple change of variables, that is $S_{xx}^{\alpha}(f) = S'^{\alpha}_{xx}(f - \alpha/2)$, where $S'^{\alpha}_{xx}(f)$ corresponds to spectral correlation function according to the definition used in (Gardner, 1987).
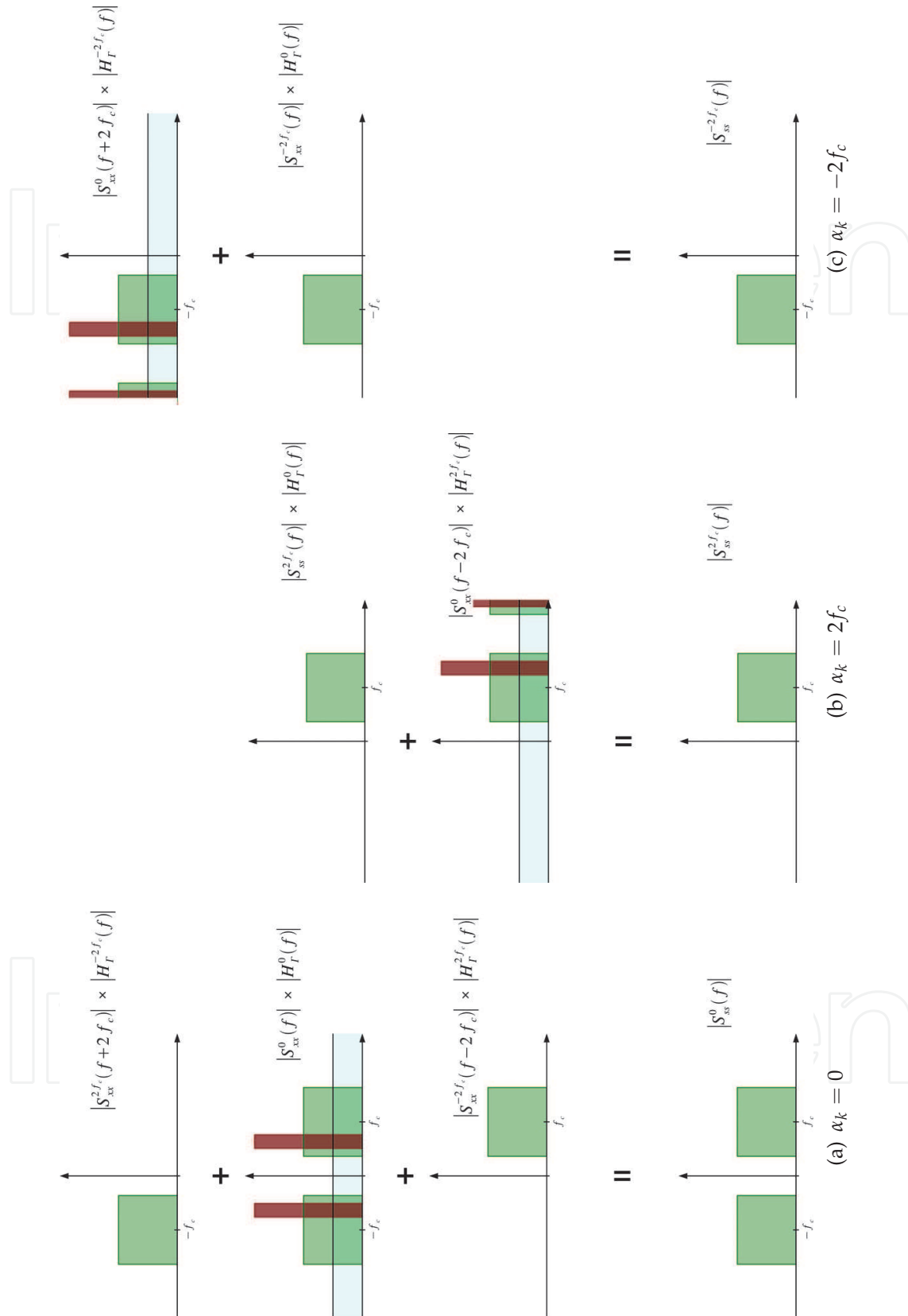
Fig. 3. Graphical representation of the design formula of LAPTV filters, which has been applied to our example. Each plot corresponds to a different value of α_k in (25).

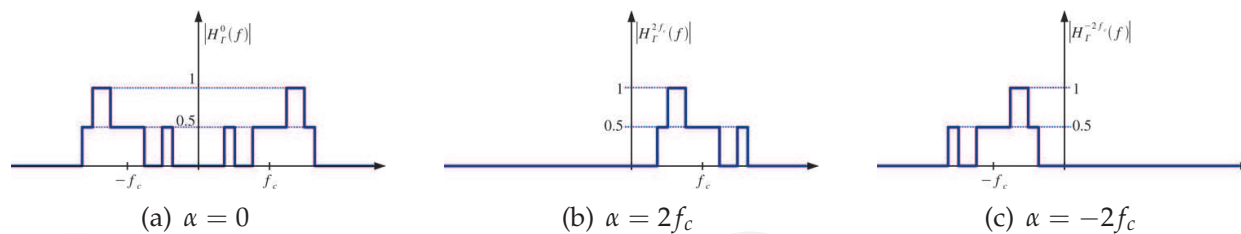(a) $\alpha = 0$ (b) $\alpha = 2f_c$ (c) $\alpha = -2f_c$

Fig. 4. Fourier coefficients of the optimal LAPTV filter.

$$\begin{cases} (\eta_s + \eta_r + \eta_u)\, H_\Gamma^0(f) + \eta_s\, H_\Gamma^{2f_c}(f) = \eta_s \\ \eta_s\, H_\Gamma^0(f) + (\eta_s + \eta_r)\, H_\Gamma^{2f_c}(f) = \eta_s \qquad , f \in B_u^+ \\ \eta_r\, H_\Gamma^0(f) = 0 \end{cases} \tag{28}$$

where B_u^+ stands for the range of positive frequencies occupied by the interference. The solution to the linear system in (28) is:

$$H_\Gamma^0(f) = \frac{\eta_s\, \eta_r}{\eta_s\,(\eta_r + \eta_u) + \eta_r\,(\eta_s + \eta_r + \eta_u)} \qquad , f \in B_u^+ \tag{29}$$

$$H_\Gamma^{2f_c}(f) = \frac{\eta_s\,(\eta_r + \eta_u)}{\eta_s\,(\eta_r + \eta_u) + \eta_r\,(\eta_s + \eta_r + \eta_u)} = H_\Gamma^0(f)\left(\frac{\eta_u}{\eta_r} + 1\right) \qquad , f \in B_u^+ \tag{30}$$

$$H_\Gamma^{-2f_c}(f) = 0 \qquad , f \in B_u^+ \tag{31}$$

The result in (31) is coherent with the fact that there are not any signal spectral components located in B_u^+ after frequency shifting the input downwards by $2f_c$. After using the approximations of high SNR and low SIR, the above results for the other filter coefficients can be approximated by:

$$H_\Gamma^0(f) \approx 0 \qquad , f \in B_u^+ \tag{32}$$

$$H_\Gamma^{2f_c}(f) \approx \frac{\text{SNR}}{\text{SNR} + 1} \approx 1 \qquad , f \in B_u^+ \tag{33}$$

The underlaying meaning of (32) is that the optimal LAPTV filter cancels the spectral components of the signal which are corrupted by the interference. But contrary to the Wiener filter, these spectral components are restored from other components which are separated $2f_c$ in frequency, which is indicated by (33).

By using a similar approach, the Fourier coefficients of the LAPTV filter are computed for the rest of frequencies (those which do not belong to B_u^+). The result is represented in Fig. 4. We can see in Fig. 4(a) that $|H_\Gamma^0(f)|$ takes three possible values, i.e. 0, 0.5, and 1. $|H_\Gamma^0(f)| = 0$ when $f \in B_u$ (as explained above) or $f \notin B_s$ (out of the signal frequency range). The frequencies at which the Fourier coefficients are plotted with value $|H_\Gamma^\alpha| = 0.5$ correspond to those spectral components of the signal which are estimated from themselves, and jointly from spectral components separated $\alpha = 2f_c$ (or alternatively $\alpha = -2f_c$), since both are only corrupted by noise. For their part, the frequencies at which $|H_\Gamma^0(f)| = 1$ match the spectral components cancelled by $H_\Gamma^{\pm 2f_c}(f)$ (see Fig. 4(b) and 4(c)). These frequencies do not correspond to the spectral band of the interference, but the resulting frequencies after shifting this band by $\pm 2f_c$. Consequently, such spectral components are estimated only from themselves.

The preceding example has been simplified in order to obtain comprehensive results of how an LAPTV filter performs, and to intuitively introduce the idea of frequency shifting and filtering. This idea will become clearer in Section 4, when describing the FRESH implementation of LAPTV filters. In our example, no reference to the cyclostationary properties of the interference has been made. The optimal LAPTV filter also exploits the interference cyclostationarity in order to suppress it more effectively. However, if we had considered the cyclostationary properties of the interference, the closed form expressions for the filter Fourier coefficients would have result more complex, which would have prevented us from obtaining intuitive results. Theoretically, the set of cycle frequencies Γ consists of an infinite number of them, which makes very hard to find a closed form solution to the design formula in (21). This difficulty can be circumvented by forcing the number of cycle frequencies of the linear estimator $h(t, \lambda)$ to be finite, at the cost of performance (the MSE increases and the filter is no longer optimal). This strategy will be described along with the FRESH implementation of LAPTV filters, in Section 4. But firstly, the expression in (22) is generalized for complex signals in the next section.

3. Extension of the study to complex signals

Complex cyclostationary processes require up to four real LAPTV filters in order to achieve optimality, that is:

1. To estimate the real part of $d(t)$ from the real part of $x(t)$,
2. to estimate the real part of $d(t)$ from the imaginary part of $x(t)$,
3. to estimate the imaginary part of $d(t)$ from the real part of $x(t)$ and
4. to estimate the imaginary part of $d(t)$ from the imaginary part of $x(t)$.

This solution can be reduced to two complex LAPTV filters whose inputs are $x(t)$ and the complex conjugate of $x(t)$, that is $x^*(t)$. As a consequence, the optimal filter is not formally a linear filter, but a Widely-Linear Almost-Periodically Time-Variant (WLAPTV) filter (Chevalier & Maurice, 1997) (also known as *Linear-Conjugate Linear*, LCL (Brown, 1987; Gardner, 1993)). Actually, the optimal WLAPTV filter reduces to an LAPTV filter when the observations and the desired signal are jointly circular (Picinbono & Chevalier, 1995). [5] The final output of the WLAPTV filter is obtained by adding together the outputs of the two complex LAPTV filters:

$$\widehat{d}(t) = \int_{-\infty}^{\infty} h(t, u) \, x(u) \, du + \int_{-\infty}^{\infty} g(t, u) \, x^*(u) \, du \tag{34}$$

Since the orthogonality principle establishes that the estimation error must be uncorrelated with the input, it applies to both $x(t)$ and $x^*(t)$, yielding the linear system:

$$\begin{cases} \mathrm{E} \left\{ \varepsilon(t) \, x^*(\lambda) \right\} = 0 \\ \mathrm{E} \left\{ \varepsilon(t) \, x(\lambda) \right\} = 0 \end{cases} \quad , \quad \forall \lambda, t \in \mathbb{R} \tag{35}$$

$$\begin{cases} R_{dx}(t, \lambda) = \int_{-\infty}^{\infty} h_\Gamma(t, u) \, R_{xx}(u, \lambda) \, du + \int_{-\infty}^{\infty} g_\Gamma(t, u) \, [R_{xx^*}(u, \lambda)]^* \, du \\ R_{dx^*}(t, \lambda) = \int_{-\infty}^{\infty} h_\Gamma(t, u) \, R_{xx^*}(u, \lambda) \, du + \int_{-\infty}^{\infty} g_\Gamma(t, u) \, R_{xx}(\lambda, u) \, du \end{cases} \quad , \quad \forall \lambda, t \in \mathbb{R} \tag{36}$$

---

[5] The observed and desired signals, respectively $x(t)$ and $d(t)$, are jointly circular when $x^*(t)$ is uncorrelated with both $x(t)$ and $d(t)$ (Picinbono & Chevalier, 1995).

Analogously to Section 2, it can be demonstrated that the condition in (36) is satisfied if both $h_\Gamma(t, \lambda)$ and $g_\Gamma(t, \lambda)$ are almost-periodic functions, and the linear system in (36) can be reformulated as the design formula (Gardner, 1993):

$$
\begin{cases}
S_{dx}^{\alpha_k}(f) = \displaystyle\sum_{\gamma_q \in \Gamma} H_\Gamma^{\gamma_q}(f)\, S_{xx}^{\alpha_k - \gamma_q}\left(f - \gamma_q\right) + G_\Gamma^{\gamma_q}(f) \left[S_{xx^*}^{\gamma_q - \alpha_k}\left(\gamma_q - f\right)\right]^* \\
S_{dx^*}^{\alpha_k}(f) = \displaystyle\sum_{\gamma_q \in \Gamma} H_\Gamma^{\gamma_q}(f)\, S_{xx^*}^{\alpha_k - \gamma_q}\left(f - \gamma_q\right) + G_\Gamma^{\gamma_q}(f) \left[S_{xx}^{\gamma_q - \alpha_k}\left(\gamma_q - f\right)\right]^*
\end{cases} , \quad \forall \alpha_k \in \Gamma, \forall f \in \mathbb{R}
$$

(37)

where the set of cycle frequencies Γ is defined in this case as the minimum set comprising A_{dx}, A_{dx^*}, B_{xx} and B_{xx^*}, which is closed in the addition and subtraction operations (with A_{dx^*} and B_{xx^*} being the sets of cycle frequencies of the cross-correlation functions of the complex conjugate of the input, $x^*(t)$, with $d(t)$ and $x(t)$, respectively.)

As it occurred for real signals in Section 2, finding a closed-form solution to the design formula in (37) may result too complicated when a large number of cycle frequencies compose the set Γ. In the next section a workaround is proposed based on the FRESH implementation of LAPTV filters and the use of a reduced set of cycle frequencies, $\Gamma_s \subset \Gamma$.

4. Implementation of WLAPTV filters as FRESH (FREquency-SHift) filters

For simplicity reasons, LAPTV filters are often implemented as FREquency SHift (FRESH) filters (Gameiro, 2000; Gardner, 1993; 1994; Gardner & Franks, 1975; Gonçalves & Gameiro, 2002; Loeffler & Burrus, 1978; Ngan et al., 2004; Reed & Hsia, 1990; Zadeh, 1950). FRESH filters consist of a bank of LTI filters, whose inputs are frequency-shifted versions of the input signal. In general, the optimum FRESH filter would require an infinite number of LTI filters. Because this is not feasible, the FRESH filters used in practice are sub-optimal, in the sense that they are limited to a finite set of frequency shifters.

Any WLAPTV filter can be implemented as FRESH filter. [6] This result emerges from using in (34) the generalized Fourier series expansion of LAPTV filters. Let $h(t, \gamma)$ and $g(t, \gamma)$ be two arbitrary LAPTV filters. Then, each of them can be expanded in a generalized Fourier series yielding:

$$
\widehat{d}(t) = \int_{-\infty}^{\infty} h(t, u)\, x(u)\, du + \int_{-\infty}^{\infty} g(t, u)\, x^*(u)\, du \tag{38}
$$

$$
= \int_{-\infty}^{\infty} \sum_k h^{\alpha_k}(t - u)\, e^{j2\pi\alpha_k u}\, x(u)\, du + \int_{-\infty}^{\infty} \sum_p g^{\beta_p}(t - u)\, e^{j2\pi\beta_p u}\, x^*(u)\, du \tag{39}
$$

$$
= \sum_k \int_{-\infty}^{\infty} h^{\alpha_k}(t - u) \left[e^{j2\pi\alpha_k u}\, x(u)\right] du + \sum_p \int_{-\infty}^{\infty} g^{\beta_p}(t - u) \left[e^{j2\pi\beta_p u}\, x^*(u)\right] du \tag{40}
$$

It can be clearly seen that the output of the WLAPTV filter, $\widehat{d}(t)$, is the result of adding together the outputs of the LTI filters $h^{\alpha_k}(t)$ and $g^{\beta_p}(t)$, whose inputs are frequency-shifted versions of the input signal $x(t)$ and its complex conjugate $x^*(t)$, respectively. This is precisely the definition of a FRESH filter.

---

[6] Formally, two FRESH filters are required respectively for the input and its complex conjugate. Nevertheless, we will refer to the set of both as a FRESH filter herein.
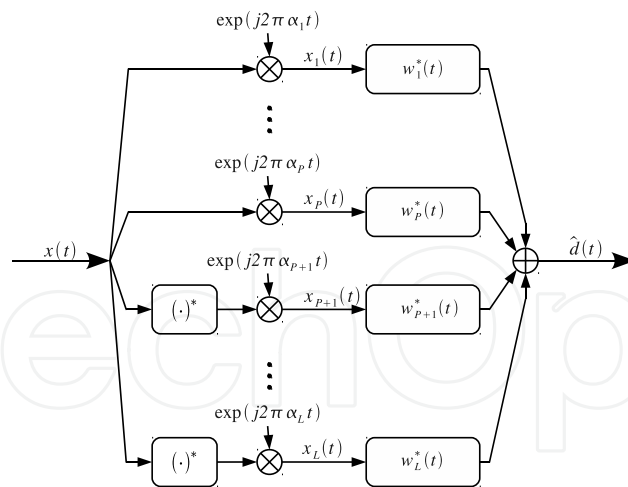
Fig. 5. Block diagram of the FRESH implementation of a WLAPTV filter.

The most difficult problem in the design of sub-optimal FRESH filters concerns the choice of the optimal subset of frequency shifts, $\Gamma_s \subset \Gamma$, under some design constraints. For instance, a common design constraint is the maximum number of branches of the FRESH filter. The optimum Γ_s becomes highly dependent on the spectral correlation function of the input signal and can change with time in nonstationary environments. However, with the aim of simplifying the FRESH implementation, Γ_s is usually fixed beforehand and the common approach to determine the frequency shifts consists in choosing those cycle frequencies at which the spectral cross-correlation functions between $d(t)$ and $x(t)$, or its complex conjugate, exhibits maximum levels (Gardner, 1993; Yeste-Ojeda & Grajal, 2008).

Once the set of frequency shifts has been fixed, the design of FRESH filters is much simpler than that of WLAPTV filters, because FRESH filters only use LTI filters. Because of the nonstationary nature of the input (it is almost-cyclostationary), the optimality criterion used in the design of the set of LTI filters is the MTAMSE criterion, in contrast to the MMSE criterion (Gardner, 1993). The resulting FRESH filter is a sub-optimum solution since it is optimum only within the set of FRESH filters using the same set of frequency shifts.

Let us formulate the output of the FRESH filter by using vector notation:

$$\widehat{d}(t) = \int_{-\infty}^{\infty} \boldsymbol{w}^{\dagger}(t-u)\, \boldsymbol{x}(u)\, du \tag{41}$$

where $\boldsymbol{w}^{\dagger}(t)$ stands for the conjugate transpose of vector $\boldsymbol{w}(t)$, and each component of the input vector $\boldsymbol{x}(t)$ represents the input of an LTI filter according to the scheme represented in Fig. 5:

$$\boldsymbol{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_P(t) \\ x_{P+1}(t) \\ \vdots \\ x_L(t) \end{bmatrix} = \begin{bmatrix} x(t)\, e^{j2\pi\alpha_1 t} \\ x(t)\, e^{j2\pi\alpha_2 t} \\ \vdots \\ x(t)\, e^{j2\pi\alpha_P t} \\ x^*(t)\, e^{j2\pi\beta_1 t} \\ \vdots \\ x^*(t)\, e^{j2\pi\beta_{L-P} t} \end{bmatrix} \tag{42}$$

where P is the number of branches used for filtering the frequency shifted versions of $x(t)$, $(L-P)$ is the number of branches used for filtering its complex conjugate $x^*(t)$, and L

is the total number of branches. Note that the first P cycle frequencies can be repeated in the next $P - L$ cycle frequencies, since using a frequency shift for the input $x(t)$ does not exclude it from being used for its complex conjugate. With the aim of computing the MTAMSE optimal set of LTI filters, a stationarized signal model is applied to both the desired signal $d(t)$ and the input vector $\boldsymbol{x}(t)$, which are jointly almost-cyclostationary processes, and therefore stationarizable (Gardner, 1978). The stationary auto- and cross-correlation functions are obtained by taking the stationary part (time-averaging) the corresponding nonstationary correlation functions. As a result, the orthogonality principle is formulated as follows:

$$\mathrm{E}\left\{\varepsilon(t+\tau)\,\boldsymbol{x}^{\dagger}(t)\right\} = \boldsymbol{0}^{T}\,, \quad \forall \tau \in \mathbb{R} \tag{43}$$

$$\boldsymbol{R}_{d\boldsymbol{x}}(\tau) = \int_{-\infty}^{\infty} \boldsymbol{w}_{\Gamma_s}^{\dagger}(\tau - u)\,\mathbf{R}_{\boldsymbol{xx}}(u)\,du\,, \quad \forall \tau \in \mathbb{R} \tag{44}$$

where $\boldsymbol{0}$ is the null vector, and the cross-correlation (row) vector $\boldsymbol{R}_{d\boldsymbol{x}}(\tau)$ and the auto-correlation matrix $\mathbf{R}_{\boldsymbol{xx}}(\tau)$ are computed by time-averaging the corresponding nonstationary correlation functions. For $\boldsymbol{R}_{d\boldsymbol{x}}(\tau)$ this yields:

$$\boldsymbol{R}_{d\boldsymbol{x}}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \mathrm{E}\left\{d(t+\tau)\,\boldsymbol{x}^{\dagger}(t)\right\}\,dt = \begin{bmatrix} R_{dx}^{\alpha_1}(\tau) \\ R_{dx}^{\alpha_2}(\tau) \\ \vdots \\ R_{dx}^{\alpha_P}(\tau) \\ R_{dx^*}^{\alpha_{P+1}}(\tau) \\ \vdots \\ R_{dx^*}^{\alpha_L}(\tau) \end{bmatrix}^{T} \tag{45}$$

where the cyclic cross-correlation functions were defined in (15). The matrix $\mathbf{R}_{\boldsymbol{xx}}(\tau)$ becomes:

$$\mathbf{R}_{\boldsymbol{xx}}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \mathrm{E}\left\{\boldsymbol{x}(t+\tau)\,\boldsymbol{x}^{\dagger}(t)\right\}\,dt \tag{46}$$

where the element in the q-th row and k-th column is:

$$[\mathbf{R}_{\boldsymbol{xx}}(\tau)]_{qk} = \begin{cases} R_{xx}^{\alpha_k - \alpha_q}(\tau)\,e^{j2\pi\alpha_q\tau} & q \le P,\ k \le P \\ R_{xx^*}^{\alpha_k - \alpha_q}(\tau)\,e^{j2\pi\alpha_q\tau} & q \le P,\ k > P \\ \left[R_{xx^*}^{\alpha_q - \alpha_k}(\tau)\right]^{*}\,e^{j2\pi\alpha_q\tau} & q > P,\ k \le P \\ \left[R_{xx}^{\alpha_q - \alpha_k}(\tau)\right]^{*}\,e^{j2\pi\alpha_q\tau} & q > P,\ k > P \end{cases} \tag{47}$$

where the cyclic auto-correlation functions were defined in (16).
Finally, the orthogonality principle leads directly to the multidimensional Wiener filter, and the frequency response of the set of filters is obtained by taking the Fourier transform in (44):

$$\boldsymbol{S}_{d\boldsymbol{x}}(f) = \boldsymbol{W}_{\Gamma_s}^{\dagger}(f)\mathbf{S}_{\boldsymbol{xx}}(f) \tag{48}$$

$$\boldsymbol{W}_{\Gamma_s}(f) = [\mathbf{S}_{\boldsymbol{xx}}(f)]^{-1}\,\boldsymbol{S}_{d\boldsymbol{x}}^{\dagger}(f) \tag{49}$$

where the Hermitian property of matrix $\mathbf{S}_{\boldsymbol{xx}}(f)$ has been used,

$$\boldsymbol{S}_{d\boldsymbol{x}}(f) = \left[S_{dx}^{\alpha_1}(f),\ S_{dx}^{\alpha_2}(f),\ \dots,\ S_{dx}^{\alpha_P}(f),\ S_{dx^*}^{\alpha_{P+1}}(f),\ \dots,\ S_{dx^*}^{\alpha_L}(f)\right] \tag{50}$$

and the element of the q-th row and k-th column of matrix $\mathbf{S}_{xx}(f)$ is:

$$[\mathbf{S}_{xx}(f)]_{qk} = \begin{cases} S_{xx}^{\alpha_k - \alpha_q}(f - \alpha_q) & q \leq P, \, k \leq P \\ S_{xx^*}^{\alpha_k - \alpha_q}(f - \alpha_q) & q \leq P, \, k > P \\ \left[S_{xx^*}^{\alpha_q - \alpha_k}(\alpha_q - f) \right]^* & q > P, \, k \leq P \\ \left[S_{xx}^{\alpha_q - \alpha_k}(\alpha_q - f) \right]^* & q > P, \, k > P \end{cases} \tag{51}$$

It is noteworthy that the expression in (37), which defines the optimal WLAPTV filter, is a generalization of (48) for the case where $\Gamma_s = \Gamma$. In addition, we want to emphasize the main differences between optimal and sub-optimal FRESH filters:

1. Optimal FRESH filters are direct implementations of optimal WLAPTV filters defined by (37), and generally consist of an infinite number of LTI filters. Contrarily, sub-optimal FRESH filters are limited to a finite set of LTI filters.

2. For jointly almost-cyclostationary input and desired signals, optimal FRESH filters are the optimal with respect to any other linear estimator. On the contrary, sub-optimal FRESH filters are defined for a given subset Γ_s and, therefore, they are optimal only in comparison to the rest of FRESH filters using the same frequency shifts.

3. Optimal FRESH filters minimize the MSE at all times (MMSE criterion). However, sub-optimal FRESH filters only minimize the MSE on average (MTAMSE criterion). This means that another FRESH filter, even by making use of the same set of frequency shifts, could exhibit a lower MSE at specific times.

Finally, the inverse of matrix $\mathbf{S}_{xx}(f)$ could not exist for all frequencies, which would invalidate the expression in (49). However, since its main diagonal represents the power spectral density of the frequency shifted versions of the input, this formal problem can be ignored by assuming that a white noise component is always present at the input.

At this point we have finished the theoretical background concerning FRESH filters. The following sections are focused on the applications of adaptive FRESH filters. Firstly, the introduction of an adaptive algorithm in a FRESH filter is reviewed in the next section.

5. Adaptive FRESH filters

The main drawback of using FRESH filters is that the phase of the LTI filters (or alternatively the phase of the frequency shifters) must be "synchronized" with the desired signal. Otherwise, the contribution of the different branches can be destructive rather than constructive. For example, the phase of the signal carrier must be known if the frequency shifts are related to the carrier frequency, or the symbol synchronism must be known if the symbol rate is involved in the frequency shifts. As a consequence, the knowledge required for the design of sub-optimal [7] FRESH filter is rarely available beforehand, which leads to the use adaptive algorithms. Since FRESH filters consist of a set LTI filters, conventional adaptive algorithms can be directly applied, such as Least-Mean-Square (LMS) or Recursive-Least-Squares (RLS). The general scheme of an adaptive FRESH filter is shown in Fig. 6.

---

[7] Hereinafter, FRESH filters are always limited to a finite set of frequency shifts. Therefore, sub-optimal FRESH filters will be referred to as "optimal" for brevity.
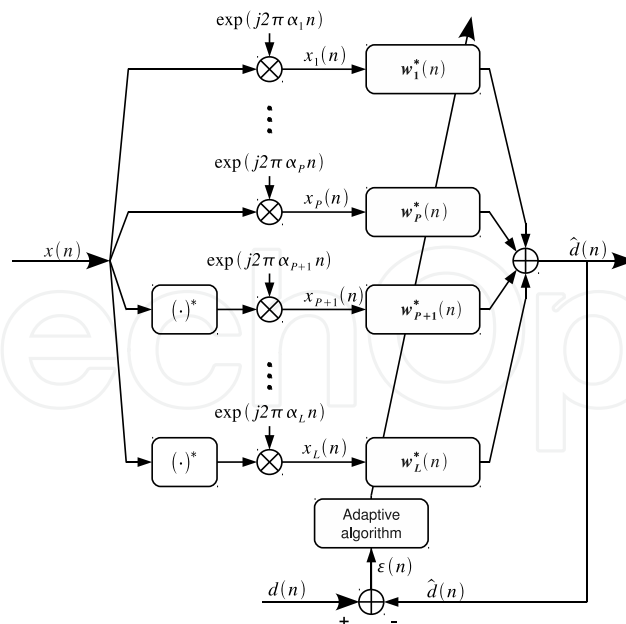
Fig. 6. Block diagram of an adaptive FRESH filter.

In order to simplify the analysis of the adaptive algorithms, the structure of FRESH filters is particularized to the case of discrete-time signals with the set of LTI filters exhibiting Finite Impulse Response (FIR filters). After filtering the received signal, $x(n)$, the output of the adaptive FRESH filter is compared to a desired (or reference) signal, $d(n)$, and the error, $\varepsilon(n)$, is used by an adaptive algorithm in order to update the filter coefficients. Commonly, the desired signal is either a known sequence (as a training sequence for equalizers), or obtained from the received signal (as for blind equalizers).

Let the output of the FRESH filter be defined as the inner product:

$$\widehat{d}(n) \triangleq \boldsymbol{w}^{\dagger}(n)\boldsymbol{x}(n) = \sum_{i=1}^{L} \boldsymbol{w}_i^{\dagger}(n)\boldsymbol{x}_i(n) \tag{52}$$

where the vectors $\boldsymbol{w}(n)$ and $\boldsymbol{x}(n)$ are the concatenation of the filter and the input vectors, respectively:

$$\boldsymbol{w}(n) = \begin{bmatrix} \boldsymbol{w}_1(n) \\ \boldsymbol{w}_2(n) \\ \vdots \\ \boldsymbol{w}_L(n) \end{bmatrix}, \qquad \boldsymbol{x}(n) = \begin{bmatrix} \boldsymbol{x}_1(n) \\ \boldsymbol{x}_2(n) \\ \vdots \\ \boldsymbol{x}_L(n) \end{bmatrix} \tag{53}$$

$\boldsymbol{w}_i(n)$ and $\boldsymbol{x}_i(n)$ are the filter and input vectors, respectively, of the i-th branch:

$$\boldsymbol{w}_i(n) = [w_i(n), w_i(n-1), \ldots, w_i(n-M_i+1)]^T \tag{54}$$

$$\boldsymbol{x}_i(n) = [x_i(n), x_i(n-1), \ldots, x_i(n-M_i+1)]^T \tag{55}$$

where M_i is the length of the i-th filter. Consequently, the total length of vectors $\boldsymbol{w}(n)$ and $\boldsymbol{x}(n)$ is $M = \sum_{i=1}^{L} M_i$.

Using the results in the previous sections, the optimal set of LTI filters is given by multidimensional Wiener filter theory:

$$\boldsymbol{w}_o = (\mathbf{R}_x)^{-1} \boldsymbol{p} \tag{56}$$

where $\mathbf{R}_x$ stands for the time-averaged auto-correlation matrix of the input vector, [8]

$$\mathbf{R}_x = \left\langle \mathrm{E}\left\{\boldsymbol{x}(n)\,\boldsymbol{x}^\dagger(n)\right\}\right\rangle \tag{57}$$

where $\langle\cdot\rangle$ denotes the time average operator:

$$\left\langle \mathrm{E}\left\{\boldsymbol{x}(n)\,\boldsymbol{x}^\dagger(n)\right\}\right\rangle = \lim_{N\to\infty}\frac{1}{2N+1}\sum_{n=-N}^{N}\mathrm{E}\left\{\boldsymbol{x}(n)\,\boldsymbol{x}^\dagger(n)\right\} \tag{58}$$

In Eq. (56), the vector $\boldsymbol{p}$ represents the time-averaged cross-correlation between the input vector $\boldsymbol{x}$ and $d(n)$:

$$\boldsymbol{p} = \langle \mathrm{E}\left\{\boldsymbol{x}(n)d^*(n)\right\}\rangle \tag{59}$$

The correlation functions in (57) and (59) are time-averaged in order to force the LTI characteristic of $\boldsymbol{w}_o$. The expressions in (56), (57), and (59) allow to follow the classical developments of adaptive algorithms (Haykin, 2001). Despite using classical algorithms, adaptive FRESH filters naturally exploit the cyclostationary properties of the signals as the inputs of the LTI filters are frequency-shifted versions of the input of the adaptive FRESH filter.

5.1 Blind Adaptive FRESH filters (BAFRESH)

One of the main reasons for using adaptive FRESH filters is their inherent ability to operate blindly (Zhang et al., 1999). Commonly, adaptive algorithms require a reference signal which is not available in many applications of adaptive filtering. When the adaptive filter works without such a reference signal, we say that it is "blind". Blind Adaptive FRESH (BAFRESH) filters can be obtained directly by simply using as frequency shifts those cycle frequencies which belong to the cyclic spectrum of the input and satisfy the condition:

$$S_{xx}^\alpha(f) = S_{dx}^\alpha(f) \tag{60}$$

for the first P branches, and

$$S_{xx^*}^\alpha(f) = S_{dx^*}^\alpha(f) \tag{61}$$

for the rest of branches (those working with $x^*(t)$). Then, the input signal itself can be used as the reference signal, as it is represented in Fig. 7, while the adaptive algorithm converges to the same solution as if a "clean" reference signal were used. [9] Intuitively, the conditions (60) and (61) mean that a BAFRESH filter must use only those cycle frequencies (among all the cycle frequencies of the input) at which only the part of the input which is correlated with the desired signal exhibits spectral correlation. It is noteworthy that (60) excludes cycle frequency zero from being used in the branches not using the complex conjugate of the input signal. The reason is that all stochastic processes exhibit spectral correlation at cycle frequency zero

---

[8] We have assumed that the matrix $\mathbf{R}_x$ is nonsingular, which is always true as long as there exists a white noise component at the input.

[9] However, using $x(t)$ instead of a "clean" reference produces an increase in the power of $\varepsilon(n)$, which could influence the adaptive FRESH filter performance.
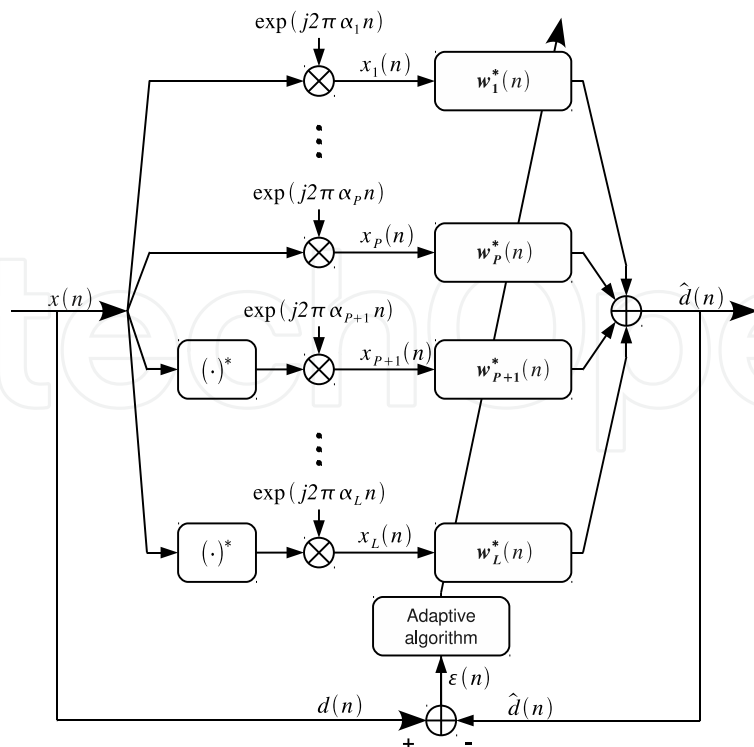
Fig. 7. Block diagram of a BAFRESH filter.

and therefore (60) could only be satisfied if the desired and the input signals are the same, which would eliminate the necessity of a filter. Contrarily, many stochastic processes do not exhibit spectral cross-correlation with their complex conjugate at cycle frequency zero, for instance the circular stochastic processes. This allows the condition (61) to be satisfied without implying that $d(t) = x(t)$.

When adaptive BAFRESH filters use a suitable set of frequency shifts, the correlation between the inputs of the LTI filters, $x_i(n)$, and the reference signal $x(n)$, is only due to the signal to be estimated. Thus, the adaptive algorithm will converge to a solution by which the desired signal is estimated at the output of the BAFRESH filter, while any other component present at the input is minimized. This is a powerful characteristic of adaptive FRESH filters which turns them into a good candidate for those applications where no reference signal is available at the receiver.

6. Adaptive FRESH filters for cycle-frequency errors compensation

This section deals with the problem of uncertainties in the periodicities of LAPTV filters. [10] The purpose in this section is to analyze the nonstationary behavior of adaptive FRESH filters in order to mitigate this problem. Since they exploit the spectral correlation of signals at specific values of frequency separation or cycle frequency, LAPTV filters are severely affected by errors in such cycle frequencies. These errors may be due to non-perfect knowledge of the cyclic spectrum, or natural effects such as the Doppler effect. In such a case, it will be demonstrated that the filter can become totally useless. The reason is that a high enough error (in comparison to the observation time) produces a loss of coherence among the outputs of the FRESH branches. This causes a destructive interference rather than a

---

[10] The main body of this section can be found in (Yeste-Ojeda & Grajal, 2010).

constructive one when the outputs of the branches are added together. Although this is a well known problem, it has been exhaustively dealt with only in the field of beamforming and direction finding, mainly regarding MUSIC (Schmidt, 1986) and SCORE (Agee et al., 1990) algorithms. Besides describing the problem, some solutions have been proposed. Schell and Gardner stated that the degradation in performance increases with the product of the observation time and the error in cycle frequency (Schell & Gardner, 1990). As a solution, they proposed two alternatives: Using multiple cycle frequencies in order to increase robustness, or adaptively estimating the cycle frequencies by maximizing the magnitude of the FFT of the lag product of the observations. Lee and Lee proposed two procedures for estimating the cycle-frequency offset (Lee & Lee, 1999). Both of them are based on the maximization of the highest eigenvalue of the product of a sample autocorrelation-related matrix. In (Lee et al., 2000), cycle frequencies are estimated iteratively through a gradient descent algorithm which maximizes the output power of an adaptive beamformer, assuming that the initial cycle frequencies are close enough. A similar approach can be found in (Jianhui et al., 2006), where the conjugate gradient algorithm is used instead of gradient descent, and after proper modifications in the cost function. A new approach is described in (Zhang et al., 2004), where the robustness of the beamformer is achieved by minimizing a cost function which is averaged over a range of cycle frequencies, instead of trying to correct the cycle frequency errors.

The approach used in this section is rather different from the cited work. The purpose of this section is to explore an additional advantage of using adaptive filters for FRESH filtering, which is their capability of working in the presence of errors in the frequency shifts. This is possible because the adaptive filter is able to track these errors, by updating the coefficients of the LTI filters in a cyclic manner. As a result, the adaptive filter at each branch of the FRESH filter behaves as a Linear Periodically Time-Variant (LPTV) filter, rather than converging to an LTI filter. This ability is strongly conditioned by the rate of convergence of the adaptive algorithm. For slow convergence, the outputs of the branches with frequency-shift errors are cancelled. As the convergence accelerates, the adaptive filters of those branches with frequency-shift errors behave as LPTV filters in order to compensate the errors. These subjects shall be dealt with later on, but firstly let us highlight the problem of cycle-frequency errors in FRESH filtering.

6.1 Effect of errors in the frequency shifts

The scheme of the adaptive FRESH filter considered in this section has been shown in Fig. 6. The problem of an error in the frequency shifts of the FRESH filter is that the correlation between the desired signal and the corresponding filter input vanishes. For a nonadaptive FRESH filter, the error can be tolerated during a limited observation time. At each branch, the contribution to the signal estimate is constructive for an interval lesser than the inverse of the error of its frequency shift. Therefore, the observation time tolerated by the whole filter is determined by the inverse of the biggest error in the frequency shifts. [11] For an infinite observation time, the optimal LTI filter is null for all branches with an error in their frequency shifts, as we shall demonstrate in this section. This is due to the discrete nature of the cyclic spectrum, which means that the cyclic correlation function is different from zero only at a countable set of cycle-frequency values. Thus, an error in a cycle frequency yields that the signal at the input of the corresponding filter is correlated neither with $d(n)$, nor with the inputs of the rest of filters.

---

[11] However, the degradation of the estimation error depends on the contribution of the corresponding branch to the signal estimate.

Let us use the definitions introduced in the previous section, specifically from (52) to (59). For convenience, let the auto-correlation matrix of the input vector be expressed in the form

$$\mathbf{R}_x = \begin{bmatrix} \mathbf{R}_{x_{11}} & \langle \mathrm{E}\left\{ \boldsymbol{x}_1(n)\boldsymbol{x}_{2L}^\dagger(n) \right\} \rangle \\ \langle \mathrm{E}\left\{ \boldsymbol{x}_{2L}(n)\boldsymbol{x}_1^\dagger(n) \right\} \rangle & \mathbf{R}_{x_{2L}} \end{bmatrix} \tag{62}$$

where $\mathbf{R}_{x_{ij}}$, with $i \le j$, stands for the auto-correlation matrix of input vectors from the i-th to the j-th branch:

$$\mathbf{R}_{x_{ij}} = \left\langle \mathrm{E}\left\{ \boldsymbol{x}_{ij}(n)\ \boldsymbol{x}_{ij}^\dagger(n) \right\} \right\rangle \tag{63}$$

Similarly, the vector $\boldsymbol{x}_{ij}(n)$ stands for the concatenation of input vectors from the i-th to the j-th branch:

$$\boldsymbol{x}_{ij}(n) = \left[\boldsymbol{x}_i^T(n),\ \ldots,\ \boldsymbol{x}_j^T(n) \right]^T \tag{64}$$

Let us assume that there is an error in the frequency shift used for the first branch. Then, the time-averaged cross-correlation between $\boldsymbol{x}_1(n)$ and the input vectors of the rest of branches is zero. Moreover, the time-averaged cross-correlation of $\boldsymbol{x}_1(n)$ with the desired signal, $d(n)$, is also zero. Thus, the optimal set of LTI filters when there exist errors in the frequency shift of the first branch results:

$$\boldsymbol{w}_o' = (\mathbf{R}_x)^{-1} \boldsymbol{p} = \begin{bmatrix} \mathbf{R}_{x_{11}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{x_{2L}}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \boldsymbol{p}_{2L} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{R}_{x_{2L}}^{-1}\, \boldsymbol{p}_{2L} \end{bmatrix} \tag{65}$$

Therefore, the optimal LTI filter of the first branch and its output are null. In addition, the optimal set of filters for the other branches is equivalent to not using the first branch. In short, what is happening is that the first frequency shift does not belong to set Γ. When all branches exhibit errors in their frequency shifts, then the optimal filter vector is zero and the FRESH filter becomes useless. It is noteworthy that the cancellation of the branches with an error in their frequency shift is caused by the LTI characteristic of the filters, which has been forced by time-averaging $\mathbf{R}_x$ and $\boldsymbol{p}$. In the next section, it is shown that the errors in the frequency shifts can be compensated by allowing the filters to be time-variant, specifically, LPTV.

6.2 LPTV solution

When there exist errors in the frequency shifts, an adaptive-FRESH filter can compensate them. In order to understand the behavior of the adaptive filter, let us assume that these errors are known. One direct solution to this problem is to substitute the corresponding LTI filters by LPTV filters consisting of a correcting frequency shift followed by an LTI filter, as shown in Fig. 8. Formally, the output of the optimal FRESH filter defined in (56), when there are not errors in the frequency shifts, is:

$$\widehat{d}(n) = \boldsymbol{w}_o^\dagger\, \boldsymbol{x}(n) = \sum_{i=1}^{L} \boldsymbol{w}_{oi}^\dagger\, \boldsymbol{x}_i(n) \tag{66}$$

where $\boldsymbol{w}_{oi}$ stands for the optimal LTI filter in the i-th branch and $\boldsymbol{x}_i(n)$ are the outputs of the frequency shifters without error. The LPTV solution can be obtained by including the errors
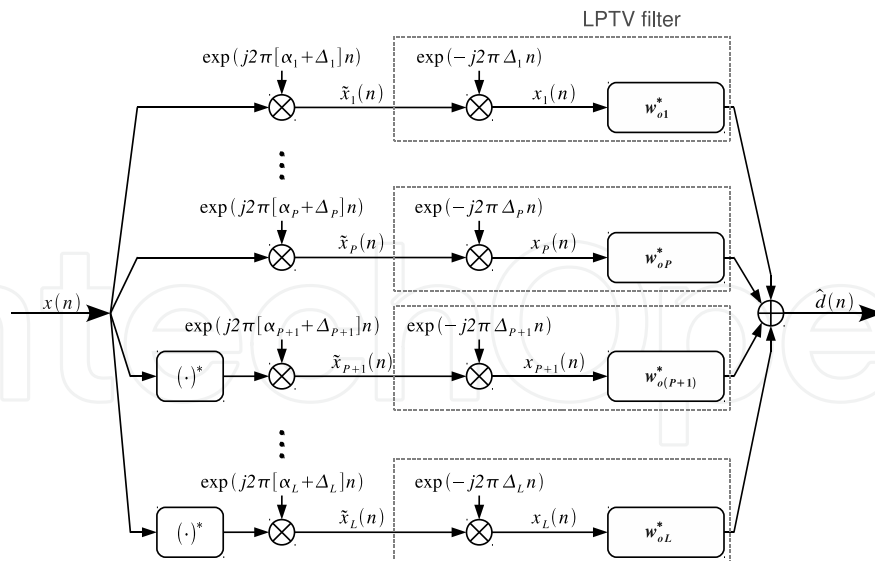
Fig. 8. Cycle frequency errors correction by means of LPTV filters.

in the frequency shifts:

$$\widehat{d}(n) = \sum_{i=1}^{L} \boldsymbol{w}_{oi}^{\dagger} \, \boldsymbol{x}_i(n) \, e^{j2\pi\Delta_i n} \, e^{-j2\pi\Delta_i n}$$

$$= \sum_{i=1}^{L} \left(e^{-j2\pi\Delta_i n} \, \boldsymbol{w}_{oi}^{\dagger} \right) \left(\boldsymbol{x}_i(n) \, e^{j2\pi\Delta_i n} \right) \tag{67}$$

$$= \sum_{i=1}^{L} \tilde{\boldsymbol{w}}_{oi}^{\dagger}(n) \, \tilde{\boldsymbol{x}}_i(n) = \tilde{\boldsymbol{w}}_o^{\dagger}(n) \, \tilde{\boldsymbol{x}}(n)$$

where $\tilde{\boldsymbol{w}}_o(n)$ is the set of LPTV filters, and $\tilde{\boldsymbol{x}}(n)$ are the input vector to these filters (the output of the first frequency shifters in Fig. 8).
By defining the diagonal matrix

$$\boldsymbol{\Phi} \triangleq \begin{bmatrix} e^{j2\pi\Delta_1}\mathbf{I}_{M_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & e^{j2\pi\Delta_2}\mathbf{I}_{M_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & e^{j2\pi\Delta_L}\mathbf{I}_{M_L} \end{bmatrix} \tag{68}$$

where $\mathbf{I}_N$ is the $N \times N$ identity matrix, the vectors $\tilde{\boldsymbol{w}}_o(n)$ and $\tilde{\boldsymbol{x}}(n)$ can be expressed in the form:

$$\tilde{\boldsymbol{w}}_o(n) = \boldsymbol{\Phi}^n \, \boldsymbol{w}_o \tag{69}$$

$$\tilde{\boldsymbol{x}}(n) = \boldsymbol{\Phi}^n \, \boldsymbol{x}(n) \tag{70}$$

Additionally, the optimal set of linear time-variant filters is obtained by minimization of the instantaneous MSE. This optimal set (whose input vector is $\tilde{\boldsymbol{x}}(n)$) is (Van Trees, 1968):

$$\tilde{\boldsymbol{w}}_o(n) = \left[\mathbf{R}_{\tilde{x}}(n) \right]^{-1} \tilde{\boldsymbol{p}}(n) \tag{71}$$

where

$$\mathbf{R}_{\tilde{x}}(n) = \mathrm{E}\left\{\tilde{\boldsymbol{x}}(n)\ \tilde{\boldsymbol{x}}^{\dagger}(n)\right\} \tag{72}$$

$$\tilde{\boldsymbol{p}}(n) = \mathrm{E}\left\{\tilde{\boldsymbol{x}}(n)\ d^{*}(n)\right\} \tag{73}$$

Therefore, since (69) is verified, a sufficient condition to obtain (69) from (71), is that:

$$\mathbf{R}_{\tilde{x}}(n) \triangleq \boldsymbol{\Phi}^{n}\left\langle \mathrm{E}\left\{\boldsymbol{x}(n)\ \boldsymbol{x}^{\dagger}(n)\right\}\right\rangle \boldsymbol{\Phi}^{-n}$$

$$= \boldsymbol{\Phi}^{n}\ \mathbf{R}_{x}\ \boldsymbol{\Phi}^{-n} \tag{74}$$

$$\tilde{\boldsymbol{p}}(n) \triangleq \boldsymbol{\Phi}^{n}\left\langle \mathrm{E}\left\{\boldsymbol{x}(n)d^{*}(n)\right\}\right\rangle$$

$$= \boldsymbol{\Phi}^{n}\ \boldsymbol{p} \tag{75}$$

where $\mathbf{R}_x$ and $\boldsymbol{p}$ have been defined in (57) and (59) respectively.

Note that (74) and (75) are not the true auto-correlation matrix and cross-correlation vector between $\tilde{\boldsymbol{x}}(n)$ and $d(n)$. On the contrary, they are based on the stationarized signal model of the input vector $\boldsymbol{x}(n)$ and the desired signal $d(n)$, so that the LTI characteristic of the filters after the second frequency shifters in Fig. 8 is forced.

An additional motivation for introducing (74) and (75) is that these expressions will serve us to develop the convergence of the LMS algorithm mathematically. We have chosen the Least Mean Squares (LMS) adaptive algorithm for the analysis because of its mathematical simplicity, which allows an analytical treatment of the problem. Nonetheless, any adaptive algorithm (for instance, Recursive Least Squares (RLS) (Haykin, 2001)) can be used instead to compensate cycle-frequency errors. The approach herein is the same as as in (Widrow et al., 1976), where the MSE (Mean-Squared-Error) of the adaptive filter is decomposed into the error due to noise in the gradient estimation, and the error due to lag between the optimal and the adaptive filters.

It is straightforward that the LPTV solution in (69) is equivalent to the optimal FRESH filter in the absence of errors. In practice, the errors in the frequency shifts are unknown. Thus, there is not any *fixing* frequency shifter (the second one, in Fig. 8), and the filters of each branch must work with input vector $\tilde{\boldsymbol{x}}(n)$. The advantage of using an adaptive filter is that it can produce the periodic variations of the filters by updating their coefficients. This is only possible for a fast enough rate of convergence. Otherwise, the adaptive filter tends to the LTI optimal solution defined by (56) after substituting the input vector $\boldsymbol{x}(n)$ with $\tilde{\boldsymbol{x}}(n)$ in (57) and (59). This solution implies the cancellation of the branches with errors in their frequency shift. This issue is addressed in next section.

6.3 Excess MSE of the LMS algorithm in the presence of errors in the frequency shifts

In this section we develop an analytical expression for the time-averaged MSE at steady-state of adaptive FRESH filters using the LMS algorithm (LMS-FRESH). The objective is to obtain an expression that accounts for errors in the frequency shifts of the LMS-FRESH filter.

Basically, the LMS algorithm consists in updating the filter weights in the direction of a gradient estimate of the MSE function, which can be thought as a stochastic version of the steepest descent algorithm (Widrow et al., 1976). Let $\tilde{\boldsymbol{w}}(n)$ be the concatenation vector defining the LMS-FRESH filter. The updating algorithm can be formulated as follows:

$$\tilde{\boldsymbol{w}}(n+1) = \tilde{\boldsymbol{w}}(n) + 2\mu\varepsilon^*(n)\tilde{\boldsymbol{x}}(n)$$
$$\varepsilon(n) = d(n) - \widehat{d}(n) \tag{76}$$
$$\widehat{d}(n) = \tilde{\boldsymbol{w}}^\dagger(n)\,\tilde{\boldsymbol{x}}(n)$$

where μ is the convergence factor (also known as step-size parameter).

The LMS-FRESH filter exhibits an excess MSE with respect to the set of filters $\tilde{\boldsymbol{w}}_o(n)$ presented in the previous section as the LPTV solution. This excess MSE can be computed as follows (Widrow et al., 1976):

$$\xi_e(n) = \mathrm{E}\left\{\left[\tilde{\boldsymbol{w}}(n) - \tilde{\boldsymbol{w}}_o(n)\right]^\dagger \mathbf{R}_{\tilde{x}}(n)\,\left[\tilde{\boldsymbol{w}}(n) - \tilde{\boldsymbol{w}}_o(n)\right]\right\} \tag{77}$$

Let us separate the excess MSE into two terms: The excess MSE due to gradient noise and the excess MSE due to lag error. The separation is possible by considering the weight error vector as the sum of two terms:

$$\tilde{\boldsymbol{w}}(n) - \tilde{\boldsymbol{w}}_o(n) = \left[\tilde{\boldsymbol{w}}(n) - \mathrm{E}\left\{\tilde{\boldsymbol{w}}(n)\right\}\right] + \left[\mathrm{E}\left\{\tilde{\boldsymbol{w}}(n)\right\} - \tilde{\boldsymbol{w}}_o(n)\right] = \tilde{\boldsymbol{u}}(n) + \tilde{\boldsymbol{v}}(n) \tag{78}$$

The first term of the sum in the right hand side, $\tilde{\boldsymbol{u}}(n) = \tilde{\boldsymbol{w}}(n) - \mathrm{E}\left\{\tilde{\boldsymbol{w}}(n)\right\}$, produces the excess MSE due to gradient noise, which is related to the gradient estimation process. The second term, $\tilde{\boldsymbol{v}}(n) = \mathrm{E}\left\{\tilde{\boldsymbol{w}}(n)\right\} - \tilde{\boldsymbol{w}}_o(n)$, produces the excess MSE due to lag, which quantifies the error due the fact that the adaptive filter cannot follow the variations of $\tilde{\boldsymbol{w}}_o(n)$ with time. By substituting (78) in (77), we obtain:

$$\xi_e(n) = \mathrm{E}\left\{\tilde{\boldsymbol{u}}^\dagger(n)\,\mathbf{R}_{\tilde{x}}(n)\,\tilde{\boldsymbol{u}}(n)\right\} + \mathrm{E}\left\{\tilde{\boldsymbol{v}}^\dagger(n)\,\mathbf{R}_{\tilde{x}}(n)\,\tilde{\boldsymbol{v}}(n)\right\} + 2\mathcal{R}e\left\{\mathrm{E}\left\{\tilde{\boldsymbol{u}}^\dagger(n)\,\mathbf{R}_{\tilde{x}}(n)\,\tilde{\boldsymbol{v}}(n)\right\}\right\} \tag{79}$$

It can be easily shown (from the definitions of $\tilde{\boldsymbol{u}}(n)$ and $\tilde{\boldsymbol{v}}(n)$) that the last term of (79) is zero. Thus, the total time-averaged MSE at steady-state of the LMS-FRESH filters can be expressed as the sum of three terms:

$$\xi_{LMS} \triangleq \left\langle \mathrm{E}\left\{|\varepsilon(n)|^2\right\}\right\rangle = \xi_{min} + \left\langle \xi_e(n)\right\rangle = \xi_{min} + \xi_\nabla + \xi_{lag} \tag{80}$$

where ξ_{min} is the minimum time-averaged MSE attained by the optimal FRESH filter, ξ_∇ is the time-averaged excess MSE due to gradient noise, and ξ_{lag} is the time-averaged excess MSE due to lag error.

The minimum time-averaged MSE results from multivariate Wiener theory (Gardner, 1993):

$$\xi_{min} = \sigma_{\tilde{d}}^2 - \left\langle \tilde{\boldsymbol{w}}_o^\dagger(n)\,\mathbf{R}_{\tilde{x}}(n)\,\tilde{\boldsymbol{w}}_o(n)\right\rangle$$
$$= \sigma_{\tilde{d}}^2 - \boldsymbol{w}_o^\dagger\,\mathbf{R}_x\,\boldsymbol{w}_o \tag{81}$$

where $\sigma_{\tilde{d}}^2 = \left\langle \mathrm{E}\left\{|d(n)|^2\right\}\right\rangle$ is the mean power of the reference signal.

At each instant, the excess MSE due to gradient noise can be computed as follows (Widrow et al., 1976):

$$\xi_\nabla(n) = \mathrm{E}\left\{\tilde{\boldsymbol{u}}^\dagger(n)\,\mathbf{R}_{\tilde{x}}(n)\,\tilde{\boldsymbol{u}}(n)\right\} \tag{82}$$

Assuming that the LMS adaptive filter converges to an LPTV filter at each branch (as shown in Fig. 8), the time-averaged excess MSE at steady-state can be approximated by[12]:

$$\xi_\nabla = \langle \xi_\nabla(n) \rangle \approx \mu \, \xi_{min} \, \text{tr} \left\{ \mathbf{R}_x \right\} \tag{83}$$

where $\text{tr} \left\{ \mathbf{R}_x \right\}$ is the trace of matrix $\mathbf{R}_x$ defined in (57). From (83), ξ_∇ can be reduced as much as desired by decreasing μ. However, the convergence factor μ controls the rate of convergence of the LMS, so that it is faster as μ increases. Additionally, it will be shown next that the excess MSE due to lag error increases as μ decreases.

The excess MSE due to lag error becomes apparent when the adaptive filter cannot follow the variations of $\tilde{w}_o(n)$ (the LPTV solution) with time. This excess MSE is a function of the weight-error vector $\tilde{v}(n)$, which represents the instantaneous difference between the expected value of the adaptive filter and $\tilde{w}_o(n)$:

$$\xi_{lag}(n) = \tilde{v}^\dagger(n) \, \mathbf{R}_{\tilde{x}}(n) \, \tilde{v}(n) \tag{84}$$

The mathematical development of the time-averaged excess MSE at steady-state due to the lag error can be found in the appendix of (Yeste-Ojeda & Grajal, 2010), and yields:

$$\xi_{lag} = v^\dagger \, \mathbf{R}_x \, v \tag{85}$$

where v, which does not depend on time, is the weight-error vector at steady-state. Under the same conditions that were assumed for (83) (see footnote 12):

$$v = \lim_{n \to \infty} \mathbf{\Phi}^{-n} \, \mathrm{E} \left\{ \tilde{w}(n) \right\} - w_o$$
$$= \left(2\mu \left[\mathbf{\Phi} - \mathbf{I}_M + 2\mu \mathbf{R}_x \right]^{-1} - \mathbf{R}_x^{-1} \right) p \tag{86}$$

6.4 Analysis of the results

The study of the previous expressions allows to infer several conclusions:

1. The lag error, ξ_{lag}, tends to zero as μ increases. This is easily seen by taking the limit of (86) as $\mu \to \infty$:

$$\lim_{\mu \to \infty} v = \left(\mathbf{R}_x^{-1} - \mathbf{R}_x^{-1} \right) p = 0 \tag{87}$$

   This result indicates that, for a high enough convergence factor, the LMS-FRESH filter is fast enough to follow the variations of the LPTV solution, and therefore the excess MSE due to lag error is zero. However, μ cannot be made as big as desired because then, the adaptive algorithm may not converge, as we shall discuss next.

2. The lag error is also zero when there are not any errors in the frequency shifts. In this case, $\mathbf{\Phi} = \mathbf{I}_M$, which substituted in (86) gives $v = 0$. This result is coherent with the concept of lag error. In the absence of errors, the LPTV solution to which the LMS-FRESH filter converges becomes LTI. Consequently, the existence of a lag is not possible.

3. On the contrary, the lag error is maximum when μ tends to zero (and there exist frequency-shift errors). Taking the limit of (86) as $\mu \to 0$ yields:

$$\lim_{\mu \to 0} v = -\mathbf{R}_x^{-1} \, p = -w_o \tag{88}$$

---

[12] Under small error conditions, and by assuming that $\varepsilon(n)$ and $\tilde{x}(n)$ are Gaussian, and $\tilde{x}(n)$ is uncorrelated over time, i.e. $\mathrm{E} \left\{ \tilde{x}(n) \, \tilde{x}(n+k) \right\} = 0$, with $k \neq 0$, (Widrow et al., 1975; 1976).

which implies that the LMS-FRESH filter converges to the null vector in a mean sense. Moreover, the filter converges to the null vector also in a mean square sense. This results is derived from the fact that the gradient noise tends to zero as μ tends to zero (see (83)), which entails that $\tilde{u}(n) = 0$, and therefore the LMS-FRESH filter vector matches its expected value, i.e. $\tilde{w}(n) = E\{\tilde{w}(n)\} = 0$. As a result, the outputs of the branches with an error in their frequency shift are null. The time-averaged excess MSE due to lag error is obtained by substituting (88) in (85), which yields:

$$\lim_{\mu \to 0} \xi_{lag} = w_o^\dagger \, \mathbf{R}_x \, w_o \tag{89}$$

Thus, the total time-averaged MSE at steady-state for μ tending to zero is

$$\lim_{\mu \to 0} \xi_{LMS} = \sigma_d^2 \tag{90}$$

This result could be obtained considering that, since the output of the adaptive-FRESH filter is null, the error signal is $\varepsilon(n) = d(n)$.

4. The optimal convergence factor which minimizes the time-averaged MSE of the LMS algorithm is obtained from (80), (83) and (85), which yields:

$$\mu_o = \arg \min_{\mu} \left\{ \mu \, \xi_{min} \, \operatorname{tr}\{\mathbf{R}_x\} + v^\dagger \, \mathbf{R}_x \, v \right\} \tag{91}$$

where vector v is defined by (86).

As regards the convergence of the LMS-FRESH filter, a sufficient condition is that the poles of all components of vector $v(z)$ are located inside the unit circle (Yeste-Ojeda & Grajal, 2010). This is equivalent to the condition that the maximum absolute value of the eigenvalues of matrix

$$(\mathbf{I}_M - 2\mu\mathbf{R}_x) \, \mathbf{\Phi}^{-1} \tag{92}$$

is lesser or equal than 1. When a single branch is used for the LMS-FRESH filter, then it is straightforward that the eigenvalues of (92) are $e^{-j2\pi\Delta}(1 - 2\mu\lambda_k)$, with $k = 1, \ldots, M$, and where Δ is the frequency-shift error of the single branch, and λ_k are the eigenvalues of matrix $\mathbf{R}_x$. As a result, the condition for the convergence of the LMS algorithm is:

$$0 < \mu < (\lambda_{max})^{-1} \tag{93}$$

where λ_{max} is the maximum eigenvalue of matrix $\mathbf{R}_x$. This is the same condition as for stationarity environments (Haykin, 2001; Widrow et al., 1976). [13]

6.5 Application

Finally, let us quantify the performance of adaptive FRESH filters for cycle-frequency error compensation through a case study where the received signal consists of a BPSK signal embedded in stationary white Gaussian noise. The receiver incorporates a FRESH filter with the purpose of extracting the BPSK signal from the noise. Two frequency shifts related to the cyclic spectrum of the BPSK signal are used, the inverse of its symbol interval $\alpha_1 = 1/T_s$, and twice its carrier frequency, $\alpha_2 = 2f_c$. These two cycle frequencies have been chosen, according to the common approach mentioned in Section 4, because they exhibit the highest values of

---

[13] Note that the value of the convergence factor used in (Haykin, 2001) is twice the value used herein.
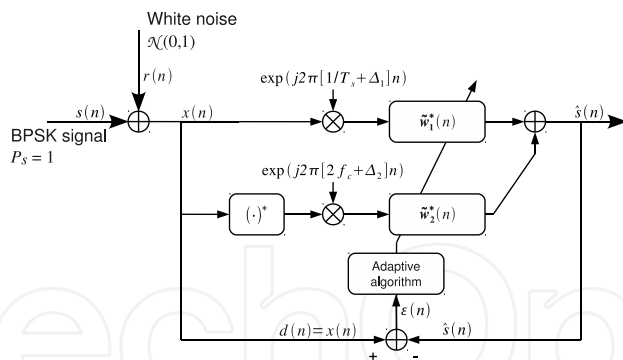
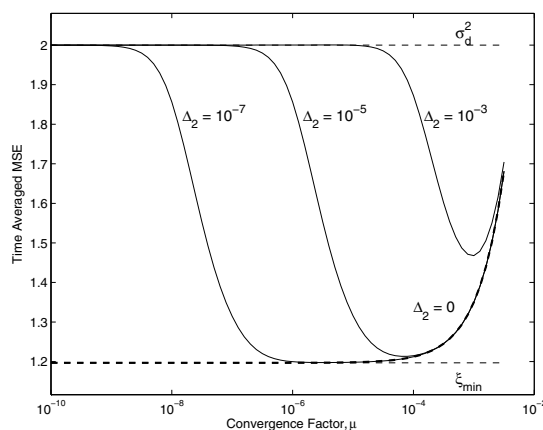Fig. 9. Blind adaptive FRESH filter for BPSK signal extraction.

Fig. 10. Analytical time-averaged MSE as a function of the convergence factor, when using only the branch corresponding to frequency shift $\alpha = 2f_c$. The thick dashed line corresponds to $\Delta_2 = 0$.

spectral correlation for a BPSK modulation (Gardner et al., 1987). The desired signal for the adaptive algorithm is the received signal, $d(n) = x(n)$, following the blind scheme described in Section 5.1 (see Fig. 9). In all cases, the carrier frequency and symbol interval have been set to $f_c = 0.3$ (normalized to the sampling frequency) and $T_s = 32$ samples. The noise power is set to $\sigma_r^2 = \mathrm{E}\left\{|r(n)|^2\right\} = 1$, and the SNR is also fixed to SNR= 0 dB, and is defined as the ratio between the mean power of the signal and the noise: SNR= $\langle \mathrm{E}\left\{|s(n)|^2\right\}\rangle / \sigma_r^2$. Both $\tilde{w}_1(n)$ and $\tilde{w}_2(n)$, are FIR filters with $M_i = 64$ taps.

In order to clarify some concepts, let us consider firstly the case where the FRESH filter is composed of only the branch associated with the frequency shift $\alpha_2 = 2f_c$. The total time-averaged MSE at steady-state (hereinafter referred to as simply "the MSE") is shown in Fig. 10, as a function of the convergence factor and for different values of the frequency-shift error, Δ_2, which are also normalized to the sampling frequency. The MSE of the LMS-FRESH filter when $\Delta_2 = 0$ is plotted with a dashed thick line as a point of reference. This is the lower bound of the MSE attainable by the LMS-FRESH filter, and converges to the MSE of the optimal FRESH filter, ξ_{min}, as μ decreases. Note that the minimum MSE is always greater than the noise power, i.e. $\xi_{min} > 1$, since even if the signal were perfectly estimated ($\hat{s}(n) = s(n)$), the error signal would match the noise ($\varepsilon(n) = r(n)$).

When there exist errors in the frequency shift, the MSE is always higher than the lower bound. Since the lower bound includes the gradient noise effect, the excess MSE over the lower bound
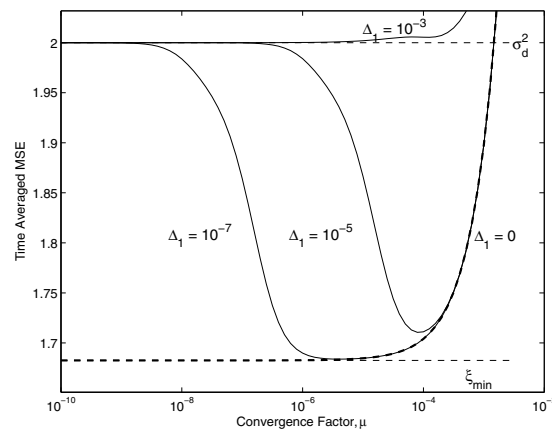
Fig. 11. Analytical time-averaged MSE as a function of the convergence factor, when using only the branch corresponding to frequency shift $\alpha = 1/T_s$. The thick dashed line corresponds to $\Delta_1 = 0$.

is due to the lag error only, and varies from zero (for high μ) up to $\boldsymbol{w}_o^\dagger \mathbf{R}_x \boldsymbol{w}_o = \sigma_d^2 - \xi_{min}$, (see (89), for small μ). Thus, the total MSE tends to σ_d^2 when μ decreases. The curves also show the dependence of the MSE with the error in the frequency shift, which increases with Δ_2 (at any given μ). Therefore, in order to obtain a small excess MSE due to lag, it is required a faster rate of convergence (bigger μ) when Δ_2 increases.

An additional result shown in Fig. 10 is that for high enough errors in the frequency shifts, the MSE does not reach ξ_{min} at any value of μ. In other words, it is impossible to simultaneously reduce the MSE terms due to the lag error and the gradient noise. In such a case, the optimal μ locates at an intermediate value as a result of the trade-off between reducing the lag error and the gradient noise. In practice, the error of the frequency-shifts is commonly unknown. Then the convergence factor should be chosen as a trade-off between the maximum cycle-frequency error which can be compensated by the adaptive filter and the increase of the excess MSE due to gradient noise.

Similar results can be extracted from Fig. 11, which corresponds to the case where only the branch of the FRESH filter with frequency shift $\alpha = 1/T_s$ has been used. The main difference between Figures 10 and 11 is that the minimum MSE, ξ_{min}, is bigger when only $\alpha = 1/T_s$ is used. The reason is that the spectral correlation level exhibited by a BPSK signal is smaller at $\alpha = 1/T_s$ than at $\alpha = 2f_c$ (Gardner et al., 1987). Furthermore, it can be seen that σ_d^2 is not an upper bound for the MSE (as could be thought from Fig. 10), but the limit of the MSE as μ tends to zero.

For the case illustrated in Fig. 12, the two branches shown in Fig. 9 are used, but there is uncertainty only in the symbol rate, that is, the error in the frequency shift related to the carrier frequency, Δ_2, is always zero. The curves show that the improvement in the MSE is not very significant when the error in the symbol rate is compensated. This occurs when the contribution to the signal estimate for a branch is more significant than for the other, which is mainly caused by the different spectral level exhibited by a BPSK signal at cycle frequencies $\alpha = 1/T_s$ and $\alpha = 2f_c$. As a consequence, the minimum MSE when only the second branch is used ($\xi_{min,2}$), and when both branches are used ($\xi_{min,12}$) are very similar. Furthermore, μ tends to zero the MSE tends to $\xi_{min,2}$ instead of σ_d^2, since there is no uncertainty in the carrier frequency.
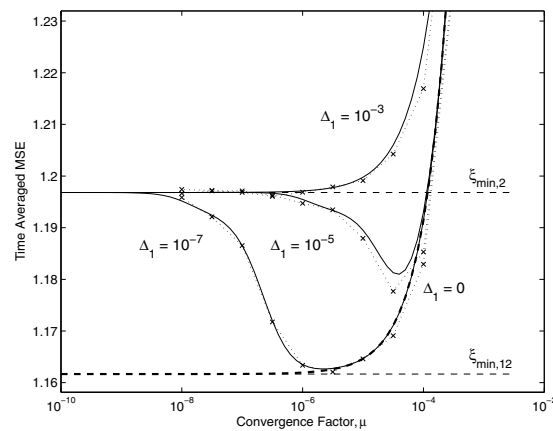
Fig. 12. Time-averaged MSE as a function of the convergence factor, when using the two branches. $\Delta_2 = 0$ in all cases, also for the thick dashed line which corresponds to $\Delta_1 = 0$. Simulation results are represented by cross marks.
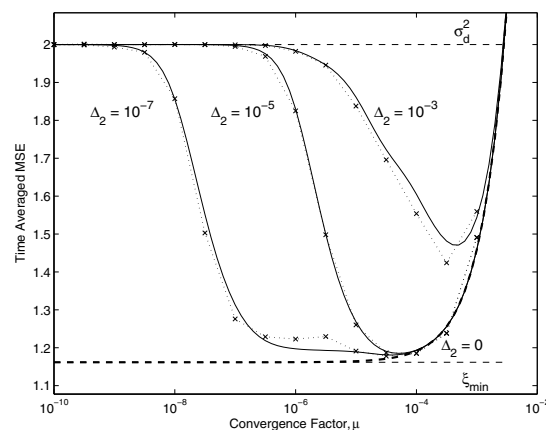

Fig. 13. Time-averaged MSE as a function of the convergence factor, when using the two branches. $\Delta_1 = 10^{-5}$ in all cases, except for the thick dashed line which corresponds to $\Delta_1 = 0$ and $\Delta_2 = 0$. Simulation results are represented by cross marks.

Fig. 13 shows the MSE when there exist errors in both frequency shifts. The curves correspond to an error in the symbol rate $\Delta_1 = 10^{-5}$ and different errors for the carrier frequency. In this case, the minimum MSE is attained at a convergence factor such that the lag error of both branches is compensated. However, compensating the error in the carrier frequency is critical, while compensating the error in the symbol rate only produces a slight improvement. This conclusion can be deduced from the curve for $\Delta_2 = 10^{-7}$. Since $\Delta_1 = 10^{-5}$, it is required a convergence factor close to $\mu = 10^{-4}$ or higher in order to compensate the frequency-shift error at the first branch (see Fig. 11). However, the excess MSE due to lag error is small for $\mu = 10^{-6}$ or higher, where only the frequency-shift error at the second branch can be compensated.

The analytical expression for the MSE has been obtained under some assumptions (Gaussianity, small error conditions and input uncorrelated over time) which in practice are only approximations, and also in the case studies presented. Therefore, in order to check its accuracy, Figures 12 and 13 also include the MSE obtained by simulation, which is represented

Parameter	Min. value	Max. value
INR (interference-to-noise ratio)	-20 dB	20 dB
T_{s2} (interference symbol interval)	1 sample	64 samples
f_{c2} (interference carrier frequency)	0	1
Δ_1	-10^{-5}	10^{-5}
Δ_2	-10^{-5}	10^{-5}

Table 1. Range of the random variables used in the simulation with a BPSK interference. The INR is defined analogously to SNR. f_{c2} is normalized with the sampling rate. Δ_1 and Δ_2 are the absolute cycle-frequency errors of the branches with cycle frequencies $\alpha_1 = 1/T_s$ and $\alpha_2 = 2f_c$, respectively.

by the lines with cross marks: 200 realizations have been used for the ensemble averages in order to obtain the instantaneous MSE. Then, the instantaneous MSE has been time-averaged over 200.000 samples after the convergence of the LMS. The agreement between theoretical and simulation results is excellent in all cases, which confirms the validity of the assumptions made.

A last case study is presented with a twofold purpose: 1) To demonstrate that the adaptive algorithm compensates cycle-frequency errors also in the presence of interferences. 2) To demonstrate that an adaptive algorithm different from the LMS exhibits a similar behavior. For this reason, we shall use the RLS algorithm in this last case study, despite the lack of an analytical expression for the MSE. The scheme for the adaptive FRESH filter presented in Fig. 9 is valid also in this case study, but a BPSK interference having random parameters has been added to the input, along with the BPSK signal and the noise defined for the previous case studies. The errors in the frequency shifts of the FRESH filter are also random. All random variables are constant at each trial, but change from trial to trial according an uniform distribution within the ranges gathered in Table 1. As regards the RLS algorithm, it is exponentially weighted with a convergence factor $\lambda = 1 - \mu$, where μ is known as the forgetting rate (Haykin, 2001). Also, we have used in this case study the BPSK signal as the reference signal, so that the MSE does not depend on the random interference power.

Fig. 14 shows the MSE obtained by simulation (using 100 trials for computing the ensemble averages and 20000 samples for the time averages). The results show the capability of the RLS algorithm for compensating errors in the frequency shifts in the presence of the interference. Otherwise, the RLS algorithm would have converged to a FRESH filter which cancels its output, and the obtained MSE would have been equal to σ_d^2. Analogously to the previous case studies using the LMS algorithm, the RLS algorithm cannot compensate errors in the frequency shifts if the forgetting rate is too small (slow convergence). In such a case, the adaptive FRESH filter tends to cancel its output and the error is σ_d^2. For moderate forgetting rates, the cycle-frequency errors are compensated, and the MSE approaches ξ_{min}. Contrarily, an excessively high forgetting rate increases the MSE as a consequence of the gradient noise.

In summary, the possible existence of cycle-frequency errors in LAPTV filtering is a serious problem that must be managed. When using the non-adaptive FRESH implementation, the minimum time-averaged MSE is obtained when the branches with uncertainties in their frequency shifts are not used (or equivalently, their output is cancelled). On the contrary, adaptive-FRESH filters can work in the presence of errors in the frequency shifts. In such a case, the adaptive-FRESH filter behaves as an LPTV filter for those branches with an error in the frequency shift. In order to be effective, the rate of convergence of the adaptive algorithm must be carefully chosen. The optimal rate of convergence results from the trade-off between
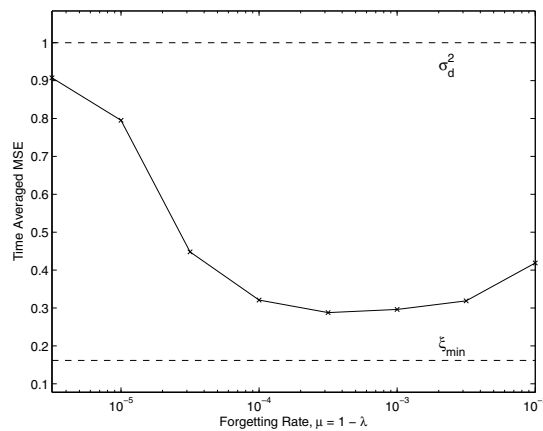
Fig. 14. Simulated time-averaged MSE as a function of the forgetting rate of the RLS algorithm, in the presence of a BPSK interference with random parameters.

decreasing the excess MSE due to gradient noise (slow rate of convergence) and decreasing the excess MSE due to lag error (fast rate of convergence). The analytical expressions in this section allow to compute the optimal convergence factor and the time-averaged MSE at steady-state of the LMS-FRESH filter.

7. Adaptive FRESH filters for interference rejection in signal detection

Finally, we end this chapter with the description of a real application example of adaptive-FRESH filters. The system developed in this section, previously presented in (Yeste-Ojeda & Grajal, 2008), finds application in fields such as electronic warfare or spectrum management. The increasing saturation of the radio-electric spectrum has led modern radar and communication systems to employ complex signal waveforms capable of operating under the presence of interferences. On the contrary, interception systems face serious problems in the detectiong unknown or partially known signals if the signal of interest (SOI) is hidden by temporally and spectrally overlapping interferences. Therefore, hostile transmitters can take advantage of this fact and reduce their probability of interception by following this strategy.

The problem dealt with in this section consists in detecting an unknown signal hidden by an interference with known parameters when only one sensor is available. The SOI and the interference are overlapped simultaneously in time and frequency, with the interference being more powerful that the SOI. As the reader will have guessed, the solution adopted to solve this problem is to exploit the cyclostationary properties of the interference in order to extract it from the received signal. This procedure requires knowing, at least, the cyclic spectrum of the interference. On the other hand, much work has been done when only the SOI cyclic spectrum is known, mainly with the aim of robust signal detection and estimation (Gardner, 1993; Mirbagheri et al., 2006; Zhang et al., 1999). The approach in this section is significantly different from these works, since it exploits the cyclostationarity of those signals to be removed, i.e. the interference.

Our approach is based on a "*divide and conquer*" strategy, by which the global problem is split into two sub-problems: A signal separation problem and a signal detection one. Firstly, the interference is separated by means of a blind adaptive FRESH filter. Once the interference has
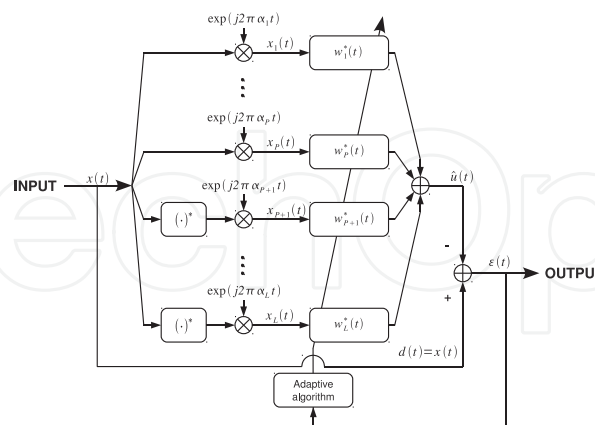
Fig. 15. Block diagram of the interference rejection system.

been removed, the detection is performed on the residual, which is assumed to consist of only the noise and the SOI.

7.1 The interference rejection system

This system is mainly based on an BAFRESH filter, following the scheme represented in Fig. 15. This scheme is the same as the one represented in Fig. 7, but after rearranging it so that the error signal is the output of the system, that is the input of the detection system.

Let the received signal be composed by the SOI, $s(t)$, the interferences, $u(t)$, and a stationary white noise, $r(t)$:

$$x(t) = s(t) + u(t) + r(t) \tag{94}$$

where $s(t)$, $u(t)$, and $r(t)$, are assumed to be independent from one another. If the interferences were perfectly estimated, that is $\hat{u}(t) = u(t)$, the estimation error would only consist of the SOI plus noise, $\varepsilon(t) = s(t) + r(t)$. Thus, the input of the detector must be the estimation error, $\varepsilon(t)$, as indicated by Fig. 15.

The next step in the design of the FRESH filter is to choose a suitable set of frequency shifts. Since the adaptive FRESH filters is blind, these cycle frequencies must belong uniquely to the cyclic spectrum of the signal to be estimated, i.e. the interference. We shall follow the strategy mentioned in Section 4 consisting in taking those cycle frequencies related with the interferences which exhibit the highest spectral correlation level, with the exception of cycle frequency zero for the branches not using the complex conjugate of the input. Since the cyclic spectrum of the SOI is unknown, we can only assume that the chosen cycle frequencies do not belong to its cyclic spectrum (which would be very infrequent in practice as the interference and the SOI are transmitted by different sources).

Given the set of frequency shifts used, Γ_s, the adaptive FRESH filter will converge to the optimum set of LTI filters, whose frequency response has been defined in (49). In our case, the desired signal is the interference, which is the only signal correlated with the inputs of the LTI
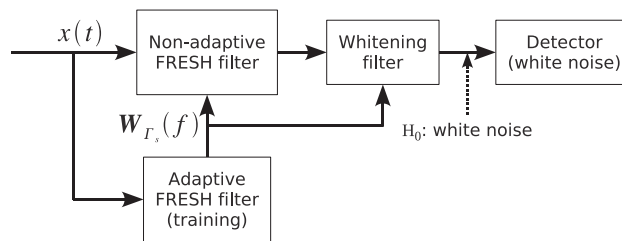
Fig. 16. Scheme for the whole interception system.

filters. As a result, the frequency response of the optimum set of LTI filters becomes:

$$\begin{aligned}
\boldsymbol{W}_{\Gamma_s}(f) &= [\mathbf{S}_{\boldsymbol{xx}}(f)]^{-1} \; \boldsymbol{S}_{d\boldsymbol{x}}^{\dagger}(f) \\
&= [\mathbf{S}_{\boldsymbol{ss}}(f) + \mathbf{S}_{\boldsymbol{uu}}(f) + \eta_r \mathbf{I_L}]^{-1} \; \boldsymbol{S}_{\boldsymbol{uu}}^{\dagger}(f)
\end{aligned} \tag{95}$$

where η_r is the noise Power Spectral Density (PSD), and $\mathbf{I_L}$ is the identity matrix of size $L \times L$. In addition, the definition of the spectral correlation vector $\boldsymbol{S}_{\boldsymbol{uu}}(f)$ is analogous to (50), while the spectral autocorrelation matrices $\mathbf{S}_{\boldsymbol{ss}}(f)$ and $\mathbf{S}_{\boldsymbol{uu}}(f)$ are defined analogously to (51).

7.2 Intermediate whitening filter and FRESH filter training

Although the separation and detection problems have been considered independently, the noise PSD at the output of the interference rejection system depends on the adaptive FRESH filter and may change with time during the convergence of the adaptive algorithm. As a result, it is advisable the use of an intermediate stage in order to effectively separate the detection problem from the FRESH filter. A whitening filter previously to the detector (as indicated in Fig. 16) can do this task, so that the detector can be designed for a white noise. The optimum detector to be used depends on the specific SOI to be detected. The whitening filter is defined as the inverse of the squared root of the noise PSD at the output of the separator. This output consists of the white noise at the input plus the noise leakage at the output of the FRESH filter, which can be expressed as a function of the FRESH filter vector. Therefore, the frequency response of the whitening filter becomes:

$$H_w(f) = \frac{1}{\sqrt{1 + |\boldsymbol{W}_{\Gamma_s}(f)|^2}} \tag{96}$$

Moreover, using an adaptive algorithm requires some training time, during which the detector decisions could be wrong. The scheme proposed in Fig. 16 consists of a branch for training the interference rejection system which incorporates the adaptive FRESH filter, and an independent non-adaptive FRESH filter which effectively performs the signal separation task and is the one connected to the detector. The coefficients of the non-adaptive FRESH filter are fixed once the adaptive one has converged. This requires some mechanism for controlling the beginning and ending of a learning interval, in which the detector decisions are considered wrong. Such a mechanism might be based on monitoring the power variations of the interferences estimate and the SOI plus noise estimate, updating the non-adaptive FRESH system when both powers stabilize.

7.3 Application

Finally, let us quantify the performance of the described interception system through a case study whose scenario is represented in Fig. 17. In that scenario, the SOI is a Continuous-Wave
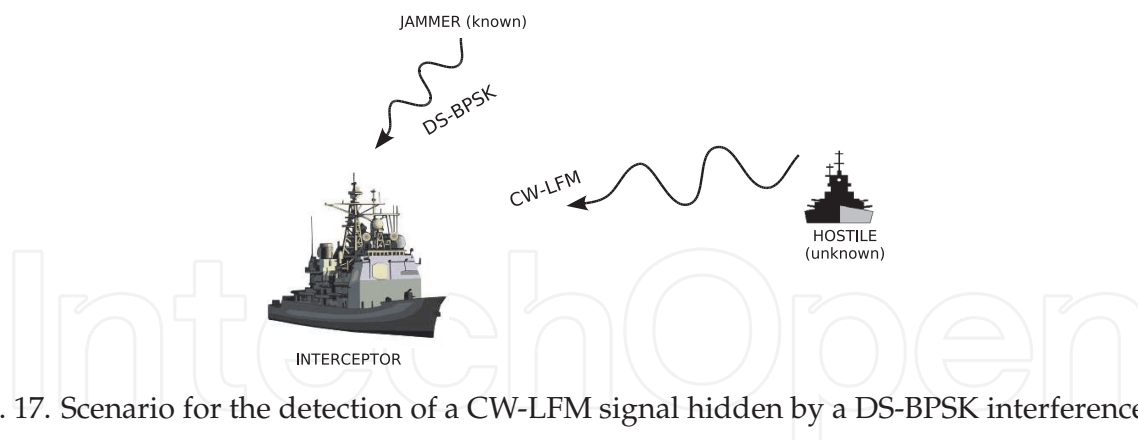
Fig. 17. Scenario for the detection of a CW-LFM signal hidden by a DS-BPSK interference.
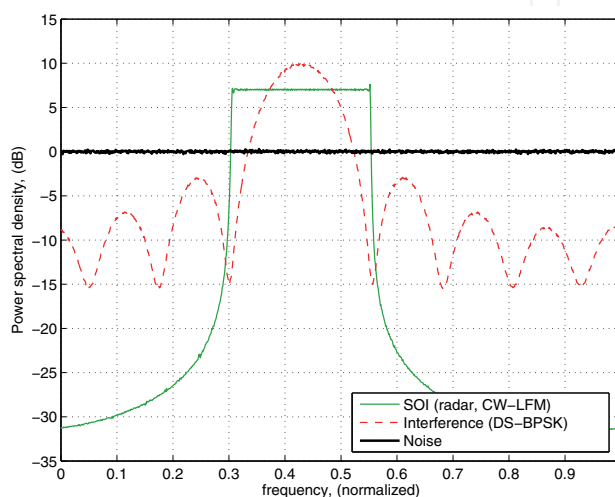

Fig. 18. PSD of noise, the CW-LFM signal (SOI) and the DS-BPSK interference (interference). SNR and INR fixed at 0 dB.

Linear Frequency Modulated (CW-LFM) radar signal which is being transmitted by a hostile equipment and, therefore, unknown. This signal is intentionally transmitted in the spectral band of a known Direct-Sequence Binary Phase Shift Keying (DS-BPSK) spread-spectrum communication signal, with the aim of hindering its detection. The SOI sweeps a total bandwidth of $BW = 20$ MHz with a sweeping time $T_p = 0.5$ ms, while the DS-BPSK interfering signal employs a chip rate $1/T_c = 10.23$ Mcps. The PSDs of the SOI, the interference and the noise are shown in Fig. 18.

The structure of the FRESH filter is designed based on a previous study of the performance of the sub-optimal FRESH filters. As a result, the interference rejection system incorporates an adaptive FRESH filter consisting of 5 branches, each one using a FIR filter of 1024 coefficients. The frequency shifts of the FRESH filter correspond to the 5 cycle frequencies of the DS-BPSK interference with the highest spectral correlation level, which are (Gardner et al., 1987): $\{\pm 1/T_c\}$ for the input $x(t)$, and $\{2f_c, 2f_c \pm 1/T_c\}$ for the complex conjugate of the input $x^*(t)$; where f_c is the carrier frequency and T_c is the chip duration of the DS-BPSK interference. The adaptive algorithm used is Fast-Block Least Mean Squares (FB-LMS) (Haykin, 2001), with a convergence factor $\mu = 1$.

Next, we present some simulation results on the interception system performance obtained after the training interval has finished. Firstly, the improvement in the SIR obtained at the output of the interference rejection system is represented in Fig. 19, as a function of the input
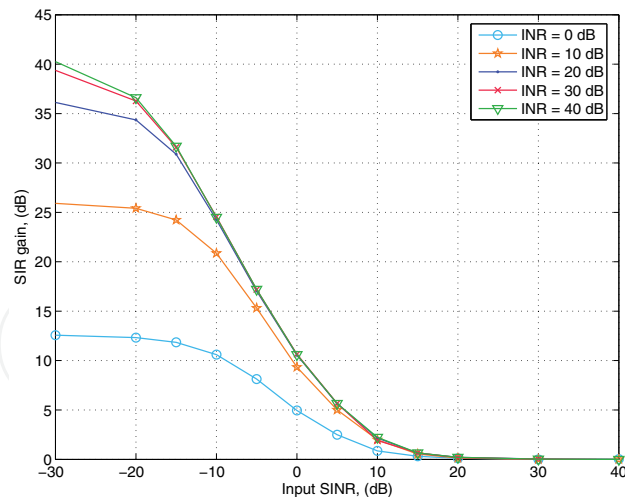
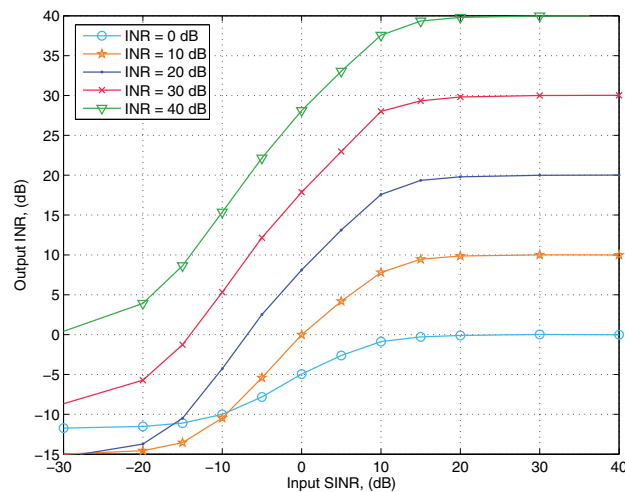Fig. 19. Simulated SIR improvement as a function of the input SINR and INR.

Fig. 20. Simulated INR at the output of the interference rejection system.

Signal-to-Interference-plus-Noise Ratio (SINR), for several values of Interference-to-Noise Ratio (INR). Two main results are revealed in Fig. 19:

1. The SIR improvement tends to zero as the SINR increases, because the SOI is powerful enough to mask the interference which makes the FRESH filter fail to estimate the interference. This is corroborated through Fig. 20, where the simulated INR at the output of the interference rejection system is shown. For high enough input SINR, the ouput INR matches the input INR, indicating that the FRESH filter cannot extract any of the interference power. This allows to define a "useful region", where the interference rejection system obtains a significant SIR improvement. In our example, the useful region comprises an input SINR$\leq$ 0 dB.

2. The SIR improvement saturates for high input INR values, which is shown by the fact that the SIR improvement for INR= 40 dB matches the curve obtained for INR= 30 dB. This is a limitation is due to the adaptive algorithm and does not appear when the optimal FRESH filter is used instead.

3. In addition, although it seems logical that the output INR increases with the input one, Fig. 20 reveals that this is not true for low input SINR and INR. The reason is that the
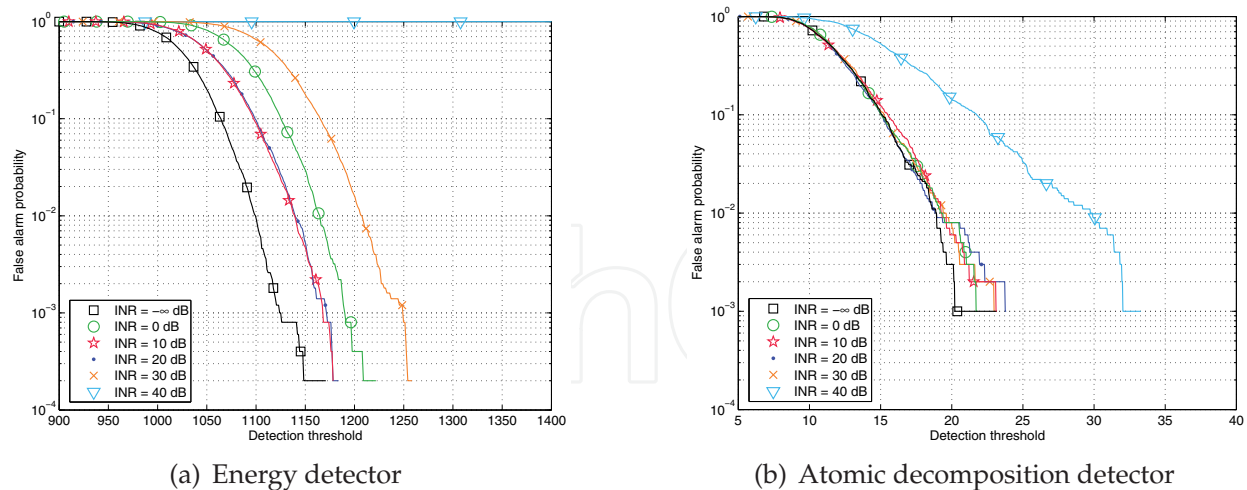
(a) Energy detector   (b) Atomic decomposition detector

Fig. 21. Degradation of the P_{FA} when the interference is present (with the interference rejection system).

> interference becomes masked by noise if the input INR is low enough (i.e. INR= 0 dB). As the input INR increases, the interference rejection system increases its effectiveness, so that the output INR decreases. That is why the output INR is lower for INR= 10 dB and INR= 20 dB than for INR= 0 dB.

The output INR can provide an idea about the degradation of the probability of false alarm (P_{FA}) of the detector (the probability of detection when the SOI is not present). However, each particular detector is affected in a different way. We shall illustrate this fact with two different detectors. The first one is an energy detector (ED), which consists of comparing the total energy to a detection threshold set to attain a desired P_{FA}. The second one is a detector based on atomic decomposition (AD), such as that proposed in (López Risueño et al., 2003). This detector exhibits an excellent performance for LFM signals, as the SOI considered in our case study. The AD-based detector can be thought as a bank of correlators or matched filters, each one matched to a chirplet (a signal with Gaussian envelope and LFM), whose maximum output is compared to a detection threshold. Both detectors process the signal by blocks, taking a decision each 1024 samples.

Fig. 21 shows the degraded P_{FA} of the whole interception system in the presence of the interference, and when the detection threshold has been determined for an input consisting of only noise. The curves clearly show the different dependence of the P_{FA} of both detectors on the input INR. The energy detector exhibits a higher sensitivity to the interference than the AD-based one. Thus, the AD-based detector visibly degrades its P_{FA} only for an input INR= 40 dB and above. On the contrary, ED always exhibits a degraded P_{FA} in the presence of the interference due to the energy excess, which is proportional to the output INR shown in Fig. 20.

Finally, we end this application example by showing the sensitivity improvement of the interception system obtained thanks to the interference rejection system. The sensitivity is defined as the SNR at the input of the interception system required to attain an objective probability of detection ($P_D = 90\%$), for a given probability of false alarm ($P_{FA} = 10^{-6}$). Thus, the detection threshold takes a different value depending on the input INR so that the P_{FA} holds for all the INR values. The simulation results are gathered in Tab. 2, with all values expressed in dB. At each INR value, the sensitivities of both detectors, AD and ED, with and

	With interf. rej. system		Without interf. rej. system		Sensitivity improvement	
INR	AD	ED	AD	ED	AD	ED
$-\infty$	-12.1	-6.1	-12.1	-6.1	0.0	0.0
0	-9.6	-4.3	-4.0	-3.5	5.6	0.8
10	-9.0	-4.3	5.8	1.1	14.8	5.4
20	-9.2	-4.3	15.9	6.8	27.1	11.1
30	-8.3	-3.4	26.0	14.3	34.3	17.7
40	-4.1	-0.5	35.8	21.6	39.9	22.1

Table 2. Sensitivity (SNR, dB) for the CW-LFM signal as a function of the INR. $P_{FA} = 10^{-6}$, $P_D = 90\%$.

without the interference rejection system, are shown. The sensitivity improvement obtained by using the interference rejection system is also shown.

As can be seen, the improvement is very significant and proves the benefit of using the interference rejection system. Moreover, the improvement is higher for increasing input INR. However, there is still a sensitivity degradation as the INR increases due to an increase in the detection threshold and/or a distortion produced by the interference rejection system to the SOI because of the signal leakage at the FRESH output (the latter only applies to AD, since ED is insensitive to the signal waveform). And, as expected, the AD-based detector outperforms ED (López Risueño et al., 2003).

8. Summary

This chapter has described the theory of adaptive FRESH filtering. FRESH filters represent a comprehensible implementation of LAPTV filters, which are the optimum filters for estimating or extracting signal information when the signals are modelled as almost-cyclostationary stochastic processes. When dealing with complex signals, both the signal and its complex conjugate must be filtered, resulting in WLAPTV filters. The knowledge required for the design optimal FRESH filters is rarely available beforehand in practice, which leads to the incorporation of adaptive scheme. Since FRESH filters consist of a set of LTI filters, classical algorithms can be applied by simply use the stationarized versions of the inputs of these LTI filters, which are obtained by time-averaging their statistics. Then, the optimal set of LTI filters is given by the multidimensional Wiener filter theory. In addition, thanks to their properties of signal separation in the cycle frequency domain, adaptive FRESH filters can operate blindly, that is without reference of the desired signal, by simply using as frequency shifts the cycle frequencies belonging uniquely to the desired signal cyclic spectrum. Furthermore, adaptive FRESH filters have the advantage of being able to compensate small errors in their frequency shifts, which can be present in practice due to non-ideal effects such as Doppler or the oscillators stability. In this case, the convergence rate of the adaptive algorithm must be carefully chosen in order to simultaneously minimize the gradient noise and the lag errors. The chapter is finally closed by presenting an application example, in which an adaptive FRESH filter is used to suppress known interferences for an unknown hidden signal interception system, demonstrating the potentials of adaptive FRESH filters in this field of application.

9. References

Adlard, J., Tozer, T. & Burr, A. (1999). Interference rejection in impulsive noise for VLF communications, *IEEE Military Communications Conference Proceedings, 1999. MILCOM 1999.*, pp. 296–300.

Agee, B. G., Schell, S. V. & Gardner, W. A. (1990). Spectral self-coherence restoral: A new approach to blind adaptive signal extraction using antenna arrays, *Proceedings of the IEEE* 78(4): 753–767.

Brown, W. A. (1987). *On the Theory of Cyclostationary Signals*, Ph.D. dissertation.

Chen, Y. & Liang, T. (2010). Application study of BA-FRESH filtering technique for communication anti-jamming, *IEEE 10th International Conference on Signal Processing (ICSP)*, pp. 287–290.

Chevalier, P. & Blin, A. (2007). Widely linear MVDR beamformers for the reception of an unknown signal corrupted by noncircular interferences, *IEEE Transactions on Signal Processing* 55(11): 5323–5336.

Chevalier, P. & Maurice, A. (1997). Constrained beamforming for cyclostationary signals, *International Conference on Acoustics, Speech, and Signal Processing, ICASSP-97*.

Chevalier, P. & Pipon, F. (2006). New insights into optimal widely linear array receivers for the demodulation of BPSK, MSK, and GMSK signals corrupted by noncircular interferences - Application to SAIC, *IEEE Transactions on Signal Processing* 54(3): 870–883.

Corduneanu, C. (1968). *Almost Periodic Functions*, Interscience Publishers.

Franks, L. E. (1994). Polyperiodic linear filtering, *in* W. A. Gardner (ed.), *Cyclostationarity in Communicactions and Signal Processing*, IEEE Press.

Gameiro, A. (2000). Capacity enhancement of DS-CDMA synchronous channels by frequency-shift filtering, *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*.

Gardner, W. A. (1978). Stationarizable random processes, *IEEE Transactions on Information Theory* 24(1): 8–22.

Gardner, W. A. (1986). *Introduction to Random Processes with Applications to Signals and Systems*, Macmillan Publishing Company.

Gardner, W. A. (1987). Spectral correlation of modulated signals: Part I – analog modulation, *IEEE Transactions on Communications* 35(6): 584–594.

Gardner, W. A. (1991). Exploitation of spectral redundancy in cyclostationary signals, *IEEE Signal Processing Magazine* 8(2): 14–36.

Gardner, W. A. (1993). Cyclic Wiener filtering: Theory and method, *IEEE Transactions on Communications* 41(1): 151–163.

Gardner, W. A. (1994). *Cyclostationarity in Communications and Signal Processing*, IEEE Press.

Gardner, W. A., Brown, W. A. & Chen, C. K. (1987). Spectral correlation of modulated signals: Part II – digital modulation, *IEEE Transactions on Communications* 35(6): 595–601.

Gardner, W. A. & Franks, L. E. (1975). Characterization of cyclostationary random signal processes, *IEEE Transactions on Information Theory* 21(1): 4–14.

Gelli, G. & Verde, F. (2000). Blind LPTV joint equalization and interference suppression, *Acoustics, Speech, and Signal Processing, 2000. ICASSP '00. Proceedings. 2000 IEEE International Conference on*.

Giannakis, G. B. (1998). Cyclostationary signal analysis, *in* V. K. Madisetti & D. Williams (eds), *The Digital Signal Processing Handbook*, CRC Press.

Gonçalves, L. & Gameiro, A. (2002). Frequency shift based multiple access interference canceller for multirate UMTS-TDD systems, *The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*.

Haykin, S. (2001). *Adaptive Filter Theory*, Prentice Hall.

Hu, Y., Xia, W. & Shen, L. (2007). Study of anti-interference mechanism of multiple WPANs accesing into a HAN, *International Symposium on Intelligent Signal Processing and Communication Systems*.

Jianhui, P., Zhongfu, Y. & Xu, X. (2006). A novel robust cyclostationary beamformer based on conjugate gradient algorithm, *2006 International Conference on Communications, Circuits and Systems Proceedings*, Vol. 2, pp. 777–780.

Lee, J. H. & Lee, Y. T. (1999). Robust adaptive array beamforming for cyclostationary signals under cycle frequency error, *IEEE Transactions on Antennas and Propagation* 47(2): 233–241.

Lee, J. H., Lee, Y. T. & Shih, W. H. (2000). Efficient robust adaptive beamforming for cyclostationary signals, *IEEE Transactions on Signal Processing* 48(7): 1893–1901.

Li, X. & Ouyang, S. (2009). One reduced-rank blind fresh filter for spectral overlapping interference signal extraction and DSP implementation, *International Workshop on Intelligent Systems and Applications, ISA*, pp. 1–4.

Loeffler, C. M. & Burrus, C. S. (1978). Optimal design of periodically time-varying and multirate digital filters, *IEEE Transactions on Acoustic, Speech and Signal Processing* 66(1): 51–83.

López Risueño, G., Grajal, J. & Yeste-Ojeda, O. A. (2003). Atomic decomposition-based radar complex signal interception, *IEE Proceedings on Radar, Sonar and Navigation* 150: 323–331.

Martin, V., Chabert, M. & Lacaze, B. (2007). Digital watermarking of natural images based on LPTV filters, *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*.

Mirbagheri, A., Plataniotis, K. & Pasupathy, S. (2006). An enhanced widely linear CDMA receiver with OQPSK modulation, *IEEE Transactions on Communications* 54(2): 261–272.

Napolitano, A. & Spooner, C. M. (2001). Cyclic spectral analysis of continuous-phase modulated signals, *IEEE Transactions on Signal Processing* 49(1): 30–44.

Ngan, L. Y., Ouyang, S. & Ching, P. C. (2004). Reduced-rank blind adaptive frequency-shift filtering for signal extraction, *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, Vol. 2.

Petrus, P. & Reed, J. H. (1995). Time dependent adaptive arrays, *IEEE Signal Processing Letters* 2(12): 219–222.

Picinbono, B. & Chevalier, P. (1995). Widely linear estimation with complex data, *IEEE Transactions on Signal Processing* 43(8): 2030–2033.

Reed, J. H. & Hsia, T. C. (1990). The performance of time-dependent adaptive filters for interference rejection, *IEEE Transactions on Acoustic, Speech and Signal Processing* 38(8): 1373–1385.

Schell, S. V. & Gardner, W. A. (1990). Progress on signal-selective direction finding, *Fifth ASSP Workshop on Spectrum Estimation and Modeling*, pp. 144–148.

Schmidt, R. O. (1986). Multiple emitter location and signal parameter estimation, *IEEE Transactions on Antennas and Propagation* 34(3): 276–280.

Van Trees, H. L. (1968). *Detection, Estimation, and Modulation Theory*, Vol. 1, John Wiley and Sons.

Whitehead, J. & Takawira, F. (2004). Low complexity constant modulus based cyclic blind adaptive multiuser detection, *AFRICON, 2004. 7th AFRICON Conference in Africa*, Vol. 1, pp. 115–120.

Whitehead, J. & Takawira, F. (2005). Blind adaptive multiuser detection for periodically time varying interference suppression [DS-CDMA system applications], *IEEE Wireless Communications and Networking Conference, 2005*, Vol. 1, pp. 273–279.

Widrow, B., Glover, Jr., J. R., McCool, J. M., Kaunitz, J., Williams, C. S., Hearn, R. H., Zeidler, J. R., Dong, Jr., E. & Goodlin, R. C. (1975). Adaptive noise cancelling: Principles and applications, *Proceedings of the IEEE* 63(12): 1692–1717.

Widrow, B., McCool, J. M., Larimore, M. G. & Johnson, Jr., C. R. (1976). Stationary and nonstationary learning characteristics of the LMS adaptive filter, *Proceedings of the IEEE* 64(8): 1151–1162.

Wong, H. E. & Chambers, J. A. (1996). Two-stage interference immune blind equaliser which exploits cyclostationary statistics, *Electronic Letters* 32(19): 1763–1764.

Yeste-Ojeda, O. A. & Grajal, J. (2008). Cyclostationarity-based signal separation in interceptors based on a single sensor, *IEEE Radar Conference 2008*, pp. 1–6.

Yeste-Ojeda, O. A. & Grajal, J. (2010). Adaptive-FRESH filters for compensation of cycle-frequency errors, *IEEE Transactions on Signal Processing* 58(1): 1–10.

Zadeh, L. A. (1950). Frequency analysis of variable networks, *Proceedings of the I.R.E.* pp. 291–299.

Zhang, H., Abdi, A. & Haimovich, A. (2006). Reduced-rank multi-antenna cyclic Wiener filtering for interference cancellation, *Military Communications Conference, 2006. MILCOM 2006*.

Zhang, J., Liao, G. & Wang, J. (2004). A novel robust adaptive beamforming for cyclostationary signals, *The 7th International Conference on Signal Processing, ICSP*, Vol. 1, pp. 339–342.

Zhang, J., Wong, K. M., Luo, Z. Q. & Ching, P. C. (1999). Blind adaptive FRESH filtering for signal extraction, *IEEE Transactions on Signal Processing* 47(5): 1397–1402.

Adaptive Filtering

Edited by Dr Lino Garcia

Adaptive filtering is useful in any application where the signals or the modeled system vary over time. The configuration of the system and, in particular, the position where the adaptive processor is placed generate different areas or application fields such as prediction, system identification and modeling, equalization, cancellation of interference, etc., which are very important in many disciplines such as control systems, communications, signal processing, acoustics, voice, sound and image, etc. The book consists of noise and echo cancellation, medical applications, communications systems and others hardly joined by their heterogeneity. Each application is a case study with rigor that shows weakness/strength of the method used, assesses its suitability and suggests new forms and areas of use. The problems are becoming increasingly complex and applications must be adapted to solve them. The adaptive filters have proven to be useful in these environments of multiple input/output, variant-time behaviors, and long and complex transfer functions effectively, but fundamentally they still have to evolve. This book is a demonstration of this and a small illustration of everything that is to come.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Omar A. Yeste Ojeda and Jesús Grajal (2011). Adaptive-FRESH Filtering, Adaptive Filtering, Dr Lino Garcia (Ed.), ISBN: 978-953-307-158-9, InTech, Available from: http://www.intechopen.com/books/adaptive-filtering/adaptive-fresh-filtering1