# State-Complexity Reduction in MLSD Receivers for Optical Communications With Direct Photodetection

Michele Franceschini, Gianluigi Ferrari, *Member, IEEE*, Riccardo Raheli, *Member, IEEE*, Fausto Meli, and
Andrea Castoldi

*Abstract*—We investigate the impact of state-complexity reduction on the performance of maximum likelihood sequence detection (MLSD) receivers for direct-photodetection long-haul optical communication systems affected by uncompensated chromatic dispersion (CD). We directly compare two possible approaches: (i) detection through a simple "brute-force" state-complexity reduction strategy and (ii) a more structured reduced-state sequence detection (RSSD) strategy. The performance of both state-complexity reduction techniques is evaluated considering two realistic optical transmission schemes, based on on-off keying (OOK) and optical duobinary modulation (ODBM), respectively. The detection algorithms are characterized considering the impact of the timing offset, the quantization scheme, and the amount of uncompensated CD. As one would expect, for a given number of states in MLSD receivers, the schemes based on RSSD exhibit better performance with respect to those based on simple brute-force state-complexity reduction. However, we show that MLSD schemes based on the use of brute-force state-complexity reduction are characterized by a better complexity/performance trade-off for low/medium CD values, whereas RSSD is the best choice for high CD values.

*Index Terms*—Maximum likelihood sequence detection (MLSD), on-off keying (OOK), optical duobinary modulation (ODBM), optical fiber communications, reduced-state sequence detection (RSSD).

## I. INTRODUCTION

DUE to the increasing availability of low-cost high-speed integrated circuits, maximum likelihood sequence detection (MLSD) for optical transmission schemes is gathering the interests of both research institutions and industry. Recently, several works have characterized various aspects of MLSD schemes for optical communications based on direct photodetection and operating in the electrical domain. In particular, in [1]–[5], MLSD schemes based on the use of the Viterbi algorithm [6] (VA) have been characterized in terms of tolerance against chromatic dispersion (CD) and polarization mode dispersion (PMD). In these papers, various models for the branch metrics of the VA have been considered. However, to account for a large system memory, due, for example, to severe values of uncompensated CD, a standard VA requires a number of states which is an exponential function of the system memory. In [7], the authors analyze the performance of MLSD schemes with reduced state-complexity. More precisely, a reduced state is obtained by selecting a subset of the "most representative" postcursor bits (selected by minimization of the bit error rate, BER). In [8], the ultimate performance limits, in terms of achievable *information rate* [9], of direct-detection optical communication systems have been presented. These results suggest that very large values of uncompensated CD can in principle be tolerated through the use of electronic signal processing based on suitable detection algorithms and forward error correction schemes. However, even if electronic signal processing is attractive from a theoretical viewpoint, the results in [8] indicate that the required state complexity increases tremendously due to a large system memory. The use of complexity reduction techniques becomes, therefore, crucial.

In this paper, we focus on complexity reduction techniques for MLSD receivers suitable for optical communication systems with direct photodetection, by expanding upon preliminary results in [10]. More precisely, we investigate and compare two different *state-complexity reduction* techniques: (i) "brute-force" state-complexity reduction techniques, which consist of a hard truncation of the system memory assumed by the VA, and (ii) classical reduced-state sequence detection (RSSD) techniques [11]–[13]. Simple brute-force state complexity reduction has previously been investigated (see, e.g., [7]). Nonetheless, in this paper, we extend previous results by giving a rigorous description on how state reduction is performed, by considering an extensive performance comparison between brute-force and RSSD state-complexity reduction techniques, and by analyzing the impact of the modulation format and several system parameters. As representative case studies for numerical analysis, we consider schemes with two modulation formats: on-off keying (OOK) with non-return-to-zero (NRZ) pulses and optical duobinary modulation (ODBM). Since our focus is on state-complexity reduction, we limit the analysis to a linear optical fiber model. The data rate is assumed to be equal to 10 Gb/s in both cases. The performance of the considered MLSD algorithms with reduced state complexity is characterized taking into account several
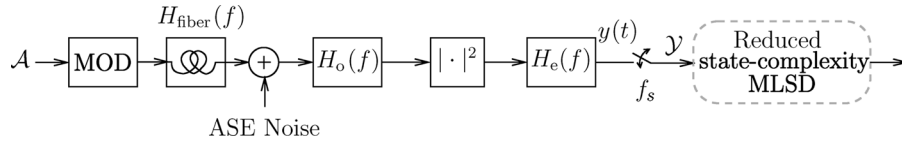
Fig. 1.  Optical communication system model with direct photodetection.

important parameters, namely: timing offset, quantization level, and uncompensated CD. The obtained results show that, whenever the effective system memory is small (i.e., low CD values), the use of brute-force state-complexity reduction techniques is an appealing choice from the viewpoint of the complexity-performance trade-off, whereas for larger values of the effective system memory (i.e., medium/high CD values) the use of RSSD techniques guarantees a good tradeoff.

This paper is structured as follows. In Section II, the considered communication system models are discussed. In Section III, preliminaries on MLSD are given. In Section IV, state-complexity reduction techniques are discussed. In Section V, the simulation setup is presented. In Section VI, the impact of the timing offset on the system performance is investigated. In Section VII, the impact of the quantization scheme of the observed received signal on the system performance is investigated. In Section VIII, the performance of the proposed reduced state-complexity MLSD algorithms is characterized in terms of tolerance against uncompensated CD. In Section IX, the obtained results are interpreted and discussed. Section X concludes the paper.

## II. COMMUNICATION SYSTEM MODEL

In Fig. 1, an optical communication system model with direct photodetection is shown. At the input of the system, a sequence of $K$ independent and equally likely information bits $\mathcal{A} = \{a_k\}_{k=1}^{K}$ is assumed. We consider two different optical modulation schemes: (i) OOK with NRZ pulses and (ii) ODBM. In the NRZ-OOK case, the modulator (MOD) comprises an NRZ pulse shaping filter, an electrical low-pass filter, and a Mach-Zehnder device in push-pull configuration. The input/output relation of the Mach–Zehnder device is as follows [14]:

$$E_{\text{out}} = E_0 \cos\left(\frac{\pi}{4} + \nu\frac{\pi}{2}V_{\text{in}}\right) \tag{1}$$

where $E_{\text{out}}$ is the output electrical field, $E_0$ is the input electrical field from the laser source, $\nu$ is the modulation index, and $V_{\text{in}}$ is the input voltage normalized in the range $(-0.5, 0.5)$. The modulated signal propagates through a linear optical fiber characterized by the following frequency response [15]:

$$H_{\text{fiber}}(f) = e^{-jDz\frac{\pi\lambda^2}{c}f^2} \tag{2}$$

where $Dz$ is the total amount of uncompensated CD (usually expressed in [ps/nm]), $\lambda$ is the optical carrier wavelength (dimension: [nm]), and $c$ is the propagation speed ($c \simeq 3 \times 10^8$ m/s). The received optical signal is corrupted by wideband amplified spontaneous emission (ASE) noise, which we assume to be dominant over all other noise sources. The ASE noise is modeled as a white Gaussian complex bidimensional vector process

(one complex dimension per polarization) [15], with mono-lateral power spectral density per complex dimension equal to $N_0$. The received optical signal, i.e., the sum of the useful signal and the noise, is filtered by an optical filter with transfer function $H_{\text{o}}(f)$ and then converted to an electrical signal by a square-law photodetector. The signal is then passed through an electrical filter with transfer function $H_{\text{e}}(f)$.

The obtained electrical signal $y(t)$ is sampled at frequency $f_{\text{s}} = \eta/T_{\text{b}}$, where $T_{\text{b}}$ is the bit interval and $\eta \in \mathbb{N}$ is the sampling rate expressed in samples per bit interval (or oversampling factor). The samples relative to the $k$th bit interval are collected in vectors

$$\boldsymbol{y}_k = \left(y_k^{(0)}, \ldots, y_k^{(\eta-1)}\right) = \left(y\left(t_k^{(0)}\right), \ldots, y\left(t_k^{(\eta-1)}\right)\right)$$

in which $t_k^{(i)} = (k + (i/\eta))T_{\text{b}} + \tau_0 + \tau, i = 0, \ldots, \eta - 1$, where $\tau_0$ is an arbitrary time reference used to align the information sequence and the output sample sequence, and $\tau$ is the sampling offset (for more details, see [8]).

If ODBM is used, the MOD block includes also a precoder, inserted between the data source and the shaping filter. The precoder is intended to allow simple threshold detection at the receiver and consists of a modulo-2 accumulator [16]. The Mach–Zehnder device input-output relation in the ODBM case is as follows:

$$E_{\text{out}} = E_0 \cos\left(\frac{\pi}{2} + \pi V_{\text{in}}\right) \tag{3}$$

where $E_{\text{out}}$ is the output electrical field, $E_0$ is the input electrical field from the laser source, and $V_{\text{in}}$ is the input voltage, normalized in the range $(-0.5, 0.5)$.

## III. PRELIMINARIES ON MLSD RECEIVER STRUCTURES

In the following, we briefly review MLSD schemes based on the VA and introduce the adopted branch metrics [1], [3]–[5], [8].

### A. Viterbi Algorithm

The MLSD strategy corresponds to choosing the most likely data sequence given an observed sequence of samples. In other words, the MLSD strategy leads to the selection of the following estimate $\hat{\boldsymbol{a}}$ of the transmitted data sequence:

$$\hat{\boldsymbol{a}} = \underset{\boldsymbol{a}}{\operatorname{argmax}} \log p(\boldsymbol{y}|\boldsymbol{a}) \tag{4}$$

where $p(\boldsymbol{y}|\boldsymbol{a})$ is the probability density function (pdf) of the observable vector $\boldsymbol{y} = \{y_k\}_{k=1}^{K}$ given the data sequence $\boldsymbol{a} = \{a_k\}_{k=1}^{K}$. We remark that the cardinality of the observable vector $\boldsymbol{y}$ is larger than $K$ if oversampling is used (i.e., $\eta > 1$).

Assume that the system satisfies a finite memory property [17], i.e.,

$$p\left(\boldsymbol{y}_k|\boldsymbol{y}_0^{k-1}, \boldsymbol{a}_0^{K-1}\right) = p\left(\boldsymbol{y}_k|\boldsymbol{y}_0^{k-1}, a_k, \boldsymbol{a}_{k-L}^{k-1}, \mu_{k-L}\right) \quad (5)$$

where the notation $\boldsymbol{x}_i^m$ stands for the sequence of elements $(x_i, x_{i+1}, \ldots, x_m)$; $\mu_{k-L}$ is the state associated with a possibly used encoder, or a modulation with memory; and $L$ is the system memory. In particular, the state $\mu_{k-L}$ is an integer function of $a_{k-L}$ and $\mu_{k-L-1}$. In other words, the observable $\boldsymbol{y}_k$ depends on $L+1$ information bits and an integer value $\mu_{k-L}$. The state $S_k$ of the system may be defined as follows:

$$S_k \triangleq \left(\boldsymbol{a}_{k-L}^{k-1}, \mu_{k-L}\right) \quad (6)$$

and can be interpreted as the state of a finite state machine (FSM) driven by the input data sequence. Considering a time-invariant FSM (as will be the case in the following), a "next-state" function which, given $a_k$ and $S_k$, returns $S_{k+1}$ can be identified. In this case, by application of chain factorization to $p(\boldsymbol{y}|\boldsymbol{a})$, (4) may be expressed as [17]

$$\begin{aligned} \hat{\boldsymbol{a}} &= \arg\max_{\boldsymbol{a}} \log p(\boldsymbol{y}|\boldsymbol{a}) \\ &= \arg\max_{\boldsymbol{a}} \log \prod_{k=0}^{K-1} p\left(\boldsymbol{y}_k|\boldsymbol{y}_0^{k-1}, a_k, S_k\right) \\ &= \arg\max_{\boldsymbol{a}} \sum_{k=0}^{K-1} \underbrace{\log p\left(\boldsymbol{y}_k|\boldsymbol{y}_0^{k-1}, a_k, S_k\right)}_{\text{branch metric}} \end{aligned} \quad (7)$$

where the branch metric will be discussed in the next subsection. The decision rule (7) can be efficiently computed by means of a dynamic programming algorithm such as the VA, which operates on a trellis diagram whose breadth is given by the cardinality of the set of possible values of $S_k = (\boldsymbol{a}_{k-L}^{k-1}, \mu_{k-L})$, i.e., the number of states.

### B. Branch Metrics

Determining the correct branch metrics is fundamental in order to successfully implement the VA. In general, the branch metrics involved in electronic dispersion compensation (EDC) cannot be formulated in closed form and can be given the following approximation:

$$\lambda_k(a_k, S_k) \simeq \sum_{i=1}^{\eta} \log p\left(y_k^{(i)}|a_k, S_k\right) \quad (8)$$

where $y_k^{(i)}$ is $i$th sample belonging to the observable at epoch $k$, $a_k$ is a hypothetical bit at epoch $k$, and $(a_k, S_k)$ identifies a *transition* in the trellis diagram. In (8), the conditional dependence of the sample $y_k^{(i)}$ on previous observables $\{\boldsymbol{y}_\ell\}_{\ell<k}$, and on the samples $\{y_k^{(j)}\}_{j\neq i}$ belonging to the same observable, is deliberately neglected. Although not strictly exact, from a practical viewpoint, this assumption is numerically accurate [8]. In fact, the conditional dependence between observables may arise due to the receiver front-end memory, which is usually kept as

small as possible in order to avoid additional intersymbol interference.

In [18], a method is proposed for computing the exact pdf of an observed sample in a fiber optic communication scheme with a standard (noiseless) front-end, formed by the cascade of (i) a linear optical filter, (ii) a square law photodetector, and (iii) an electrical filter. More precisely, an analytic expression for the moment generating function of the sample is obtained, from which the pdf can be derived through an inverse Laplace transform.

In [8], we have introduced the following simple best-fit approximate expression of the exact branch metrics:

$$\log p\left(y_k^{(i)}|a_k, S_k\right) = \alpha_i + \beta_i y_k^{(i)} + \gamma_i \sqrt{y_k^{(i)}} + \delta_i \log y_k^{(i)} \quad (9)$$

where $\alpha_i$, $\beta_i$, $\gamma_i$, and $\delta_i$ are functions of $(a_k, S_k)$ and can be obtained by numerical least squares fitting of the exact metric in [18]. This branch metric expression allows a very accurate approximation of the exact metric, while retaining a low computation complexity.

### C. Quantization of the Observables

Quantization is carried out by applying a nondecreasing monotone staircase function which maps real-valued quantities onto integer numbers. A quantization function is completely described by the number and positions of the so-called quantization thresholds. A *uniform quantization* is a quantization function which is characterized by equally spaced quantization thresholds. The quantization thresholds partition the observable domain into regions associated with an integer number. If the number of bits representing the quantized value is $m$, then $2^m$ regions are considered and, correspondingly, $2^m - 1$ thresholds.

A uniform quantization function is uniquely identified by the number of bits $m$ and a suitably defined quantization range $Q_r$. In particular, we assume that adjacent thresholds are placed at a distance $\Delta = Q_r/2^m$, with the minimum threshold placed at $\Delta$ and the maximum threshold placed at $(2^m - 1)\Delta$. An illustrative example with $m = 3$ is shown in Fig. 2. If we denote the quantized interval to which the observable $y_k^{(i)}$ belongs by an index $z_k^{(i)}$ and, as a consequence, the quantized observable vector by $\boldsymbol{z}_k = (z_k^{(0)}, \ldots, z_k^{(\eta-1)})$, the VA branch metric (8) becomes

$$\lambda_k(a_k, S_k) = \log p(\boldsymbol{z}_k|a_k, S_k) = \log \prod_{i=0}^{\eta-1} p\left(z_k^{(i)}|a_k, S_k\right)$$

in which the conditional probability mass function $p(z_k^{(i)}|a_k, S_k)$ is obtained as follows:

$$\begin{aligned} p\left(z_k^{(i)}|a_k, S_k\right) &= P\left\{y_k^{(i)} \in \mathcal{I}\left(z_k^{(i)}\right)|a_k, S_k\right\} \\ &= \int_{\mathcal{I}\left(z_k^{(i)}\right)} p\left(y_k^{(i)}|a_k, S_k\right) \mathrm{d}y_k^{(i)} \end{aligned}$$

where $\mathcal{I}(z_k^{(i)})$ denotes the quantization interval which corresponds with the index $z_k^{(i)}$ and $p(y_k^{(i)}|a_k, S_k)$ is obtained as described in the previous subsection.
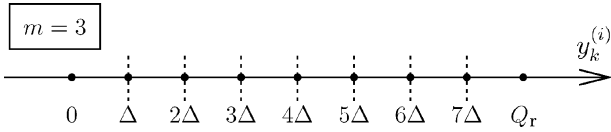
Fig. 2.  Uniform 3-bit quantization in the interval $(0, Q_r)$, where $Q_r$ is the quantization range and the number of thresholds is $2^3 - 1 = 7$.

## IV. STATE-COMPLEXITY REDUCTION TECHNIQUES

In this section, we introduce the state reduction techniques considered in this work. As mentioned in Section I, the first method is a straight memory truncation in the FSM that models the channel, whereas the second approach corresponds to a more refined state reduction technique. As a preliminary observation, recall the definition (6) of the system state, where $\mu_k$ is the encoder/modulator state and $L$ denotes the channel memory. For conciseness, in this section we will assume that no encoder is present at the transmitter, therefore, the encoder/modulator state $\mu_k$ will be omitted. The extension to account for the possible presence of an encoder is possible following the guidelines in the literature, e.g., [11]–[13], [17], [19]. Accordingly, the state definition (6) becomes

$$S_k \triangleq (a_{k-1}, \ldots, a_{k-L}). \qquad (10)$$

### A. Brute-Force State-Complexity Reduction

A possible state-complexity reduction approach consists of using an FSM model accounting for a memory $Q$ smaller than the actual memory $L$. The branch metrics at epoch $k$ will be based on a reduced state

$$w_k \triangleq (a_{k-1}, \ldots, a_{k-Q}) \qquad (11)$$

shorter than the full state $S_k$ and, therefore, will not be able to correctly represent the statistical properties of the observed sample. The reduced number of states is then $\zeta' = 2^Q \le \zeta = 2^L$. Since this approach is suboptimal, the branch metrics can be defined according to any reasonable criterion. In particular, it seems reasonable to base the definition of the reduced state metrics on the exact branch metrics accounting for the full system memory. We define the new branch metric as follows:

$$\tilde{\lambda}_k(a_k, w_k) = \sum_{i=1}^{\eta} \log p\left(y_k^{(i)} \middle| \chi_k^{\mathrm{BF}}(a_k, w_k)\right) \qquad (12)$$

where a *pseudo*-transition $\chi_k^{\mathrm{BF}}(a_k, w_k)$ is a function of the reduced transition $(a_k, w_k)$ and represents the vector of $L + 1$ bits that $\boldsymbol{y}_k$ depends on. In particular, the pseudo-transition $\chi_k^{\mathrm{BF}}(a_k, w_k)$ is defined as follows:

$$\chi_k^{\mathrm{BF}}(a_k, w_k) \triangleq \underbrace{\tilde{a}_{k+L_1}, \ldots, \tilde{a}_{k+1}}_{L_1 \text{ precursor bits}}, a_k, \overbrace{\underbrace{a_{k-1}, \ldots, a_{k-Q}}_{Q \text{ state bits}}}^{w_k}$$
$$\underbrace{\tilde{a}_{k-Q-1}, \ldots, \tilde{a}_{k-Q-L_2}}_{L_2 \text{ postcursor bits}}) \qquad (13)$$

where the precursor bit sequence $\tilde{a}_{k+L_1}, \ldots, \tilde{a}_{k+1}$ and the postcursor bit sequence $\tilde{a}_{k-Q-1}, \ldots, \tilde{a}_{k-Q-L_2}$ are time invariant functions of $(a_k, w_k)$ and $L_1 + L_2 = L - Q$. The source of inaccuracy is therefore given by the fact that precursor and postcursor sequences may differ from the actual transmitted bits and the detection algorithm has no control of this mismatch. The numerical results will be obtained choosing the precursors and postcursors sequences based on a pseudo-random bit sequence (PRBS). In particular, the $(Q + 1)$-bit pattern $a_k, a_{k-1}, \ldots, a_{k-Q}$ uniquely identifies a position $n^*$ in an $(Q + 1)$-PRBS periodic process $\{\xi_n\}$ such that $(a_k, a_{k-1}, \ldots, a_{k-Q}) = (\xi_{n^*}, \xi_{n^*-1}, \ldots, \xi_{n^*-Q})$. The pseudo-transition $\chi_k^{\mathrm{BF}}(a_k, w_k)$ is then defined as follows:

$$\chi_k^{\mathrm{BF}}(a_k, w_k) = \underbrace{\xi_{n^*+L_1}, \ldots, \xi_{n^*+1}}_{\text{precursors}}, \underbrace{\xi_{n^*}, \xi_{n^*-1}, \ldots, \xi_{n^*-Q}}_{\text{equal to}(a_k, w_k)},$$
$$\underbrace{\xi_{n^*-Q-1}, \ldots, \xi_{n^*-Q-L_2}}_{\text{postcursors}}. \qquad (14)$$

The period of $\{\xi_n\}$ is $2^{Q+1}$ and, in general, should be larger than $L + 1$.

### B. Reduced State Sequence Detection

The state-complexity of MLSD schemes can be naturally decoupled from the memory parameter $L$ by means of RSSD techniques [11]–[13], [20]. Let $Q < L$ denote the memory parameter to be taken into account in the definition of a reduced trellis state

$$\omega_k = (a_{k-1}, a_{k-2}, \ldots, a_{k-Q}).$$

The branch metric can be obtained by defining the pseudo-transition as follows:

$$\chi_k^{\mathrm{RSSD}}(a_k, \omega_k) = \left( \underbrace{\tilde{a}_{k+L_1}, \ldots, \tilde{a}_{k+1}}_{L_1 \text{ precursor bits}}, a_k, \overbrace{\underbrace{a_{k-1}, \ldots, a_{k-Q}}_{Q \text{ state bits}}}^{\omega_k} \right.$$
$$\underbrace{\breve{a}_{k-Q-1}(\omega_k), \ldots, \breve{a}_{k-Q-P}(\omega_k)}_{P \text{ postcursor bits}}$$
$$\left. \underbrace{\hat{\tilde{a}}_{k-Q-P-1}, \ldots, \hat{\tilde{a}}_{k-Q-L_2}}_{L_2 - P \text{ postcursor bits}} \right) \qquad (15)$$

where the first $L_1$ bits, i.e., the precursor bit sequence, may be chosen as an arbitrary time invariant function of $(a_k, \omega_k)$, $P$ bits of the $L_2$ postcursor bits may be chosen by a per-survivor processing technique [20] as a time varying function of $\omega_k$, i.e., the bits $\breve{a}_{k-Q-1}(\omega_k), \ldots, \breve{a}_{k-Q-P}(\omega_k)$ are the information bits associated with the survivor of $\omega_k$. The remaining $L_2 - P$ postcursor bits can be defined either as tentative (or preliminary) decisions or as an arbitrary time invariant function of $(a_k, \omega_k)$, as described in Section IV-A. Note that the use of tentative decisions for postcursor bits is possible in the case of brute-force state complexity reduction, as well.

## C. RSSD Versus Brute-Force State-Complexity Reduction

In both state-complexity reduction techniques described in the previous subsections, the branch metric $\tilde{\lambda}_k(a_k, \omega_k)$ in the reduced-state trellis can be defined in terms of a pseudo-transition $\chi_k$ according to the following:

$$\tilde{\lambda}_k(a_k, \omega_k) = \lambda_k(\chi_k) = \sum_{i=1}^{\eta} \log p\left(y_k^{(i)}|\chi_k\right)$$

where $\chi_k$ is defined as

$$\chi_k = \begin{cases} \chi_k^{\mathrm{BF}}(a_k, w_k), & \text{for brute-force state-complexity} \\ & \qquad\qquad \text{reduction} \\ \chi_k^{\mathrm{RSSD}}(a_k, \omega_k), & \text{for RSSD.} \end{cases}$$

In order to better understand the difference between RSSD and brute-force state-complexity reduction, we refer to a *linear* intersymbol interference (ISI) channel. We remark that this is *not* the case for a fiber-optic communication system, which, because of the presence of the MZ device and the photodetector, is a *nonlinear* system. However, the considered example allows to grasp the basic ideas behind the application of state-complexity reduction techniques to nonlinear systems, as well.

As an illustrative example, we consider the case of uncoded transmission over a linear ISI channel. The observable at the output of the channel can typically be expressed as follows:

$$r_k = \sum_{i=0}^{L} h_i a_{k-i} + n_k. \tag{16}$$

In Fig. 3, a pictorial description of how state reduction works in a scenario with ISI ($L = 4$) and triangular channel impulse response is shown, with both RSSD and brute-force state-complexity reduction techniques.

- In Fig. 3(a), the full-complexity case is depicted. More precisely, the 5 tap weights of the ISI channel impulse response are shown. Under each tap weight, the corresponding information symbol, according to (16), is indicated. The full state is $S_k = (a_{k-1}, a_{k-2}, a_{k-3}, a_{k-4})$ and the full transition is $T_k = (a_k, a_{k-1}, a_{k-2}, a_{k-3}, a_{k-4})$.
- In Fig. 3(b), a pictorial description for RSSD state-complexity reduction is shown, considering $L_1 = 0$ and $L_2 = P = 2$. The reduced transition is $(a_k, w_k) = (a_k, a_{k-1}, a_{k-2})$, and the missing information symbols $\{a_{k-3}, a_{k-4}\}$ for the construction of the pseudo transition $\chi_k(a_k, w_k)$ are recovered through the survivor history as described in Section IV-B.
- Finally, in Fig. 3(c), a pictorial description for the brute-force state-complexity reduction is shown. In this case, the information symbols corresponding to the channel taps $h_3$ and $h_4$ are selected as described in Section IV-A, assuming $L_1 = 0$ and $L_2 = 2$.[1] As one can intuitively understand, this brute-force state-complexity reduction should lead to a performance worse than that of RSSD, since the missing information bits are associated with the reduced state in a (temporally) *fixed* fashion. On the contrary, use of RSSD

allows to *dynamically* associate the reduced states with the "best" (i.e., the most reasonable) missing information bits, i.e., those of the corresponding survivor.

## D. Fine-Tuning of State-Complexity Reduction Techniques

The application of state-complexity reduction techniques to MLSD for transmission over optical fiber systems may lead to significant complexity savings with limited performance degradation. As previously observed, it is not possible to simply visualize the response of a *nonlinear* channel of an optical fiber transmission system, as done in Fig. 3 in the case of a linear ISI channel response. However, one can refer to a linear channel to better understand how to apply state-complexity reduction techniques. In particular, the application of state-complexity reduction techniques to MLSD assumes a reduced state concentrated over the precursor taps of the impulse response [11]–[13]: for instance, in Fig. 3 $w_k = (a_{k-1}, a_{k-2})$. In a scenario with a long nonconstant impulse response (e.g., the triangular response exemplified in these figures), it might happen that the leading tap weights convey little energy of the channel impulse response. In this case, it may be more convenient to consider a reduced state corresponding to the bits with higher energy. This can be obtained by a proper choice of the parameters $L_1$ and $L_2$.

In Fig. 4, a linear ISI triangular 7-tap channel impulse response is considered. The full state $S_k = (a_{k-1}, \ldots, a_{k-6})$ is shown, together with the reduced state $w_k = (a_{k-1}, a_{k-2}, a_{k-3})$, i.e., with $L_1 = 0$. As one can see, the tap weights "covered" by the reduced state (i.e., $h_1$, $h_2$, and $h_3$) are *not* the highest-energy ones. Considering $L_1 = 1$, the resulting reduced state $w_{k-1}^{(+1)} = (a_{k-2}, a_{k-3}, a_{k-4})$, referred to as "shift+1," is associated with the highest-energy channel impulse response tap weights (i.e., $h_2$, $h_3$, and $h_4$). The $L_1$ precursor bits are then obtained as described in Section IV-A. Intuitively, the reduced-state trellis based on the latter shifted reduced state "captures" most of the channel impulse response energy and the performance of the corresponding VA-based MLSD receiver is expected to improve.

## V. SIMULATION SETUP

In the following sections, we will consider two 10 Gb/s schemes with OOK and ODBM.

The OOK system setup is as follows. With reference to Fig. 1, the low-pass modulator electrical filter is a third-order Bessel filter with bandwidth 9.5 GHz, the modulation index $\nu$ in (1) is equal to 0.93, corresponding to an extinction ratio of 24.3 dB at the transmitted side, the optical filter is a third-order Bessel filter with bandwidth 32.5 GHz, and the electrical filter is a fifth-order Bessel filter with bandwidth 7.7 GHz.

The ODBM system setup is as follows. With reference to Fig. 1, the low-pass modulation electrical filter is a fifth-order Bessel filter with bandwidth 3 GHz, the optical filter is a third-order Bessel filter with bandwidth 33 GHz, and the electrical filter is a fifth-order Bessel filter with bandwidth 7.7 GHz.

Unless otherwise stated, the sampling offset $\tau$ is chosen to minimize the BER of the communication scheme.

---

[1]Note, however, that the impulse response of the simulated channel, obviously, remains the 5-tap original one.
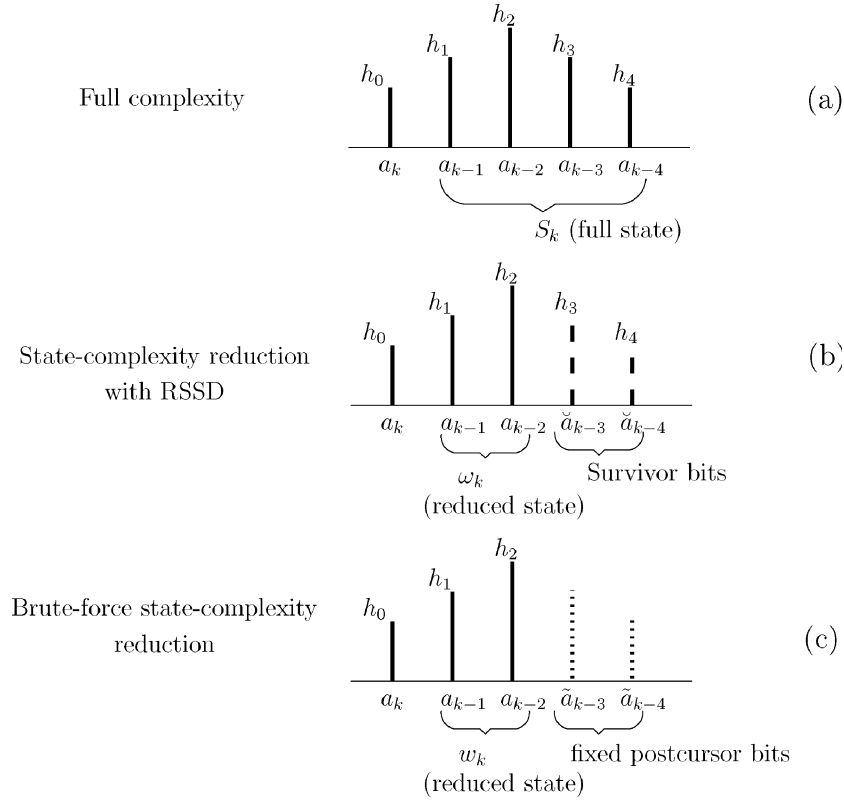
Fig. 3. Pictorial description of state reduction in the case of a linear intersymbol interference (ISI) channel. From top to bottom: (a) full complexity case; (b) reduced-state case with RSSD; (c) brute-force state-complexity reduction.
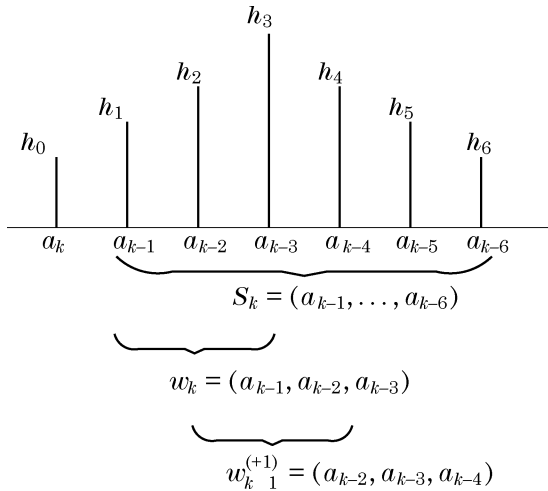


Fig. 4. State reduction in the case of a linear ISI channel. The full state $S_k = (a_{k-1}, \ldots, a_{k-6})$ is shown, together with the reduced state with $L_1 = 0$ $w_k = (a_{k-1}, a_{k-2}, a_{k-3})$ and the "shift+1" reduced state $w_{k-1}^{(+1)} = (a_{k-2}, a_{k-3}, a_{k-4})$, i.e., with $L_1 = 1$.

## VI. IMPACT OF SAMPLING OFFSET

In this section, we investigate the impact of the sampling offset considering both brute-force state-complexity reduction and RSSD. As it will be shown in the following, care has to be taken in selecting the proper sampling offset in order to optimize the performance of the communication systems at hand.

In Fig. 5, the BER is shown, as a function of the sampling offset $\tau$ within a bit period, in a NRZ-OOK scenario. The optical signal-to-noise ratio (OSNR) is defined as follows:

$$\text{OSNR} \triangleq \frac{P_{\text{signal}}}{2N_0 B_{0.1 \text{ nm}}}$$

where $P_{\text{signal}}$ is the average received noiseless optical power, $N_0$ is the mono-lateral noise power spectral density, and $B_{0.1}$ nm is the 0.1 nm bandwidth, at a carrier wavelength of 1550 nm. The uncompensated CD is 2550 ps/nm, the OSNR is set to 15 dB, and the sampling rate is 1 samp/symb. The sampling offset resolution is set to $T_b/64$, i.e., we consider 64 possible sampling offset values within a bit period $T_b$. As one can see, the sensitivity with respect to the sampling offset is non-negligible. In fact, one can easily observe the presence of an absolute minimum which represents an optimized operating condition. This sensitivity is greatly reduced by using oversampling, i.e., $\eta > 1$.

In Fig. 6, the impact of the sampling offset on the BER is analyzed in a scenario with NRZ-OOK and a sampling rate equal to $\eta = 2$ samples per bit interval. The considered sampling offset spans over a range which is larger than a bit interval. Adding an integer value $s$ to the sampling offset corresponds to shifting by $s$ bits the state vector (i.e. "shift+$s$"), as described in Section IV-D. The offset value $\tau = 0$ corresponds to the optimal sampling time in back-to-back (B2B) for threshold detection. Three values for the CD are considered. The MLSD receiver adopts RSSD and the number of states is reduced from $\zeta = 256$
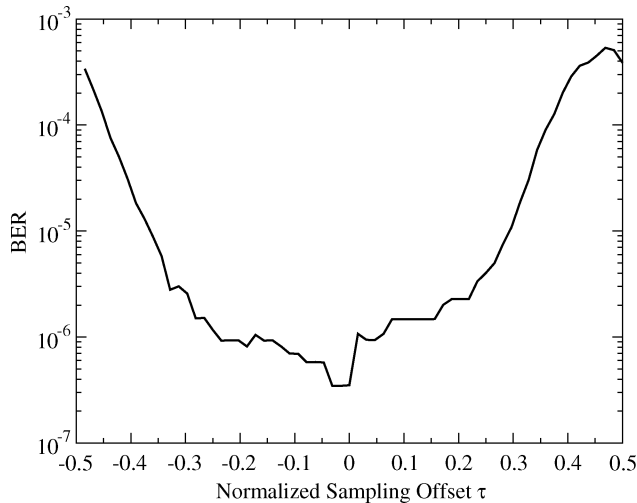
Fig. 5. BER, as a function of the sampling offset $\tau$ within a bit period, in an NRZ-OOK scenario. The CD is 2550 ps/nm and the OSNR is set to 15 dB. The sampling rate is 1 samp/symb.



Fig. 6. BER, as a function of the sampling offset $\tau$, in a scenario with NRZ-OOK. Various values of CD and OSNR are considered. The sampling rate is $\eta = 2$ samples per bit interval. An RSSD strategy, reducing the number of states from $\zeta = 256$ to $\zeta' = 16$ is considered.



Fig. 7. BER, as a function of the sampling offset $\tau$, in a scenario with NRZ-OOK. Various values of CD and OSNR are considered. The sampling rate is $\eta = 2$ samples per bit interval. A brute-force state-complexity reduction to $\zeta' = 16$ is considered.

to $\zeta' = 16$. The impact of the sampling offset depends on the particular CD value.

- At a CD equal to 0 ps/nm (back-to-back, B2B) with OSNR equal to 10 dB, the lowest possible BER is obtained for a sampling offset between 0 and 3.5 $T_b$. In other words, there is a large range over which the BER is minimized. This is due to the short memory of the channel in B2B.
- At a CD equal to 3400 ps/nm with OSNR equal to 12 dB, the optimal sampling offset is around 1.5 $T_b$. Unlike the previous case, there is a shorter range over which the BER is minimized (approximately, between 0.5 $T_b$ and 2 $T_b$). Moreover, one can observe that the curve is not symmetrical, i.e., it does not admit any axis of symmetry. This is due to the use of a state-complexity reduction technique.
- At a CD equal to 6800 ps/nm with OSNR equal to 13 dB, the optimal sampling offset is around $T_b$ and the region width is around $T_b$. One can observe that, like the case with CD equal to 3400 ps/nm, the BER curve does not admit any axis of symmetry.

In Fig. 7, an analysis similar to that in Fig. 6 is carried out, the only difference being the fact that a brute-force state-complexity reduction to $\zeta' = 16$ is considered. As one can observe from the results in Fig. 7, the optimal sampling offset at 3400 ps/nm and 6800 ps/nm is 0, i.e., as in B2B. In this case, unlike the case with RSSD, the BER curves have a vertical symmetry axis at about $\tau = -0.25$. This is due to the fact that brute-force state-complexity reduction treats postcursor and precursor bits in the same fashion.

The analysis of ODBM schemes yields similar results which, therefore, are not shown. The general conclusion of the analysis carried out in this section is that the sampling offset becomes more important for increasing values of the CD, although, in general, small mismatches on the sampling offset cause a negligible performance loss. Moreover, if RSSD techniques are used, the optimized sampling offset significantly differs from that of a system using brute-force state-complexity reduction or no state-complexity reduction at all. In fact, as described in Section IV,
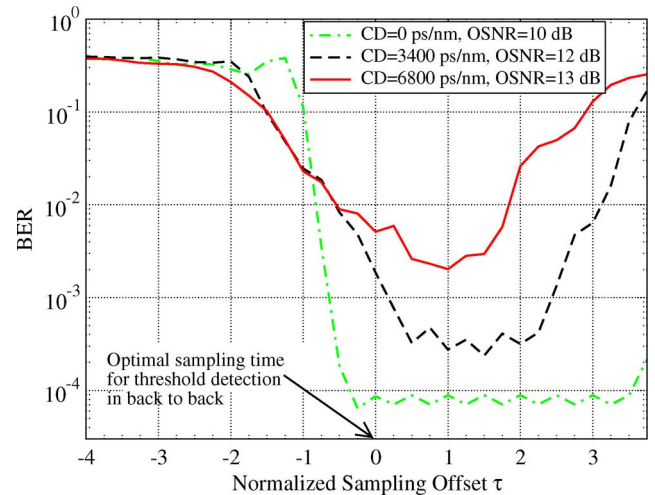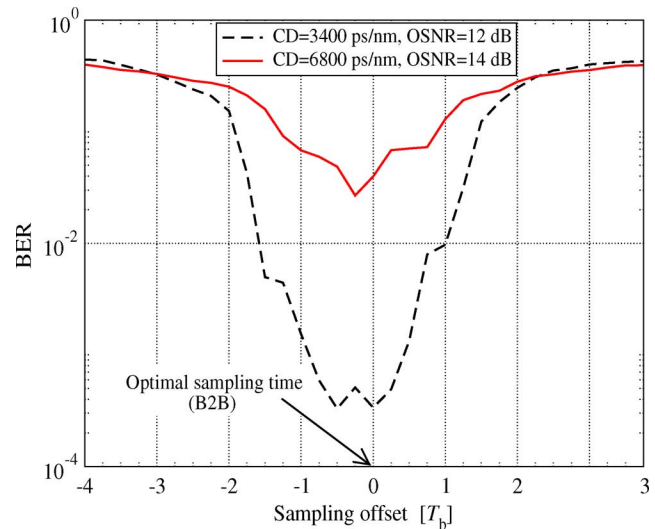
while a brute force state-complexity reduction neglects both precursor bits and postcursor bits in the computation of the metrics, RSSD enables to accurately select postcursor bits and only neglects precursor bits. This fact is responsible for the asymmetry of RSSD curves in Fig. 6, as opposed to the marked symmetry of the brute-force state reduction curves in Fig. 7.

## VII. IMPACT OF THE QUANTIZATION RANGE

In this section, we investigate the impact of the quantization range assuming uniform quantization. The theoretical predictions in [8], based on the computation of the information rate between the transmitter modulator input and the output of the electrical filter of the receiver, show that a 3-bit quantization allows to exploit almost all the information embedded in the received electrical signal with an OSNR penalty smaller than 1 dB, whereas fewer bits entail higher penalties and more bits
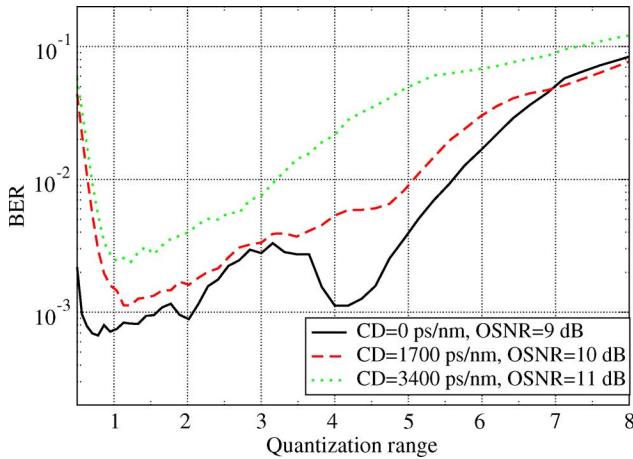
Fig. 8. BER, as a function of the quantization range (with uniform quantization), in an NRZ-OOK scenario. Various values of dispersion and OSNR are considered. The sampling rate is equal to 2 samples per bit interval.

enable limited performance improvement. Extensive investigations based on BER analysis of MLSD receivers (not shown here, for the sake of conciseness) have substantially confirmed the results in [8]. In this paper, therefore, we considered 3 bits per quantized sample. We preliminarily remark that the input signal is normalized such that the noiseless transmission of the all-"1" signal corresponds to a received electrical signal amplitude equal to 1. Uniform quantization with $2^3 - 1 = 7$ equally spaced thresholds, as shown in Fig. 2 is considered. The first threshold is placed at $\theta_1 = Q_r/8$ and the last one is placed at $\theta_7 = 7Q_r/8$, where $Q_r$ is the quantization range. Equivalently, instead of controlling the quantization range, one could control the front-end gain and perform uniform quantization over a fixed range.

In Fig. 8, the BER is shown as a function of the quantization range, in a scenario with NRZ-OOK. Values of CD equal to 0 ps/nm (B2B), 1700 ps/nm, and 3400 ps/nm are considered. The sampling rate is set to $\eta = 2$ samp/symb. A brute-force state reduction technique with 16 states is used at the receiver. As one can see from the results in Fig. 8, it is important to optimize the quantization range, since too small or too high ranges lead to unacceptable performance. Nevertheless, from the results in Fig. 8, one can conclude that there is no need for high accuracy in quantization range optimization, i.e., the sensitivity on quantization range is limited. Moreover, the optimum $Q_r$ value is close to 1. This implies that all the quantization thresholds fall between 0 and 1. In other words it seems that there is no need for an accurate representation of observables beyond the value 1, even if such observables are statistically frequent. One can further observe that in B2B there is an oscillatory behavior of the BER versus the quantization range. This can be easily understood since, in B2B, simple threshold detection offers good performance. The local minima in the BER versus the quantization range curve correspond to one of the 7 thresholds passing the optimum value for threshold detection.

The impact of the quantization range on the performance of ODBM schemes is similar. Therefore, these results are not shown here. Interestingly enough, the obtained results suggest that the quantization thresholds should be concentrated in the

relatively small region corresponding to the region where the all-"1" and all-"0" pdfs cross. This may be due to the fact that in order to perform effective detection, it is important to discriminate between different bit sequences (each associated with a different pdf). Two different bit sequences can be effectively distinguished if a threshold falls in the proximity of the crossing point of the corresponding pdfs.

## VIII. ROBUSTNESS AGAINST UNCOMPENSATED CD: RSSD VERSUS BRUTE-FORCE STATE-COMPLEXITY REDUCTION

In this section, we investigate the robustness of the considered reduced-states MLSD receivers against uncompensated CD. The system performance will be evaluated in terms of OSNR, as a function of the CD, required to achieve a desired BER. As expected, by reducing the number of states (i.e., increasing the degree of state-complexity reduction) the performance worsens. However, our results allow to conclude that simple and "well-tuned" state-complexity reduction techniques can significantly limit the performance degradation with respect to full state-complexity cases.

### A. NRZ-OOK

In the following, the performance results of NRZ-OOK schemes with quantization will refer to a quantization range $Q_r$ equal to 1.2, which represents a good compromise according to the analysis carried out in Section VII. In Fig. 9, the OSNR required to achieve a BER[2] equal to $10^{-3}$ is shown, as a function of the CD, for various MLSD system configurations, corresponding to different quantization settings and reduced-state configurations. The sampling rate is $\eta = 2$ samp/symb for all MLSD receivers (the sampling rate is 1 samp/symb for the threshold detector). We now comment explicitly on each of the 5 performance curves shown in Fig. 9.

- The first (from the bottom) MLSD curve, namely the one with best performance, i.e., lowest OSNR, corresponds to a receiver using a VA with 256 states and no quantization (denoted "MLSD(256)"). This OSNR curve represents a bound on the performance of practical receivers using VAs with a reduced number of states and/or quantization. Comparing the results in Fig. 9 with OSNR curves shown in [8] obtained considering a given information rate, rather than a target BER, one can observe a substantial agreement in the overall shape, although the IR analysis in [8] guarantees that an additional gain of more than 2 dB is possible with a proper coding and detection scheme.
- The second MLSD curve provides the performance obtained with an MLSD receiver whose number of states is reduced from 256 to 16 (denoted "RSSD(256 → 16)") using RSSD techniques, with no input quantization. As one can observe from Fig. 9, there is basically no performance loss, with respect to the performance of an MLSD receiver with 256 states and no quantization, for CD values lower than 4000 ps/nm. For higher CD values, the OSNR tends to increase and diverges around 8000 ps/nm.

[2]The choice of a BER value equal to $10^{-3}$ is motivated by the fact that recent high performance error correction codes for optical communications, decoded after the MLSD block, allow to recover almost completely the detection errors for BER equal to or lower than $10^{-3}$.
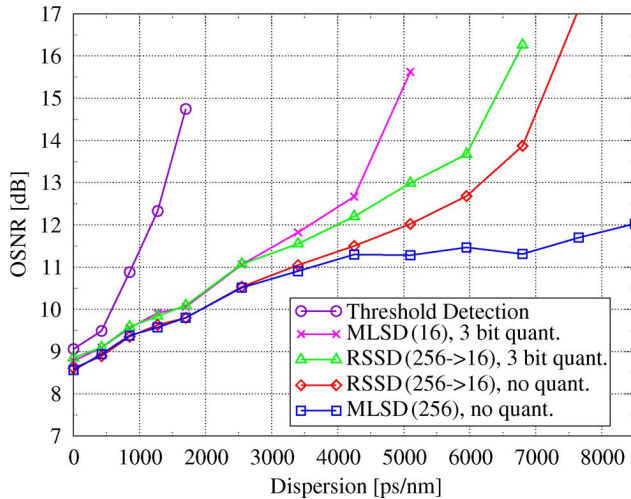
Fig. 9. OSNR, as a function of the dispersion, in an NRZ-OOK scenario at a BER equal to $10^{-3}$. Various receiver structures are considered, with RSSD or brute-force state reduction, with or without quantization.

- The third MLSD curve is associated with the same RSSD receiver considered at the previous point, in the presence of 3-bit quantization of the observables. As one can see from Fig. 9, the OSNR loss with respect to the second MLSD curve, i.e., RSSD without quantization, increases from 0.2 dB (at a CD equal to 0 ps/nm) to approximately 1 dB (at a CD equal to 6000 ps/nm). In particular, the OSNR loss due to quantization is always smaller than 0.7 dB for CD values lower than 4000 ps/nm.
- The fourth MLSD curve is associated with an MLSD receiver with brute-force state-complexity reduction to $\zeta' = 16$ states and 3-bit quantization (denoted "MLSD(16)"). For CD values lower than 2550 ps/nm, there is no performance loss with respect to the third MLSD receiver (RSSD-based, with state reduction from 256 to 16). For higher values of the CD, the OSNR loss increases: it remains modest (at most 0.3 dB) till 4200 ps/nm, whereas it becomes significant (higher than 2 dB) at 5000 ps/nm.
- For comparison, the performance of a threshold detector is also shown. As one can see, the performance of a threshold detector is very close (within 0.5 dB) to that of MLSD receivers for CD values lower than 500 ps/nm. The performance of the threshold detector, however, degrades rapidly for larger values of the CD.

Summarizing, our results suggest that below 3 dB penalty with respect to B2B, brute-force state reduction and RSSD yield similar performance, whereas above 3 dB, RSSD has greater robustness against CD. In other words, if a penalty of more than 3 dB can be tolerated by the system, RSSD outperforms brute-force state reduction by about 1500 ps/nm.

### B. ODBM

In the case of ODBM schemes, by trial and error a good value of the quantization range was derived as $Q_r = 1.35$. In Fig. 10, the OSNR required to achieve a BER equal to $10^{-3}$ is shown, as a function of the CD, for various MLSD system configurations, corresponding to different quantization settings and reduced-state configurations. The sampling rate is $\eta = 2$ samp/symb for

all MLSD receivers, whereas the sampling rate is 1 samp/symb for the threshold detector. We now comment explicitly on each of the five OSNR curves shown in Fig. 10.

- The first (from the bottom) MLSD curve corresponds to a receiver using a VA with 256 states and no quantization. As for the case with NRZ-OOK, this curve represents a bound on the OSNR performance of practical receivers using VAs with a smaller number of states and/or quantization. Unlike a scenario with NRZ-OOK, where the OSNR curve is increasing with the CD, in this case the OSNR curve is first decreasing (from 0 ps/nm to 4250 ps/nm), then increasing (from 4250 ps/nm to 6000 ps/nm), and then decreasing again (from 6000 ps/nm to 8200 ps/nm). As in the NRZ-OOK case, comparing the shape of the just described MLSD curve in Fig. 10 with the equivalent one in [8], one can observe a substantial agreement between "practical" (BER-based) and theoretical (information rate-based) results.
- The second MLSD curve corresponds to the performance obtained with an MLSD receiver whose number of states is reduced from 256 to 16 using RSSD techniques, with no quantization. As one can observe from Fig. 10, there is a very limited OSNR loss, with respect to the performance of an MLSD receiver with 256 states and no quantization, for CD values lower than 4250 ps/nm. For higher values of the CD, the OSNR tends to increase. For CD values between 4250 ps/nm and 7600 ps/nm, the OSNR curve is a sort of up-shifted version of that with 256 states and no quantization; the OSNR tends to diverge around 8000 ps/nm.
- The third MLSD curve is associated with the same RSSD receiver considered at the previous point, in presence of 3-bit quantization of the observables. As one can see from Fig. 10, the OSNR loss with respect to the best MLSD curve remains approximately equal to 0.2 dB in the CD range from 0 ps/nm to 4250 ps/nm. The OSNR loss becomes approximately 1.2 dB at CD values between 6000 ps/nm and 7000 ps/nm, and tends to increase for higher CD values.
- The fourth MLSD curve is relative to an MLSD receiver with brute-force state-complexity reduction to 16 states and 3-bit quantization. For CD values lower than 4250 ps/nm, there is no performance loss with respect to the third MLSD receiver (RSSD-based quantized MLSD receiver, with state reduction from 256 to 16). For higher values of the CD, the OSNR loss increases, but remains modest (at most 0.6 dB) till 5000 ps/nm.
- For comparison, the performance of a threshold detector is also shown. As one can see, the performance of a threshold detector is identical to that of the 256-state unquantized MLSD receiver in B2B, i.e., slightly better than that of quantized 16-state MLSD receivers. The set of quantization thresholds, in fact, does not comprise the optimum value for standard threshold detection, since the quantization range was chosen as a good compromise over all considered CD range. The performance of a threshold detector starts degrading around 1500 ps/nm.

Similarly to the OOK-NRZ case, the RSSD approach enables to cope with higher CD values (in the considered ODBM case,
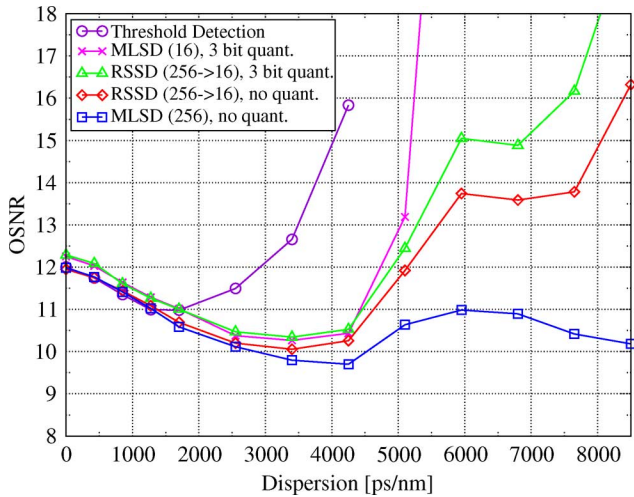
Fig. 10. OSNR, as a function of the CD, in a scenario with ODBM at a BER equal to $10^{-3}$. Various receiver structures are considered, with RSSD or brute-force state-complexity reduction, with or without quantization.

the difference is about 1000 ps/nm). On the other hand, this performance improvement can be exploited at high CD values and relatively high penalty values (i.e., on the order of 2.5 dB with respect to B2B).

## IX. DISCUSSION

The results obtained in Section VI show that, whenever 2 samples per bits interval are considered, there is limited sensitivity on the fractional portion of the sampling offset. The use of brute-force state-complexity reduction entails an optimal sampling offset equal to that of a standard threshold detection system, whereas the use of RSSD techniques calls for ad hoc sampling offset optimization.

In Section VII, we have investigated the impact of quantization range on system performance. Our results confirm that a quantization strategy with 3 bits per sample allows good performance, as predicted by the theoretical investigation of the ultimate performance limits in [8]. Moreover, they suggest that all the quantization thresholds should be placed in a relatively short range. This is not obvious, since by placing all thresholds in a short range one obtains a coarse representation of the signal (in terms of mean square error). On the other hand, from a detection-theoretic point of view, the obtained result is intuitive since the goal of the detection strategy is not to estimate the received signal but to recover the embedded data sequence.

In Section VIII, we have shown performance results of the considered MLSD schemes, in terms of OSNR required to achieve a BER equal to $10^{-3}$ versus CD. Two possible modulation formats have been considered: NRZ-OOK and ODBM. The obtained results can be summarized as follows.

- Changing the modulation format from NRZ-OOK to ODBM leads to different and characteristic *shapes* for the OSNR curves. More precisely, for NRZ-OOK schemes, the OSNR is an increasing function of the CD, whereas for ODBM schemes, there is a relative minimum around 4000 ps/nm followed by a relative maximum around 6000 ps/nm. This is in excellent agreement with the theoretical findings presented in [8].

- For a given modulation format, the relative behavior of possible MLSD schemes (with or without quantization, with brute-force state-complexity reduction or RSSD) is similar, regardless of the target BER. For the sake of conciseness, we have not reported here our performance results relative to a BER equal to $10^{-4}$, but the shapes of the obtained OSNR curves are very similar to those obtained with BER equal to $10^{-3}$. Moreover, the performance degradation incurred by considering a reduced-state quantized MLSD receiver, with respect to a full-state unquantized MLSD receiver, is relatively similar for both modulation formats.

- From a practical viewpoint, using an MLSD receiver, with 3-bit quantization and brute-force state-complexity reduction to 16 states, is practically optimal for CD values lower than 4000 ps/nm.

We remark that although the results in this paper are obtained using a linear optical fiber model, the underlying finite memory approach does no rely on the system linearity, therefore, a *nonlinear* optical fiber model should not require a different receiver structure. Possible system nonlinearities, both deterministic and nonlinear noise-signal interaction, can be taken into account using a suitable set of branch metrics.

Furthermore, we remark that the presence of electrical Gaussian noise can be also taken into account by using proper branch metrics.

## X. CONCLUDING REMARKS

In this paper, we investigated state complexity reduction for MLSD for direct-photodetection optical communication schemes affected by uncompensated CD. We considered both NRZ-OOK and ODBM, optimizing the sampling offset and the quantization strategy. Simple brute-force state-complexity reduction and RSSD have been compared, considering a target number of MLSD states equal to 16, which represents a reasonable number of states for state of the art very high speed VA implementations. The results show that, if small OSNR penalty (i.e., lower than 3 dB) is required, a good performance/complexity tradeoff is given by the brute-force state-complexity reduction solution. If higher penalty values may be tolerated RSSD enables improved performance and higher CD tolerance than brute-force state-complexity reduction. The improvement in CD tolerance is on the order of 1500 ps/nm for NRZ-OOK and 1000 ps/nm for ODBM.

## REFERENCES

[1] 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 *Proc. IEEE Global Telecommun. Conf.*, Dallas, TX, Dec. 2004, pp. 299–303.

[2] A. Fäbert, S. Langenbach, N. Stojanovic, C. Dorschky, T. Kupfer, C. Schulien, J.-P. Elbers, H. Wernz, H. Grisser, and C. Glingener, "Performance of a 10.7 Gb/s receiver with digital equaliser using maximum likelihood sequence estimation," presented at the Proc. European Conf. Optical Commun., Stockholm, Sweden, Sep. 2004, Th.4.1.5, unpublished.

[3] O. E. Agazzi, M. R. Hueda, H. S. Carrer, and D. E. Crivelli, "Maximum-likelihood sequence estimation in dispersive optical channels," *J. Lightw. Technol.*, vol. 23, no. 2, pp. 749–763, Feb. 2005.

[4] T. Foggi, E. Forestieri, G. Colavolpe, and G. Prati, "Maximum likelihood sequence detection with closed-form metrics in OOK optical systems impaired by GVD and PMD," *J. Lightw. Technol.*, vol. 24, no. 8, pp. 3073–3087, Aug. 2006.

[5] 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," *J. Lightw. Technol.*, vol. 25, no. 3, pp. 957–975, Mar. 2007.

[6] G. D. Forney, Jr., "The Viterbi algorithm," *Proc. IEEE*, vol. 61, no. 3, pp. 268–278, Mar. 1973.

[7] D. E. Crivelli, H. S. Carrer, and M. R. Hueda, "On the performance of reduced-state Viterbi receivers in IM/DD optical transmission systems," presented at the Proc. European Conf. Optical Commun. (ECOC), Stockholm, Sweden, Sep. 2004, Paper We4.P.083, unpublished.

[8] M. Franceschini, G. Bongiorni, G. Ferrari, R. Raheli, F. Meli, and A. Castoldi, "Fundamental limits of electronic signal processing in direct-detection optical communications," *J. Lightw. Technol.*, vol. 25, no. 7, pp. 1742–1753, Jul. 2007.

[9] D. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavcic, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3498–3508, Aug. 2006.

[10] M. Franceschini, G. Ferrari, R. Raheli, F. Meli, and A. Castoldi, "State-complexity reduction in MLSD receivers for direct-detection optical communications," presented at the European Conf. Optical Commun., Berlin, Germany, Sep. 2007.

[11] M. V. Eyuboğlu and S. U. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Trans. Commun.*, vol. 36, no. 1, pp. 13–20, Jan. 1988.

[12] A. Duel-Hallen and C. Heegard, "Delayed decision feedback estimation," *IEEE Trans. Commun.*, vol. 37, no. 5, pp. 428–436, May 1989.

[13] P. R. Chevillat and E. Eleftheriou, "Decoding of trellis-encoded signals in the presence of intersymbol interference and noise," *IEEE Trans. Commun.*, vol. 36, no. 7, pp. 669–676, Jul. 1989.

[14] C. Lawetz, J. Cartledge, C. Rolland, and J. Yu, "Modulation characteristics of semiconductor mach-zehnder optical modulators," *J. Lightw. Technol.*, vol. 15, no. 4, pp. 697–703, Apr. 1997.

[15] G. Agrawal, *Fiber-Optic Communications Systems*. New York: Wiley, 2002.

[16] K. Yonenaga and S. Kuwano, "Dispersion-tolerant optical transmission system using duobinary transmitted and binary receiver," *J. Lightw. Technol.*, vol. 15, no. 8, pp. 1530–1537, Aug. 1997.

[17] G. Ferrari, G. Colavolpe, and R. Raheli, "A unified framework for finite-memory detection," *IEEE J. Select. Areas Commun.*, vol. 23, no. 9, pp. 1697–1706, Sep. 2005.

[18] E. Forestieri, "Evaluating the error probability in lightwave systems with chromatic dispersion, arbitrary pulse shape and pre- and postdetection filtering," *J. Lightw. Technol.*, vol. 18, no. 11, pp. 1493–1503, Nov. 2000.

[19] G. Ferrari, G. Colavolpe, and R. Raheli, *Detection Algorithms for Wireless Communications*. Chichester, U.K.: Wiley, 2004.

[20] R. Raheli, A. Polydoros, and C. Tzou, "Per-survivor processing: A general approach to MLSE in uncertain environments," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 354–364, Feb./Apr. 1995.

**Gianluigi Ferrari** (S'97–M'03) was born in Parma, Italy, in November 1974. He received the "Laurea" degree (five-year program) (*summa cum laude*) and the Ph.D. degree in electrical engineering from the University of Parma in October 1998 and January 2002, respectively.

From July 2000 to December 2001, he was a Visiting Scholar at the Communication Sciences Institute, University of Southern California, Los Angeles. Since 2002, he has been a Research Professor with the Department of Information Engineering, University of Parma, where he is now the Coordinator of the Wireless Ad-hoc and Sensor Networks (WASN) Laboratory. Between 2002 and 2004, he visited several times, as a Research Associate, the Electrical and Computer Engineering Department, Carnegie Mellon University, Pittsburgh, PA. He has published more than 90 papers in leading international conferences and journals. He is coauthor of a few books, including *Detection Algorithms for Wireless Communications, with Applications to Wired and Storage Systems* (Wiley, 2004) and *Ad Hoc Wireless Networks: A Communication-Theoretic Perspective* (Wiley, 2006). His main research interests include wireless ad hoc and sensor networking, and adaptive signal processing.

Dr. Ferrari is a co-recipient of a best student paper award at the 2006 International Workshop on Wireless Ad hoc Networks (IWWAN'06). He acts as a technical program member for several international conferences and also as a frequent reviewer for many international journals and conferences. He is currently on the editorial boards of *The Open Electrical and Electronic Engineering Journal* (Bentham Publishers), the *International Journal of RF Technologies: Research and Applications* (Taylor & Francis), and the *International Journal of Future Generation Communication and Networking* (SERSC).

**Riccardo Raheli** (M'87) received the Dr. Ing. degree (Laurea) in electrical engineering (*summa cum laude*) from the University of Pisa, Pisa, Italy, in 1983, the M.S. degree in electrical and computer engineering from the University of Massachusetts, Amherst, in 1986, and the Ph.D. degree (Perfezionamento) in electrical engineering (*summa cum laude*) from the Scuola Superiore di Studi Universitari e di Perfezionamento (now "S. Anna"), Pisa, in 1987.

From 1986 to 1988, he was with Siemens Telecomunicazioni, Cassina de' Pecchi, Milan, Italy. From 1988 to 1991, he was a Research Professor at the Scuola Superiore di Studi Universitari e di Perfezionamento S. Anna, Pisa. In 1990, he was a Visiting Assistant Professor at the University of Southern California, Los Angeles. Since 1991, he has been with the University of Parma, Parma, Italy, where he is currently a Professor of communications engineering. His scientific interests are in the general area of statistical communication theory, with application to wireless, wired and storage systems, and special attention to data detection in uncertain environments, iterative information processing and adaptive algorithms for communications. His research work has lead to numerous scientific publications in leading international journals and conference proceedings, as well as a few industrial patents. He is coauthor of the book *Detection Algorithms for Wireless Communications, with Applications to Wired and Storage Systems* (Wiley, 2004). In 1990, he conceived (with A. Polydoros) the principle of *Per-Survivor Processing*.

Dr. Raheli served on the Editorial Board of the IEEE TRANSACTIONS ON COMMUNICATIONS as an Editor for Detection, Equalization, and Coding from 1999 to 2003. He also served as a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Special Issue on Differential and Noncoherent Wireless Communications, published in September 2005. Since 2003, he has been on the Editorial Board of the *European Transactions on Telecommunications* as an Editor for Communication Theory.

**Michele Franceschini** was born in Milan, Italy, in 1977. He received the Dr. Ing. degree ("Laurea," 5-year program) in electrical engineering (*summa cum laude*) from the University of Parma in 2002 and the Ph.D. degree from the the University of Parma, where he held a postdoctorate position in 2006 and 2007.

Since March 2008, he has held a postdoctorate position at the IBM T. J. Watson Research Center, Yorktown Heights, NY. His research interests lie in the area of communication and information theory, with particular emphasis on low-density parity check code design, theoretical aspects of optical communication, advanced signal processing techniques, synchronization, and low-complexity implementation of digital communication systems.

Dr. Franceschini received the "Paolo Conti Award" in 2003 as the best graduate in information engineering at the University of Parma in the academic year 2002.

**Fausto Meli** was born in Piacenza, Italy , in 1956. He received the doctorate degree in nuclear engineering (optoelectronic specialization) in 1981 from the Politecnico di Milano, Milano, Italy.

In 1983, he joined Pirelli Cavi e Sistemi R&D Labs, working in the area of new optical fibers development and characterization. Starting in 1989, he was involved in a joint research program on optical amplification with Southampton University, U.K. He then became responsible for optical development with Pirelli Optical Systems, contributing to the design of the first field deployed optically amplified system and the first and next generations DWDM optical systems. In the same period, he was also engaged in EC projects like RACE II/MWTN (Multi Wavelength Transport Network) and ACTS/FAST (Fluoro-aluminate Amplifier for Second Telecom Window). In 2000, he joined Cisco Photonics, Italy, as a result of Pirelli Optical Systems acquisition by Cisco Systems, where he was responsible for the Optical Development and Advanced Technology teams. He is coauthor of more than 45 international patents filed and of several scientific journals and conference papers in the different fields covered. During the last two years, as part of the engineering technical staff, he was involved in future technology developments for optical systems, also leading the optical network design support activity.

**Andrea Castoldi** was born in Monza, Italy, on November 24, 1969. He received the Laurea degree (M.S.) in physics from the Universita' Statale degli Studi of Milano, Milano, Italy, in 1995.

From 1995 to 1999, he was a Designer of integrated optical components with Pirelli Optical Systems, Italy. Since 2000, he has been an optical designer with Cisco Photonics, working in the field of fiber nonlinearities for optical network planning. He is the author of four patents on optical components technology.