Computationally Efficient Monitoring of PON Fiber Link Quality Using Periodic Coding

Mohammad M. Rad, Habib A. Fathallah, Sophie LaRochelle, and Leslie A. Rusch

***Abstract*—We investigate experimentally and via simulation, the monitoring of fiber link quality in a PON using optical coding technology. We discuss design issues for periodic coding and the optimal detection criteria. We develop a reduced-complexity maximum-likelihood sequence estimation (RC-MLSE) algorithm. We conduct experiments to validate our detection algorithm using four periodic encoders that we designed and fabricated. Error-free detection is confirmed for encoders with separation as small as 1 m. Using the experimental data for the encoder impulse responses, we conduct Monte Carlo simulations for realistic PON geographical distributions with randomly located customers, again with error-free detection. We investigate the effect of coverage area and network size (number of subscribers) on the computational efficiency of our algorithm. We provide a bound on the probability that any given network will cause the algorithm to take exorbitant time to monitor the network, i.e., the time-out probability. Finally, we highlight the importance of averaging to remedy the power/loss budget limitations in our monitoring system to support higher network sizes and longer fiber reaches.**

***Index Terms*—Detection probability (PD); Maximum-likelihood sequence estimation (MLSE); Network monitoring; Periodic coding; PON; Optical time domain reflectometry (OTDR).**

I. Introduction

Monitoring of a passive optical network (PON) can significantly reduce the operational expense (OPEX) for both operators and customers and so plays an important role in the deployment of future access networks [1–4]. The growth of PONs (supporting more customers per fiber with higher data rates and better quality of service) increases both the importance and the complexity of network monitoring.

Optical-time-domain reflectometry (OTDR) is efficient for testing optical devices and monitoring of point-to-point (PTP) networks [1]. For the installation and maintenance of tree architecture PONs, testing of each distribution drop fiber (DDF) segment separately requires the OTDR be run from the remote node to the customer premises and vice versa. OTDR is less suitable for centralized monitoring (i.e., from the central office rather than the remote node) of point-to-multipoint (PMP) networks like fiber-to-the-home (FTTH) PONs [1,4]. The backscattering signal of each branch in a PMP network is partially masked by that of the other branches.

A few solutions have been proposed to adapt the standard OTDR to a PON, among which the most well-known is the technique based on reference reflectors [1,4]. A reference reflector at each branch end provides information on the integrity of that branch. The technique is not viable when branches are located very closely or at the same distance; thus drops are required to have distinct lengths. While feasible at system installation, this approach is unmanageable during incremental expansion and routine maintenance. As the network size increases, the reliability and manageability of this technique decreases.

In [5], we introduced a modified optical code division multiplexing (OCDM) scheme for centralized monitoring of PONs that is architecture agnostic. The reference reflector is replaced by a coding mirror. The multiple reflections from subscribers are distinguishable by proper decoding, and customers no longer need to be connected to the central office (CO) with unequal fiber length. The geographical distribution of the customers has significant impact on the performance, and we developed a Monte Carlo simulation that statistically averaged over subscriber locations [6].

We analyzed optical orthogonal codes (first proposed for coded data) for the monitoring application, examining both one-dimensional and two-dimensional codes [6,7]. In [8] we developed codes better adapted to location detection (instead of using codes developed for data detection), whose implantation costs were also significantly reduced from those studied in [5–7]. These new, simple, and cost-effective codes were periodic in nature. While [8] examined these codes to explore the autocorrelation and cross correlation, in this paper we examine optimal detection algorithms that far outperform correlation for periodic codes (PCs).

Manuscript received April 13, 2010; accepted September 21, 2010; published December 27, 2010 (Doc. ID 126818).

M. M. Rad is with the Electrical and Computer Engineering Department at the University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1.

H. A. Fathallah is with the Electrical Engineering Department, College of Engineering, King Saud University, Riyadh, Saudi Arabia, and the Electrical and Computer Engineering Department, Université Laval, 1065 Avenue de la Médecine, Quebec, Canada.

S. LaRochelle and L. A. Rusch (e-mail: larusch@ulaval.ca) are with the Department of Electrical and Computer Engineering and The Center for Optics, Photonics, and Lasers (COPL), Université Laval, 1065 Avenue de la Médecine, Quebec, Canada.

Digital Object Identifier 10.1364/JOCN.3.000077

1943-0620/11/010077-10/$15.00

A. Principle of Operation

Our monitoring system is illustrated in Fig. 1. A U band (i.e., 1625–1675 nm devoted for PON monitoring [1,2]) short pulse with peak power P_s and duration T_s is transmitted through the feeder and split into N subpulses at the remote node (RN). Each pulse is encoded and reflected back to the CO by a dual-function device: a coder and mirror that we refer to as a coding mirror (CM_n, $n=1,\ldots,N$) [5]. Note that the monitoring signals are carried on λ_m while the data wavelength is λ_d in Fig. 1. Each DDF is terminated by a CM with a unique code and is located physically close to the optical network terminal (ONT). The CO monitoring equipment receives the accumulation of all sequences coming from CMs. Information on an individual DDF is discernable at the CO by exploiting the code structure.

Features of the encoded sequences could be used to assess the link quality for individual fiber branches. For instance, for a break in a DDF, no corresponding autocorrelation peak is observed; similarly, the lack of autocorrelation peaks for all CMs identifies a fault in the feeder [5]. Decreasing autocorrelation indicates that the link between the central office and the ONT is degrading. When all autocorrelation peaks decrease, the problem originates in the feeder segment. In practical PON installations, network segments consist of bundles of fibers with different counts. When an external event occurs (i.e., breaks or bends), all channels in that cable are affected.

B. Periodic Coding

Periodic codes can be implemented using only two fiber Bragg gratings (FBGs) written for the same waveband; the first is partially reflective, while the second acts as a frequency-selective mirror. By inserting a patch cord between the gratings, an optical cavity is formed [8]; see Fig. 1(b). A single incident pulse generates an infinite-length periodic sequence of multilevel subpulses from the cavity. A code is determined by the number of silent pulse intervals separating the multilevel pulses, or equivalently the period of the code, p_n. The length ℓ_n patch cord determines the period p_n, which in turn distinguishes different users. For a pulse width of T_s, the periodicity p_n and ℓ_n are related by

$$\ell_n = c\frac{p_n T_s}{2n_g}, \tag{1}$$

where c and n_g are the speed of light and the fundamental mode group index that determines the time of flight. The grating pair is the same for all CMs and only their separation ℓ_n (equivalently the period number p_n) changes with the code. As explained in [8], the second grating of the encoder, tuned to reflect the monitoring wavelength, eliminates the requirement for a separate wavelength selector per subscriber to demultiplex the data band [5]. Periodic CMs are small in size and easy to handle.

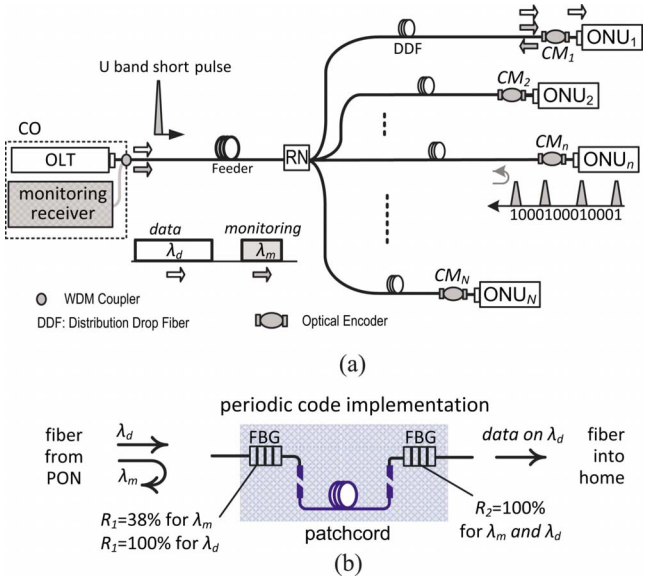

Fig. 1. (Color online) (a) Optical-coding-based PON monitoring, (b) the structure of the periodic encoder.

II. Periodic Encoder Design

This section focuses on the design of the gratings required to make periodic encoders. Further details can be found in [8].

A. Monitoring Wavelength

The most common wavelengths for out-of-band monitoring are 1625 nm or 1650 nm [9–11]. At 1650 nm we have higher sensitivity to macrobends than at 1625 nm, and more strict isolation is required at 1625 nm as it is closer to the data band (L band, i.e., 1565–1625 nm). For these reasons, the ITU-T recommends 1650 nm [1].[1] Although 1625 nm has yet to be adopted in the standard, it is more economical and readily available than 1650 nm lasers. In our experimental setup, we operate at 1550 nm due to equipment availability; however, coder design is valid at 1625 nm as well. Our experimental objective is capturing traces for our simulations. The difference in wavelength should only affect attenuation, which is corrected in simulations to reflect values for 1625 nm.

B. Reflectivity

As explained in [8], to concentrate the total energy in a few reflected pulses (shorter codes) the grating reflectivities are found to be $R_1=38\%$ and $R_2=100\%$ for all encoders. This fixes the transmission loss of the gratings, i.e., $T_1=2.1$ dB and $T_2=\infty$, respectively, in the wavelength interval around the monitoring wavelength. Typical values of $T_2=50$ dB can be achieved via conventional FBG fabrication [9,11]. Again recall that all the encoders have the same specifications for FBG1 and FBG2. The implemented spectral characteristics of the FBGs are given in Fig. 2.

C. Bandwidth

In choosing the encoder bandwidth, d, we must take into account: 1) the effect on the isolation of data and monitoring bands and 2) the availability of bandwidth and spectral ef-

[1]International Telecommunication Union (www.itu.int)

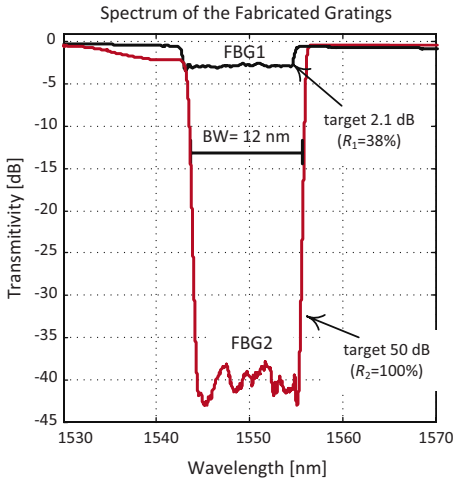

Fig. 2. (Color online) Experimental implementation of FBG1 and FBG2 forming the periodic encoder.

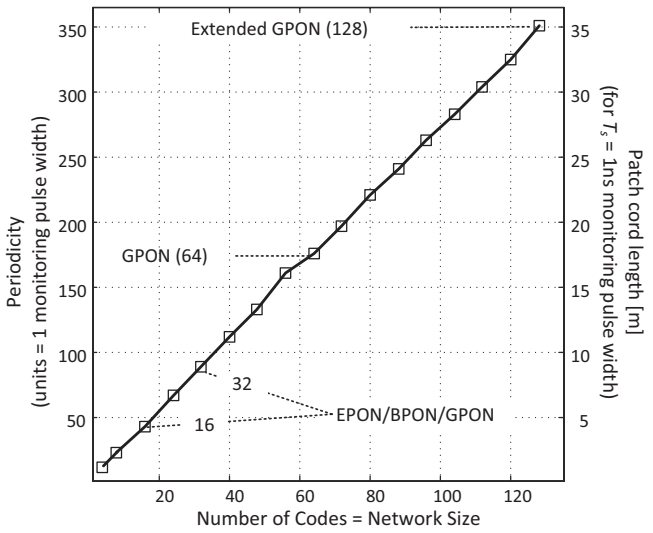

Fig. 3. Code periodicity (patch cord length) versus the network size for periodic code.

ficiency trade-offs. We can choose a rejection bandwidth d that is spectrally efficient, so long as it also returns a large portion of the energy from the source and is tolerant to some source drift. While the U band is 50 nm wide, we may wish to create several bands, either for monitoring or for other information bands. In our experiments we fix $d=12$ nm as a reasonable compromise [11].

D. Fiber Length Optimization

Once the pulse duration, T_s, of the monitoring signal is fixed, the length of the fiber joining FBG1 and FBG2 is determined per Eq. (1). The pulse width thus impacts the compactness of our encoders. It also impacts the sampling rate required at the digital receiver. For our simulations and experiments we have fixed $T_s=1$ ns, although other system designers may seek another compromise.

For the results reported in this paper, we generate codes using the simple algorithm developed in [8] for code weight (to $w=4$) and cross correlation. The periodicity of each code p_n $(n=1,\ldots,N)$ determines the corresponding fiber length ℓ_n per Eq. (1). Any desired number of codes can be generated. In Fig. 3, we plot the periodicity versus the network capacity for a $T_s=1$ ns pulse. From this we can also read the maximum required patch cord length for a given number of subscribers. Periodic coding can support current capacities of PONs as well as future expansion [1,2]. For a GPON with $N=32$ branches, the first periodic code CM requires a $\ell_1=0.6$ m patch cord, while the last one requires a $\ell_{32}=8.3$ m patch cord. New CMs (for new subscribers) are simple to design and fabricate; hence our solution is flexible and scalable.

III. Transmitter and Receiver Considerations

A. Transmitter Module

The transmitter module has a U band laser that launches a short pulse of duration T_s into the network. For our simulations we assume a 1650 nm monitoring wavelength [9–11] and a light source sideband suppression ratio below −80 dB for good data isolation [11]. The repetition rate T_r for pulse transmission is constrained by the distance between the CO and the farthest ONT. Repetition rates on the order of kilohertz would be typical. Both direct and indirect modulation schemes can be used to generate the U band monitoring pulses. Experimentally, we consider an externally modulated spectrum-sliced broadband light source (SS-BBS) operating at 1543 nm. Note that a BBS has previously been used for the monitoring applications [12]. Direct modulation would incur lower losses and reduce cost.

B. Receiver Module

All decoding functions are implemented in electronics following simple photodetection, as seen in Fig. 4. No programmable optical decoding is required, significantly reducing the cost (and size) of receivers as compared to other optical coding solutions. In previous proposals, the CO needs programmable optical decoders or a bank of all-optical decoders to decode the signal of each decoder [5–7]. By implementing the decoding in electronics, we obviate this expensive, complex, and bulky all-optical module. In addition, electronic processing reduces power insertion and decoding loss [8]. The sampling rate is determined by the pulse width. For our choice of $T_s=1$ ns, gigahertz sampling is required. The analog-to-digital converter (ADC) should provide good quantization (8–10 bits per sample) for our detection algorithm; such ADCs are a mature, commercially available technology.

Digital signal processing will accomplish multiple functions. First among them is the averaging of received waveforms for noise reduction [13]. As no training sequence is

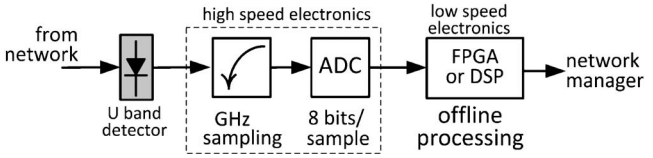

Fig. 4. (Color online) Receiver architecture.

used, the processing must perform pulse synchronization. These functions are not addressed in this article, as we focus on fiber fault detection beginning in the next section.

IV. Decoding Algorithm for Error-Free Detection

Detection requires information on encoder impulse responses, network loss budget, and characteristics of the transmitted pulse. We assume this information is known. Our detection algorithm has three objectives:

1) *Fault detection*: identify the presence or not of all subscribers without prior knowledge of subscriber location (i.e., fiber length between the remote node and the subscriber location).
2) *Network discovery*: estimate the fiber length of each subscriber.
3) *Channel condition monitoring*: estimate signal features for each subscriber.

The primary mission (objective one) is to identify unhealthy subscriber lines; a binary decision is made for each subscriber. The secondary mission is to gain information on the quality of each subscriber line. For this, the estimation of the subscriber position (or equivalently the fiber length) is essential so that signal processing can be directed at this portion of the returned signal. At a minimum, objective two will provide information on signal strength. Additional signal processing can be used to determine additional features and accomplish the third objective, although this is not explored in this article.

A. Maximum-Likelihood Sequence Estimation

Optimal detection of a single subscriber coded signal in additive white Gaussian noise requires a simple matched filter or correlator. Due to the simultaneous presence of multiple subscriber signals, the optimal detection algorithm must consider all possible states of the network before deciding whether each particular subscriber signal is present [14]. The received signal must be compared in turn with an expected return signal for each possible network configuration; this constitutes a classic maximum-likelihood sequence estimation (MLSE) algorithm. A network configuration refers to the presence or not of a subscriber signal, as well as the lengths of fiber between the remote node and each subscriber.

The search space of the MLSE algorithm depends on 1) network size (number of subscribers), 2) coverage area (maximum fiber length from remote node to subscriber), and 3) sampling rate of the monitoring system (samples per pulse duration). Even moderate values of network parameters yield an exceedingly large search space. For example, a PON with only $N=4$ subscribers with a maximum fiber run of $\Delta l=12.6$ m (500 m^2 coverage area) and a receiver sampling rate of $R_s=1$ Gsps requires 126^4 comparisons for MLSE.

A brute force algorithm for MLSE is prohibitively complex, especially as network size grows. The most commonly employed algorithm for MLSE is the Viterbi algorithm that associates a finite state machine with the sequences to estimate. The Viterbi algorithm exploits the state transitions to eliminate at each symbol interval sequences with low probability that need not be retained for consideration. The Viterbi algorithm is not only less computationally complex, it is also very efficient in terms of memory usage. The continuous transmission of symbols whose values are influenced by preceding symbols (as with intersymbol interference and forward error correction) can be described by a finite state machine and is therefore compatible with the Viterbi algorithm.

Not only is the Viterbi algorithm applicable only to finite-state Markov systems, the transition probabilities should also only depend on the corresponding transition states. For our monitoring application, constructing the Markov system is very difficult due to the random distribution of the customers in the network. In addition, the complexity of the state diagram depends on network size, sampling rate, coverage area, and symbol size. For instance, for an eight customer network using periodic coding technology, the corresponding state diagram has more than 200 states. Therefore, the Viterbi solution is not appropriate for our PON monitoring application.

Note that the reduced memory requirements for the Viterbi algorithm are due to the processing of sequences at each symbol interval and retention of only potentially "winning" sequences. This is an extremely powerful and important aspect of the Viterbi algorithm for data communications. The PON monitoring problem is very different. For monitoring we need to use signal averaging to improve the signal-to-noise ratio, as is common in OTDR methods [13]. For this reason, memory is reserved for the sampled return signal; moderate delay for averaging and processing (milliseconds to seconds) is acceptable for this application. Therefore the MLSE algorithm can process the entire sampled signal return without an overhead for memory. We next present an algorithm for MLSE computation that is adapted to the PON monitoring problem and to periodic codes in particular [14].

B. Reducing MLSE Complexity for PON Monitoring

There are two characteristics of the PON monitoring return signal that we will exploit to accelerate the MLSE. As in the Viterbi algorithm, we seek to eliminate low probability network configurations, so that only highly probable configurations need be compared in the MLSE search. The first characteristic we will exploit is the correlation distance of a particular code in the network, and the second is the structure of the periodic codes.

The correlation distance of a code is defined by

$$l_{CD}^{(n)}=\frac{p_n T_s(w-1)c}{n_g},\quad n=1,\ldots,N, \tag{2}$$

where p_n is the periodicity of the corresponding code and $w-1$ is the number of periods between the first and last pulse [6]. Only subscribers whose fiber lengths differ from that of user n by less than $l_{CD}^{(n)}$ may potentially interfere with user n. The geographical distribution of the ONTs will impact the number of subscribers falling within this correlation distance.

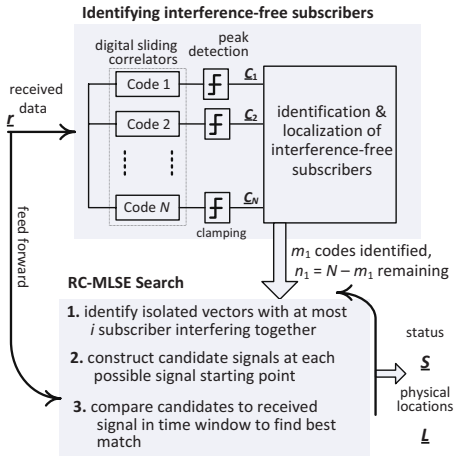

Fig. 5. (Color online) Reduced-complexity MLSE algorithm.

Due to the correlation distance, in PON monitoring we expect that only a relatively small subset of subscribers will contribute to multiple access interference (MAI), in contrast with data communications, where all active subscribers contribute to MAI. In our algorithm we will process signals in the order of increasing MAI. Signal returns without interference will be identified first and eliminated from the subsequent searches. Signals with one interferer will be identified next, etc. The periodic structure of the codes allows for a simple test for the presence of a code. Once a single pulse has been identified, we search in time to see whether other $w-1$ pulses are present at time intervals corresponding to the code period under test. We will use a digital sliding correlator to identify possible locations for specific subscriber returns.

C. RC-MLSE Algorithm

The algorithm can be divided in two parts as illustrated in Fig. 5. The first part uses N parallel sliding (matched-filter) correlators of the periodic codes to generate our test statistic. Interference-free subscribers are detected and localized without requiring an MLSE search. This step determines the uncertainty space for MLSE searches of remaining subscriber status and location. The second part performs a serial search on possible status/locations; at each iteration, progressively more computationally complex searches are made.

We make the following definitions. We have N codes with periodicities $\{p_n\}$ corresponding to N subscribers; $|p_n|$ is the length of the nth code; the Nth code is the longest code; and a vector of length N_s is formed from the sampled return.

Identifying Interference-Free Users

1) Run N simple (digital) sliding matched-filter correlators against the sampled return signal. Each sliding correlator outputs a vector $\underline{C}_n$ with a one as an element if the correlator output exceeds a threshold, i.e., if it is possible that code p_n begins at that time sample.
   - For vectors $\underline{C}_n$ with Hamming weight zero the nth subscriber line is declared faulty.
   - The vectors $\underline{C}_n$ with Hamming weight one (with only one nonzero element) are identified.
     ◇ The fiber length to subscriber n is deduced from the position of the nonzero entry in vector $\underline{C}_n$.
     ◇ The nth link is declared healthy and its location is noted.
     ◇ The nth code is removed from the set of codes to be searched, so it is not considered in future iterations.
2) Find among (the now reduced set of) vectors $\underline{C}_n$ any window of length $2\{p_n\}$ where only $\underline{C}_n$ (and no other $\underline{C}_n$) has a single nonzero element, i.e., an isolated peak.
   - The nth fiber length is deduced from this window; it is declared healthy and its location is noted.
   - The nth code is removed from the set of codes to be searched.

Iterative RC-MLSE Search

At iteration i, the remaining collection of $n_i \in \{1,\ldots,N\}$ vectors $\underline{C}_n$ are examined to form the mutual interference matrix (MIM_i). Recall that an individual subscriber j may have many "false positive" hits in the vector $\underline{C}_j$. This matrix indicates where subscribers may possibly overlap.

The MIM_i is an $n_i \times n_i$ matrix; each row (column) corresponds to one of the remaining n_i subscribers to be localized through the MLSE search. Entries of the MIM are binary. A zero entry ($a_{h,k}=0$) indicates no mutual interference between the subscribers represented by the hth and kth row/column. MIM is a symmetric matrix by construction. The diagonal entries are by definition 1. In the minimum interference case, MIM is an identity matrix. For the highly interfering case, MIM is a unitary matrix, i.e., $a_{h,k}=1$ for all h and k.

1) The matrix MIM_i is formed from the remaining unidentified subscribers.
2) The MIM_i is examined to identify all sets of i subscribers that interfere only among one another.
   - If no such sets exist, i is incremented.
   - If such sets exist, for each set an exhaustive search algorithm is executed for the set; for each set,
     ◇ a candidate signal is generated for each nonzero position of the correlator output for each subscriber in the set,
     ◇ all possible combinations of candidate signals are combined to produce a collection of candidate received vectors,
     ◇ the candidate received vector closest to the measured received signal $\underline{r}$ is declared the winning combination of candidate signals,
     ◇ the detected subscribers are removed from the set of codes to be searched at the next iteration.
3) The algorithm continues looking for increasingly larger sets of mutually interfering users. Once such sets are no longer identifiable, an exhaustive search is made of the remaining unidentified $\underline{C}_n$.

D. Challenges of the RC-MLSE Algorithm

Clearly the RC-MLSE offers significant computational savings by exploiting the PON monitoring environment and periodic code structure. Analysis is problematic as the complexity is a function of the geographic distribution of subscribers. Bounds on the performance are addressed in Section VI. In the next section, we validate our algorithm

experimentally, testing a severe case where up to four subscribers are virtually collocated; the difference in fiber lengths is minimal. As we will see, the RC-MLSE algorithm detects and localizes all four encoders, enabling live, in-service PON monitoring.

The clamping threshold (in the first stage of the algorithm) can be fixed under normal healthy conditions. When new subscribers are added, the uncertainty in the threshold increases. The receiver can determine numerically (via iteration) the correct threshold by gauging performance on already-known subscribers. Similarly, an iterative approach would be necessary to set the threshold on a new network installation. Thus we anticipate that computation requirements would be greater at network setup as opposed to steady state. The increase, however, would be linear in the number of thresholds tested.

V. Experimental Validation

We fabricated four periodic encoders having the description given in Section II. Typical FBG characteristics are given in Fig. 2(b): a $d=12$ nm rejection bandwidth centered at 1550 nm. Using code weight of $w=4$ and cross correlation of one, we produced codes with periodicity of $p_1=6$ (60 cm), $p_2=7$ (70 cm), $p_3=13$ (130 cm), and $p_4=16$ (160 cm). We excluded the code with periodicity of 1 due to difficulty in fabrication.

A. Experimental Setup

The experimental setup is illustrated in Fig. 6. The setup has two main sections, the central office and a network of four subscribers with distinct CMs. For experimental convenience, the monitoring source is a BBS with optical bandwidth of 40 nm. The BBS is filtered (spectrally sliced) by a narrowband tunable filter with a 0.2 nm optical passband. The filtered BBS is preamplified by an erbium-doped fiber amplifier (EDFA). A Mach–Zehnder (MZ) external modulator is used to generate the monitoring pulse with duration of $T_s=1$ ns (1 GHz). The MZ is driven with a repetition interval of 500 μs.[2] To compensate the loss of the MZ we postamplified the modulated signal. Optical filters are used after EDFAs to suppress the amplified spontaneous emission (ASE) noise.

[2]This corresponds to unambiguous reception for a maximum distance of 50 km between the CO and CMs.

Using the setup in Fig. 6, in Fig. 7 we plot the time domain response of the encoders to the 1 ns probe pulse. The center wavelength of the monitoring source was tuned for the best possible impulse response and in our case it was $\lambda_m=1545.3$ nm. As is illustrated, the multilevel codes are nearly identical except for differences in code period. The DC value of each code does vary slightly due to variations in the grating spectra and contributions from the filtered ASE. In particular, the impulse response of the encoders is very sensitive to the reflectivity of the first grating.

The maximum transmitted power was kept low enough not to induce any fiber nonlinearity. The monitoring pulse is sent to the network using a circulator; a 20 km roll of fiber plays the role of feeder, and a 1×4 coupler serves as the RN. The fibers to each home are of roughly equal lengths (within 45 cm as seen in Fig. 7) corresponding to a very dense PON. Note that due to the round trip of the monitoring pulse through the network, the total loss is 2×20 km $\times0.2$ dB/km (feeder) $+2\times6$ dB (splitter)$=20$ dB. A real-time oscilloscope sampling at 10 Gs/s was used to capture the measured trace. The oscilloscope provides both sampling and averaging.

Recall that filtering a BBS increases the relative intensity noise (RIN) [13]. As a result, our monitoring light source suffers from significant RIN. We were able to use averaging to effectively remove the RIN of the source; the RIN power was reduced by 24 dB for averaging over 256 traces. Similar reductions are experienced in beat noise, shot noise, and thermal noises. Our measurement is therefore interference limited. Averaging is an important strategy even in deployments that will use a coherent source, as discussed in Section VII.

Figure 7(e) shows the total received signal when all four encoders are connected to the network. For a 17 ps/km·nm dispersion factor, the total amount of dispersion is 136 ps (0.2 nm optical bandwidth), which is negligible compared with the transmitted pulse width.

B. Simulation Results

The data captured in Fig. 7(e) was sent to our RC-MLSE

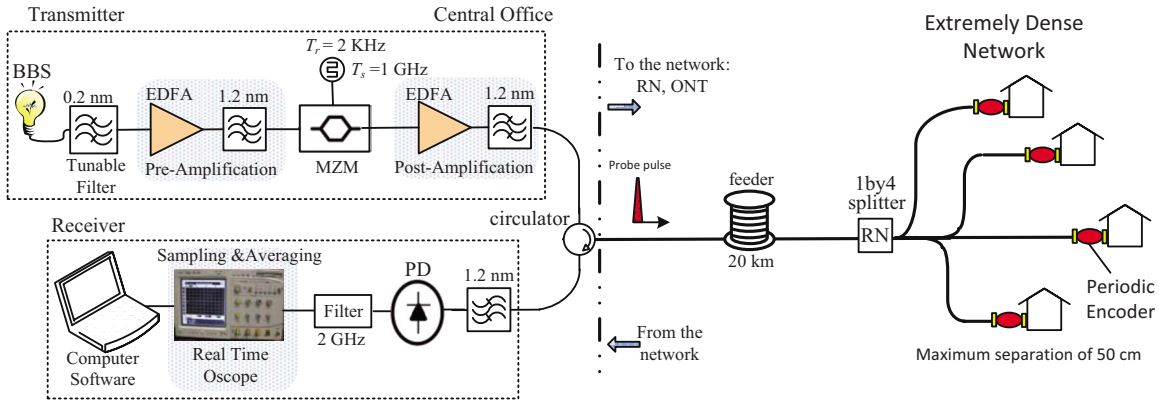

Fig. 6. (Color online) Experimental setup for the monitoring of a high-density 4 customer PON.

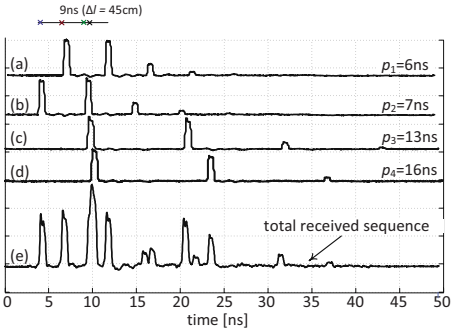

Fig. 7. Experimental traces of (a)–(d) isolated subscribers and (e) when all four subscribers are present; signals averaged over 256 traces to reduce noise.

algorithm. While having encoders (equivalently subscribers) less than a meter apart is not realistic, this represents a worst-case scenario in that the probability of interference is high. In a larger network, close proximity may occur for a subset of subscribers; hence we tested experimentally this extreme case in our detection algorithm. Our RC-MLSE algorithm correctly determined the fiber length to all CMs to within a pulse width (1 ns or ±10 cm).

These experimental results were also used as realistic traces of individual returns (i.e., taking encoders returns one by one). These returns were used in our simulator of larger, more realistic networks. Using the measured traces for the encoder impulse responses from our experimental setup, we completed Monte Carlo simulation (10^5 realizations) over a 10^4 m^2 coverage area. The subscriber locations followed a uniform radial (UR) distribution [6]. In all cases, we were able to detect and localize correctly all encoders; error-free detection was achieved ($P_D=1$).

VI. Impact of Coverage Area and Network Size

The previous experimental validation confirmed that for a given network topology with four users we were able to determine to very fine accuracy the positions of all subscribers without *a priori* information other than knowledge of the codes. Clearly even the exhaustive MLSE algorithm could have been used for this or any configuration; some configurations, however, would require exorbitant calculation times. In this section we seek to quantify the efficiency of the RC-MLSE algorithm as a function of network size and coverage area. It is not enough to be efficient on a specific configuration; we would like to have confidence that any configuration encountered could be monitored in a reasonable processing time.

We begin by presenting statistics on the effectiveness of the first part of the RC-MLSE algorithm. If most subscribers will not experience interference, our algorithm will have an easy time. What happens when there is heavy interference? This is addressed in a section on the time-out probability, the probability that our RC-MLSE algorithm will not produce estimates for all users during a set time limit. We run Monte Carlo simulations to bound this probability as a function of coverage area and network size, N.

We will see that the threshold for the sliding correlators will play a large part in the time-out occurrence. For problematic installations (that small percentage of installations whose geographical distribution causes the RC-MLSE algorithm to time out) we will propose a variable threshold. Therefore we begin with a brief discussion of the threshold setting.

A. Impact of the Threshold

The impact of the coverage area will be twofold. The smaller the coverage area, the more tightly packed the subscribers, and we expect more interference. However, for larger coverage areas we will also see resurgence in the interference due to the thresholding effect. As explained in Section IV, the threshold after the sliding correlator is adjusted to detect the autocorrelation peak of all encoders. The farthest subscriber determines the maximum height of the threshold. A lower threshold will lead to more false positive hits in the clamped correlator vectors $\{\underline{C}_n\}$. Therefore, increasing the coverage area causes more perceived interference and needless searches in the second stage of the RC-MLSE algorithm.

Let Δl be the maximum possible separation between the RN and a subscriber. The size of the coverage area will, of course, impact Δl. Assuming a uniform radial distribution of subscribers, we have $\Delta l=\sqrt{Area}/\pi$ [6]. In the following, we investigate the effect of coverage area (parameterized by Δl) and network size (N) on the efficiency of our proposed algorithm.

B. Percentage of Interference-Free Customers

We performed extensive Monte Carlo simulations to gather statistics on the RC-MLSE algorithm. We fixed the coverage area and network size per Table I. For 10^6 iterations, we randomly draw a location for each subscriber assuming a uniform radial (UR) distribution [6]. Using experimental traces from Section V we constructed a received signal and run the first section of our RC-MLSE. We record the percentage of interference-free subscribers and report the results in Table I.

For small network sizes, increasing the coverage area improves the percentage of interference-free users. However, for very large coverage areas the threshold is lowered, so vectors $\{\underline{C}_n\}$ contain more nonzero entries. The uncertainty (number of candidate locations) per user increases and

TABLE I
Percentage of the Users Identified as Interference-Free for Different Network Sizes (N) and Coverage Areas ($\pi\Delta l^2$) for $R_s=1$ GHz

		Network Size (N)			
		8	16	32	64
Coverage Area ($\pi\Delta l^2$) [km^2]	0.01	96%	82%	18%	1%
	0.1	99%	98%	77%	11%
	1	99%	99%	96%	72%
	10	99%	98%	89%	58%

fewer users are identified as interference free. Consequently more users pass to the second stage of the RC-MLSE search.

C. Time-Out Probability (P_{TO})

A reduced-complexity search is required to identify customer locations in a timely manner. The processing time should be limited to at most a few minutes to detect faults and validate new installations without undue delay. This value is somewhat subjective; however, clearly there is some limit to tolerable processing delay.

In the previous section we quantified the efficiency of our algorithm by determining with what frequency exhaustive searches in the second stage were required. In this section we quantify the processing delay of the second stage. We define a time-out to be a network realization that would result in processing time exceeding 2 minutes in our Matlab implementation of the RC-MLSE algorithm.

The time-out probability is a complex function of the code correlation, network size (N), coverage area (or Δl), sampling rate (R_s), processor speed (PC, FPGA, or DSP), algorithm implementation, and the specific choice of the acceptable processing time (T_{TO}). We fix several of these parameters ($T_{TO}=2$ min, $R_s=2$ Gsps, 2.3 GHz dual core Pentium) and use Monte Carlo techniques to estimate an upper bound, P_{TO}, to the time-out probability. Details of the estimation method are provided in Appendix A.

Figure 8 shows the simulation results for the upper bound of the time-out probability P_{TO} versus the maximum separation length Δl for different network sizes. The estimate is based on 10^5 Monte Carlo trials. As can be seen, Δl has a similar effect on P_{TO} as it does on the percentage of interference-free users. For both very high network densities (left side of the plots) and very large coverage areas (right side of the plots), P_{TO} is high. Moderate values of the coverage areas result in minimal P_{TO}. For very small coverage areas, the interference is very high; so the uncertainty per user is high; this results in a very large search space and high P_{TO}. Increasing the coverage area reduces the interference probability and more users are identified as interference free. Consequently, the search space reduces and P_{TO} improves significantly for moderate coverage areas. For very large coverage areas, the threshold adjustment leads to an increase in cases of suspected mutual interference. In this region, the search space increases and directly degrades P_{TO}. In Fig. 8 we have identified two regions that are problematic: to the left an interference-dominated region exists, while to the right is a region we call threshold limited.

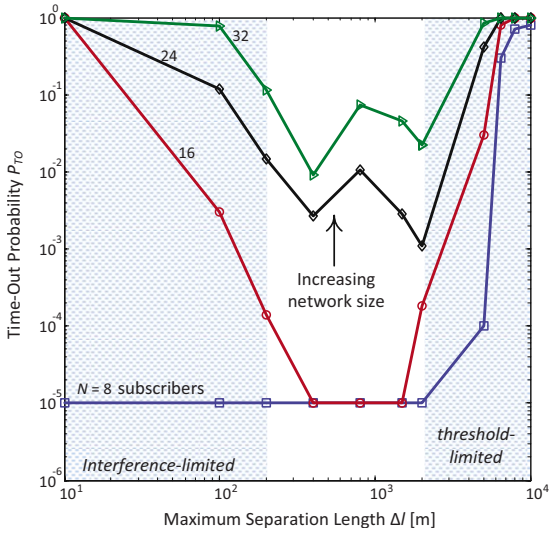

Fig. 8. (Color online) Time-out probability (P_{TO}) versus the maximum separation length between the customers (Δl) for different network sizes.

From Fig. 8, we see that for an $N=8$ customer PON, P_{TO} is high only for very large Δl ($\Delta l \geq 5$ km). For $N=16$, P_{TO} is high for both very small and very large Δl; for moderate values of Δl, P_{TO} is minimal. Note that for practical coverage areas (such as $\Delta l=1$ km) and a 16 customer network, no time-out is observed [15].

D. Variable Threshold to Counteract Time-Out

Recall that no prior information about the customer geography is used in our proposed algorithm. In fact, the only information is the maximum separation length Δl. Therefore, increasing Δl increases the uncertainty about each customer location. However, in practical PON deployment, customers are typically distributed in different geographical tiers. Each tier supports a local grouping of subscribers. Depending on the deployment technology, tiers are separated from each other from 2 to 6 km.

For tiered deployments, information about the network configuration, such as relative distance between tiers or the proximity of some users to the CO, can greatly reduce the time-out probability. Subscribers in different tiers can be distinguished in the total received monitoring signal by their corresponding separations in physical distance. Hence, instead of processing all the monitoring data, a window related to a specified tier is processed separately and independently from the others. The thresholds would be determined independently, with significant improvement in the time-out probability.

VII. Loss Budget Limits: Importance of Averaging

As mentioned earlier, due to the round-trip path of the monitoring pulses, the insertion loss of passive elements is doubled, imposing strict limitations on the total loss budget of our monitoring system. In particular, a higher number of customers leads to higher splitting ratios at the RN and a more critical power/loss budget. For instance, for a 16 (32) customer PON, the total insertion loss only due to RN splitting is 24 (30) dB.

Note that increasing the transmitted power to mitigate the loss budget increases nonlinear effects and cannot be employed without restraint. To avoid nonlinear impairments the launched power should be kept smaller than approximately 10 dBm for a single-mode fiber. Our strategy for higher loss budget costs relies on signal averaging. By averaging, i.e., repeating the measurement many times, we reduce noise and increase our signal-to-noise ratio (SNR) [13]. The dynamic range and sensitivity of the monitoring re-

ceiver significantly increases. As we are monitoring and not transmitting data, the temporal overhead for averaging is not burdensome.

When employing signal averaging, the measurement becomes interference limited; quantum-limited detection is achieved. Consequently the minimum detectable power is limited only by detector dark current. Therefore, the monitoring receiver is able to detect very small powers comparable to dark current. As an example, consider a CO that uses a PIN detector with 1 nA dark current and unitary responsivity [13], a total 10 dBm launched power for the monitoring signal, and a modulator with infinite extinction ratio. In this case, the DC current in the absence of monitoring pulses is 1 nA, with a minimum detectable power of 1 nW. Now consider a 10 dB margin for the received monitoring pulse power compared to the dark current DC level; i.e., set the minimum acceptable power for monitoring the pulse to 10 nW. The total permissible (round-trip) loss is 60 dB. Thus we are able to support high network sizes such as 32 (30 dB loss at the RN) with long fiber lengths such as 20 km (12 dB loss). Recall that the loss/budget can be further improved by using high-gain avalanche photodiodes (APDs) with very low dark currents [13].

VIII. Conclusion

In this paper, we experimentally investigated optical-coding-based PON monitoring. We addressed the design issues of recently proposed periodic encoders, fabricated them, and used them for fiber link quality monitoring of a PON. Four customers at extreme proximity, simulating a high-density PON, were examined in our experiments. We developed a reduced-complexity maximum-likelihood sequence detection (RC-MLSE) algorithm that distinguishes, detects, and localizes each encoder in the network. The measured data were fed into our algorithm. All users were detected/localized correctly, demonstrating error-free detection performance. We discussed the coverage area impact on the efficiency of our proposed algorithm. We simulated the performance of our algorithm in terms of time-out probability. Our simulation results validate our algorithm in detecting and localizing all encoders. The importance of averaging is also highlighted to remedy the loss/budget limitations.

Appendix A

Time-Out Upper Estimation via Monte Carlo Simulation

To explain in detail our estimate of the time-out probability reported in Fig. 5, we begin by examining the performance of the second stage of our algorithm in Matlab, version 7.6. The time of calculation is a function of the size of the received vector, $\underline{r}$, and the number of candidates $\hat{\underline{r}}$ to be examined. The size of the received vector can be parameterized by Δl, i.e., the larger the coverage area, the longer the received vector. For several Δl, we find the maximum number of $\hat{\underline{r}}$ that can be searched without a time-out, $R_{max}(\Delta l)$. We ran simulations on an Intel dual-core CPU operating at 2.13 GHz with 2 GB RAM.

During the mth iteration of the second stage, a reconstructed return signal is calculated for subsets of mutually interfering subscribers. The statistics of these subsets are difficult to ascertain. Therefore we use a worst-case bound to find the time-out probability. We assume that during the mth iteration, all m users must be examined in the exhaustive search. Let the random variable u_i $\{i=1,\ldots,m\}$ be the number of candidate locations (time indexes) per user to be examined by the MLSE search. These candidates $\{u_i\}$ are extracted from the clamped correlator output vectors $\{\underline{C}_n\}$ during the first stage. As we assume all subscribers are mutually interfering, the second stage would need to calculate $u_1 \times u_2 \times \cdots \times u_m$ candidate $\hat{\underline{r}}$ to compare with $\underline{r}$.

To estimate the time-out probability we run Monte Carlo trials to determine with what frequency $u_1 \times u_2 \times \cdots \times u_m$ exceeds $R_{max}(\Delta l)$. A coverage area is fixed, a network is generated randomly, and the vector $\underline{r}$ is constructed and passed through the bank of correlators. From the correlator output vectors $\{\underline{C}_n\}$ we find $\{u_i\}$, compute $u_1 \times u_2 \times \cdots \times u_m$, and compare it to $R_{max}(\Delta l)$. This gives us the time-out probability (P_{TO}) as a function of coverage area.

Acknowledgment

The authors would like to thank G. Tremblay and S. Doucet for their contribution in encoder fabrication. This paper was presented in part at IEEE GLOBECOM, Honolulu, Hawaii, USA, Dec. 2009.

References

[1] A. Girard, "FTTx PON technology and testing," EXFO Electro-Optical Engineering Inc., 2005.
[2] http://www.ftthcouncil.org/.
[3] R. Davey and D. Payne, "The future of optical transmission in access and metro networks—an operator's point of view," in *Proc. of the 31st European Conf. on Optical Communication*, 2005, paper We 2.1.3.
[4] F. Caviglia and V. C. Biase, "Optical maintenance in PONs," in *Proc. of the 24th European Conf. on Optical Communication*, 1998, pp. 621–625.
[5] H. Fathallah and L. A. Rusch, "Code division multiplexing for in-service out-of-band monitoring," *J. Opt. Netw.*, vol. 6, no. 7, pp. 819–829, July 2007.
[6] M. M. Rad, H. Fathallah, and L. A. Rusch, "Effect of PON geographical distribution on monitoring by optical coding," in *Proc. of the 33rd European Conf. on Optical Communication*, 2007, paper 7.6.6.
[7] M. M. Rad, H. Fathallah, and L. A. Rusch, "Fiber fault monitoring for passive optical networks using hybrid 1D/2D coding scheme," *IEEE Photon. Technol. Lett.*, vol. 24, pp. 2054–2056, 2008.
[8] H. Fathallah, M. M. Rad, and L. A. Rusch, "PON monitoring: periodic encoders with low capital and operational cost," *IEEE Photon. Technol. Lett.*, vol. 24, pp. 2039–2041, 2008.
[9] N. Gagnon, A. Girard, and M. Leblanc, "Considerations and recommendations for in-service out-of-band testing on live FTTH networks," in *Proc. of the IEEE Optical Fiber Communication Conf.*, 2006, paper NWA3.
[10] N. Nakao, H. Izumita, T. Inoue, Y. Enomoto, N. Araki, and N. Tomita, "Maintenance method using 1650 nm wavelength

band for optical fiber cable networks," *J. Lightwave Technol.*, vol. 10, pp. 1513–1520, 2001.
[11] N. Honda, H. Izumita, and M. Nakamura, "Spectral filtering criteria for U-band test light for in-service line monitoring in optical fiber networks," *J. Lightwave Technol.*, vol. 6, pp. 2328–2335, 2006.
[12] S. B. Park, D. K. Jung, H. S. Shin, D. J. Shin, S. Hwang, Y. Oh, and C. Shim, "Optical fault monitoring method using broadband light source in WDM-PON," *Electron. Lett.*, vol. 4, pp. 239–241, 2006.
[13] D. Deickson, *Fiber Optic Test and Measurements*. Prentice Hall, 1997.
[14] S. Benedetto and E. Biglieri, *Principles of Digital Transmission: With Wireless Application*. Springer, 1999.
[15] M. Vaughn, D. Kozischek, D. Meis, A. Boskovic, and R. E. Wagner, "Value of reach-and-split ratio increase in FTTH access networks," *J. Lightwave Technol.*, vol. 11, pp. 2617–2622, 2004.

Mohammad M. Rad received his B.S.E.E. and M.S.C., both from Sharif University of Technology, in 2003 and 2005, respectively, and his Ph.D. from The Center for Optics, Photonics and Lasers (COPL), Université Laval, in 2010. He is currently working as a Researcher at the University of Waterloo. His research interests include fiber-optic telecommunications, network monitoring, passive optical networks (PONs), free-space optics (FSO), and fiber-optic sensors.

Habib A. Fathallah (S'96, M'01) received the B.S.E.E. degree (with honors) from the National Engineering School of Tunis in 1994 and the M.S.C. and Ph.D. degrees in electrical engineering from Université Laval, Québec, Canada, in 1997 and 2001, respectively. He initiated the use of Bragg gratings technology for all-optical/all-fiber coding/decoding in optical CDMA systems. He was the founder of Access Photonic Networks (2001–2006). He is currently with the Electrical Engineering Department, College of Engineering, King Saud University (Riyadh, Saudi Arabia) and an Adjunct Professor with the Electrical and Computer Engineering Department of Laval University (Quebec, Canada). His research interests include optical communications systems and technologies, metro and access networks, optical CDMA, PONs and long-reach PONs, FTTH, network monitoring, and hybrid fiber wireless (FiWi) systems.

Sophie LaRochelle received her B.Eng. degree in engineering physics from Université Laval, Sainte-Foy, Québec, Canada, in 1987 and her Ph.D. degree in optics from the University of Arizona, Tucson, in 1992. From 1992 to 1996, she was a Research Scientist at Defense Research and Development Canada-Valcartier, where she worked on electro-optical systems. She is currently a Professor in the Department of Electrical and Computer Engineering, Université Laval, where she holds a Canada Research Chair in Optical Fibre Communications and Components. She is a member of the Center for Optics, Photonics and Lasers (COPL). Her current research activities are focused on active and passive fiber optics components for optical communication systems, including fiber Bragg gratings, optical amplifiers, and multiwavelength and pulsed fiber lasers. Her other research interests include packet-switched networks with photonic code processing, transmission of radio-over-fiber signals, and optical code-division multiple access. Dr. LaRochelle is a member of the Optical Society of America and the IEEE Lasers and Electro-Optics Society.

Leslie A. Rusch (S'91-M'94-SM'00-F'10) received the B.S.E.E. degree (with honors) from the California Institute of Technology, Pasadena, in 1980 and the M.A. and Ph.D. degrees in electrical engineering from Princeton University, Princeton, NJ, in 1992 and 1994, respectively. Dr. Rusch has experience in defense, industrial, and academic communications research. She was a Communications Project Engineer for the Department of Defense from 1980–1990. While on leave from Université Laval she spent 2 years (2001–2002) at Intel Corporation creating and managing a group researching new wireless technologies. She is currently a Professor in the Department of Electrical and Computer Engineering at Université Laval, Québec, Canada, performing research on wireless and optical communications. Prof. Rusch's research interests include wavelength division multiple access using incoherent sources for metropolitan-area networks; analysis of optical systems using coherent detection; semiconductor and erbium-doped optical amplifiers and their dynamics; and in wireless communications, optical pulse shaping for high-bit rate ultrawideband systems (UWB), as well as performance analysis of reduced-complexity receivers for UWB. She has served as an associate editor for *IEEE Communications Letters* and on several IEEE technical program committees.