Research Article

Maximum Likelihood Sequence Detection Receivers for Nonlinear Optical Channels

Gabriel N. Maggio,[1] Mario R. Hueda,[1] and Oscar E. Agazzi[2]

[1]*Laboratorio de Comunicaciones Digitales, Universidad Nacional de Córdoba, CONICET, Avenida Vélez Sarsfield 1611,
 Córdoba X5016GCA, Argentina*
[2]*ClariPhy Communications, Inc., 7585 Irvine Center Drive, Suite 100, Irvine, CA 92618, USA*

Correspondence should be addressed to Gabriel N. Maggio; gmaggio@efn.uncor.edu

The space-time whitened matched filter (ST-WMF) maximum likelihood sequence detection (MLSD) architecture has been recently proposed (Maggio et al., 2014). Its objective is reducing implementation complexity in transmissions over nonlinear dispersive channels. The ST-WMF-MLSD receiver (i) drastically reduces the number of states of the Viterbi decoder (VD) and (ii) offers a smooth trade-off between performance and complexity. In this work the ST-WMF-MLSD receiver is investigated in detail. We show that the *space compression* of the nonlinear channel is an instrumental property of the ST-WMF-MLSD which results in a major reduction of the implementation complexity in intensity modulation and direct detection (IM/DD) fiber optic systems. Moreover, we assess the performance of ST-WMF-MLSD in IM/DD optical systems with chromatic dispersion (CD) and polarization mode dispersion (PMD). Numerical results for a 10 Gb/s, 700 km, and IM/DD fiber-optic link with 50 ps differential group delay (DGD) show that the number of states of the VD in ST-WMF-MLSD can be reduced ~4 times compared to an oversampled MLSD. Finally, we analyze the impact of the imperfect channel estimation on the performance of the ST-WMF-MLSD. Our results show that the performance degradation caused by channel estimation inaccuracies is low and similar to that achieved by existing MLSD schemes (~0.2 dB).

1. Introduction

Maximum likelihood sequence detection (MLSD) receivers for nonlinear channels have been extensively investigated in the literature (e.g., [1, 2] and references therein). Their ability to achieve optimal performance in the presence of additive white Gaussian noise (AWGN) has always been of great theoretical and practical interest. The theoretical interest lies in that it provides a performance bound for any reception scheme. The practical interest stems from its ability to actually achieve optimal or nearly optimal performance in transmissions over nonlinear channels.

Much of the early work in this area has been focused on the compensation of nonlinearities in satellite communications [1–5]. A traditional architecture for the optimal nonlinear receiver consists of a matched-filter bank (MFB) followed by a maximum likelihood sequence detector (MLSD) [1].

Owing to the correlation among the *spatial* noise components at the MFB outputs, an MLSD with non-Euclidean metrics must be used. The use of oversampling in combination with MLSD (OS-MLSD) has also been proposed to implement the optimal receiver in the presence of nonlinearities (see [2] and references therein). Since the complexity of both MFB and OS MLSD-based receivers grows exponentially with the channel memory, their practical application in transmissions over highly dispersive channels has been limited. Despite this fact, MLSD-based receivers are still preferred over decision feedback equalizers (DFE) in applications such as multigigabit intensity modulation/direct detection (IM/DD) fiber optic systems for the two following reasons.

(i) *Performance:* DFE suffers from severe performance limitations in *moderate to high* dispersion single-mode fiber links as a result of its inability to compensate nonlinear ISI [6, 7]. Although nonlinear

DFE (NL-DFE) structures have also been considered in the literature (e.g., [8, 9]), their performances still degrade significantly at long fiber lengths (e.g., ≥500 km). On the other hand, MLSD can operate with a constant penalty around 3 dB with respect to back-to-back (B2B) at virtually any distance [10, 11].

(ii) *High-speed implementation:* future generation of communication systems will operate at multigigabit per second data rates on highly dispersive channels [12]. In commercial applications, the digital receiver is often implemented as a monolithic chip in CMOS technology [12]. Maximum clock frequency of state-of-the-art complex digital signal processors in 28 nm CMOS technology is limited to frequencies lower than ~1 GHz. Therefore, in order to achieve multigigabit per second data rates, parallel processing techniques are required [12]. Although the complexity of serial implementations of DFE grows linearly with channel memory, all presently known parallel processing implementations require that the bottleneck created by the feedback loop be broken using techniques such as the ones proposed by [13–15], whose complexity *grows exponentially with the channel memory.* Therefore, in high-speed applications, complexity of both DFE and MLSD increases with the channel memory in a similar way.

From the above it is clear that complexity reduction of MLSD is crucial for many practical applications. In IM/DD optical channels, the receiver must compensate the linear fiber dispersion as well as nonlinearities caused by lasers, optical modulators, the fiber Kerr effect, photo-detectors, and other components of the link. Chromatic dispersion (CD) and polarization-mode dispersion (PMD), in combination with the quadratic response of the photo-detector, are major factors that limit the reach and drive the complexity of optical-transmission systems at data rates ≥10 Gb/s [16, 17]. With traditional implementations based on oversampling techniques, an 8192-state MLSD is required to compensate 700 km of fiber at 10 Gb/s [10]. This is prohibitive in current CMOS technology[1].

A new MLSD receiver architecture for nonlinear channels has been proposed in [18]. The major breakthrough of this proposal consists in a novel representation of the received signal obtained by a Gram-Schmidt-like orthogonalization of the kernels of a Volterra series expansion of the channel. This procedure yields a special form of space-time whitened matched filter (ST-WMF) [19] whose baud-rate-sampled outputs are sufficient statistics with independent noise components in both space and time. Combined with the minimum phase property of the response of each branch, the ST-WMF provides an effective way to reduce the complexity of MLSD in nonlinear channels. The ST-WMF MLSD technique offers a smooth trade-off between performance and complexity. As complexity is progressively reduced, performance degrades in a graceful manner. Numerical results in [18] demonstrate that the number of states of the VD in ST-WMF-MLSD required on a 10 Gb/s, 700 km, and IM/DD fiber-optic link

without PMD can be reduced 8 times compared to an oversampled MLSD.

Further contributions on the ST-WMF-MLSD receiver and its performance on IM/DD optical links are provided in this work. First, the space channel compression achieved by the ST-WMF-MLSD is analyzed in detail. Space channel compression is important because it is the property that enables the major complexity reductions of the proposed architecture. Second, the performance evaluation of the ST-WMF-MLSD receiver is extended by addressing the combined effect of CD and PMD. Numerical results confirm that the ST-WMF-MLSD remains an attractive solution in the presence of these combined impairments. Finally, the impact of channel estimation inaccuracies on the ST-WMF-MLSD performance is assessed. Accuracy and speed of channel estimation are particularly important when tracking nonstationary channels. Nonstationarity in optical channels results from PMD and random rotations of the laser state of polarization. Our results show that the performance degradation caused by an imperfect channel estimation is low and similar to that achieved by existing MLSD schemes (~0.2 dB).

This paper is organized as follows. The nonlinear channel model and the ST-WMF-MLSD architecture are described in Sections 2 and 3, respectively. Performance evaluation of the ST-WMF-MLSD under different channel conditions, as well as its robustness in the presence of imperfect channel knowledge, are presented and discussed in Sections 4 and 5. Finally, conclusions are drawn in Section 6.

2. Nonlinear Channel Model

The noisy received signal is given by

$$r(t) = s(t) + z(t), \tag{1}$$

where $s(t)$ is the noise-free signal and $z(t)$ is the noise component, which is assumed to be a white Gaussian process with power spectral density N_0. Component $s(t)$ can be expressed in terms of its Volterra-series expansion [20, 21]. For example, neglecting the DC term of the expansion we get

$$s(t) = \sum_k a_k f_0(t - kT) + \sum_{m=1}^{N-1} \sum_k a_k a_{k-m} f_m(t - kT), \tag{2}$$

where $f_0(t)$ is the linear kernel, $f_m(t)$ with $m > 0$ is the mth second-order kernel, a_k is the kth symbol at the input of the nonlinear channel, $1/T$ is the symbol rate, and N is the total number of kernels[2].

2.1. Channel Model Orthogonalization. Next we present an alternative representation of the nonlinear signal $s(t)$ [18]. Without loss of generality, we consider here the Volterra-series with a dominant linear kernel $f_0(t)$, and we select the first *pivoting* response as $h_0(t) = f_0(t)$. Let $\mathcal{H}_0$ be the *signal space* spanned by the set $\{h_0(t - kT)\}$ [19]. We assume that signal spaces are Hilbert spaces with inner product defined as $\int_{-\infty}^{\infty} x(t) y^*(t) dt$, where superscript $*$ denotes complex

conjugate. From the *projection theorem*, the nonlinear kernels $f_m(t)$ can be uniquely expressed as

$$f_m(t) = \sum_n \lambda_n^{(0,m)} h_0(t - nT) + g_m^{(0)}(t), \tag{3}$$

where $g_m^{(0)}(t)$ is *orthogonal* to the signal space $\mathscr{H}_0$; that is,

$$\int_{-\infty}^{\infty} g_m^{(0)}(t) h_0^*(t - jT)\, dt = 0, \quad m = 1, \ldots, N - 1, \ \forall j \in \mathscr{Z}, \tag{4}$$

while $\int_{-\infty}^{\infty} |g_m^{(0)}(t)|^2 dt$ is minimum [19], and $\mathscr{Z}$ denotes the set of all integers. We highlight that the first summation in (3) is the *projection* of $f_m(t)$ onto $\mathscr{H}_0$. Define $\mathscr{G}_m^0$ as the signal space spanned by $\{g_m^{(0)}(t-kT)\}$. For $x(t) \in \mathscr{H}_0$ and $y(t) \in \mathscr{G}_m^0$, from (4) note that $\int_{-\infty}^{\infty} x(t) y^*(t)\, dt = 0$; therefore $x(t)$ and $y(t)$ are orthogonal signals [19]. Replacing (3) in (2) and operating, we obtain

$$s(t) = s_0(t) + \bar{s}_0(t), \tag{5}$$

where

$$s_0(t) = \sum_k \left[a_k + \sum_{m=1}^{N-1} a_k a_{k-m} \otimes \lambda_k^{(0,m)} \right] h_0(t - kT), \tag{6}$$

$$\bar{s}_0(t) = \sum_k \sum_{m=1}^{N-1} a_k a_{k-m} g_m^{(0)}(t - kT), \tag{7}$$

with operator $\otimes$ denoting convolution. Notice that $s_0(t) \in \mathscr{H}_0$ and $\bar{s}_0(t) \in \cup_{m=1}^{N-1} \mathscr{G}_m^0$; therefore signals $s_0(t)$ and $\bar{s}_0(t)$ are orthogonal (see (4)). Next we focus on $\bar{s}_0(t)$. Without loss of generality we select the second pivoting response as $h_1(t) = g_1^{(0)}(t)$; then (7) can be rewritten as

$$\bar{s}_0(t) = \sum_k a_k \left[a_{k-1} h_1(t - kT) + \sum_{m=2}^{N-1} a_{k-m} g_m^{(0)}(t - kT) \right]. \tag{8}$$

Similarly to (5)–(7), $\bar{s}_0(t)$ can be expressed as

$$\bar{s}_0(t) = s_1(t) + \bar{s}_1(t), \tag{9}$$

where

$$s_1(t) = \sum_k \left[a_k a_{k-1} + \sum_{m=2}^{N-1} a_k a_{k-m} \otimes \lambda_k^{(1,m)} \right] h_1(t - kT), \tag{10}$$

$$\bar{s}_1(t) = \sum_k \sum_{m=2}^{N-1} a_k a_{k-m} g_m^{(1)}(t - kT), \tag{11}$$

with $\lambda_n^{(1,m)}$ chosen to satisfy

$$\int_{-\infty}^{\infty} g_m^{(1)}(t) h_1^*(t - jT)\, dt = 0, \quad \text{for } m = 2, \ldots, N - 1, \tag{12}$$

$$\forall j \in \mathscr{Z}.$$

Thus, note that $\bar{s}_1(t)$ is orthogonal to the signal spaces $\mathscr{H}_0$ and $\mathscr{H}_1$, spanned by $\{h_0(t - kT)\}$ and $\{h_1(t - kT)\}$, respectively. Repeating the processing on (11) and generalizing, we get

$$s(t) = \sum_{n=0}^{N-1} s_n(t) = \sum_{n=0}^{N-1} \sum_k b_k^{(n)} h_n(t - kT), \tag{13}$$

where $h_n(t)$ is the response of the nth channel *path* and with $b_k^{(n)}$ given by

$$b_k^{(n)} = \begin{cases} a_k + \sum_{m=1}^{N-1} a_k a_{k-m} \otimes \lambda_k^{(0,m)} & \text{if } n = 0 \\[2mm] a_k a_{k-n} + \sum_{m=n+1}^{N-1} a_k a_{k-m} \otimes \lambda_k^{(n,m)} & \text{if } 0 < n < N - 1 \\[2mm] a_k a_{k-N+1} & \text{if } n = N - 1, \end{cases} \tag{14}$$

with

$$\int_{-\infty}^{\infty} h_m(t) h_n^*(t - jT)\, dt = 0 \quad m \neq n, \ \forall j \in \mathscr{Z}. \tag{15}$$

From (13) and (15) note that signal components from different paths are orthogonal; that is,

$$\int_{-\infty}^{\infty} s_m(t) s_n^*(t)\, dt = 0, \quad m \neq n. \tag{16}$$

Figure 1 shows two multiple input-single output (MISO) representations of the nonlinear channel. Figure 1(a) shows the signal represented by the traditional Volterra model of (2); Figure 1(b) depicts the orthogonal representation given by (13) and (14).

2.2. Space Channel Compression. The model orthogonalization procedure described in Section 2.1 can be performed in $N!$ different ways with N being the number of kernels. For example, consider a Volterra model with $N = 3$. One way to obtain the orthogonalized model may be realized by selecting the kernels as shown in Figure 2. In this case, the pulse $f_0(t)$ is selected as the pivoting $h_0(t)$ for the first orthogonalization step. Then, $g_1^{(0)}$ and $g_2^{(0)}$ are obtained as the components of $f_1(t)$ and $f_2(t)$ orthogonal to $h_0(t) = f_0(t)$. For the second step, $g_1^{(0)}(t)$ is selected as the pivoting response, so $g_2^{(1)}(t)$ results as the component of $g_2^{(0)}(t)$ orthogonal to $h_1(t) = g_1^{(0)}(t)$. This unique ordering could be identified by the following set of indexes $\mathscr{I} = \{0, 1, 2\}$.

A different result is achieved by selecting the pivoting responses as illustrated in Figure 3. In this situation, $f_1(t)$ is selected as the pivoting response for the first step, and $g_1^{(0)}(t)$ and $g_2^{(0)}(t)$ are the components of $f_0(t)$ and $f_2(t)$ orthogonal to $h_0(t) = f_1(t)$. According to Figure 3, $g_2^{(0)}(t)$ (which is different from $g_2^{(0)}(t)$ shown in Figure 2) is selected as the pivoting response for the second orthogonalization step. Then, $g_2^{(1)}(t)$ will be obtained as the component of $g_1^{(0)}(t)$ orthogonal to $h_1(t) = g_2^{(0)}(t)$. The index set corresponding to the procedure depicted in Figure 3 is $\mathscr{I} = \{1, 2, 0\}$.
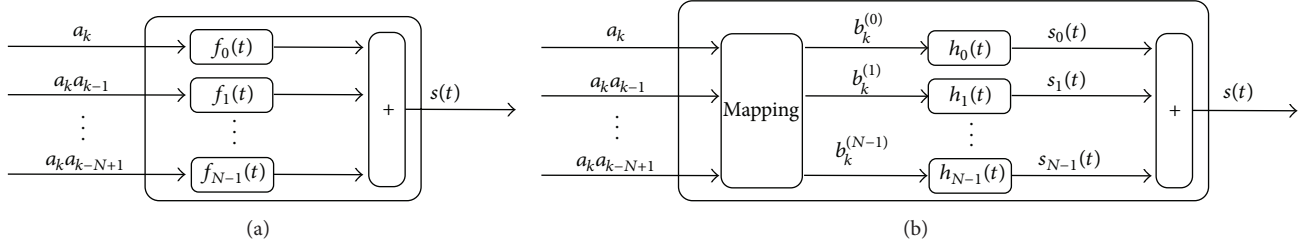
FIGURE 1: MISO model of the nonlinear channel: (a) based on a Volterra representation; (b) based on the spatially compressed orthogonal representation.
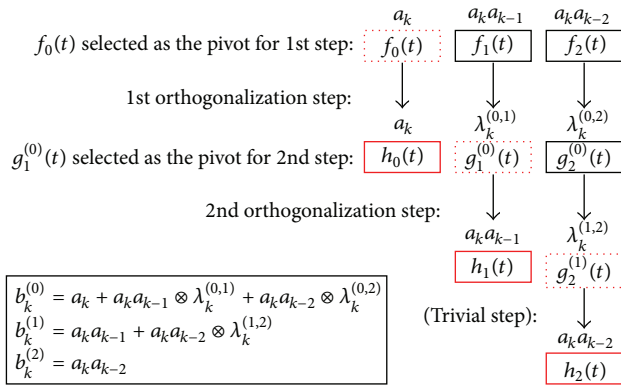

FIGURE 2: Orthogonalization procedure following the ordering $\mathcal{I} = \{0, 1, 2\}$.
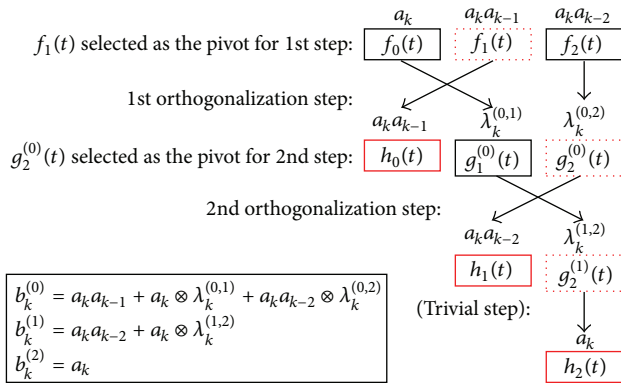

FIGURE 3: Orthogonalization procedure following the ordering $\mathcal{I} = \{1, 2, 0\}$.

Note that the set of responses $\{h_i(t)\}_{i=0}^2$ and sequences $\{b_k^{(i)}\}_{i=0}^2$ resulting from the procedures described in Figures 2 and 3 are two different expansions of the same signal $s(t)$. Any one of the $N!$ possible sets can be selected in the orthogonalization process. *Space compression* of the channel can be achieved concentrating the maximum of the channel energy on a reduced set of paths, independent of the distribution of their individual energy, as explained in the following discussion.

Let P be a given number of paths used to model the channel (e.g., if $P = N$, note that any orthogonalization process

$\mathcal{I}$ shall represent exactly the behavior of the channel). To reduce the complexity of the receiver, it is desirable for P to be as small as possible. On the other hand, to minimize the performance degradation (or the inaccuracy of the channel representation), the orthogonalization process should be carried out in such a way that the most part of the signal energy is concentrated on these P paths. From the above, *for a given* value of P, the *optimum* set $\mathcal{I}_{\text{op}}$ for space compression (i.e., minimal channel distortion modeling) should meet the following condition:

$$\sum_{i=0}^{P-1} \mathcal{E}_{\mathcal{I}_{\text{op}}}^{(i)} \geq \sum_{i=0}^{P-1} \mathcal{E}_{\widehat{\mathcal{I}}}^{(i)}, \quad \forall \widehat{\mathcal{I}} \neq \mathcal{I}_{\text{op}}, \qquad (17)$$

where

$$\mathcal{E}_{\mathcal{I}}^{(i)} = E\left\{\left|b_k^{(i)}\right|^2\right\} \int_{-\infty}^{\infty} \left|h_i(t)\right|^2 dt \qquad (18)$$

is the energy in the ith path for the set $\mathcal{I}$, while $E\{\cdot\}$ denotes expectation. If $P = N$, notice that any orthogonalization process $\mathcal{I}$ will satisfy condition (17). Otherwise, if $P < N$ the criterion guarantees that the signal energy contained in *the complete P paths shall be maximum*, independent of the individual energy distribution among them.

The optimum channel expansion that satisfies (17) can be found by an exhaustive search. Instead, we propose to achieve the orthogonalization process to meet the following criteria:

$$\widehat{\mathcal{E}}^{(0)} \geq \widehat{\mathcal{E}}^{(1)} \geq \cdots \geq \widehat{\mathcal{E}}^{(N-1)}, \qquad (19)$$

with $\widehat{\mathcal{E}}^{(j)} = \int_{-\infty}^{\infty} |h_n(t)|^2 dt$. Condition (19) can be achieved by selecting at each orthogonalization step (ith step, e.g., with $i \in \{1, 2, \ldots, N\}$) the pivoting response ($h_{i-1}(t)$) as the response with highest energy among the *remaining* responses (those orthogonal to $\cap_{n=0}^{i-2} \mathcal{H}_n$ for $t = kT$, $\forall k$). That is, at the first step we select the pivoting response $h_0(t)$ as the Volterra kernel with the highest energy in $\{f_n(t)\}_{n=0}^{N-1}$. At the second step we select the pivoting response $h_1(t)$ as the response with highest energy in the set of responses orthogonal to $\mathcal{H}_0$ (i.e., $g_m^{(0)}(t)$, with $m = 1, 2, \ldots, N-1$). This procedure is repeated in the same way at each step, where the best pivoting response at each stage is selected by an exhaustive search among all the pivot candidates at that stage. The minimization of the energy of the orthogonal responses (see (3), (4), and associated discussion) ensures that condition (19) is met. As we shall show later, condition (19) gives rise to *space* channel compression, which can be exploited to reduce complexity.
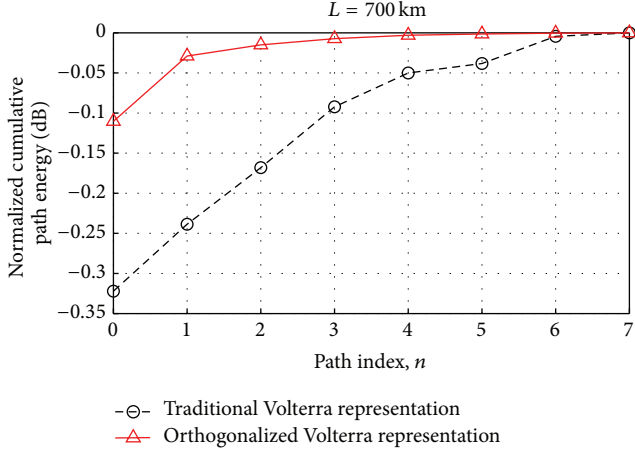
FIGURE 4: Normalized cumulative path energy of the nonlinear IM/DD fiber-optic link with $L = 700$ km and $1/T = 10$ Gb/s.

2.3. Example. As an example, we present results for a 10 Gb/s, 700 km, and IM/DD fiber-optic link (details about the simulated system are given in Section 4.1).

Without loss of generality, we assume that $a_k \in \{\pm 1\}$. Let $\mathcal{S}_n$ be the normalized cumulative path energy of the *traditional* Volterra-series expansion (see (2)) given by

$$\mathcal{S}_n = \frac{1}{E_s} \sum_{m=0}^{n} \int_{-\infty}^{\infty} |f_m(t)|^2 \, dt, \quad n = 1, 2, \ldots, N-1, \quad (20)$$

where E_s is the total energy of the signal component $s(t)$:

$$E_s = \sum_{m=0}^{N-1} \int_{-\infty}^{\infty} |f_m(t)|^2 \, dt. \quad (21)$$

On the other hand, the normalized cumulative path energy of the *orthogonalized* Volterra-series expansion is defined as

$$\mathcal{S}_n = \frac{1}{E_s} \sum_{m=1}^{n} \mathcal{E}_{\mathcal{I}}^{(m)}, \quad n = 0, 1, \ldots, N-1, \quad (22)$$

where $\mathcal{E}_{\mathcal{I}}^{(m)}$ is given by (18) with $\mathcal{I}$ obtained according to the criteria defined by (19). In all cases, note that $\mathcal{S}_n \leq 1 \; \forall n$, with $\mathcal{S}_{N-1} = 1$.

Figure 4 shows the normalized cumulative path energy $\mathcal{S}_n$ for the traditional (20) and orthogonalized (22) Volterra-series representation with $N = 8$. For the orthogonal representation, the pivoting responses are selected according to the criteria in (19). We observe that most energy of the nonlinear signal is concentrated on the first two paths for the orthogonalized Volterra-series expansion (i.e., $\mathcal{E}_{\mathcal{I}}^{(0)} + \mathcal{E}_{\mathcal{I}}^{(1)}$ represents ~99.4% of the total signal energy). The traditional Volterra model requires five paths to capture a similar level of nonlinear signal energy. We highlight that the increase of $\mathcal{S}_n$ from -0.33 dB to -0.12 dB at $n = 0$ is due to the factor $E\{|b_k^{(0)}|^2\} \geq 1$ in (18). Notice that this factor is a measure of the correlation between all nonlinear kernels $f_m(t)$ with $m > 0$ and the linear kernel $f_0(t)$ (e.g., $E\{|b_k^{(0)}|^2\} = 1 \; (N)$ if $f_0(t)$ and

$f_m(t)$ are orthogonal (same) pulses $\forall m$). Thus, we conclude the following for the channel analyzed in the example.

(i) Most energy of the nonlinear kernels with the traditional Volterra model is contained in their projection onto the *linear* signal space spanned by the set $\{h_0(t - kT)\} = \{f_0(t - kT)\}$.

(ii) A channel model with 2 paths captures practically all the signal energy.

(iii) Since $\mathcal{E}_{\mathcal{I}}^{(0)}$ is ~97.5% of the total signal energy (see Figure 4). Therefore, as we shall show later, a receiver that considers the first channel path should be able to achieve a *good* performance.

3. MLSD Receiver for Nonlinear Channels with AWGN

The MLSD receiver chooses the sequence $\{a_k\}$ that minimizes the metric

$$J = \int_{-\infty}^{\infty} |r(t) - s(t)|^2 \, dt. \quad (23)$$

Using (13) and (16) in (23), we get

$$J \propto -2 \sum_{n=0}^{N-1} \Re \left\{ \sum_k b_k^{(n)} \int_{-\infty}^{\infty} r(t) h_n^*(t - kT) \, dt \right\}$$
$$+ \sum_{n=0}^{N-1} \int_{-\infty}^{\infty} |s_n(t)|^2 \, dt \quad (24)$$
$$= -2 \sum_{n=0}^{N-1} \Re \left\{ \sum_k b_k^{(n)} \bar{r}_k^{(n)} \right\} + \sum_{n=0}^{N-1} \int_{-\infty}^{\infty} |s_n(t)|^2 \, dt,$$

where $\Re$ denotes the real part, while

$$\bar{r}_k^{(n)} = \int_{-\infty}^{\infty} r(t) h_n^*(t - kT) \, dt, \quad n = 0, 1, \ldots, N-1. \quad (25)$$

From (24) notice that the computation of the cost function J for every candidate symbol sequence $\{a_k\}$ can be achieved from the samples at the outputs of the matched filters given by (25).

From (1), equation (25) can be rewritten as

$$\bar{r}_k^{(n)} = \bar{s}_k^{(n)} + \bar{z}_k^{(n)}, \quad n = 0, 1, \ldots, N-1, \quad (26)$$

where

$$\bar{s}_k^{(n)} = \int_{-\infty}^{\infty} s(t) h_n^*(t - kT) \, dt,$$
$$\bar{z}_k^{(n)} = \int_{-\infty}^{\infty} z(t) h_n^*(t - kT) \, dt. \quad (27)$$

From (15) note that $E\{\bar{s}_l^{(n)} [\bar{s}_{l-k}^{(m)}]^*\} = 0$ if $m \neq n$ while

$$E\left\{ \bar{z}_l^{(n)} \left[\bar{z}_{l-k}^{(m)} \right]^* \right\} = \begin{cases} N_0 \rho_k^{(n)}, & \text{if } n = m \\ 0, & \text{otherwise,} \end{cases} \quad (28)$$

where

$$\rho_k^{(n)} = \int_{-\infty}^{\infty} h_n(t) h_n^*(t - kT) \, dt. \tag{29}$$

Since the noise is assumed Gaussian, from (28) we conclude that the noise components at the output of the proposed MFB are *spatially* independent. Therefore, the bank of matched filters $\{h_n^*(-t)\}_{n=0}^{N-1}$ followed by a baud rate sampler is called the *space-whitened matched filter* (S-WMF) bank.

3.1. Space-Time Whitened Matched Filter MLSD Receiver.
From the above, we observed that the matched filter bank derived from the new expansion of the nonlinear channel gives rise to *spatially* independent noise components. In order to simplify the implementation of the sequence detector, a space and time whitening filter bank is derived.

Let $\phi_n(t)$ be the impulse response of the filter with Fourier transform (FT) given by

$$\Phi_n(\omega) = \frac{H_n(\omega)}{\gamma^{(n)} M_n(e^{j\omega T})}, \quad n = 0, \ldots, N-1, \tag{30}$$

where $H_n(\omega)$ is the FT of $h_n(t)$ and $\gamma^{(n)}$ and $M_n(z)$ are defined by the folded spectrum factorization $S_n(z) = (\gamma^{(n)})^2 M_n(z) M_n^*(1/z^*)$ with $S_n(z)$ being the Z-transform of the sequence $\rho_k^{(n)}$ given by (29). The set $\{\phi_n(t - kT)\}$ forms an orthonormal basis for the signal space spanned by $\{h_n(t - kT)\}$. Furthermore, we choose $M_n(z)$ to be minimum phase [19]. Define $\tilde{r}_n(t)$ as the projection of $r(t)$ onto the signal space spanned by $\{h_n(t - kT)\}$; that is,

$$\tilde{r}_n(t) = \sum_k \tilde{r}_k^{(n)} \phi_n(t - kT), \tag{31}$$

where $\tilde{r}_k^{(n)} = \int_{-\infty}^{\infty} r(t) \phi_n^*(t - kT) \, dt$. From (15) and following the procedure used in [19, Section 10.2.4], metric (23) can be expressed as

$$J = \int_{-\infty}^{\infty} \left\| \tilde{\mathbf{r}}(t) - \sum_k \mathbf{H}(t - kT) \mathbf{b}_k \right\|^2 dt, \tag{32}$$

where $\mathbf{H}(t)$ is an $N \times N$ diagonal matrix given by

$$\mathbf{H}(t) = \begin{pmatrix} h_0(t) & 0 & \cdots & 0 \\ 0 & h_1(t) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_{N-1}(t) \end{pmatrix}, \tag{33}$$

while

$$\mathbf{b}_k = \begin{bmatrix} b_k^{(0)} & b_k^{(1)} & \cdots & b_k^{(N-1)} \end{bmatrix}^T, \tag{34}$$

$$\tilde{\mathbf{r}}(t) = \begin{bmatrix} \tilde{r}_0(t) & \tilde{r}_1(t) & \cdots & \tilde{r}_{N-1}(t) \end{bmatrix}^T, \tag{35}$$

with $b_k^{(n)}$ and $\tilde{r}_n(t)$ given by (14) and (31), respectively.

On the other hand, from (30) it is possible to show that

$$\sum_k \mathbf{H}(t - kT) \mathbf{b}_k = \sum_k \mathbf{\Phi}(t - kT) \left[\mathbf{M}_k \otimes \mathbf{b}_k \right], \tag{36}$$

where $\mathbf{\Phi}(t)$ and $\mathbf{M}_k$ are $N \times N$ diagonal matrices given by

$$\mathbf{\Phi}(t) = \begin{pmatrix} \phi_0(t) & 0 & \cdots & 0 \\ 0 & \phi_1(t) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \phi_{N-1}(t) \end{pmatrix}, \tag{37}$$

$$\mathbf{M}_k = \begin{pmatrix} \gamma^{(0)} m_k^{(0)} & 0 & \cdots & 0 \\ 0 & \gamma^{(1)} m_k^{(1)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma^{(N-1)} m_k^{(N-1)} \end{pmatrix}, \tag{38}$$

with $m_k^{(n)}$ being the inverse FT of $M_n(e^{j\omega T})$. From (31) notice that (35) can be expressed as

$$\tilde{\mathbf{r}}(t) = \sum_k \mathbf{\Phi}(t - kT) \tilde{\mathbf{r}}_k, \tag{39}$$

where $\tilde{\mathbf{r}}_k = [\tilde{r}_k^{(0)} \cdots \tilde{r}_k^{(N-1)}]^T$ (see (31)). From (32), (36), and (39) it can be shown that MLSD reduces to minimize

$$J = \left\| \tilde{\mathbf{r}}_k - \mathbf{M}_k \otimes \mathbf{b}_k \right\|^2. \tag{40}$$

Let $\tilde{\mathbf{z}}_k$ be the N-dimensional vector with the noise components of the baud rate samples $\tilde{\mathbf{r}}_k$. From (15) and (30), the power spectral density of $\tilde{\mathbf{z}}_k$ results in $\mathbf{S}_{\tilde{z}} = N_0 \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. Therefore, the minimization of (40) can be easily implemented by using a Viterbi detector with *multidimensional* Euclidean branch metrics. Figure 5 shows a block diagram of the ST-WMF-MLSD receiver with *dimension P* (i.e., P is the number of filters in the bank). If $P = N$, all the paths of the nonlinear channel are used by the receiver.

3.2. Complexity Considerations.
The ST-WMF-MLSD consists of a filter bank followed by a Viterbi decoder (VD) and a channel estimation stage. Although the computational load of the latter may be important, its complexity is not critical since it can be implemented at a low frequency rate[3]. On the other hand, the dimension of the front-end (P) is expected to be very low in general as a result of the spatial energy compression achieved by the ST-WMF-MLSD. Then, $P \leq 2$ is a reasonably good estimation for applications in IM/DD optical systems and the computational complexity of the filter bank is reduced to the one of 2 *linear* filters. Therefore, *the implementation complexity of ST-WMF-MLSD is dominated by the VD*. Practical aspects related to high-speed implementations of VD have been widely investigated in the past literature (e.g., see [12, 22] for more details). The computational load (usually measured in number of multiplications and comparisons) and storage requirements of the VD depend directly on the number of states. Therefore, we adopt the number of states of the VD as the measure of complexity in order to compare the MLSD architectures investigated in this work.
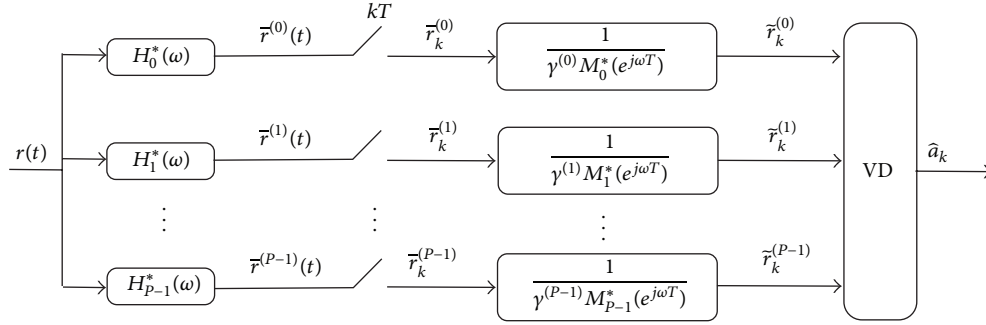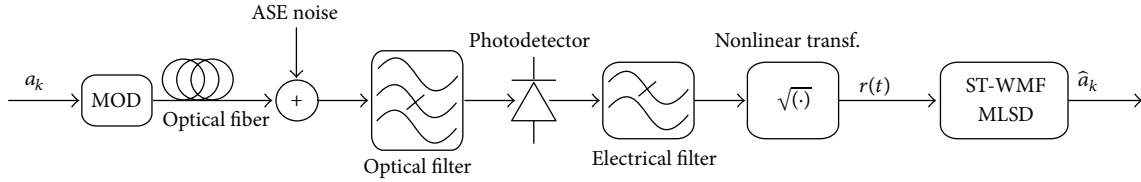
FIGURE 5: Block diagram of the P-dimensional ST-WMF-MLSD receiver.

FIGURE 6: IM/DD fiber-optic system with ST-WMF-MLSD receiver.

4. Performance Evaluation in IM/DD Optical Systems

Next we analyze the proposed ST-WMF-MLSD receiver in transmissions over IM/DD fiber-optic systems with on-off keying (OOK) modulation. We focus on two key aspects of ST-WMF-MLSD: its performance (in comparison with current solutions based on OS-MLSD), and its ability to reduce complexity (e.g., number of states of VD). Complexity reduction is possible owing to (i) the minimum-phase property of the equivalent channel response provided by ST-WMF and (ii) condition (19). The latter gives rise to *space compression*, which reduces the ST-WMF dimension, P. This is achieved by using the *most significant P paths* of the nonlinear channel.

Figure 6 depicts the optical system under consideration. The transmitter modulates the intensity of the transmitted signal using NRZ-OOK modulation. The standard single mode fiber (SMF) introduces CD and PMD, as well as attenuation. Optical amplifiers are deployed periodically along the fiber to compensate the attenuation, also introducing amplified spontaneous emission (ASE) noise in the signal. ASE noise is modeled as AWGN in the optical domain. The received optical signal is filtered by an optical filter and then converted to a current with a PIN diode or avalanche photodetector. The resulting photocurrent is filtered by an electrical filter. The noise component after the electrical filtering is non-Gaussian and signal-dependent [16]. Therefore, the electrical signal is first processed by a memoryless nonlinear transformation. It has been found that after a square root transformation, the noise can be assumed Gaussian and signal-independent [23, 24]. Furthermore, channel nonlinearities can also be reduced by using the square root transformation [25], which improves the space compression used to reduce the receiver dimension (i.e., most

of the channel energy is concentrated on the linear kernel). The split-step Fourier method [26] is used to compute the propagation of optical signals through the fiber. Oversampled linear and nonlinear kernels are extracted from the electrical signal after the square root transformation. The oversampling factor $R = T/T_s$ depends on various parameters of the communication system (e.g., optical power, fiber length, etc.). In our case, we have found that an oversampling factor of $R = 16$ is good enough to accurately model the system; that is, no improvement was appreciated by increasing the sampling rate for the entire set of conditions considered for this work. Then, we compute $h_n(kT_s)$ and $\lambda_k^{(n,m)}$ according to (17), and the symbol rate channel response matrix $\mathbf{M}_k$ can be easily obtained from (38). Since the noise after the square root transformation is approximately Gaussian and signal-independent [27], the theory proposed in [28] is used to evaluate the bit error probability[4]. All the kernels of the nonlinear channel are used to compute the error probability, independent of the receiver dimension, P. Data rate is $1/T = 10$ Gb/s and the transmitted pulse shape has an unchirped Gaussian envelope $e^{-t^2/2T_0^2}$ with $T_0 = 36$ ps. We use a Lorentzian optical filter and a fourth-pole Butterworth electrical filter with bandwidths of 15 and 10 GHz, respectively. The fiber dispersion is $D = 17$ ps/(nm-km).

4.1. IM/DD Systems with CD and PMD. Typically, the performance of equalization stages in fiber optic transmissions is evaluated by using the optical signal-to-noise ratio (OSNR) required to achieve a given BER (e.g., see [10, 11]). The target BER is around 10^{-3}, which corresponds to the value required at the input of the forward error correction (FEC) code to achieve the error rate expected by the application (e.g., ~ 10^{-15}) [29].
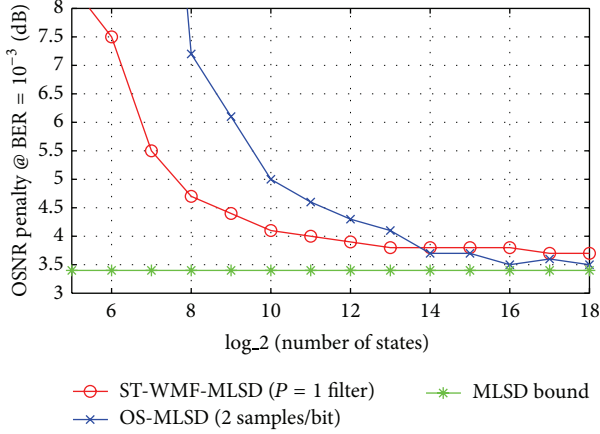
FIGURE 7: OSNR penalty with respect to B2B at BER = 10^{-3} versus number of states of the VD for $L = 700$ km and DGD = 0 ps.

Figure 7 depicts the OSNR penalty with respect to a B2B system at a bit-error rate (BER) of 10^{-3} versus the number of states of the VD for $L = 700$ km. We present results for ST-WMF-MLSD with $P = 1$ (i.e., only one filter in the bank). For OS-MLSD with 2 samples/bit, the reduction of states is achieved by truncation and optimization of the sampling phase in order to minimize BER (8 uniformly distributed phases in the interval $T/2$ were tested). From Figure 7, we verify that the number of states of the VD at a penalty of 4.6 dB can be reduced from 2048 to 256 with ST-WMF-MLSD and $P = 1$. Notice that this performance is achieved by using a VD with one sample per bit [11]. Furthermore, we emphasize that these benefits greatly outperform the extra complexity required by the linear filter and the channel estimator (implementation details are omitted here due to space limitations). From Figure 7 we can also see that the performance degradation achieved with $P = 1$ with respect to the ideal MLSD is only ~0.3 dB. This result can be understood from the analysis of Section 2.2, where it has been observed that 97.5% of the total signal energy is contained in the first channel path.

The performance of ST-WMF-MLSD in the presence of CD and first-order PMD (i.e., differential group delay or DGD) is investigated. Figure 8 depicts the OSNR penalty versus the number of states of the VD for a 700 km fiber link with two values of DGD: 25 and 50 ps. We evaluate the OS-MLSD with 2 samples/bit and ST-WMF-MLSD with $P = 2$. As expected, both receivers tend to the same performance as the number of states of the VD increases. We also note that the benefits of the ST-WMF-MLSD reduce when the DGD increases[5]. Nevertheless, from Figure 8 notice that the number of states of the VD at a penalty of 6 dB can be reduced 8 and 4 times with ST-WMF-MLSD at 25 and 50 ps DGD, respectively. These results show that the ST-WMF-MLSD is still an attractive solution to reduce complexity in transmissions over fiber-optic channels in the presence of PMD.

5. Impact of the Imperfect Channel Knowledge

Performance evaluation of the ST-WMF-MLSD has been achieved by assuming a perfect knowledge of the channel. In the following, we analyze the impact of the channel estimation inaccuracy on the performance of the ST-WMF-MLSD architecture in transmissions over IM/DD optical systems. This study will show that the performance degradation in ST-WMF-MLSD receivers, with $P = 1$, caused by an imperfect channel estimation, is low (~0.2 dB) and similar to that achieved by oversampled OS-MLSD receivers in the $L = 700$ km fiber link used in the example of Figure 7.

The estimation of the oversampled linear and nonlinear kernels is required to implement both MLSD-based receivers. Let $R = T/T_s$ be the oversampling factor. Based on the polyphase filter representation of the oversampled channel response (see Figure 9), the received samples can be expressed as

$$r_n^{(i)} = \sum_k a_k f_0^{(i)}[n-k] + \sum_k \sum_{m=1}^{N-1} a_k a_{k-m} f_m^{(i)}[n-k] + z_n^{(i)}, \quad (41)$$

where $r_n^{(i)} = r(nT + iT_s)$, $z_n^{(i)} = z(nT + iT_s)$, and $f_m^{(i)}[n] = f_m(nT + iT_s)$ with $i = 0, \ldots, R-1$. Notice that there are R different sequences sampled at the baud rate, corresponding to R different sampling phases.

Equation (41) can be rewritten as

$$r_n^{(i)} = \sum_k a_{n-k} f_0^{(i)}[k] + \sum_k \sum_{m=1}^{N-1} a_{n-k} a_{n-k-m} f_m^{(i)}[k] + z_n^{(i)}. \quad (42)$$

Since symbols a_k are assumed zero-mean and i.i.d. real random variables with $\sigma_a^2 = E\{|a_k|^2\}$, we can verify that

$$f_0^{(i)}[k] = \frac{1}{\sigma_a^2} E\left\{r_n^{(i)} a_{n-k}\right\},$$

$$f_m^{(i)}[k] = \frac{1}{\sigma_a^4} E\left\{r_n^{(i)} a_{n-k} a_{n-k-m}\right\}, \quad (43)$$

$$i = 0, \ldots, R-1.$$

From (43), a simple estimator of the oversampled linear and nonlinear kernels can be implemented with an averaging filter as follows:

$$\hat{f}_0^{(i)}[k] = \frac{1}{\sigma_a^2 L_A} \sum_{n=n_0}^{n_0+L_A-1} r_n^{(i)} \hat{a}_{n-k},$$

$$\hat{f}_m^{(i)}[k] = \frac{1}{\sigma_a^4 L_A} \sum_{n=n_0}^{n_0+L_A-1} r_n^{(i)} \hat{a}_{n-k} \hat{a}_{n-k-m}, \quad (44)$$

where L_A is the length of the averaging filter, n_0 is an arbitrary time index, and $\hat{a}_k$ is the detected symbol[6]. The accuracy of the channel estimation given by (44) depends on the precision of the decisions $\hat{a}_k$, the length of the averaging filter L_A, and the channel noise power. We consider that decisions provided by the forward error correction (FEC) decoder are available; therefore the effect of decision errors can be
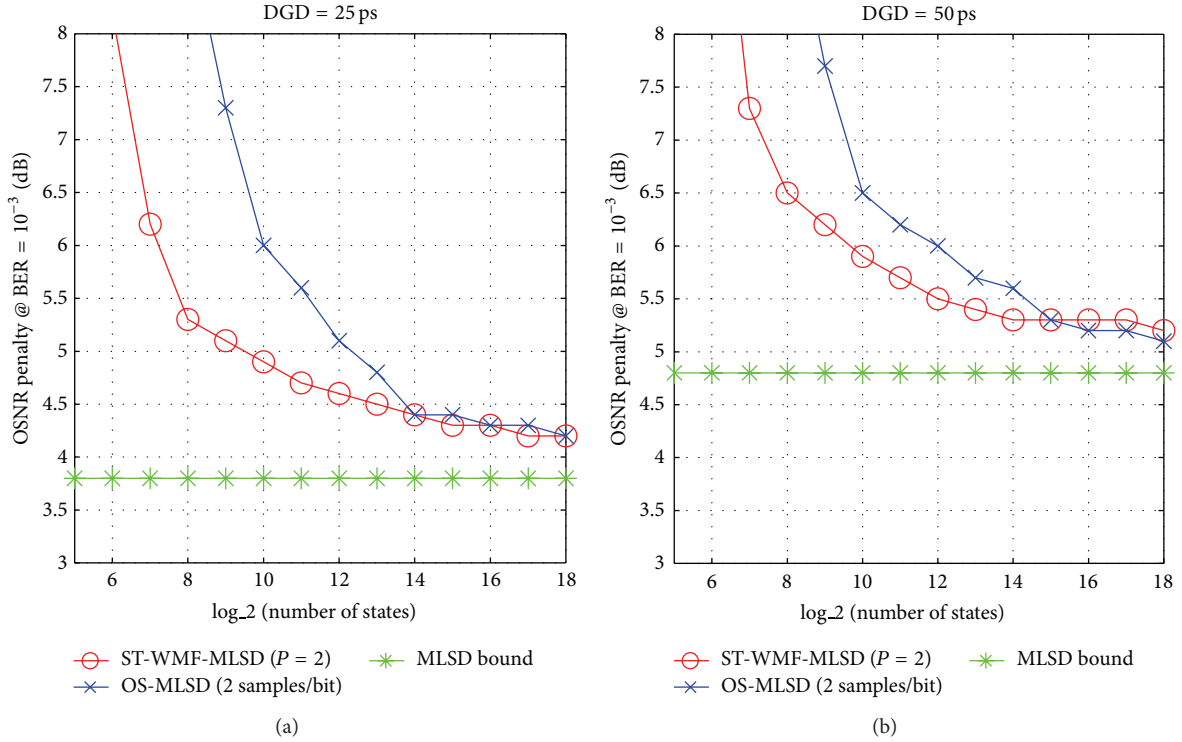
FIGURE 8: OSNR penalty respect to B2B at BER = 10^{-3} versus number of states of the VD for $L = 700$ km. DGD = 25 ps (a) and 50 ps (b).
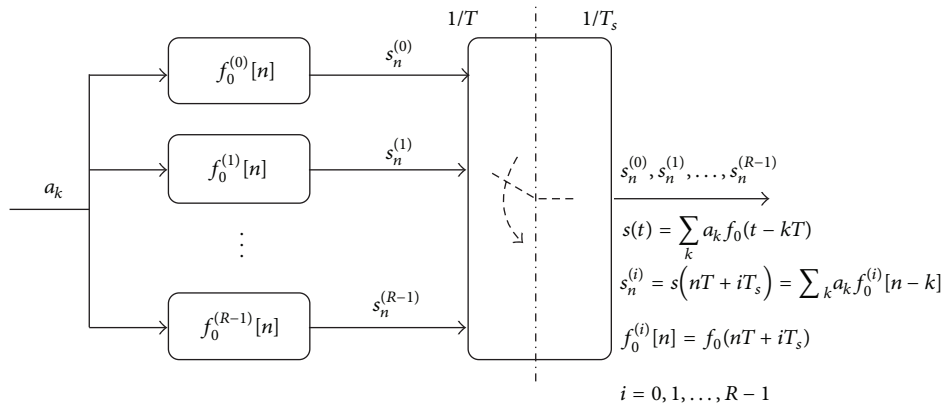

FIGURE 9: Example of the polyphase filter representation of an oversampled linear channel with $R = T/T_s$.

neglected (i.e., $\hat{a}_k = a_k$). We highlight that this assumption is still valid if pre-FEC decisions are used as a result of the *low* bit error rates experienced in this link (e.g., ~10^{-3}). From the above, we conclude that the goodness of the estimates (44) shall mainly depend on the filter length L_A and the channel noise power.

The precision of (44) improves as the value of L_A increases. On the other hand, the maximum value of L_A shall be imposed by the speed of temporal variations of the fiber optic channel. As a result of its dependence on stress and vibrations, as well as on random changes in the state of polarization of the laser, PMD is nonstationary. Fluctuations with a time scale of a hundreds of microseconds have been considered in previous works (e.g., [30]). Therefore, the

response time of channel estimation algorithms for PMD mitigation must be less than 1 ms (in practice a response time less than 100 μs is required [27]). This imposes, for example, that the bandwidth of the averaging filter (~$1/(4L_A T)$) should be ≥20 kHz in order to efficiently track the channel variation.

The received signal *seen* by an MLSD receiver in the presence of imperfect knowledge of the channel dispersion can be expressed as

$$
\begin{aligned}
\tilde{r}_n^{(i)} &= \sum_k a_{n-k} \left(f_0^{(i)}[k] + \sum_{m=1}^{N-1} a_{n-k-m} f_m^{(i)}[k] \right) \\
&\quad + z_n^{(i)} + \hat{z}_n^{(i)} \\
&= s_n^{(i)} + z_n^{(i)} + \hat{z}_n^{(i)},
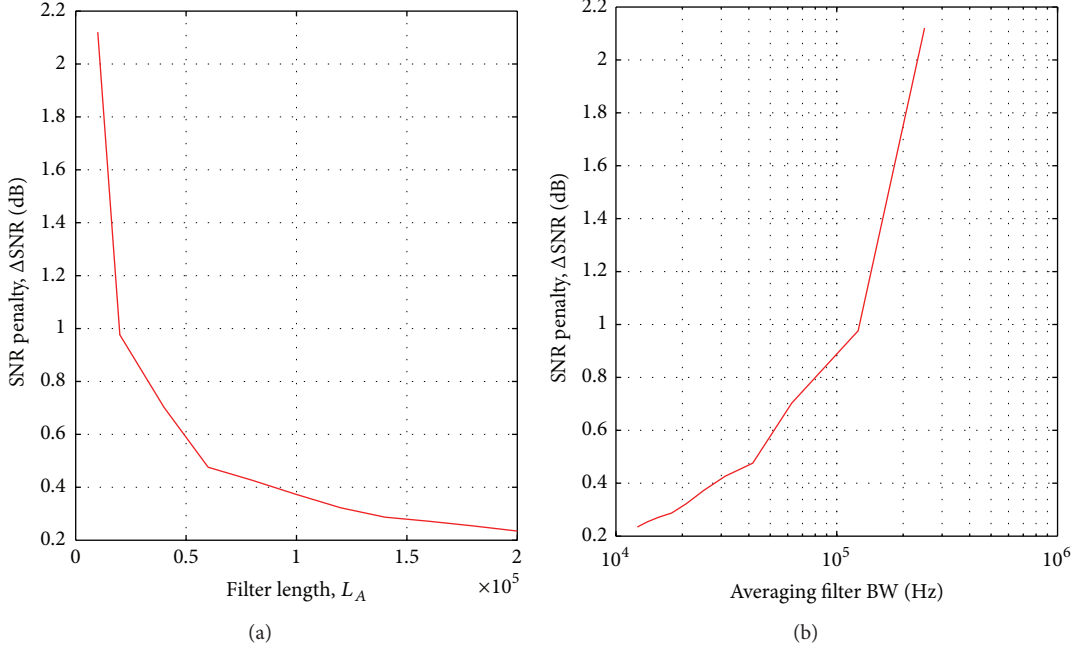\end{aligned}
\tag{45}
$$

(a)

(b)

FIGURE 10: SNR penalty caused by imperfect channel estimation versus length of the averaging filter (a) and the bandwidth of the averaging filter (b).

where $s_n^{(i)} = s(nT + iT_s)$ and $\hat{z}_n^{(i)} = \hat{s}(nT + iT_s) - s(nT + iT_s)$ is the *estimation error* component, while $\hat{s}(nT + iT_s)$ is the synthesized signal obtained from (44).

Figure 10(a) shows the SNR penalty caused by the imperfect channel estimation as a function of L_A obtained from computer simulations. We consider $1/T = 10$ GHz, $R = 16$, and the fiber link with $L = 700$ km, as used in Figure 7. The SNR penalty caused by an imperfect channel estimation is computed as

$$\Delta\text{SNR} = \frac{\sigma_z^2 + \sigma_{\hat{z}}^2}{\sigma_z^2}, \qquad (46)$$

where σ_z^2 is the channel noise power required to achieve a BER $= 10^{-3}$ with an unconstrained complexity OS-MLSD receiver (i.e., the VD uses as many states as required) and $\sigma_{\hat{z}}^2$ is the variance of the estimation error component[7]. Notice that the penalty is ~0.4 dB for $L_A = 10^5$. This value of L_A represents a BW of ~$1/(4L_A T) = 25$ kHz (see Figure 10(b)). Assuming that the estimation error is white Gaussian noise with power $\sigma_{\hat{z}}^2$, from Figure 10(a) we infer that the SNR penalty caused by an imperfect channel knowledge in OS-MLSD receivers with $L_A = 10^5$ should be $\lesssim \Delta\text{SNR} \sim 0.4$ dB.

Figure 11 depicts the OSNR penalty at BER $= 10^{-3}$ versus the number of states of the Viterbi detector (VD) with $L = 700$ km. We present results with perfect knowledge of the fiber dispersion (denoted as $L_A = \infty$), and for imperfect channel estimation with $L_A = 10^5$. We see that the mean penalty caused by inaccuracies of channel estimation agrees with that expected from Figure 10 with $L_A = 10^5$ (i.e., ~0.14 dB < ΔSNR ~ 0.4 dB). Furthermore, we observe that the impact of imperfect channel knowledge on the

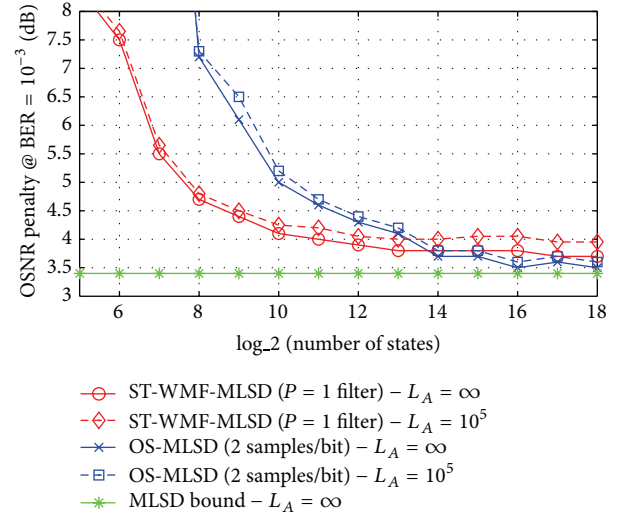
FIGURE 11: OSNR penalty at BER $= 10^{-3}$ versus number of states of the VD with $L = 700$ km.

performance is similar in both MLSD receivers (i.e., ~0.14 dB and ~0.18 dB for OS and ST-WMF, resp.). This result can be understood from the fact that the filters $h_0^*(-t)$ and $m_k^{(0)}$ are computed from the samples of the estimated linear kernel $\hat{f}_0(t)$. Taking into account that the energy of the linear component is significantly higher than the nonlinear kernels [25] (see Section 2.2), an accurate estimation of $h_0(t)$ can be achieved for the channel considered. Then, we infer that the energy loss of the signal component at the output of $\hat{m}_0(t)$ will be small. Therefore, and based on (45), notice that the

performance of OS-MLSD and ST-WMF-MLSD with $P = 1$ should degrade in a similar way.

6. Conclusions

New results on the recently proposed ST-WMF-MLSD non-linear receiver have been presented in this paper. These results are the following: (1) the *space compression* property of the factorization introduced in [18] has been analyzed in detail; (2) the performance of the ST-WMF-MLSD in IM/DD fiber optic systems in the combined presence of CD and PMD has been evaluated, and (3) it has been shown that the performance degradation caused by an imperfect channel estimation and tracking is low and similar to that achieved by existing MLSD schemes. These features make the ST-WMF-MLSD a good architecture for receivers for long distance IM/DD fiber-optic links.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

Endnotes

1. Some commercial implementations of VD at 10 Gb/s with 4, 8, and 16 states are available in 90 nm CMOS technology [31–33]. It should be possible to implement VD's with 64 or 128 states by using the available 28 nm technology.

2. The Volterra model can be easily extended to include higher order terms (e.g., third order kernels $f_m(t)$, with input sequence $a_k a_{k-m} a_{k-n}$ and $n > 0$, could be included in the expansion). To keep the notation simple, only second order kernels are used in the derivations throughout this paper. However, numerical results incorporate nonlinear kernels of order higher than two.

3. As a result of its dependence on stress and vibrations, as well as on random changes in the state of polarization of the laser, PMD is nonstationary. Fluctuations with a time scale of a few milliseconds have been observed in PMD measurements [27]. Thus, the response time of the channel estimation schemes for PMD mitigation must be less than 1 ms. In practice, a response time less than 100 μs is required. Therefore, the channel estimation stage could be easily implemented by using current technology.

4. We did not run Monte Carlo simulations of the communication systems. The numerical results are semianalytic, in the sense that they are theoretic estimations assisted by numeric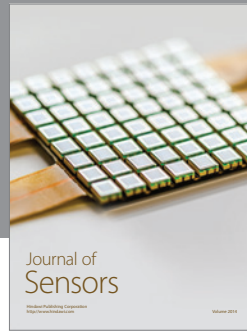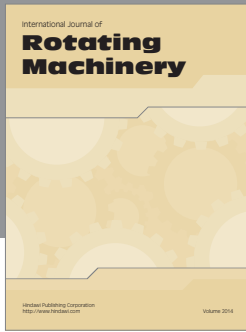al simulations. The probability of error is estimated according to the theory proposed in [28]; that is, the first terms of the union bound (those with the lowest distance) are used to estimate the probability of error. Numerical simulations are used to find the distances of the error-events, their Hamming weights, and a priori probabilities. These results are then introduced in the formulas for the probability of error estimation as reported in [28].

5. In order to explain this result, consider a fiber channel with DGD = 100 ps only (i.e., no CD). Since DGD $\sim T$, the IM/DD system approximately behaves as a duobinary channel; therefore the *time* compression achieved by the ST-WMF-MLSD will be *negligible*.

6. In order to improve the channel tracking capability, an efficient implementation of the channel estimator with the well-known LMS algorithm may be preferred [12]. Nevertheless, the objective is to shed light on the impact of imperfect knowledge of the fiber dispersion on the orthogonalized Volterra model. Therefore, practical aspects of the receiver architecture (e.g., buffers, number of taps of the WF, finite precision arithmetic effects, etc.) are not considered.

7. The mean penalty of 20 runs with different seeds of the random number generator is presented.

References

[1] S. Benedetto and E. Biglieri, *Principles of Digital Transmission with Wireless Applications*, Kluwer Academic, 1999.

[2] A. Vannucci and R. Raheli, "Sequence detection in nonlinear channels: a convenient alternative to analog predistortion," *IEEE Transactions on Communications*, vol. 50, no. 9, pp. 1515–1524, 2002.

[3] M. F. Mesiya, P. J. M. Lane, and L. L. Campbell, "Maximum likelihood sequence estimation of binary sequences transmitted over bandlimited nonlinear channels," *IEEE Transactions on Communications*, vol. 25, no. 7, pp. 633–643, 1977.

[4] V. K. Dubey and D. P. Taylor, "Maximum likelihood sequence detection for QPSK on nonlinear, band-limited channels," *IEEE Transactions on Communications*, vol. 34, no. 12, pp. 1225–1235, 1986.

[5] T. E. Stern, "Viterbi algorithm receivers for nonlinear satellite channels," in *Proceedings of the 11th Asilomar Conference on Circuits, Systems and Computers, Conference Record*, pp. 219–224, IEEE, November 1977.

[6] O. E. Agazzi and V. Gopinathan, "The impact of nonlinearity on electronic dispersion compensation of optical channels," in *Proceedings of the Optical Fiber Communication Conference (OFC '04)*, vol. 1, pp. 434–436, February 2004.

[7] Q. Yu and A. Shanbhag, "Electronic data processing for error and dispersion compensation," *Journal of Lightwave Technology*, vol. 24, no. 12, pp. 4514–4525, 2006.

[8] C. Xia and W. Rosenkranz, "Electrical dispersion compensation for different modulation formats with optical filtering," in *Proceedings of the National Fiber Optic Engineers Conference, Optical Fiber Communication Conference (OFC '06)*, p. OWR2, March 2006.

[9] C. Xia and W. Rosenkranz, "Nonlinear electrical equalization for different modulation formats with optical filtering," *Journal of Lightwave Technology*, vol. 25, no. 4, pp. 996–1001, 2007.

[10] G. Bosco and P. Poggiolini, "Long-distance effectiveness of MLSE IMDD receivers," *IEEE Photonics Technology Letters*, vol. 18, no. 9, pp. 1037–1039, 2006.

[11] S. J. Savory, Y. Benlachtar, R. I. Killey et al., "IMDD transmission over 1,040 km of standard single-mode fiber at 10 Gbit/s using a one-sample-per-bit reduced-complexity MLSE receiver," in *Proceedings of the Optical Fiber Communication Conference (OFC '07)*, March 2007.

[12] O. E. Agazzi, M. R. Hueda, D. E. Crivelli et al., "A 90 nm CMOS DSP MLSD transceiver with integrated AFE for electronic dispersion compensation of multimode optical fibers at 10 Gb/s," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 12, pp. 2937–2957, 2008.

[13] S. Kasturia and J. H. Winters, "Techniques for high-speed implementation of nonlinear cancellation," *IEEE Journal on Selected Areas in Communications*, vol. 9, no. 5, pp. 711–717, 1991.

[14] K. K. Parhi, "Pipelining in algorithms with quantizer loops," *IEEE transactions on circuits and systems*, vol. 38, no. 7, pp. 745–754, 1991.

[15] K. K. Parhi and D. G. Messerschmitt, "Pipeline interleaving and parallelism in recursive digital filters I. Pipelining using scattered look-ahead and decomposition," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 1099–1117, 1989.

[16] O. E. Agazzi, M. R. Hueda, H. S. Carrer, and D. E. Crivelli, "Maximum-likelihood sequence estimation in dispersive optical channels," *Journal of Lightwave Technology*, vol. 23, no. 2, pp. 749–763, 2005.

[17] W. Yan, T. Tanaka, B. Liu et al., "100 Gb/s optical IM-DD transmission with 10G-class devices enabled by 65 GSamples/s CMOS DAC core," in *Proceedings of the Optical Fiber Communication Conference and Exposition*, March 2013.

[18] G. N. Maggio, M. R. Hueda, and O. E. Agazzi, "Reduced complexity MLSD receivers for nonlinear optical channels," *IEEE Photonics Technology Letters*, vol. 26, no. 4, pp. 398–401, 2014.

[19] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communications*, KAP, 2004.

[20] O. Agazzi, D. G. Messerschmitt, and D. A. Hodges, "Nonlinear echo cancellation of data signals," *IEEE Transactions on Communications*, vol. 30, no. 11, pp. 2421–2433, 1982.

[21] W. Chung, "Channel estimation methods based on volterra kernels for MLSD in optical communication systems," *IEEE Photonics Technology Letters*, vol. 22, no. 4, pp. 224–226, 2010.

[22] P. J. Black and T. H.-Y. Meng, "A 1-Gb/s, four-state, sliding block Viterbi decoder," *IEEE Journal of Solid-State Circuits*, vol. 32, no. 6, pp. 797–805, 1997.

[23] M. R. Hueda, D. E. Crivelli, and H. S. Carrer, "Performance of MLSE-based receivers in lightwave systems with nonlinear dispersion and amplified spontaneous emission noise," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '04)*, vol. 1, pp. 299–303, November 2004.

[24] H. S. Carrer, D. E. Crivelli, and M. R. Hueda, "Long-distance effectiveness of MLSE IMDD receivers," *Latin American Applied Research*, vol. 35, pp. 99–104, 2005.

[25] J. Prat, A. Napoli, J. M. Gene, M. Omella, P. Poggiolini, and V. Curri, "Square root strategy: a novel method to linearize an optical communication system with electronic equalizers," in *Proceedings of the 31st European Conference on Optical Communication (ECOC '05)*, vol. 3, pp. 713–714, September 2005.

[26] G. P. Agrawal, *Nonlinear Fiber Optics*, Academic Press, New York, NY, USA, 2007.

[27] M. R. Hueda, D. E. Crivelli, H. S. Carrer, and O. E. Agazzi, "Parametric estimation of IM/DD optical channels using new closed-form approximations of the signal PDF," *Journal of Lightwave Technology*, vol. 25, no. 3, pp. 957–975, 2007.

[28] O. E. Agazzi and N. Seshadri, "On the use of tentative decisions to cancel intersymbol interference and nonlinear distortion (with application to magnetic recording channels)," *IEEE Transactions on Information Theory*, vol. 43, no. 2, pp. 394–408, 1997.

[29] D. A. Morero, M. A. Castrillon, F. A. Ramos, T. A. Goette, O. E. Agazzi, and M. R. Hueda, "Non-concatenated FEC codes for ultra-high speed optical transport networks," in *Proceedings of the 54th Annual IEEE Global Telecommunications Conference: "Energizing Global Communications" (GLOBECOM '11)*, pp. 1–5, December 2011.

[30] S. M. R. Motaghiannezam, J. Cho, D. Tauber et al., "Single chip 46 Gb/s DP-QPSK digital clock recovery and channel equalization performance in the presence of CD, PMD, and ultra-fast SOP rotation rates exceeding 20 krad/s," in *Proceedings of the Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC '13)*, pp. 1–3, Anaheim, Calif, USA, March 2013.

[31] D. Fritzsche, D. Breuer, L. Schurer, A. Ehrhardt, H. Oeruen, and C. G. Schaffer, "Experimental investigation of real time 10 Gbit/s MLSE equalizer using 4-states and 16-states Viterbi detector," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '09)*, pp. 1–5, November 2009.

[32] O. Agazzi, D. Crivelli, M. Hueda et al., "A 90nm CMOS DSP MLSD transceiver with integrated AFE for electronic dispersion compensation of multi-mode optical fibers at 10Gb/s," in *Proceedings of the IEEE International Solid-State Circuits Conference (ISSCC '08)*, pp. 232–609, IEEE, San Francisco, Calif, USA, February 2008.

[33] T. Kupfer, C. Dorschky, M. Ene, and S. Langenbach, "Measurement of the performance of 16-states MLSE digital equalizer with different optical modulation formats," in *Proceedings of the Optical Fiber Communication/National Fiber Optic Engineers Conference (OFC/NFOEC '08)*, p. PDP13, February 2008.