Article

Channel Identification Based on Cumulants, Binary Measurements, and Kernels

Hicham Oualla [1,2], Rachid Fateh [1,*], Anouar Darif [1], Said Safi [1], Mathieu Pouliquen [2] and Miloud Frikel [2]

[1] Department of Mathematics and Informatics, Sultan Moulay Slimane University, Po. Box 592, 23000 Beni Mellal, Morocco; hicham.oualla@unicaen.fr (H.O.); anouar.darif@usms.ac.ma (A.D.); said.safi@usms.ac.ma (S.S.)

[2] Laboratoire d'Automatique de Caen, UNICAEN, ENSICAEN, Normandie University, 6, B. Marchal Juin, 14050 Caen, France; mathieu.pouliquen@unicaen.fr (M.P.); miloud.frikel@ensicaen.fr (M.F.)

* Correspondence: rachid.fateh@usms.ac.ma

Abstract: In this paper, we discuss the problem of channel identification by using eight algorithms. The first three algorithms are based on higher-order cumulants, the next three algorithms are based on binary output measurement, and the last two algorithms are based on reproducing kernels. The principal objective of this paper is to study the performance of the presented algorithms in different situations, such as with different sizes of the data input or different signal-to-noise ratios. The presented algorithms are applied to the estimation of the channel parameters of the broadband radio access network (BRAN). The simulation results confirm that the presented algorithms are able to estimate the channel parameters with different accuracies, and each algorithm has its advantages and disadvantages for a given situation, such as for a given SNR and data input. Finally, this study provides an idea of which algorithms can be selected in a given situation. The study presented in this paper demonstrates that the cumulant-based algorithms are more adequate if the data inputs are not available (blind identification), but the kernel- and binary-measurement-based methods are more adequate if the noise is not important ($SNR \geq 16$ dB).

1. Introduction

The identification of single-input single-output (SISO) systems in the blind case (without knowledge of the input system data) has been well studied by using higher-order statistics [1–5], without employing any restrictive hypotheses about channel zeros, channel-order overestimation errors, or additive noise color, as well as without increasing the data stream transmission rate. The shared characteristic of this category of techniques [4] is that they use a periodically time-varying precoder to provoke cyclo-stationary statistics at the transmitter (i.e., a simple periodic modulator or a filter bank) and take advantage of the received samples' second-order cyclo-stationary statistics. Those algorithms have been called transmitter-induced multi-stationary-based techniques. Iterative strategies are used in some types of algorithms for blind channel identification. In these algorithms, an initial channel estimation is employed by using a symbol predictor to obtain provisional soft estimates of the symbol sequence as it is transmitted. A channel estimator uses these estimates to optimize the channel parameters. The symbol estimator then uses the improved channel estimates to optimize the symbol parameters, and so on. The second technique that will be discussed here deals with the identification of the system based on quantized measurement and is used in a wide range of applications, such as microfabricated devices [6–8].

The channels under investigation have a non-minimal phase and a selective frequency (i.e., normalized channels for MC-CDMA: BRAN). In the majority of wireless areas, there

are numerous hitches in the channels, such as mountains, buildings, and barriers between the transmitter and the receiver. Many different propagation paths are created by reflections from these obstacles. Multi-path propagation, or a multi-path channel, is the term for this. Due to the reverberations and reflection between the transmitter and receiver, the frequency impulse response of this channel is not leveled (ideal case), but contains some cavities and bumps. The synchronization required between the transmitter and the receiver is also a problem in the communication process. To handle this phase estimation problem, some authors have employed higher-order cumulants (HOCs) in order to investigate the reliability of these techniques if the channel is influenced by a colored noise [2,9–16]. HOCs are a significant topic with multiple applications in the field of systems theory, such as in digital signal recognition [17], automatic modulation classification of real signals [18], and space-time block code identification [19].

HOCs are only used for non-linear and non-Gaussian processes because the cumulants of a Gaussian process are identical to zero [2,3,20]. There are many real-world applications that are really non-Gaussian [1–3,21]. In addition, the Fourier transforms of HOCs, which are called higher-order spectra (or polyspectra), provide an excellent technique for resolving the problem of equalization technology encountered in communication. The main attribute of HOCs in terms of equalization is that the channel phase information is available [4,5,22,23]. As a result, HOCs can be used to estimate the channel model parameters, even if the phase property (non-minimum phase (NMP) or minimum phase (MP)) of the channel or of the transmitted data is unknown (under the assumption of a non-Gaussian distribution) [1–3,24].

A technique founded on binary observation has been employed in a wide range of applications in industries and telecommunication, such as in chemical process sensors for a vacuum, the asynchronous transfer mode (ATM), pressure switches, liquid levels, and industrial sensors for brushless DC motors, shifting using a wire in automotive applications, and switching sensors for exhaust-gas oxygen ABS [7,25–27]. One of the most important problems in this wide range of applications is parameter estimation for systems with binary (or quantized) outputs [7,8,28–32]. Some examples of this field include sensor networks, telecommunications, and networked control systems. The third type of technique discussed here in terms of system identification is based on kernel adaptive filtering, which represents an emerging technique for nonlinear signal processing [33], classification, and machine learning (regularization networks [34], support vector machines (SVMs) [35]); the most emblematic attribute is the use of kernel functions, thus permitting the inexpensive operation of the algorithms in high-dimensional spaces.

The kernel methods are based on a central principle called the kernel trick, which gives a non-linear characteristic to many traditionally linear methods, as well as on the theory of reproducing kernel Hilbert spaces (RKHSs); such spaces provide an important framework for elegantly developing adaptive nonlinear filters. So far, several algorithms, including the kernel least mean square (KLMS) [36], kernel affine projection algorithm (KAPA) [37], kernel principal component analysis (KPCA) [38], and kernel recursive least square (KRLS) [39], have been proposed and effectively applied to nonlinear regression problems and classification. In addition, to increase the efficiency of kernel adaptive filtering techniques, some subtypes of these algorithms have also been introduced [40–43] for channel identification [44] and nonlinear system equalization.

In this paper, we focused on a comparison of the results of channel identification using some algorithms based on HOCs, binary measurement, or kernel-positive methods. We selected three algorithms that use the first two techniques (cumulants and binary measurement) and two algorithms that use the last technique (kernel methods). In order to study the efficiency of these algorithms, we used a practical—i.e., measurable—frequency-selective fading channel called the broadband radio access network (BRAN) (in particular, BRAN A, which is for transmission in indoor scenarios). This model radio channel was normalized by the European Telecommunications Standards Institute (ETSI) in [45,46]. Simulation results were obtained for different values of SNR and different data inputs.

This paper is arranged as follows. In Section 2, we discuss the identification problem. In Section 3, we present three algorithms for channel identification based on higher-order cumulants. In Section 4, we describe the presentation of the binary output measurement algorithms. In Section 5, we illustrate the concept of kernel methods for channel identification and the description of the KRLS and KLMS algorithms. Some simulations for evaluating the performance of the presented algorithms are shown in Section 6 and, finally, Section 7 concludes the paper.

2. Problem Statement

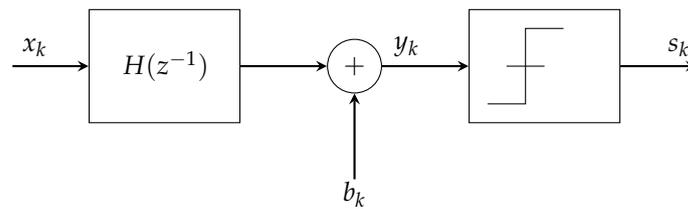Let us consider the finite impulse response (FIR) system described in Figure 1.

Figure 1. Block diagram of the system model.

Where, x_k and s_k are, respectively, the input and the output, y_k is an internal signal (unknown for algorithms based on binary output measurement), and b_k is an additive noise in this internal signal. We suppose that H has a finite impulse response of length n, where:

$$y_k = \sum_{i=1}^{n} h_i x_{k-i} + b_k \tag{1}$$

y_k can be rewritten as:

$$y_k = \theta^\top \phi_k + b_k \tag{2}$$

or

$$y_k = \phi_k^\top \theta + b_k \tag{3}$$

where $\phi_k = \begin{pmatrix} x_{k-1} \\ x_{k-2} \\ \vdots \\ x_{k-n} \end{pmatrix} \in \mathbb{R}^n$ and $\theta = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} \in \mathbb{R}^n$

The binary-valued output, which has a fixed threshold $C \in \mathbb{R}$, can be represented by:

$$s_k = I(y_k) = \begin{cases} 1, & \text{if } y_k \geq C \\ -1, & \text{otherwise.} \end{cases} \tag{4}$$

3. Channel Identification Based on Cumulants

In this section, three algorithms based on higher-order cumulants (HOCs) are presented. These algorithms can blindly identify the BRAN channel.

3.1. First Algorithm: Alg1

Safi and Zeroual (2004) demonstrated that they could identify the FIR parameters blindly based only on fourth-order cumulants.

In this paper, we briefly present this algorithm. The mth-order cumulants are [9]:

$$C_{my}(t_1, \ldots, t_{m-1}) = \gamma_{mx} \sum_{i=0}^{q} h(i)h(i+t_1) \ldots h(i+t_{m-1}) \tag{5}$$

where q is the FIR system order and γ_{mx} is the mth-order cumulant—at the origin—of the input FIR system. Based on Equation (5), they demonstrated that the cumulants of order m and n of the FIR output, $\{y(n)\}$, and the the impulse response coefficients, $\{h(i)\}$, with $n > m$, are linked by the following relationship:

$$\sum_{j=0}^{q} h(j) C_{ny}(j + t_1, \ldots, j + t_{m-1}, t_m, \ldots, t_{n-1}) \tag{6}$$

$$= \frac{\gamma_{nx}}{\gamma_{mx}} \sum_{i=0}^{q} h(i) \left[\prod_{k=m}^{n-1} h(i + t_k) \right] C_{my}(i + t_1, \ldots, i + t_{m-1})$$

Based on Equations (5) and (6), Safi and Zeroual (2004) developed an algorithm (Alg1) based on fourth-order cumulants, which are given as follows:

$$h(0)h^2(q)C_{4y}(0, 0, i + \alpha) = \sum_{j=0}^{q} h^3(j) C_{4y}(q, q, j + \alpha) \tag{7}$$

Based on the hypothesis that the system is causal of order q, $j + \alpha$ is in the interval $[0, q]$. For this, we have: $0 \le j + \alpha \le q \Rightarrow -j \le \alpha \le q - j$, with $0 \le j \le q$. From this, two we obtain two inequalities:

$$-q \le \alpha \le q \tag{8}$$

From Equations (7) and (8), they obtained the following system of equations:

$$\begin{pmatrix} 0 & \cdots & 0 & C_{4y}(q,q,0) \\ \vdots & \ddots & & \vdots \\ 0 & & & \\ C_{4y}(q,q,0) & \cdots & & C_{4y}(q,q,q) \\ \vdots & \ddots & & 0 \\ & & & \vdots \\ C_{4y}(q,q,q) & 0 & \cdots & 0 \end{pmatrix} \times \begin{pmatrix} \frac{1}{h^2(q)} \\ \vdots \\ \frac{h^3(i)}{h^2(q)} \\ \vdots \\ \frac{h^3(q)}{h^2(q)} \end{pmatrix} = \begin{pmatrix} C_{4y}(0,0,-q) \\ \vdots \\ C_{4y}(0,0,0) \\ \vdots \\ C_{4y}(0,0,q) \end{pmatrix} \tag{9}$$

In a compact form, the system of Equations (9) can be written as:

$$Mb_q = d \tag{10}$$

where M, b_q, and d are defined in the system of Equations (9). The least squares (*LS*) solution, which is found by using the least squares of the system in Equation (10), is given by

$$\hat{b}_q = (M^\top M)^{-1} M^\top d \tag{11}$$

This solution gives us an estimation of the quotient of the parameters $h^3(i)$ and $h^3(q)$, i.e., $b_q(i) = \left(\widehat{\frac{h^3(i)}{h^3(q)}} \right)$, $i = 1, \ldots, q$. So, in order to estimate the parameters $\hat{h}(i)$, $i = 1, \ldots, q$, they proceeded as follows:

- The parameters $h(i)$ for $i = 1, \ldots, q - 1$ are estimated from the values of $\hat{b}_q(i)$ using the following solution:

$$\hat{h}(i) = sign\left[\hat{b}_q(i)(\hat{b}_q(q))^2 \right] \left\{ abs(\hat{b}_q(i))(\hat{b}_q(q))^2 \right\}^{1/3} \tag{12}$$

with $sign(x) = \begin{cases} 1, & \text{if} \quad x > 0; \\ 0, & \text{if} \quad x = 0; \\ -1, & \text{if} \quad x < 0; \end{cases}$

and $abs(x) = |x|$ indicates the absolute value of x.

- The parameters $\hat{h}(q)$ are estimated by:

$$\hat{h}(q) = \frac{1}{2} sign\left[\hat{b}_q(q)\right] \left\{ abs(\hat{b}_q(q)) + (\frac{1}{\hat{b}_q(1)})^{1/2} \right\} \tag{13}$$

3.2. Second Algorithm: Alg2

Based on Equation (5), the authors of [3] demonstrated the following equation:

$$\sum_{j=0}^{q} h(j) C_{4y}(t_1 - j, t_2 - j, t_3 - j)$$

$$= \frac{\gamma_{4x}}{\gamma_{2x}} \epsilon \sum_{i=0}^{q} h(i) h(i + t_2 - t_1) h(t_3 - t_1) C_{2y}(t_1 - i) \tag{14}$$

where $\epsilon = \left\{ \frac{\gamma_{4x}}{\gamma_{2x}} \right\}$. If $t_3 = t_1$, the Equation (14) becomes (with $h(0) = 1$):

$$\sum_{j=0}^{q} h(j) C_{4y}(t_1 - j, t_2 - j, t_1 - j) = \epsilon \sum_{i=0}^{q} h(i) h(i + t_2 - t_1) C_{2y}(t_1 - i) \tag{15}$$

By applying some algebra to Equations (13)–(15), they obtained the following equation:

$$\begin{pmatrix} C_{4y}(2q, 2q, q) & \cdots & C_{4y}(q, q, 0) \\ \vdots & \ddots & \vdots \\ C_{4y}(2q, 2q, 2q) & \cdots & C_{4y}(q, q, q) \end{pmatrix} \times \begin{pmatrix} h(0) \\ \vdots \\ h(q) \end{pmatrix} = \epsilon' \begin{pmatrix} h(0) \\ \vdots \\ h(q) \end{pmatrix}$$

where $\epsilon' = \dfrac{C_{4y}(q, q, q) C_{4y}(q, 0, 0)}{C_{4y}(q, q, 0)}$. Taking into account the properties of cumulants in FIR systems, i.e., $C_{my}(t_1, \ldots, t_k, \ldots, t_{m-1}) = 0$, then $t_k > q$ for all $k = 1, \ldots, m - 1$. Thus, the previous system of Equations (16) becomes:

$$\begin{pmatrix} C_{4y}(2q - 1, 2q - 1, q - 1) & \cdots & C_{4y}(q, q, 0) \\ C_{4y}(2q - 1, 2q - 1, q) - \epsilon' & \cdots & C_{4y}(q, q, 0) \\ 0 & \ddots & \vdots \\ \vdots & & \vdots \\ 0 & \cdots 0 & C_{4y}(q, q, q) - \epsilon' \end{pmatrix} \times \begin{pmatrix} h(1) \\ \vdots \\ h(q) \end{pmatrix} = \begin{pmatrix} \epsilon' - C_{4y}(2q, 2q, q) \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

In more compact form, this can be written as:

$$M h_q = d \tag{16}$$

The parameters $\{h(i)\}, i = 1, \ldots, q$, can be estimated using the least squares (*LS*) as follows:

$$h_q = (M^\top M)^{-1} M^\top d \tag{17}$$

3.3. Third Algorithm: Alg3

In this algorithm, the relationships between the third- and fourth-order cumulants and, eventually, the system parameters are established. Safi and Zeroual (2008) demonstrated the following relationship:

$$\sum_{i=0}^{q} \epsilon h(i) C_{4y}(t_1, t_2 - i, t_3 - i) = \sum_{i=0}^{q} h(i) h(i + t_3 - t_2) C_{3y}(t_1, t_2 - i) \tag{18}$$

where $\epsilon = \dfrac{\gamma_{3x}}{\gamma_{4x}}$ for different values of t_1, t_2, and $t_3 \in \mathbb{Z}$. Based on this equation, they developed some algorithms.

In order to identify the channel parameters, $\{h(i)\}, i = 1, \ldots, q$, they resolved the system of Equation (18) using the least squares (LS) method based on the following procedure.

In order to simplify the estimation of the channel parameters, they supposed that $\lambda = t_1$ and $t = t_2 = t_3$. Thus, for each λ, based on the LS method, they could estimate the channel parameters from the relationship:

$$\sum_{i=0}^{q} \epsilon h(i) C_{4y}(\lambda, t-i, t-i) = \sum_{i=0}^{q} h^2(i) C_{3y}(\lambda, t-i) \tag{19}$$

For each λ, they could estimate the parameters $h(i)$ using the following equation:

$$\hat{h}_{(\lambda)}(i) = \frac{1}{2}\left\{\frac{\widehat{\epsilon h}(i)}{\hat{\epsilon}} + sign(\frac{\widehat{\epsilon h}(i)}{\hat{\epsilon}})(\widehat{h^2}(i))^{1/2}\right\} \tag{20}$$

where $\lambda \in [-q, q]$.

4. Binary Output Measurement Algorithms

4.1. LIMBO Method

The LMS-based identification Method using Binary Observations (LIMBO) [30] is a recursive parameter estimation method that uses binary observations based on the LMS algorithm. The algorithm is based on the following assumption:

- A.1: $\theta_1^\top \theta > 0$.
- A.2: ϕ_k is a random process such that:

  - The probability density function (pdf) of $\frac{\phi_k}{\|\phi_k\|}$ is non-zero on the unit sphere.
  - ϕ_k verifies the α-mixing condition [47].

- A.3: b_k represents the noise, which is uncorrelated with the input sequence.

The binary output defined in Equation (4) is used in the LIMBO algorithm with threshold $C = 0$, and the unknown internal signal y_k is defined in Equation (3).

The estimated output is:

$$\widehat{s}_k = I(\phi_k^\top \widehat{\theta}_k) \tag{21}$$

where $\widehat{\theta}_k$ represents the estimation of θ at the time k. The LIMBO algorithm is based on the concept of minimizing the instantaneous error:

$$E_k = \frac{1}{2}|s_k - \widehat{s}_k|\widehat{y}_k \tag{22}$$

$\widehat{\theta}_{k+1}$ is given by:

$$\widehat{\theta}_{k+1} = \frac{\widehat{\theta}_k - \mu\alpha_k|s_k - \widehat{s}_k|\widehat{y}_k\phi_k}{\omega_k} \tag{23}$$

where

- $0 < \mu < 1$
- $\omega_k = \sqrt{1 - 2\mu(1-\mu)\alpha_k|s_k - \widehat{s}_k|\widehat{y}_k^2}$
- $\alpha_k = \frac{1}{\phi_k^\top \phi_k}$

4.2. Recursive Least Squares (RLS) Method

The recursive identification algorithm presented in [28] is based on recursive least squares with a forgetting factor (RLS-FF), and it uses the binary output defined in Equation (4) and the threshold $C = 0$. The algorithm is based on the following assumptions:

- B.1: $\|\theta\| = 1$, where $\|.\|$ is the l_2 norm.
- B.2: At any time k, $\phi_k^\top \theta \neq 0$.
- B.3: The noise sequence b_k is an i.i.d. sequence of random variables with a mean of zero and finite covariance, and it is uncorrelated with the input sequence.

The identification is realized such that:

$$|s_k - \phi_k^\top \theta| < \delta. \tag{24}$$

with $\delta < 1$.

The estimation of θ is given by:

$$\widehat{\theta}_k = \widehat{\theta}_{k-1} + \Gamma_k \varepsilon_{k/k-1} \tag{25}$$

with

$$
\begin{cases}
\Gamma_k = \frac{P_{k-1}\phi_k\sigma_k}{\lambda + \phi_k^\top P_{k-1}\phi_k\sigma_k} \\[2mm]
P_k = \frac{1}{\lambda}(I_{n+1} - \Gamma_k\phi_k^\top)P_{k-1} \\[2mm]
\varepsilon_{k/k-1} = s_k - \phi_k^\top \widehat{\theta}_{k-1}.
\end{cases}
\tag{26}
$$

where

$$
\sigma_k = \begin{cases}
\frac{\lambda}{\phi_k^\top P_{k-1}\phi_k}(|\frac{\varepsilon_{k/k-1}}{\delta}| - 1), & \text{if } (|\varepsilon_{k/k-1}| > \delta) \text{ and } (\phi_k^\top P_{k-1}\phi_k > 0) \\
0, & \text{otherwise.}
\end{cases}
\tag{27}
$$

Finally, $0 < \lambda < 1$ is the forgetting factor.

4.3. Method Based on SVM

The method presented in [29] is based on the use of a supervised learning algorithm—a support vector machine (SVM)—through the formulation of the identification problem as a classification problem.

We assume that:

- C.1: The static gain of the system is known.
- C.2: y_k is such that at any time k, $y_k \neq C$.

The algorithm has two steps:

Step 1: Find C and the solution θ of

$$
\begin{cases}
\min_{\theta, C, e} \frac{1}{2}\|\theta\|^2 + \gamma \sum_{k=1}^{N} e_k, \\
\text{s.t } s_k(\theta^\top \phi - C) \succeq 1 - e_k, \\
\text{and } e_k \succeq 0; k \in [1, N].
\end{cases}
\tag{28}
$$

Step 2: Normalization of the static gain according to assumption C.1, where $\gamma \in \mathbb{R}_0^+$ is referred to as the regularization constant or penalty parameter.

5. Kernel-Based Channel Identification

The identification of nonlinear systems using functional estimates based on the notion of Hilbert spaces has led to very interesting results in the last decade. Note that linear problems with a finite order can be taken into consideration to address nonlinear adaptive filtering problems in which the measured input $\mathcal{X}$ has been mapped to a nonlinear Hilbert space $\mathcal{H}$ of infinite dimensions with a Mercer kernel [48]:

$$\kappa(x_i, x_j) = \langle \Psi(x_i), \Psi(x_j) \rangle_{\mathcal{H}}, \quad \forall(x_i, x_j) \in \mathcal{X}^2 \tag{29}$$

where Ψ maps $\mathcal{X}$ into a higher-dimensional space $\mathcal{H}$ with an inner product $\langle ., . \rangle_{\mathcal{H}}$.

The sequence of samples is transformed by means of a feature map Ψ:

$$
\begin{aligned}
\Psi \ : \ \mathcal{X} &\longrightarrow \mathcal{H} \\
x &\longrightarrow \kappa(x, .)
\end{aligned}
\tag{30}
$$

To construct the reproducing kernel Hilbert space model, we will use the Gaussian kernel, which is typically a default selection due to its universal approximation property:

$$\kappa(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right), \quad \forall x_i, x_j \in \mathcal{X}. \tag{31}$$

where $\sigma > 0$ denotes the kernel bandwidth.

In order to represent the elements of $\mathcal{H}$ by using their coordinates, we should give a Hilbert space $\mathcal{H}$ an orthonormal base. The associated kernel must be a continuous, symmetric, normalized, and positive definite function $\kappa : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, where the subset $\mathcal{X} \subset \mathbb{R}^N$ is compact.

5.1. Kernel Recursive Least Squares Algorithm

In this subsection, we present the kernel RLS algorithm [39,49–51]. It is an algorithm that belongs to the category of recursive least squares algorithms. The basic idea is to run the linear RLS algorithm in the kernel feature space, which is associated with a positive definite kernel κ, via the feature map $\Psi(.)$, which is defined in Equation (30), to generate the input data:

$$\{(\Psi(x_1), s_1), (\Psi(x_2), s_2), ... (\Psi(x_n), s_n), ...\}. \tag{32}$$

The optimization problem of the kernel RLS algorithm can be expressed as:

$$\min_{\Psi \in \mathcal{H}} \sum_{k=0}^{n-1} \lambda^{n-k} |s_k - \Psi(x_k)|^2 + \lambda^k \|\Psi\|_{\mathcal{H}}^2 \tag{33}$$

where λ is a positive regularization parameter, s_k is the desired response of the system, and $\Psi(x_k)$ designates the corresponding output of the system for x_k.

According to the representation theorem [52], the solution to the optimization problem (33) can be written as:

$$\Psi(.)_n = \sum_{i=1}^{m} \theta_{n,k} \kappa(., x_k) \tag{34}$$

where $\theta_n = (\theta_{n,1}, \theta_{n,2}, ..., \theta_{n,n})^\top$ is the model coefficient vector. By introducing Equation (34) into Equation (33), the optimal solution to the problem yields the following:

$$\min_{\theta} \mathbf{\Lambda} \|s_k - \mathbf{H_k}\theta\|^2 + \lambda^n \theta^\top \mathbf{K_k}\theta \tag{35}$$

where $(\mathbf{H_k})_{k,i} = \kappa(x_k, x_i)$, for $k, i = 1, 2, ..., n$, $\mathbf{K_k} = (\kappa(x_1, x_n), \kappa(x_2, x_n), ..., \kappa(x_m, x_n))^\top$, and $\mathbf{\Lambda}$ is a diagonal matrix of size (n, n), of which the (k, k)th element is λ^{n-k}. If we consider that $\mathbf{H_k}$ is invertible, the solution of the problem (35) becomes:

$$\theta_n = \mathbf{P_n}\mathbf{H_k}^\top \mathbf{\Lambda_n} s_n \tag{36}$$

with $\mathbf{P_n} = (\mathbf{H_k}^\top \mathbf{\Lambda_n}\mathbf{H_k} + \lambda^n \mathbf{K_k})^{-1}$. At the instant $n + 1$, a new x_{n+1} observation is detected at the input of the system, and we update $\mathbf{H_{n+1}}$ and s_{n+1} with Equations (37) and (38), respectively:

$$\mathbf{H_{n+1}} = \begin{bmatrix} \mathbf{H_n} \\ h_{n+1}^\top \end{bmatrix} \tag{37}$$

$$s_{n+1} = (s_0, s_2, ..., s_n) \tag{38}$$

where $h_{n+1} = (\kappa(x_{n+1}, x_1), \kappa(x_{n+1}, x_2), ..., \kappa(x_{n+1}, x_m))^\top$. Then, the equations of the kernel RLS algorithm are updated according to:

$$\mathbf{G_{n+1}} = \frac{\lambda^{-1}\mathbf{P_{n+1}}h_{n+1}}{1 + \lambda^{-1}h_{n+1}^\top \mathbf{P_{n+1}}h_{n+1}} \tag{39}$$

$$\theta_{n+1} = \theta_n + \mathbf{G_{n+1}}(s_{n+1} - h_{n+1}^\top \theta_n) \tag{40}$$

$$\mathbf{P_{n+1}} = \lambda^{-1}\mathbf{P_n}\left[1 - \mathbf{G_{n+1}}h_{n+1}^\top\right] \tag{41}$$

where the term $s_{n+1} - h_{n+1}^\top \theta_n$ represents the estimation error.

5.2. Kernel Least Mean Square Algorithm

In this paragraph, we describe the kernel LMS algorithm [36,53]. This algorithm belongs to the category of stochastic gradient algorithms. The fundamental concept is the execution of the linear least mean square algorithm in the kernel feature map, which is associated with the positively defined kernel κ via the feature map $\Psi(.)$ to generate the input data. The update equations of the LMS algorithm are given by [54]:

$$e_n = s_n - \theta_{n-1}^\top x_n \tag{42}$$

$$\theta_n = \theta_{n-1} + \mu e_n x_n \tag{43}$$

where e_n denotes the estimation error, θ_n represents the weight vector estimate, and μ stands for the step parameter.

Let us now suppose that the Ψ function map is used to transform the sample sequence; then, we perform the least mean square procedure for the processed data sequence defined in Equation (32), so we obtain the kernel LMS:

$$e_n = s_n - \theta_{n-1}^\top \Psi(x_n) \tag{44}$$

$$\theta_n = \theta_{n-1} + \mu e_n \Psi(x_n) \tag{45}$$

The biggest distinction of LMS is when it is used in a space with potentially infinite-dimensional characteristics in Equation (45), which makes direct updating virtually impossible. Rather, we will use every θ_n to relate to the initialization θ_0:

$$\theta_n = \theta_0 + \mu \sum_{i=1}^{n} e_i \Psi(x_i). \tag{46}$$

$$\theta_n = \mu \sum_{i=1}^{n} e_i \Psi(x_i). \quad \text{(if we suppose that } \theta_0 = \theta(0) = 0) \tag{47}$$

The prediction solution is computed by using the kernel trick:

$$\langle \theta_n, \Psi(x_n) \rangle_{\mathcal{H}} = \mu \sum_{i=1}^{n} e_i \langle \Psi(x_i), \Psi(x_n) \rangle \tag{48}$$

$$= \mu \sum_{i=1}^{n} e_i \kappa(x_i, x_n),$$

where $\kappa(x_i, x_n)$ is a Mercer kernel and n is the number of training samples.

6. Simulation Results

In this section, we will evaluate the efficiency of the presented algorithms for various SNR values and for a fixed data input N.

6.1. Impulse Response Parameter Estimation

In Figure 2, it can be seen that Alg1, Alg2, and Alg3 gave good results for $SNR = 0$ dB and a data input of $N = 2000$. These estimations of the channel parameters were closer to the measured values if the number of units of data was important and for $SNR > 16$ dB; this is due to an important property of the cumulant: The cumulant of a Gaussian process is zero. In the following (Figures 3 and 4), we present the estimated parameters in the cases of $SNR = 16$ dB and 32 dB, respectively, with a fixed data input of $N = 2000$.
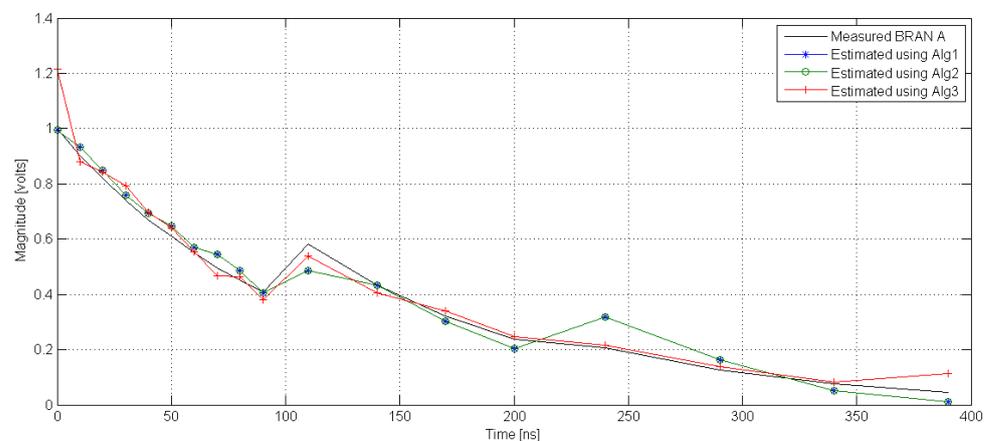

Figure 2. Parameters of BRAN A estimated using algorithms based on cumulants for $SNR = 0$ dB and $N = 2000$.
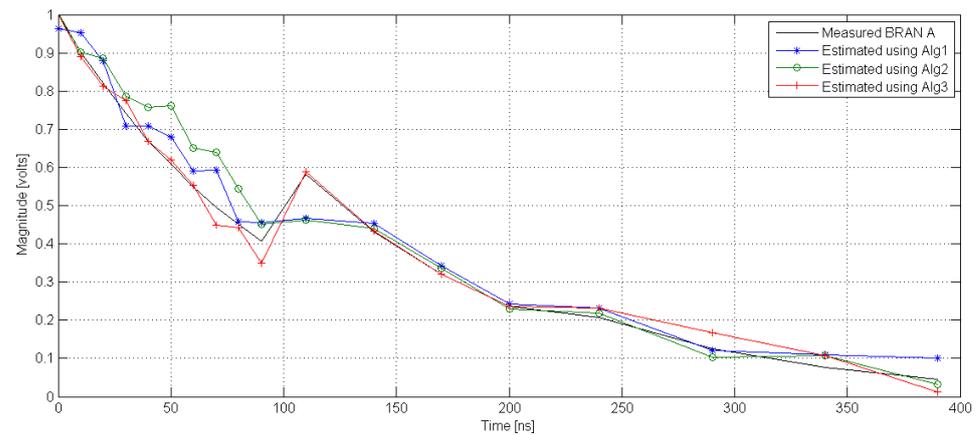
Figure 3. Parameters of BRAN A estimated using algorithms based on cumulants for $SNR = 16$ dB and $N = 2000$.
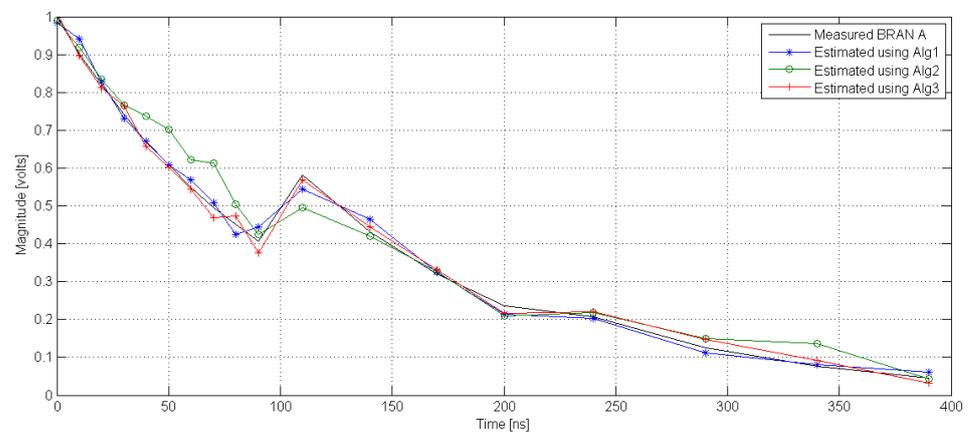

Figure 4. Parameters of BRAN A estimated using algorithms based on cumulants for $SNR = 32$ dB and $N = 2000$.

Next, we tested the performance of the algorithms based on binary measurement for different values of SNR and for a fixed data input of $N = 2000$ in order to make a comparison with the algorithms based on cumulants.

In Figure 5, we present the results obtained for $SNR = 0$ dB and $N = 2000$ when using the methods based on binary measurement. We can observe in Figure 5 that the channel parameters estimated based on the SVM method were closer to the measured values than those obtained by using the RLS and LIMBO methods, which gave values that were very far from those measured.

In Figure 6, we show the estimations of the BRAN A parameters by using the RLS, LIMBO, and SVM methods in the case of $SNR = 16$ dB and with a data input of $N = 2000$. In this case where the noise is not very important, we remark that the results obtained using the RLS and LIMBO methods were closer to those obtained in the case in which noise was important ($SNR = 0$ dB); this result can be explained by the fact that the rounding errors due to the noise were minimal, which implies an improvement in the results obtained.
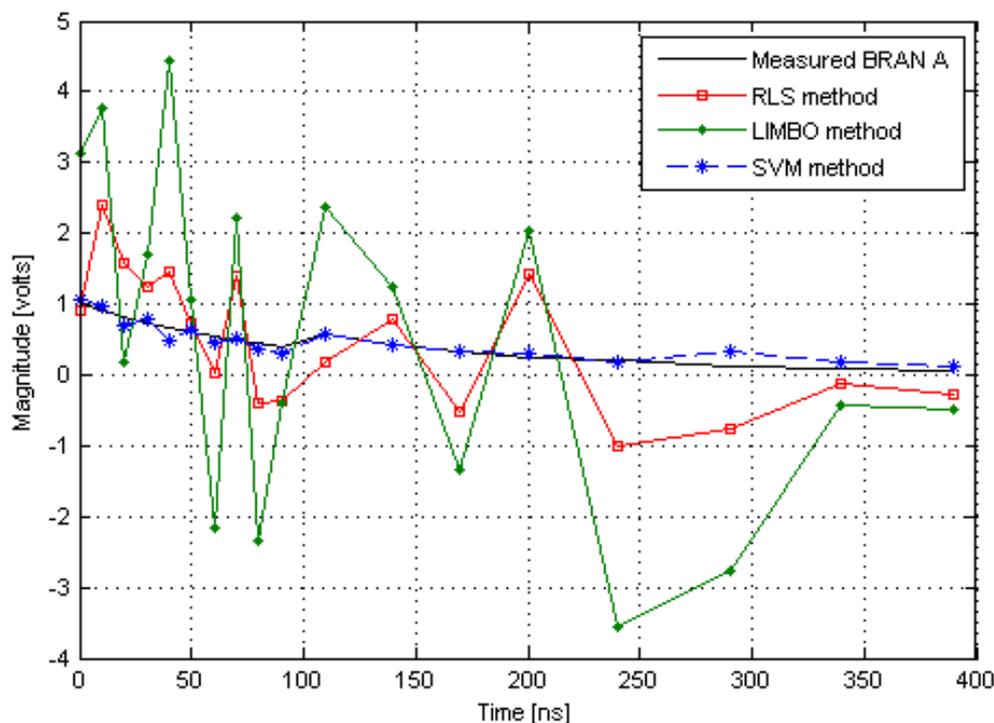
Figure 5. Parameters of BRAN A estimated using RLS ($\lambda = 0.999, \delta = 0.9999$), LIMBO ($\mu = 0.2$) and SVM ($\gamma = 0.1$) method for $SNR = 0$ dB and $N = 2000$.
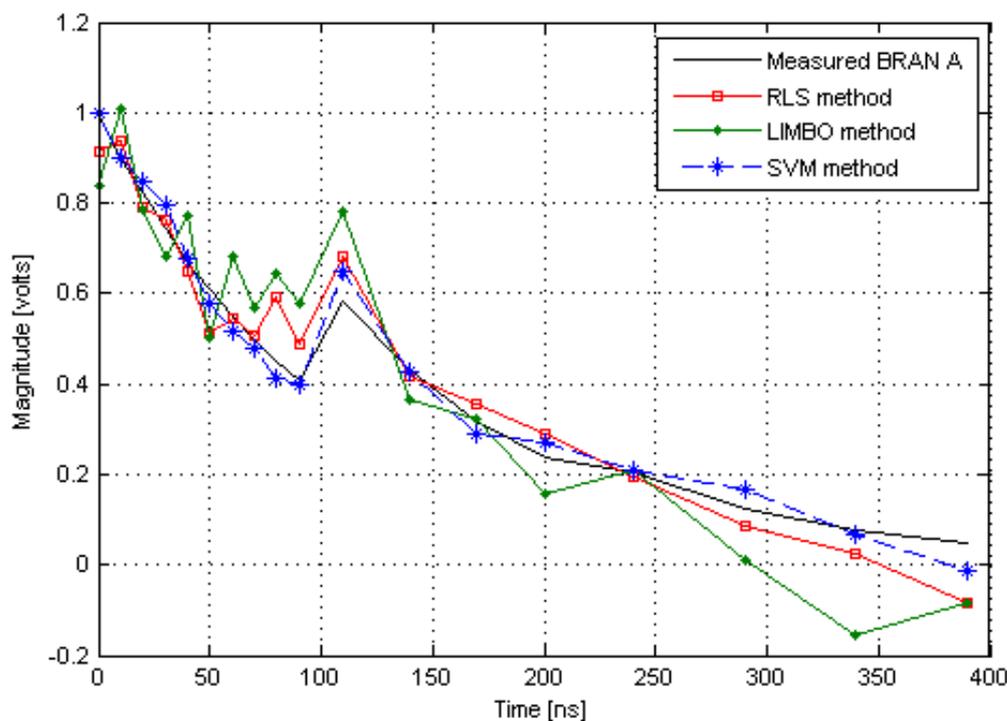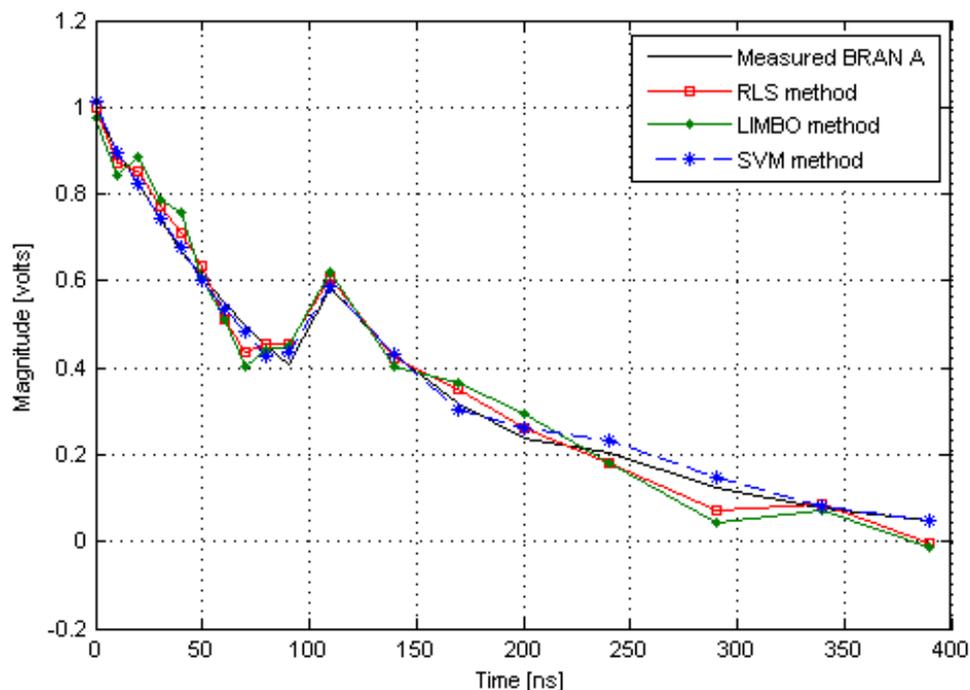

Figure 6. Parameters of BRAN A estimated using the RLS ($\lambda = 0.999, \delta = 0.9999$), LIMBO ($\mu = 0.2$), and SVM ($\gamma = 0.1$) methods for $SNR = 16$ dB and $N = 2000$.

If the noise was lower (i.e., the SNR was high), such as with $SNR = 32$ dB, we observed that the estimated channel parameters were closer to the measured values (Figure 7), but if the noise was more important, the estimated channel parameters were significantly different from the measured values. Note that if the noise was equal to the power of the information signal (i.e., $SNR = 0$ dB),

we could see that the estimated channel parameters were very different from the measured values; this is because of the lack of a filtering operation for the noise, which implies some rounding errors. However, the SVM method presented good results for different values of SNR; this was due to the improved classification performance, as well as the ability to manipulate a large volume of data with a reasonable computational load. Furthermore, this strategy was one that was perfect for minimizing prediction error. However, due to its simpler and more efficient structure and learning procedure, the SVM has a second characteristic: It can learn efficiently with small amounts of data.

Figure 7. Parameters of BRAN A estimated using the RLS ($\lambda = 0.999, \delta = 0.9999$), LIMBO ($\mu = 0.2$), and SVM ($\gamma = 0.1$) methods for $SNR = 32$ dB and $N = 2000$.

Figures 8–10 show the estimation of the impulse response parameters of the BRAN A channel as a function of the path delays by using the kernel-based algorithms for different values of SNR and $N = 2000$.
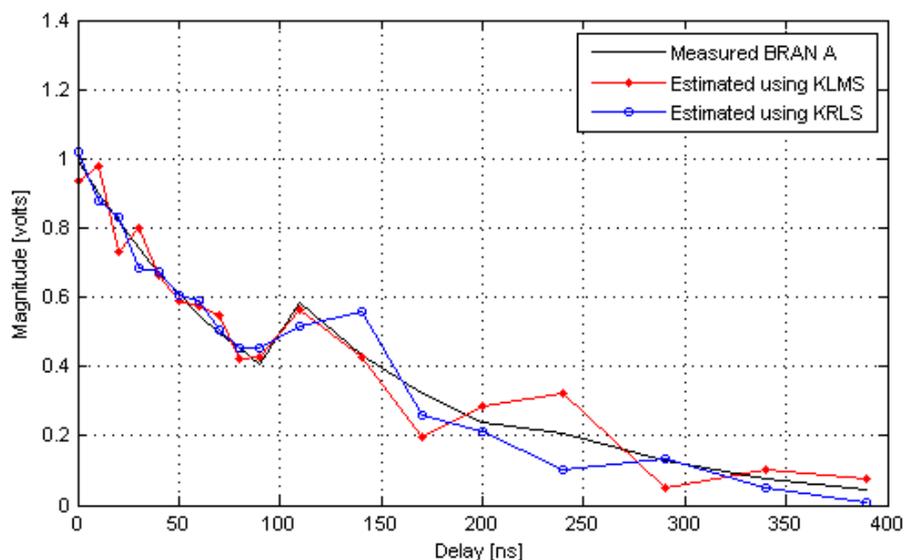

Figure 8. Parameters of BRAN A estimated using KRLS ($\lambda = 0.99$) and KLMS ($\mu = 0.05$) for $SNR = 0$ dB and $N = 2000$ with the Gaussian kernel ($\sigma = 0.5$).
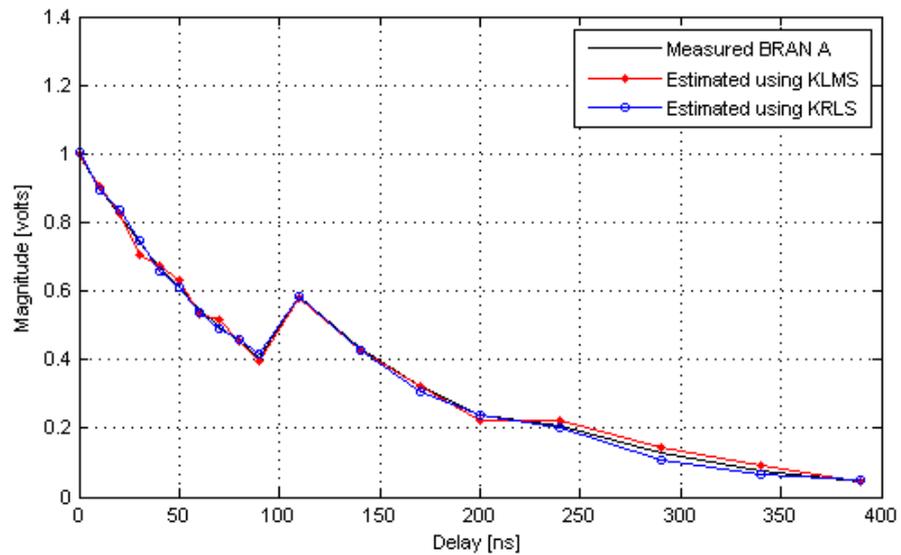
Figure 9. Parameters of BRAN A estimated using KRLS ($\lambda = 0.99$) and KLMS ($\mu = 0.05$) for $SNR = 16$ dB and $N = 2000$ with the Gaussian kernel ($\sigma = 0.5$).
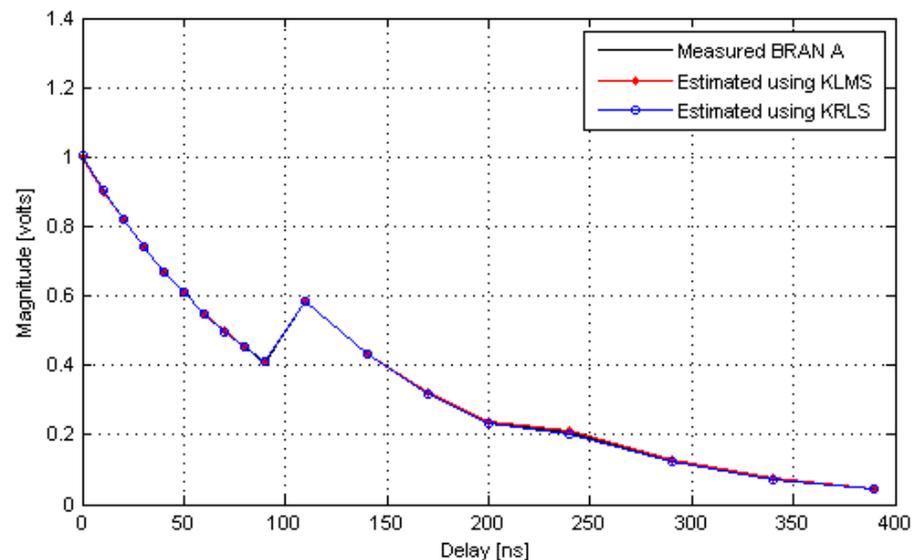

Figure 10. Parameters of BRAN A estimated using KRLS ($\lambda = 0.99$) and KLMS ($\mu = 0.05$) for $SNR = 32$ dB and $N = 2000$ with the Gaussian kernel ($\sigma = 0.5$).

From Figures 8–10, we can conclude that the estimated amplitude of the BRAN A channel was very close to the true value, especially for $SNR \geqslant 16$ dB (see Figures 9 and 10). Regarding the impulse response of the channel in a very noisy environment (Figure 8), the Gaussian noise influenced the estimated model parameters when using KLMS, and the noise slightly influenced the estimation of the impulse response parameters, especially when KRLS was used.

6.2. Magnitude and Phase Estimation

Next, we identified the amplitude and phase of the ETSI BRAN A channel for different values of SNR by setting the data input to $N = 2000$.

Figures 11–13 represent the results of the estimation of the amplitude and phase by comparing the measured parameters of the BRAN A channel and the parameters estimated using the algorithms based on cumulants (Alg1, Alg2, and Alg3). It should be noted that the algorithms based on cumulants provided remarkable performance. Indeed, if $SNR < 16$ dB, we obtained good results with the algorithms based on cumulants, but if $SNR \geqslant 16$ dB, we observed that the algorithms based on binary measurement gave us better results than the algorithms based on cumulants. It must be emphasized that the algorithms based on cumulants are simple and have the advantage of

identifying the parameters with a low calculation cost and without knowledge of the input (blind identification) in comparison with the methods based on binary measurement, which require the use of the input and output of the channel.
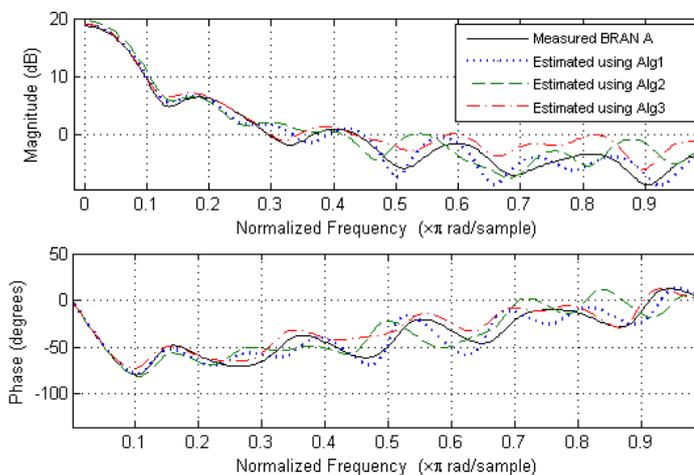

Figure 11. Estimations of the BRAN A magnitude and phase using algorithms based on cumulants for $SNR = 0$ dB and $N = 2000$.
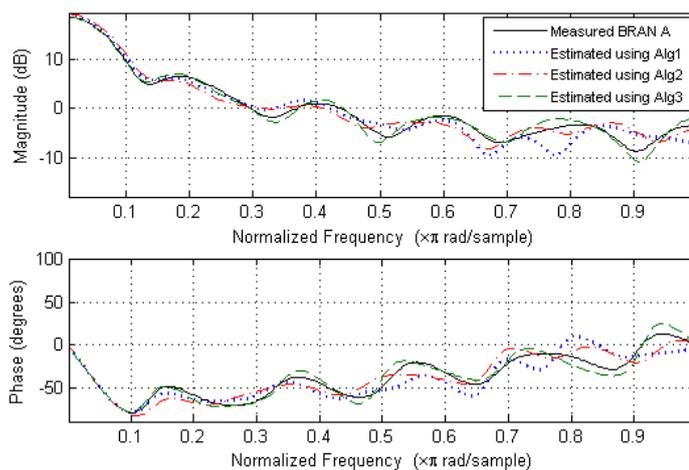

Figure 12. Estimations of the BRAN A magnitude and phase using algorithms based on cumulants for $SNR = 16$ dB and $N = 2000$.
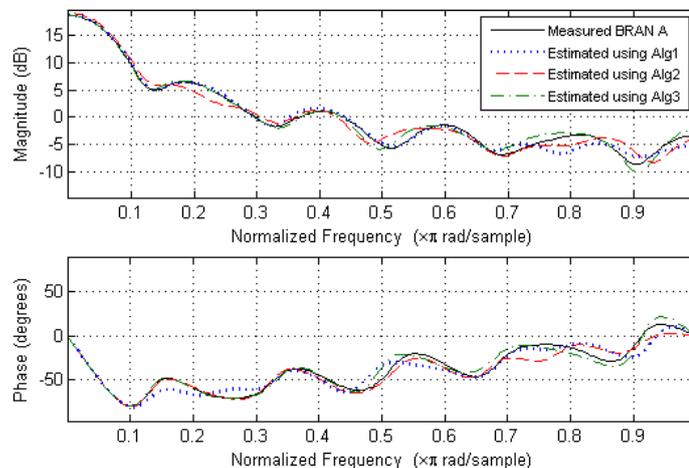

Figure 13. Estimations of the BRAN A magnitude and phase using algorithms based on cumulants for $SNR = 32$ dB and $N = 2000$.

The estimations of the amplitude and phase of the BRAN A channel using the methods based on binary measurement with different SNR values are presented in Figures 14–16. From these figures, we observe that the shapes of the amplitude and phase estimated with the SVM algorithm are the same as those of the measured data (Figure 16). We noticed a minor difference between the estimated and measured values for $SNR < 16$ dB (Figures 14 and 15). These results are very important because we are using a very quickly fading channel.
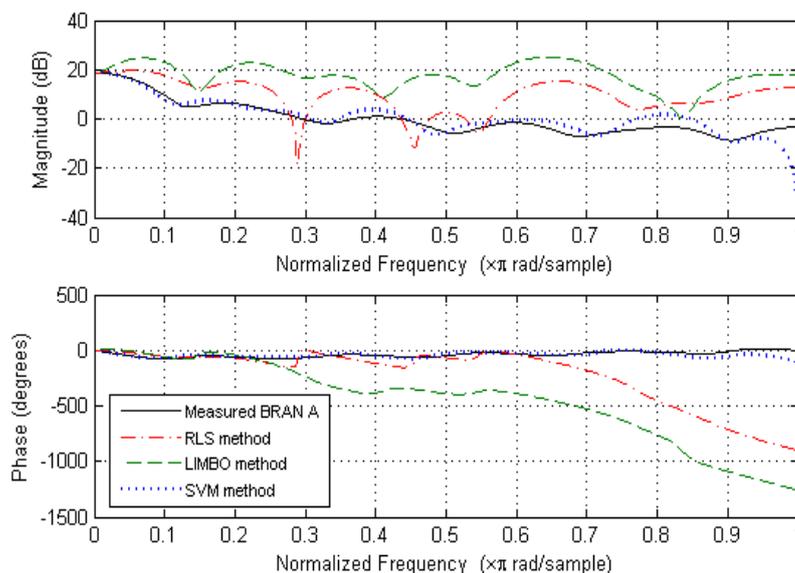

Figure 14. Estimations of the BRAN A magnitude and phase using the RLS ($\lambda = 0.999, \delta = 0.9999$), LIMBO ($\mu = 0.2$), and SVM ($\gamma = 0.1$) methods for $SNR = 0$ dB and $N = 2000$.

Figure 15. Estimations of the BRAN A magnitude and phase using the RLS ($\lambda = 0.999, \delta = 0.9999$), LIMBO ($\mu = 0.2$), and SVM ($\gamma = 0.1$) methods for $SNR = 16$ dB and $N = 2000$.
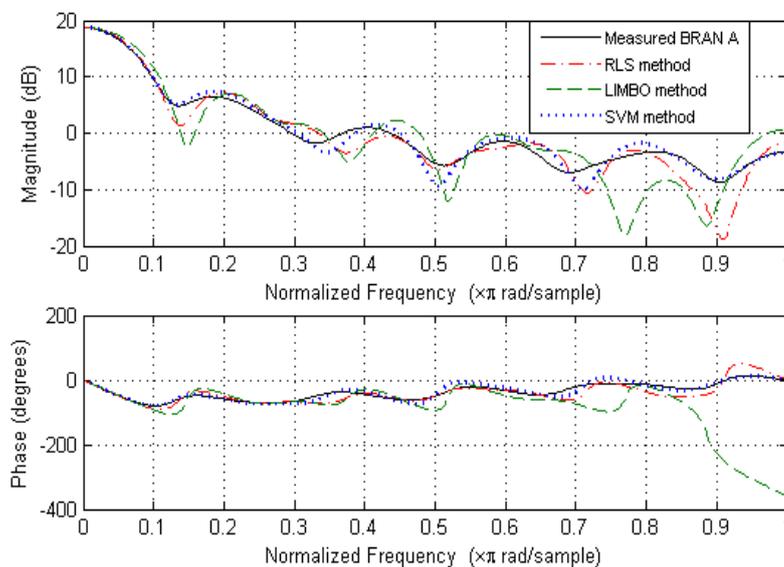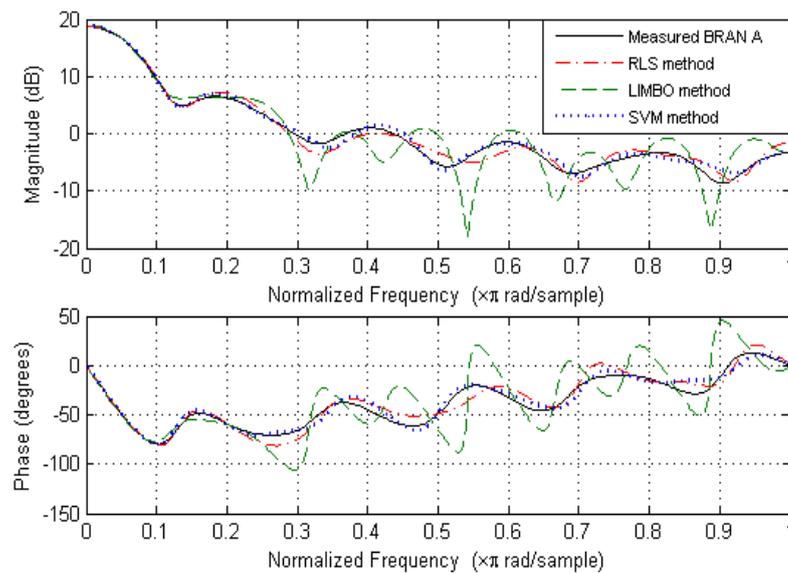
Figure 16. Estimations of the BRAN A magnitude and phase using the RLS ($\lambda = 0.999, \delta = 0.9999$), LIMBO ($\mu = 0.2$), and SVM ($\gamma = 0.1$) methods for $SNR = 32$ dB and $N = 2000$.

The curves in Figures 17–19 show the estimations of the amplitude and phase of the BRAN A channel by using kernel methods (KRLS and KLMS) for a data length of $N = 2000$ and for different values of SNR. In Figures 17–19, we present the performance of the KRLS and KLMS algorithms. We observed that the KRLS algorithm performed better than the KLMS algorithm, which produced quite large errors. We can conclude from Figure 17 that the phase of the BRAN A channel's impulse response estimated by the KRLS algorithm was closer to the true values. However, when using the KLMS algorithm, the difference between the measured and estimated phase was significantly larger.

The estimated parameters shown in Figures 12, 13, 15, 16, 18, and 19 evidently show that both KLMS and KRLS achieved higher prediction accuracy than the binary-measurement-based algorithms and cumulant-based algorithms for all values of SNR. Meanwhile, the performance of KRLS was also marginally better than that of KLMS, although it had a greater computational complexity.
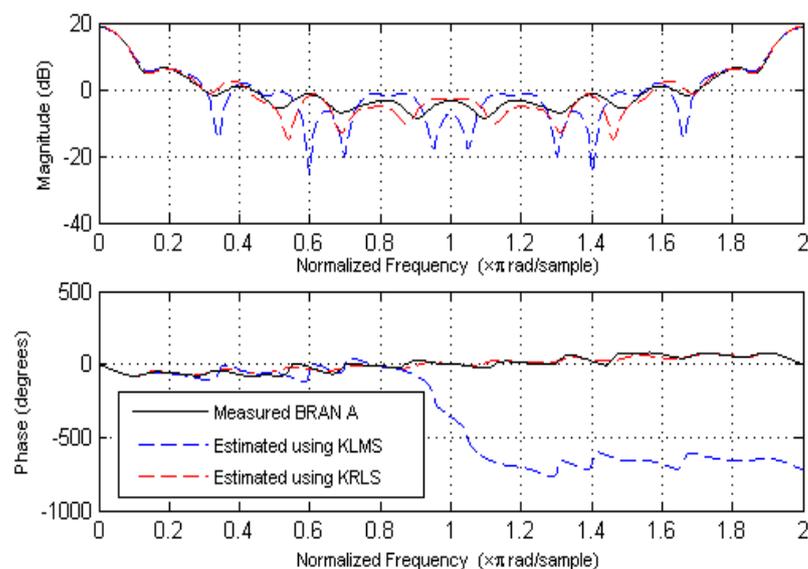

Figure 17. Estimations of the BRAN A magnitude and phase using KRLS ($\lambda = 0.99$) and KLMS ($\mu = 0.05$) for $SNR = 0$ dB and $N = 2000$ with the Gaussian kernel ($\sigma = 0.5$).

Figure 18. Estimations of the BRAN A magnitude and phase using KRLS ($\lambda = 0.99$) and KLMS ($\mu = 0.05$) for $SNR = 16$ dB and $N = 2000$ with the Gaussian kernel ($\sigma = 0.5$).

Figure 19. Estimations of the BRAN A magnitude and phase using KRLS ($\lambda = 0.99$) and KLMS ($\mu = 0.05$) for $SNR = 32$ dB and $N = 2000$ with the Gaussian kernel ($\sigma = 0.5$).
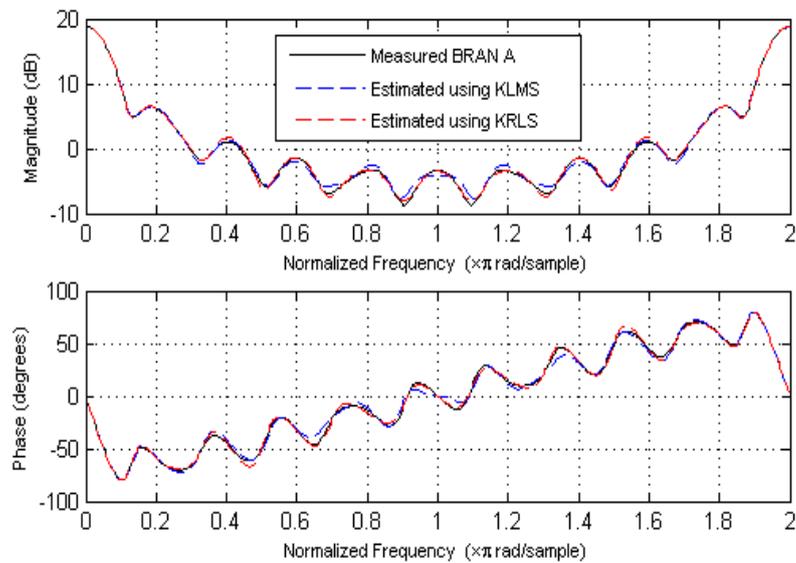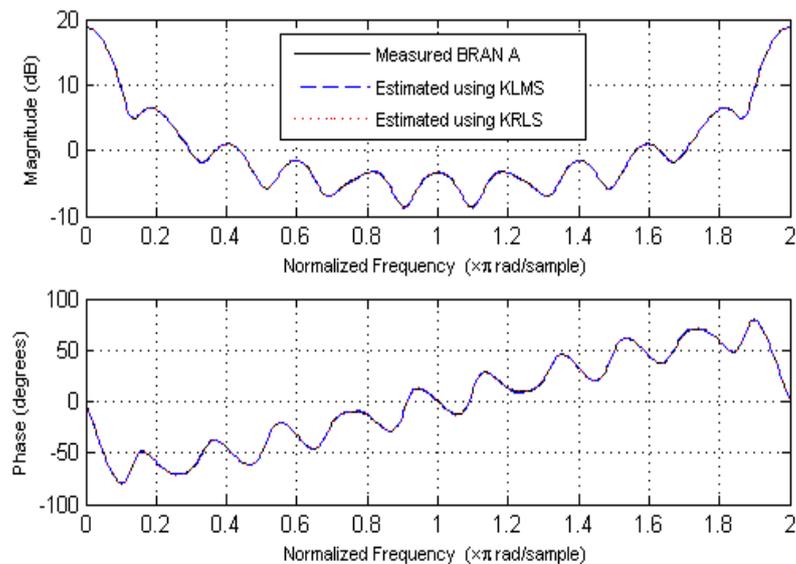
7. Conclusions

In this paper, we presented comparisons among three algorithms based on binary measurements, three algorithms based on higher-order cumulants, and two algorithms based on reproducing kernel Hilbert spaces. These algorithms were used to estimate the parameters of a broadband radio access network channel (BRAN A). Results were obtained for different SNR values and for different data inputs. From this study, we can conclude that both of the kernel algorithms are efficient and effective in identifying the amplitude and phase of the BRAN A channel compared to the other algorithms (those based on higher-order cumulants and binary measurements). In general, the algorithms based on binary measurement gave better results in the estimation of parameters if the SNR is higher than 16 dB (i.e., $SNR \geq 16$ dB), but if the noise was more present (i.e., $SNR < 16$ dB), the algorithms based on cumulants gave better results than those obtained by algorithms based on binary measurement; this was because the cumulants of the Gaussian process (here, we supposed that the noise had a Gaussian distribution) were null. We can add that if the size of the data input is important, the estimation of the cumulants will be approximately without bias, which implies a

good estimation of the channel parameters. Future work will focus on developing techniques for identifying communication channels by using kernel methods.

Author Contributions: Conceptualization, H.O., R.F. and A.D.; Data curation, R.F. and A.D.; Formal analysis, H.O. and R.F.; Investigation, R.F., A.D. and S.S.; Methodology, H.O., R.F., A.D. and S.S.; Resources, M.P. and M.F.; Software, R.F. and S.S.; Supervision, M.P. and M.F.; Validation, A.D. and S.S.; Visualization, H.O., R.F., M.P. and M.F.; Writing—original draft, R.F., H.O., A.D. and S.S.; Writing—review & editing, S.S., M.P. and M.F.; All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Safi, S.; Zeroual, A. MA system identification using higher order cumulants: Application to modelling solar radiation. *Int. J. Stat. Comput. Simul.* **2002**, *72*, 533–548. [CrossRef]
2. Safi, S.; Zeroual, A. Blind parametric identification of non-Gaussian FIR systems using higher order cumulants. *Int. J. Syst. Sci.* **2004**, *35*, 855–867. [CrossRef]
3. Safi, S.; Zeroual, A. Blind non-minimum phase channel identification using 3rd and 4th order cumulants. *Int. J. Signal Process.* **2008**, *4*, 158–168.
4. Sadler, B.M.; Giannakis, G.B.; Lii, K.S. Estimation and detection in non-Gaussian noise using higher order statistics. *IEEE Trans. Signal Process.* **1994**, *42*, 2729–2741. [CrossRef]
5. Wang, J.; He, Z. Criteria and algorithms for blind source separation based on cumulants. *Int. J. Electron.* **1996**, *81*, 1–14. [CrossRef]
6. Yin, G.G.; Zhao, Y.; Zhang, J.F. Identification input design for consistent parameter estimation of linear systems with binary-valued output observations. *IEEE Trans. Autom. Control* **2008**, *53*, 867–880.
7. Zhang, J.F.; Yin, G.G. System identification using binary sensors. *IEEE Trans. Autom. Control* **2003**, *48*, 1892–1907.
8. Godoy, B.I.; Goodwin, G.C.; Agüero, J.C.; Marelli, D.; Wigren, T. On identification of FIR systems having quantized output data. *Automatica* **2011**, *47*, 1905–1915. [CrossRef]
9. Safi, S.; Zeroual, A. Blind identification in noisy environment of nonminimum phase finite impulse response (FIR) system using higher order statistics. *Syst. Anal. Model. Simul.* **2003**, *43*, 671–681. [CrossRef]
10. Antari, J.; Iqdour, R.; Safi, S.; Zeroual, A.; Lyhyaoui, A. Identification of quadratic non linear systems using higher order statistics and fuzzy models. In Proceedings of the IEEE International Conference On Acoustic, Speech and Signal Process (ICASSP), Toulouse, France, 14–19 May 2006; pp. 712–715.
11. Zidane, M.; Safi, S.; Sabri, M. Extending HOC-based methods for identifying the diagonal parameters of quadratic systems. *Signal Image Video Process.* **2018**, *12*, 125–132. [CrossRef]
12. Zidane, M.; Safi, S.; Sabri, M.; Boumezzough, A. Identification and equalization using higher order cumulants in MC-CDMA systems. In Proceedings of the 2014 5th Workshop on Codes, Cryptography and Communication Systems, WCCCS 2014, El Jadida, Morocco, 27–28 November 2014; pp. 81–85.
13. Zidane, M.; Safi, S.; Sabri, M.; Boumezzough, A.; Frikel, M. Broadband radio access network channel identification and downlink MC-CDMA equalization. *Int. J. Energy Inform. Commun.* **2014**, *5*, 13–34. [CrossRef]
14. Frikel, M.; Targui, B.; Hamon, F.; M'Saad, M. Adaptive equalization using controlled equal gain combining for uplink/downlink MC-CDMA systems. *Int. J. Signal Process.* **2008**, *4*, 230–237.
15. Zidane, M.; Safi, S.; Sabri, M. Measured and estimated data of non-linear BRAN channels using HOS in 4G wireless communications. *Data Brief* **2018**, *17*, 1136–1148. [CrossRef]
16. Nikias, C.L.; Mendel, J.M. Signal processing with higher-order spectra. *IEEE Signal Process. Mag.* **1993**, *10*, 10–37. [CrossRef]
17. Wang, A.; Li, R. Research on digital signal recognition based on higher order cumulants. In Proceedings of the 2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), Changsha, China, 12–13 January 2019; pp. 586–588.
18. Simic, M.; Stankovic, M.; Orlic, V.D. Automatic Modulation Classification of Real Signals in AWGN Channel Based on Sixth-Order Cumulants. *Radioengineering* **2021**, *30*, 205. [CrossRef]
19. Khosraviyani, M.; Kalbkhani, H.; Shayesteh, M.G. Higher order statistics for modulation and STBC recognition in MIMO systems. *IET Commun.* **2019**, *13*, 2436–2446. [CrossRef]
20. Anderson, J.M.; Giannakis, G.B. Noisy input/output system identification using cumulants and the Steiglitz-McBride algorithm. *IEEE Trans. Signal Process.* **1996**, *44*, 1021–1024. [CrossRef]
21. Kaiser, T. MA-model identification using modulated moment sequences. *Signal Process.* **1995**, *47*, 85–103. [CrossRef]
22. Proakis, J.; Salehi, M. *Digital Communications*, 5th ed.; McGraw-Hill: New York, NY, USA, 2007.
23. Ju, L.; Zhenya, H. Blind identification and equalization using higher-order cumulants and ICA algorithms. In Proceedings of the International Conference Neural Networks and Brain (ICNN B'98), Beijing, China, 13–15 October 1998.
24. Zhang, X.D.; Zhang, Y.S. FIR system identification using higher order statistics alone.*IEEE Trans. Signal Process.* **1994**, *42*, 2854–2858. [CrossRef]

25.	Guo, J.; Zhao, Y.; Sun, C.Y.; Yu, Y. Recursive identification of FIR systems with binary-valued outputs and communication channels. *Automatica* **2015**, *60*, 165–172. [CrossRef]

26.	Guo, J.; Yin, G.; Zhao, Y.; Zhang, J.F. Identification of Wiener systems with quantized inputs and binary-valued output observations. *Automatica* **2017**, *78*, 280–286. [CrossRef]

27.	Yin, G.G.; Zhang, J.F. Joint identification of plant rational models and noise distribution functions using binary-valued observations. *Automatica* **2006**, *42*, 535–547.

28.	Pouliquen, M.; Menard, T.; Pigeon, E.; Gehan, O.; Goudjil, A. Recursive system identification algorithm using binary measurements. In Proceedings of the 2016 European Control Conference (ECC), Aalborg, Denmark, 29 June–1 July 2016; pp. 1353–1358.

29.	Goudjil, A.; Pouliquen, M.; Pigeon, E.; Gehan, O.; M'Saad, M. Identification of systems using binary sensors via support vector machines. In Proceedings of the 2015 54th IEEE Conference on Decision and Control (CDC), Osaka, Japan, 15–18 December 2015; pp. 3385–3390.

30.	Jafari, K.; Juillard, J.; Roger, M. Convergence analysis of an online approach to parameter estimation problems based on binary observations. *Automatica* **2012**, *48*, 2837–2842. [CrossRef]

31.	Guo, J.; Zhao, Y. Recursive projection algorithm on FIR system identification with binary-valued observations. *Automatica* **2013**, *49*, 3396–3401. [CrossRef]

32.	Oualla, H.; Pouliquen, M.; Frikel, M.; Safi, S. Comparison of algorithms for identification of IIR systems from binary measurements on the output. In Proceedings of the E3S Web of Conferences, EDP Sciences, Guizhou, China, 23–25 July 2021; p. 01049.

33.	Liu, W.; Principe, J.C.; Haykin, S. *Kernel Adaptive Filtering: A Comprehensive Introduction*; John Wiley & Sons: Hoboken, NJ, USA, 2011.

34.	Girosi, F.; Jones, M.; Poggio, T. Regularization theory and neural networks architectures. *Neural Comput.* **1995**, *7*, 219–269. [CrossRef]

35.	Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–6297. [CrossRef]

36.	Liu, W.; Pokharel, P.P.; Principe, J.C. The kernel least-mean-square algorithm. *IEEE Trans. Signal Process.* **2008**, *56*, 543–554. [CrossRef]

37.	Liu, W.; Principe, J.C. Kernel affine projection algorithms. *EURASIP J. Adv. Signal Process.* **2008**, *2008*, 1–12. [CrossRef]

38.	Schölkopf, B.; Smola, A.; Müller, K.R. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput.* **1998**, *10*, 1299–1319. [CrossRef]

39.	Engel, Y.; Mannor, S.; Meir, R. The kernel recursive least-squares algorithm. *IEEE Trans. Signal Process.* **2004**, *52*, 2275–2285. [CrossRef]

40.	Chen, B.; Zhao, S.; Zhu, P.; Príncipe, J.C. Quantized kernel least mean square algorithm. *IEEE Trans. Neural Netw. Learn. Syst.* **2011**, *23*, 22–32. [CrossRef]

41.	Wang, S.; Zheng, Y.; Ling, C. Regularized kernel least mean square algorithm with multiple-delay feedback. *IEEE Signal Process. Lett.* **2015**, *23*, 98–101. [CrossRef]

42.	Chen, B.; Liang, J.; Zheng, N.; Príncipe, J.C. Kernel least mean square with adaptive kernel size. *Neurocomputing* **2016**, *191*, 95–106. [CrossRef]

43.	Slavakis, K.; Bouboulis, P.; Theodoridis, S. Online learning in reproducing kernel Hilbert spaces. In *Signal Processing Theory and Machine Learning (Series Academic Press Library in Signal Processing)*; Chellappa, R., Theodoridis, S., Eds.; Academic: New York, NY, USA, 2014; pp. 883–987.

44.	Fateh, R.; Darif, A. Mean Square Convergence of Reproducing Kernel for Channel Identification: Application to Bran D Channel Impulse Response. In *International Conference on Business Intelligence*; Springer: Cham, Switzerland, 2021; pp. 284–293.

45.	Broadband Radio Acces Network (BRAN). Hyperlan type 2. In *Requirements and Archtiectures for Wireless Broadband Access*; European Telecommunications Standards Institute (ETSI): Sophia Antipolis, France, 1999.

46.	Broadband Radio Acces Network (BRAN). Hyperlan type 2. In *Physical Layer*; European Telecommunications Standards Institute (ETSI): Sophia Antipolis, France, 2001.

47.	Rosenblatt, M. A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. USA* **1956**, *42*, 43. [CrossRef] [PubMed]

48.	Aronszajn, N. Theory of reproducing kernels. *Trans. Am. Math. Soc.* **1950**, *68*, 337–404. [CrossRef]

49.	Tobar, F. Kernel-Based Adaptive Estimation: Multidimensional and State-Space Approaches. Ph.D. Thesis, Imperial College London, London, UK, 2014.

50.	Façanha, T.S.; Barreto, G.A.; Costa Filho, J.T. A Novel Kalman Filter Formulation for Improving Tracking Performance of the Extended Kernel RLS. *Circuits Syst. Signal Process.* **2021**, *40*, 1397–1419. [CrossRef]

51.	Saide, C.; Lengelle, R.; Honeine, P.; Richard, C.; Achkar, R. Nonlinear adaptive filtering using kernel-based algorithms with dictionary adaptation. *Int. J. Adapt. Control. Signal Process.* **2015**, *29*, 1391–1410. [CrossRef]

52.	Schölkopf, B.; Herbrich, R.; Smola, A.J. A generalized representer theorem. In *International Conference on Computational Learning Theory*; Springer: Berlin/Heidelberg, Germany, 2001; pp. 416–426.

53.	Li, K.; Principe, J.C. No-Trick (Treat) Kernel Adaptive Filtering Using Deterministic Features. 2019. Available online: https://arxiv.org/abs/1912.04530 (accessed on 16 June 2021).

54.	Martinek, R.; Rzidky, J.; Jaros, R.; Bilik, P.; Ladrova, M. Least Mean Squares and Recursive Least Squares Algorithms for Total Harmonic Distortion Reduction Using Shunt Active Power Filter Control. *Energies* **2019**, *12*, 1545. [CrossRef]