Rate-Adaptive Coding for Optical Fiber Transmission Systems

Gwang-Hyun Gho, Lauren Klak, and Joseph M. Kahn, *Fellow, IEEE*

***Abstract*—We propose a rate-adaptive transmission scheme using variable-rate forward error correction (FEC) codes with a fixed signal constellation and a fixed symbol rate, quantifying how achievable bit rates vary with distance in a long-haul fiber system. The FEC scheme uses serially concatenated Reed–Solomon (RS) codes with hard-decision decoding, using shortening and puncturing to vary the code rate. An inner repetition code with soft combining provides further rate variation. While suboptimal, repetition coding allows operation at very low signal-to-noise ratio (SNR) with minimal increase in complexity. A rate adaptation algorithm uses the SNR or the FEC decoder input bit-error ratio (BER) estimated by a receiver to determine the combination of RS-RS and repetition codes that maximizes the information bit rate while satisfying a target FEC decoder output BER and providing a specified SNR margin. This FEC scheme is combined here with single-carrier polarization-multiplexed quadrature phase-shift keying (PM-QPSK) and digital coherent detection, achieving 100-Gbit/s peak information bit rate in a nominal 50-GHz channel bandwidth. We simulate variable-rate single-channel transmission through a long-haul system incorporating numerous optical switches, evaluating the impact of fiber nonlinearity and bandwidth narrowing. With zero SNR margin, achievable information bit rates vary from 100 Gbit/s at 2000 km, to about 60 Gbit/s at 3000 km, to about 35 Gbit/s at 4000 km. Compared to an ideal coding scheme achieving information-theoretic limits on an AWGN channel, the proposed coding scheme exhibits a performance gap ranging from about 5.9 dB at 2000 km to about 7.5 dB at 5000 km. Much of the increase in the gap arises from the inefficiency of the repetition coding used beyond 3280 km. Rate-adaptive transmission can extend reach when regeneration sites are not available, helping networks adapt to changing traffic demands. It is likely to become more important with the continued evolution toward optically switched mesh networks, which make signal quality more variable.**

***Index Terms*—Coherent detection, forward error correction (FEC), information rates, optical fiber communication, quadrature phase-shift keying, variable-rate codes.**

I. Introduction

EARLY wireline and wireless communication systems transmitted typically at fixed information bit rate, using a fixed forward error correction (FEC) code, modulation order and transmitted signal power. In such schemes, which do not exploit any channel knowledge, either link throughput is limited by the worst-case channel state or a robust link connection is not guaranteed. Many recent communication systems adapt link parameters based on channel conditions (often time-varying), trading off bit rate (and thus spectral efficiency) for robustness. For example, asymmetric digital subscriber lines transmit data over discrete multi-tones, maximizing link capacity by dynamically loading bits per frequency bin depending on the varying level of intersymbol interference (ISI), crosstalk, and signal-to-noise ratio (SNR) in each bin [1]. Similarly, many mobile wireless communication systems, e.g., WiMAX (IEEE 802.16) and high-speed downlink packet access, use adaptive coding and modulation to enhance performance over fading channels, based on channel quality feedback from the receiver or channel estimation in the transmitter [2], [3]. These systems also use hybrid automatic repeat request techniques to improve the throughput of the radio link, by combining variable-rate FEC codes and automatic repeat query functionalities in the physical layer [2], [3].

Previous optical transmission systems, such as those based on synchronous optical networking and synchronous digital hierarchy, employed fixed bit rates. As Ethernet-based systems are deployed in large-scale IP networks [4], transponders using coherent detection and digital signal processing will enable higher bit rates and spectral efficiencies, and could also support rate-adaptive transmission, with bit rates negotiated between routers and transponders. Optical switches, such as reconfigurable optical add-drop multiplexers (ROADMs), make it possible to route signals optically over longer distances but add loss and limit signal bandwidth, making signal quality more variable. As traffic demands evolve, it may become desirable to transmit signals over long routes on which regeneration sites are not available. Rate-adaptive transmission could improve network robustness, flexibility and throughput. Although present standardization efforts for optical transport networks are concentrated on transmitting client data at fixed rates [5], rate-adaptive coding may merit consideration for future networks.

To our knowledge, variable-rate transmission in optical fiber has been studied to date only in the context of code-division multiple access [6], or with limited available rate variation [7], [8]. In this paper, we describe a rate-adaptive transmission scheme for long-haul systems using a fixed modulation format (PM-QPSK) at a fixed symbol rate, with a family of hard-decision codes permitting rate variation over a wide range. Using simulation, we study achievable information bit rates as a function of transmission distance in a model terrestrial long-haul network.

This paper is organized as follows. In Section II, we review the evolution of FEC techniques in optical fiber systems. We

Manuscript received March 22, 2010; revised June 07, 2010, September 01, 2010, October 26, 2010; accepted December 01, 2010. Date of publication December 13, 2010; date of current version January 21, 2011. This work was supported by Cisco Systems and by a Stanford Graduate Fellowship.

The authors are with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305-9515 USA (e-mail: ggho@stanford.edu; klak@stanford.edu; jmk@ee.stanford.edu).

Digital Object Identifier 10.1109/JLT.2010.2099208

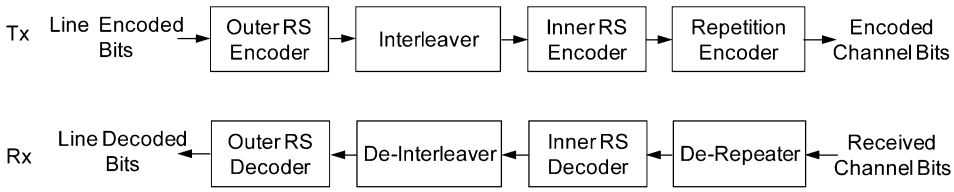

Fig. 1. FEC chain for encoding and decoding serially concatenated RS-RS and repetition codes.

then describe our variable-rate FEC scheme and a rate adaptation algorithm that uses measured SNR or FEC decoder input BER to determine the maximum FEC code rate that can be supported. In Section III, we describe simulations of the rate-adaptive scheme in a model terrestrial network. In Section IV, we present simulation results, including achievable information bit rates as a function of distance. In Section V, we discuss the observed trend of SNR versus distance and compare the performance of the proposed scheme to information-theoretic limits. We present conclusions in Section VI.

II. Rate-Adaptive Coding Scheme

A. Evolution of FEC in Optical Fiber Systems

The default FEC scheme of ITU-T G.709 is a Reed–Solomon (RS) code with parameters $(n, k) = (255, 239)$ [5], where n and k indicate the length in bytes of the code word and message word, respectively. This code, adopted during the 1990s, is sometimes classified as first-generation FEC [9], and yields a net coding gain[1] of approximately 5.8 dB at a decoded BER of 10^{-13}.

Short component codes can be concatenated in a serial or parallel manner to build a longer code and hence achieve a higher coding gain [10]. Linear block codes, e.g., RS and Bose/Ray-Chaudhuri/Hocquenghem (BCH) codes, are commonly used as component codes. Concatenated codes with several different types of component codes were recommended by ITU-T G.975.1 during the early 2000s [11], and are sometimes classified as second-generation FEC. The net coding gains of these codes are typically greater than 7.5 dB at a decoded BER of 10^{-13}.

In recent years, intensive research has addressed iterative soft-decision decoding of various codes, e.g., block turbo codes [9] and low-density parity-check codes [12], seeking further increases in coding gain. These types of codes are sometimes called third-generation FEC, and their target net coding gain is above 10 dB at a decoded BER of 10^{-13}.

B. Variable-Rate FEC Coding Scheme

We consider serially concatenated systematic RS codes over GF(8) that offer coding gains comparable to second-generation FEC codes. We refer to these as RS-RS codes throughout the remainder of this paper. Fig. 1 shows a high-level block diagram of the encoding and decoding chain for RS-RS codes. The encoder chain consists of an outer RS code, a linear block interleaver, an inner RS code, and a repetition code as the innermost code. The RS codes are nonbinary cyclic linear block codes that have strong capability to correct both random and burst errors, and are decoded using the Berlekamp-Massey algorithm or the generalized Euclidean algorithm [13], [14]. An innermost repetition code is added to compensate for the diminishing coding gain of RS-RS codes as the code rate is decreased in the low-SNR region. In the receiver, de-repetition is performed first, involving soft accumulation of repeated symbols, followed by a hard decision. De-repeated hard decisions are fed into the inner RS decoder, followed by a de-interleaver and the outer RS decoder. It would be possible to replace the repetition code by a stronger innermost code to enhance performance at low rates. However, our goal is to keep the scheme as simple as possible while ensuring that it works well in the high-rate regime near 100 Gbit/s, and reasonably well at much lower rates.

The proposed rate-adaptive coding scheme is intended to be used with a fixed constellation and fixed symbol rate, which is intended to simplify implementation of analog and digital hardware. As the SNR changes, the rates of the RS-RS and repetition codes are adjusted to maximize the overall rate. There are several options for constructing variable-rate block codes. One approach is simply to use entirely different codes at different rates, but this comes at the expense of increased encoder and decoder complexity. The alternative approaches are to puncture or shorten a single code, called a mother code, to generate derived codes of various rates. Puncturing removes parity bits from a codeword after encoding, thus increasing the code rate. The decoder treats the punctured parity bits as erasures. In puncturing, performance loss increases as the code rate increases, as compared to codes having the same code rate but having codeword length equal to that of the mother code. Shortening deletes information bits from a codeword, decreasing the code rate. The decoder knows the locations and pattern of the deleted bits. In shortening, performance loss increases as the code rate decreases.

We employ a hybrid method of creating multi-rate codes, which represents a tradeoff between complexity and performance. First, a mother code is selected, having a rate lying in the middle of the range of interest. Higher-rate codes are obtained by puncturing the mother code, and lower-rate codes are generated by shortening the mother code. This hybrid code construction method is shown schematically in Fig. 2. The numbers of bits punctured or shortened in the various derived codes are shown in Table I. The parameter $r_C = k/n$ is the rate of the concatenated RS-RS code, with the mother code having a rate $r_C = 0.6359$. Fig. 3 shows the performance of the five

[1]We define *net coding gain* as the reduction in SNR per bit enabled by a code at a specified decoder output BER as compared to an uncoded system, assuming an additive white Gaussian noise (AWGN) channel. The net coding gain takes account of an increased noise variance in the coded system.

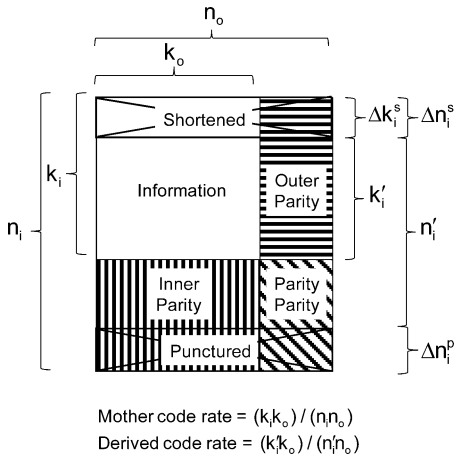

Fig. 2. Hybrid method of constructing variable-rate codes by puncturing and shortening a single mother code, which is selected from near the middle of the code rate range.

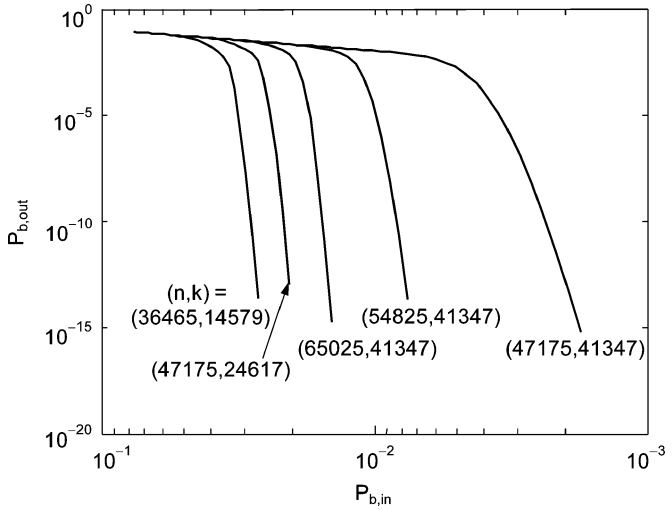

Fig. 3. Performance of serially concatenated RS-RS codes with five different rates, in terms of input and output BERs of the RS-RS decoder. When combined with an innermost repetition code, the $P_{b,\mathrm{in}}$ is measured at the output of the de-repeater.

TABLE I
PARAMETERS FOR PUNCTURING AND SHORTENING A MOTHER CODE

n_o	255
k_o	239
n_i	255
k_i	173

r_C	Δn_i^p	$\Delta n_i^s (= \Delta k_i^s)$	n_i'	k_i'
0.8765	70	0	185	173
0.7542	40	0	215	173
0.6359	0	0	255	173
0.5218	0	70	185	103
0.3998	0	112	143	61

constructed RS-RS codes, in terms of the relationship between $P_{b,\mathrm{in}}$ and $P_{b,\mathrm{out}}$, the BERs at the input and output of the RS-RS decoder, respectively (when combined with repetition coding, $P_{b,\mathrm{in}}$ is measured at the output of the de-repeater). The curves in Fig. 3 have been calculated analytically assuming independent errors and no decoding failures [15]. The highest-rate code, with code rate $r_C = 0.8765$, offers a net coding gain of about 7.6 dB at a decoded BER of 10^{-13}.

Since we use repetition coding as an innermost code, the overall FEC code rate is $r_C r_R = r_C/f_R$, where $r_R = 1/f_R$ is the rate of the repetition code. The repetition factor f_R denotes the number of repetitions of each bit from the output of the inner RS encoder, and ranges from 1 to 4. Given a symbol rate R_s, a line code rate r_L, and a modulation order M, the information bit rate R_b can be calculated as

$$R_b = 2 r_L r_C r_R R_s \log_2 M, \tag{1}$$

assuming PM transmission.

C. Rate Adaptation Algorithm

We assume that there exists a feedback link by which the receiver can communicate channel state information (CSI) to the transmitter, and that a controller at the transmitter can adjust the FEC code rate depending on the reported CSI. We consider two possible choices for CSI: (a) *SNR*, which is the SNR per symbol estimated at the de-repeater soft input, or (b) $P_{b,\mathrm{in}}$, which is the BER of hard decisions measured at the de-repeater output (RS-RS decoder input). A rate adaptation algorithm using *SNR* as CSI is easiest to understand, but may not reliably yield the highest possible information bit rate when *SNR* is not an accurate predictor of $P_{b,\mathrm{in}}$. On the other hand, an algorithm using $P_{b,\mathrm{in}}$ as CSI is perhaps harder to understand (though not necessarily harder to implement), but can reliably yield high information bit rates even when *SNR* is not an accurate predictor of $P_{b,\mathrm{in}}$.

We first describe rate adaptation using *SNR* as CSI. Assuming AWGN, the SNR per symbol per two polarizations estimated empirically at the de-repeater soft input is related to the empirically estimated Q factor[2] as

$$SNR = Q^2. \tag{2}$$

The BER of received channel bits (corresponding to hard decoding the de-repeater input) is [16]

$$P_{b,\mathrm{in}} = \frac{1}{2}\mathrm{erfc}\left(\frac{Q}{\sqrt{2}}\right) = \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{SNR}{2}}\right), \tag{3}$$

where $\mathrm{erfc}(\cdot)$ is the complementary error function. Also assuming AWGN, with repetition factor f_R and soft decision decoding of the repetition code, the SNR and Q factor at the de-repeater output become [17]

$$SNR_{\mathrm{rep}} = f_R SNR = f_R Q^2 = Q_{\mathrm{rep}}^2, \tag{4}$$

and the BER of hard decisions at the de-repeater output (RS-RS decoder input) becomes

$$P_{b,\mathrm{in}} = \frac{1}{2}\mathrm{erfc}\left(\frac{Q_{\mathrm{rep}}}{\sqrt{2}}\right) = \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{SNR_{\mathrm{rep}}}{2}}\right). \tag{5}$$

[2]The SNR per symbol is estimated empirically as the ratio between estimates of the received energy per symbol and the noise variance, both measured in two polarizations. The Q factor is estimated empirically as the ratio between the mean level separation and sum of the estimated noise standard deviations, averaging ratios computed for the inphase and quadrature components in two polarizations.

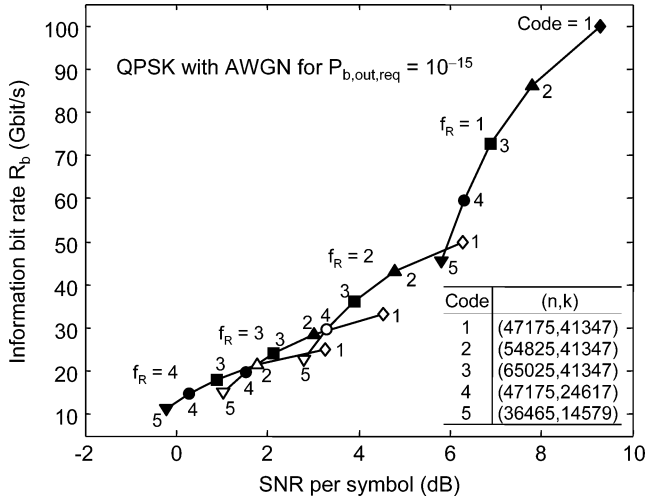

Fig. 4. Achievable information bit rate R_b versus SNR per symbol using RS-RS codes and repetition factor f_R for PM-QPSK on AWGN channel. Information bit rates are computed using (1), while SNR values are computed by combining (4), (5) and Fig. 3. The set of 13 filled (f_R, code) combinations represent a possible choice of modes for rate-adaptive transmission.

In order to determine the transmission rate, the controller requires information on the minimum *SNR* required for each (f_R, r_C) pair to achieve a target decoder output BER $P_{b,\text{out,req}}$ and the resulting R_b. This information is shown in Fig. 4, which assumes AWGN. Information bit rates are computed using (1) and required *SNR* values are computed by combining (4), (5) and Fig. 3 with $P_{b,\text{out,req}} = 10^{-15}$. The curves for $f_R \geq 2$ in Fig. 4 are computed from the curve for $f_R = 1$ by dividing R_b by f_R and subtracting $10\log_{10}(f_R)$ from *SNR*. A practical rate adaptation algorithm would quantize *SNR* values into discrete segments, and select a single combination (f_R, r_C) that achieves the highest R_b within each quantized segment. These predefined transmission parameters are commonly referred to as modes. In Fig. 4, the 13 filled-in (f_R, r_C) combinations can be used as transmission modes in a rate adaptation algorithm.

We propose an algorithm that uses reported *SNR* as CSI to determine transmission rate. Assuming there are N_s segments of quantized *SNR*, i.e., there are N_s modes, we define $(f_R, r_C)_i$ as the optimal code combination for segment i, and $SNR_{\text{th},i}$ as the minimum *SNR* (in dB) with which $(f_R, r_C)_i$ can achieve $P_{b,\text{out,req}}$ on an AWGN channel, where $SNR_{\text{th},i} < SNR_{\text{th},i+1}, i = 1, 2, \ldots, N_s - 1$. The information bit rate of segment i is denoted by $R_{b,i}$, where $R_{b,i} < R_{b,i+1}, i = 1, 2, \ldots, N_s - 1$. For example, in Fig. 4, the number of transmission modes is 13, $(f_R, r_C)_9 = (1, 14579/36465), SNR_{\text{th},9} = 5.8$ dB, and $R_{b,9} = 45.6162$ Gbit/s.

In the presence of fiber nonlinearity, residual ISI and laser phase noise, the receiver noise is not necessarily AWGN, and thus (2)–(5) are not necessarily valid, i.e., *SNR* may not be an accurate predictor of $P_{b,\text{in}}$. In this case, we could modify all of the threshold SNRs (the $SNR_{\text{th},i}$). In the interest of simplicity, we introduce penalty parameters Δ_{f_R} (in dB) for $f_R = 1, 2, 3, 4$, which quantify the increase in *SNR* required to achieve the desired $P_{b,\text{in}}$ at repetition factor f_R, as compared to the AWGN case described by (4) and (5). We have introduced different Δ_{f_R} for each f_R, which can be necessary when performing repetition in a suboptimal manner, but is unnecessary when performing repetition optimally (see Section IV). Also, to allow for performance margin, we introduce SNR margin parameters M_{up} and M_{down} (in dB) that are added to the SNR_{th} when changing the rate upward and downward, respectively. Choosing $M_{\text{up}} > M_{\text{down}}$ provides hysteresis, such that SNR fluctuations of magnitude less than $M_{\text{up}} - M_{\text{down}}$ will not cause rate changes. Transmission modes with M dB margin, and incorporating SNR penalty parameter Δ_{f_R}, can be defined by shifting the curves of Fig. 4 to the right by $M + \Delta_{f_R}$ dB, as illustrated in Fig. 5(a).

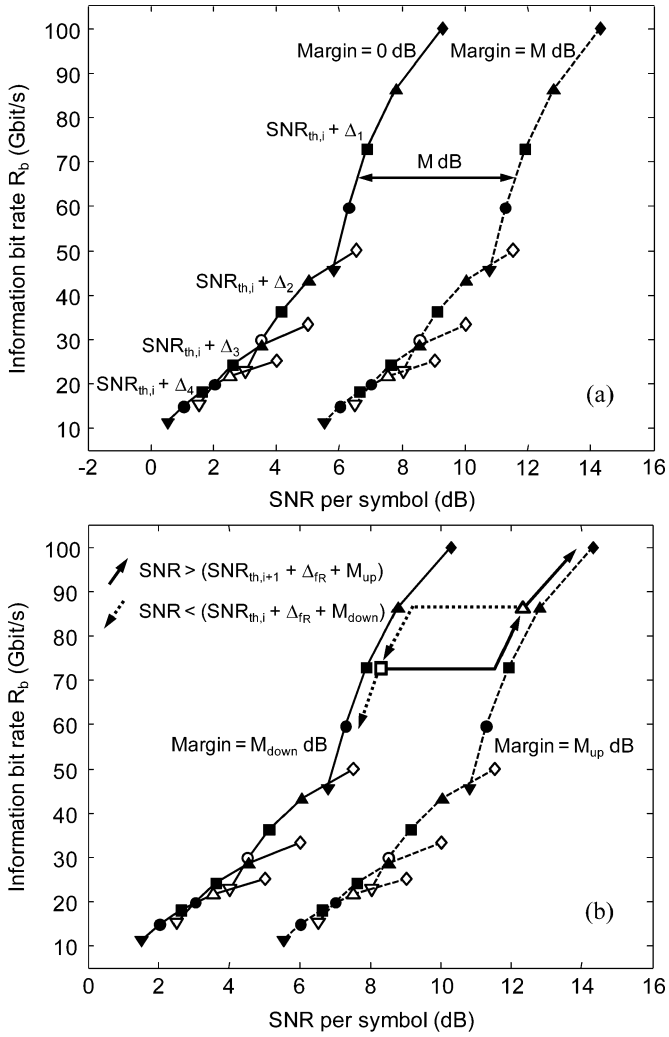

Fig. 5. (a) Transmission modes with SNR margin of $M = 5$ dB using *SNR* as CSI, taking into account SNR penalties, $\Delta_{f_1} = 0, \Delta_{f_2} = 0.25, \Delta_{f_3} = 0.5, \Delta_{f_4} = 0.75$ in dB. (b) Example of rate adaptation algorithm using *SNR* as CSI. The transmission modes for margins $M_{\text{up}} = 5$ dB and $M_{\text{down}} = 1$ dB are computed using (a).

In order to determine the transmission rate, starting from the current SNR_{th}, the controller monitors the reported *SNR*. If SNR exceeds the SNR_{th} corresponding to the next-higher rate for N_{up} consecutive observations, then the controller raises (f_R, r_C) by one level. A similar logic is applied to when lowering the rate, using a maximum downward threshold counter N_{down}. This counter-based approach can prevent unnecessary rate changes when the channel undergoes a brief degradation.

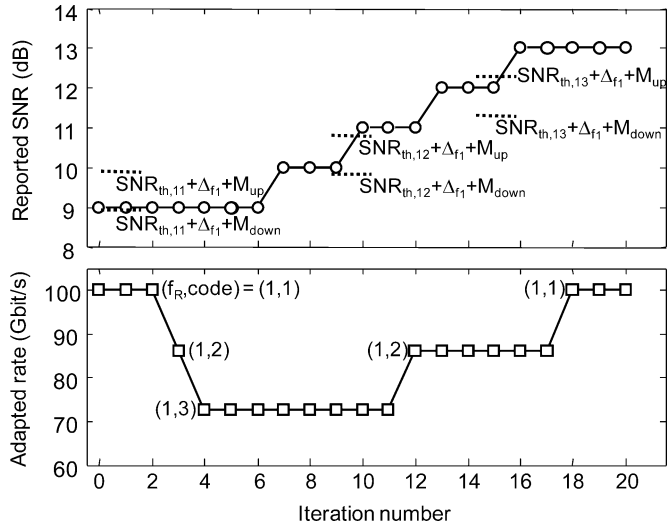

Fig. 6. An example of applying the rate adaptation algorithm in Section II-C. In this example, we use the 13 modes (f_R, code) indicated in Fig. 4, and assume $M_{\rm up} = 3$ dB, $M_{\rm down} = 2$ dB, $N_{\rm up} = N_{\rm down} = 3, \Delta_{f_1} = 0, \Delta_{f_2} = 0.25, \Delta_{f_3} = 0.5, \Delta_{f_4} = 0.75$ in dB. As the reported SNR fluctuates, the chosen (f_R, code) starts from the highest rate (1,1), decreases to the supportable rate (1,3) at iteration 4, and later recovers to (1,1) as the channel state improves.

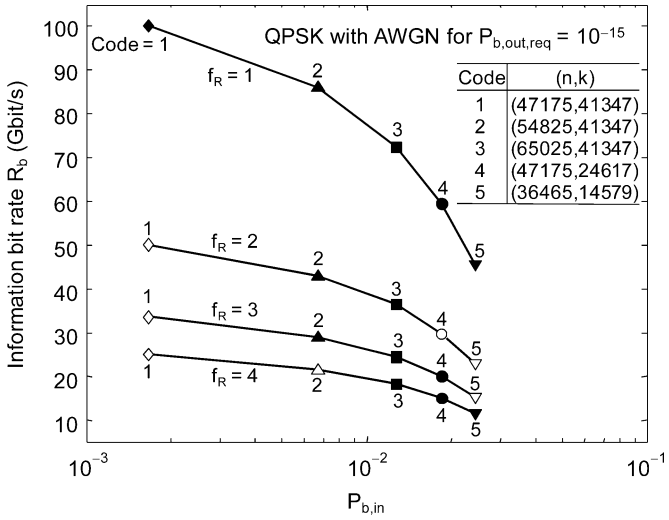

Fig. 7. Achievable information bit rate R_b versus $P_{b,\rm in}$ using RS-RS codes and repetition factor f_R for PM-QPSK on AWGN channel. Information bit rates are computed using (1), while $P_{b,\rm in}$ values are computed by combining (4), (5) and Fig. 3. The set of 13 filled (f_R, code) combinations represent a possible choice of modes for rate-adaptive transmission.

A pseudocode for the algorithm is given as follows.

1) Initialize parameters to the highest-rate mode.
   - Initialize codes: $(f_R, r_C)_i \rightarrow (f_R, r_C)_{N_s}$
   - Initialize up/down counters: $C_{\rm up} = C_{\rm down} = 0$
2) Check if rate change is necessary.
   - if $SNR \geq (SNR_{{\rm th},i+1} + \Delta_{f_R} + M_{\rm up})$
     — $C_{\rm up} = C_{\rm up} + 1$
     — if $C_{\rm up} \geq N_{\rm up}$, $(f_R, r_C)_i \rightarrow (f_R, r_C)_{i+1}$
   - elseif $SNR < (SNR_{{\rm th},i} + \Delta_{f_R} + M_{\rm down})$
     — $C_{\rm down} = C_{\rm down} + 1$
     — if $C_{\rm down} \geq N_{\rm down}$, $(f_R, r_C)_i \rightarrow (f_R, r_C)_{i-1}$
   - else
     — $C_{\rm up} = C_{\rm down} = 0$
     — $(f_R, r_C)_i \rightarrow (f_R, r_C)_i$
3) Go to step 2.

The described algorithm is illustrated by Fig. 5(b), which shows the hysteresis obtained when $M_{\rm up} > M_{\rm down}$. An example of rate adaptation using the proposed algorithm is shown in Fig. 6, which assumes $M_{\rm up} = 3$ dB, $M_{\rm down} = 2$ dB, $N_{\rm up} = N_{\rm down} = 3$, $\Delta_{f_1} = 0$ dB, $\Delta_{f_2} = 0.25$ dB, $\Delta_{f_3} = 0.5$ dB, $\Delta_{f_4} = 0.75$ dB, and the mode quantization shown in Fig. 4. It is observed that the algorithm tracks the reported *SNR* and adapts the rate as desired.

We now describe an algorithm using the de-repeater output BER (RS-RS decoder input BER) $P_{b,\rm in}$ as CSI. This BER can be estimated by the decoder from the number of bit errors corrected in decoding. In this case, the receiver would report the measured $P_{b,\rm in}$, and the controller would run the algorithm given above using BER thresholds $P_{b,{\rm in},{\rm th},i}$, computed by substituting $SNR_{{\rm th},i}$ into (4) and (5). The resulting transmission modes are shown in Fig. 7. Measuring $SNR_{{\rm th},i}$ in dB and including an SNR margin of M dB, the threshold $P_{b,{\rm in},{\rm th},i}$ can be computed as

$$P_{b,{\rm in},{\rm th},i,M} = \frac{1}{2}{\rm erfc}\left(\sqrt{\frac{f_R 10^{(SNR_{{\rm th},i}+M)/10}}{2}}\right), \quad (6)$$

and the corresponding transmission modes with margin are defined as in Fig. 8(a). Although (4) and (5) assume AWGN, once the $P_{b,{\rm in},{\rm th},i,M}$ are computed, the algorithm is intended to work even on non-AWGN channels without the need for SNR penalty parameters. A pseudocode for the algorithm using $P_{b,\rm in}$ as CSI is given as follows.

1) Initialize parameters to the highest-rate mode.
   - Initialize codes: $(f_R, r_C)_i \rightarrow (f_R, r_C)_{N_s}$
   - Initialize up/down counters: $C_{\rm up} = C_{\rm down} = 0$
2) Check if rate change is necessary.
   - if $P_{b,\rm in} \leq P_{b,{\rm in},{\rm th},i+1,M_{\rm up}}$
     — $C_{\rm up} = C_{\rm up} + 1$
     — if $C_{\rm up} \geq N_{\rm up}$, $(f_R, r_C)_i \rightarrow (f_R, r_C)_{i+1}$
   - elseif $P_{b,\rm in} > P_{b,{\rm in},{\rm th},i,M_{\rm down}}$
     — $C_{\rm down} = C_{\rm down} + 1$
     — if $C_{\rm down} \geq N_{\rm down}$, $(f_R, r_C)_i \rightarrow (f_R, r_C)_{i-1}$
   - else
     — $C_{\rm up} = C_{\rm down} = 0$
     — $(f_R, r_C)_i \rightarrow (f_R, r_C)_i$
3) Go to step 2.

The described algorithm is depicted in Fig. 8(b).

Note that, in either algorithm, as an alternative to using counter parameters $N_{\rm up}$ and $N_{\rm down}$, there exist other possible mechanisms to guard against rate changes due to short transients in CSI, e.g., comparing a moving average of CSI to a threshold.

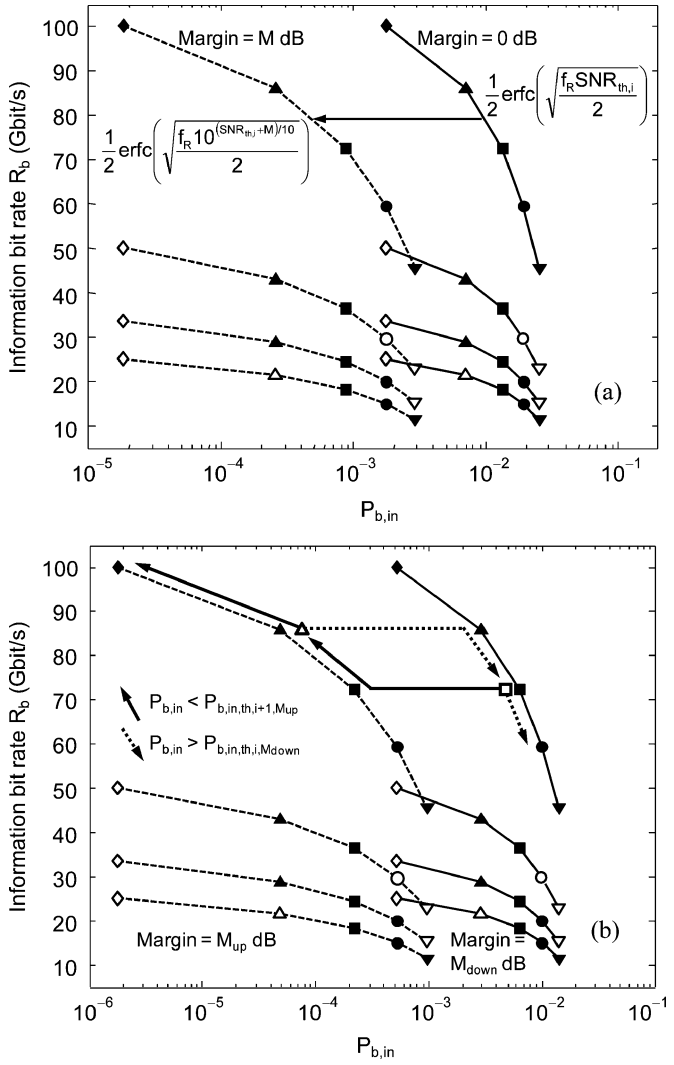

Fig. 8. (a) Transmission modes with SNR margin of $M = 3$ dB, using $P_{b,\mathrm{in}}$ as CSI. (b) Example of rate adaptation algorithm using $P_{b,\mathrm{in}}$ as CSI. The transmission modes for margins $M_{\mathrm{up}} = 4$ dB and $M_{\mathrm{down}} = 1$ dB are computed using (a).

Also note that, when using either algorithm, a controller at the transmitter should set the launched power P_t to optimize the CSI. Optimal launched power values may be stored in a look-up table or determined on the fly, based on CSI reported by the receiver. When using *SNR* as CSI, at each transmission distance, by definition, a single P_t is optimal for all values of the repetition factor f_R. When using $P_{b,\mathrm{in}}$ as CSI, at each transmission distance, the optimal P_t may or may not depend on f_R, as illustrated in Section IV.

D. Implementation Considerations

In simulating our model long-haul system, we use a repetition code block size of 66, i.e., repetitions are separated by 66 symbols, to make the scheme compatible with the 64b/66b line coding used in 10 Gbit Ethernet and proposed for 100 Gbit Ethernet [18]. In order to be consistent with these standards, we also assume that line encoding occurs first and FEC encoding follows.

We assume that line, RS-RS, and repetition encoding are performed on a single serial bit stream, after which, blocks of $66f_R$ bits are permuted and mapped in round-robin fashion to the four dimensions of the PM-QPSK signal (x and y polarizations, inphase and quadrature components). This particular implementation of encoding, permutation, and symbol mapping is illustrated in Fig. 9. In order to minimize "noise" correlation between repeated symbols caused by fiber nonlinearity, when $f_R > 1$, the second, third and fourth repetition of a block are permuted using the different matrices shown. We observed that these permutations do not significantly modify the desirable correlation properties of pseudorandom binary sequence, such as those used for line coding in [18]. The performance benefits obtained by permutation will be described in Section IV.

If it is necessary in practice to implement parallel FEC encoders, it would seem natural to parallelize the input information bit stream into four tributaries, perform line and RS-RS encoding on each tributary independently, and map each encoded tributary to one particular signal dimension. We expect that in practice, performance should essentially be independent of the implementation of encoding and symbol mapping, provided that residual ISI and "noise" correlation caused by fiber nonlinearity are not significant.

III. System Simulations

We have evaluated the rate-adaptive transmission scheme in the model long-haul system shown in Fig. 10. System design parameters are summarized in Table II.

We have simulated a single channel in a wavelength-division-multiplexed system with nominal 50-GHz channel spacing. The modulation is single-carrier PM-QPSK using non-return-to-zero pulses at a symbol rate $R_s = 29.4152$ Gsymb/s. Each modulator is a quadrature Mach–Zehnder device. Each drive waveform is a train of rectangular pulses filtered by a five-pole Bessel LPF having a 3-dB bandwidth 1.4 $R_s = 41.2$ GHz. Percentages 90, 95 and 99% of the modulated signal energy are contained in bandwidths of 28.9, 33.2 and 40.4 GHz, respectively. We assume line coding at rate $r_L = 64/66$, yielding an information bit rate $R_b = 100$ Gbit/s for the highest overall code rate $r_C r_R$.

The fiber network comprises multiple 80-km spans of standard single-mode fiber, using pre-compensation of chromatic dispersion (CD) by -510 ps/nm and inline CD compensation to 42.5 ps/nm (3.1%) residual dispersion per span. Each span uses a two-stage inline amplifier to compensate span loss, with the ratio between first- and second-stage gains optimized via simulation.

In order to simulate transmission through a large-scale mesh network, a ROADM is inserted at every third span. Filtering by each wavelength-selective switch is simulated by a super-Gaussian bandpass filter of order 2.5 and 3-dB bandwidth of 40 GHz, consistent with typical devices [19]. Over long distances, the ROADMs cause significant channel narrowing, as shown in Fig. 11. Typical networks would not have ROADMs at such precisely periodic intervals and at such high density throughout, but we choose a homogeneous, high-density distribution of ROADMs to be conservative and to obtain a simple trend of achievable bit rate versus transmission distance.

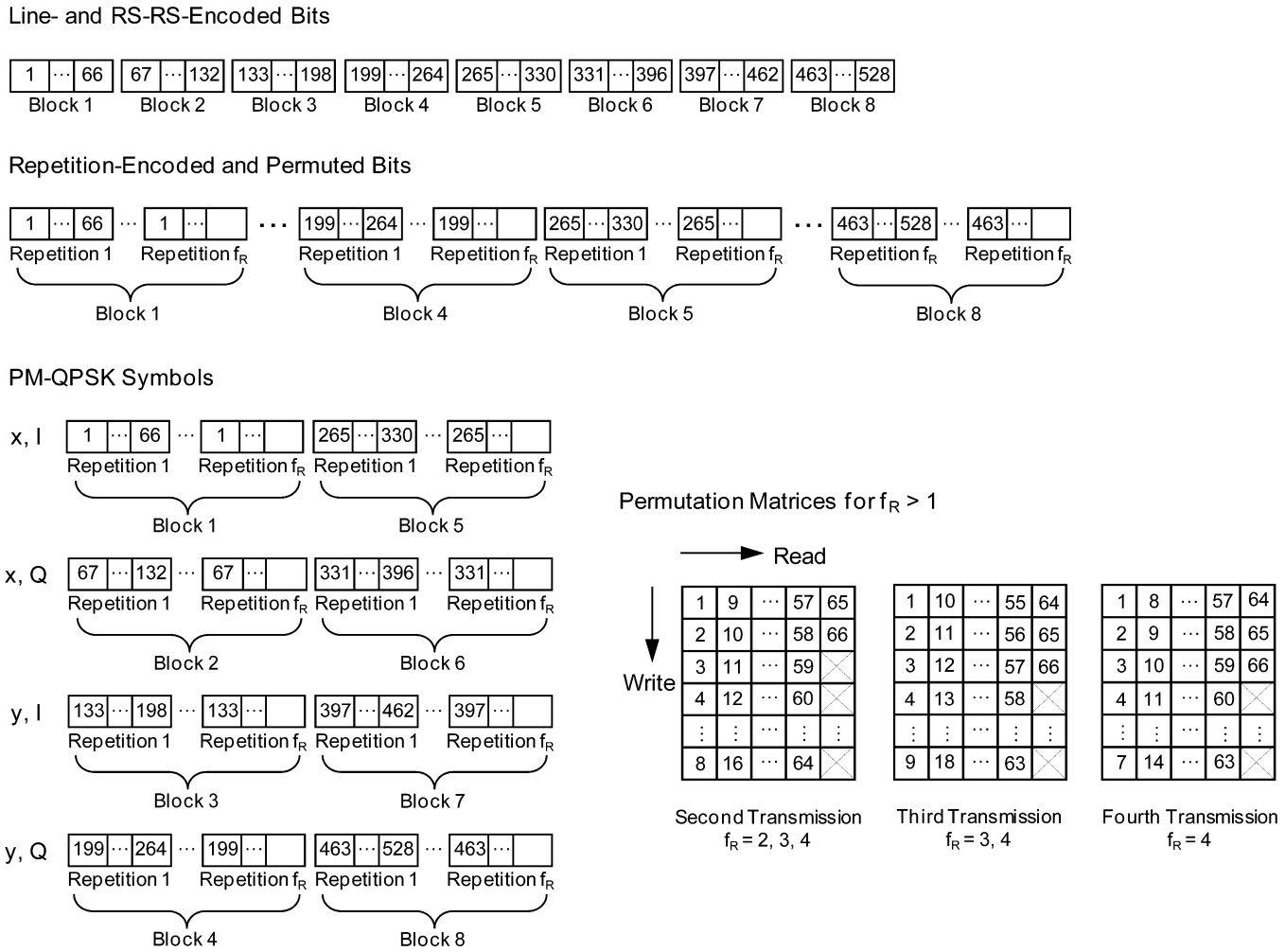

Fig. 9. Repetition encoding and mapping of encoded bits to PM-QPSK symbols assumed in model system.

The receiver employs a fifth-order Butterworth antialiasing filter of 3-dB bandwidth R_s, samples at a rate of $2R_s$ complex samples per polarization and performs digital compensation of CD and polarization-mode dispersion (PMD) using finite impulse response time-domain filtering, as described in [20]. At each transmission distance, the number of filter taps is optimized to make residual ISI negligible. The number of taps employed is about twice that given by the formula in [20], which would result in a 2-dB penalty from residual ISI. The equalizer is adapted using the least mean squares algorithm with optimized gear shifting of the step size parameter.

Signal propagation is simulated by numerical integration of the vector nonlinear Schrödinger equation by the split-step Fourier method [21].

At each transmission distance, and for each repetition factor f_R, the launched power is optimized to maximize $SNR_{\rm rep}$, the SNR at the de-repeater output (RS-RS decoder input), which we show in Section IV is equivalent to minimizing $P_{b,\rm in}$, the corresponding BER. After simulation of a sufficient number of symbols to obtain at least 100 bit errors at the de-repeater output (at least for values of $P_{b,\rm in}$ down to those corresponding to $P_{b,\rm out} = 10^{-15}$), we record $P_{b,\rm in}$, $SNR_{\rm rep}$ and $Q_{\rm rep}$ (at the de-repeater output) and *SNR* and Q (at the de-repeater input). We note that, in principle, all of this information could be determined by a real receiver.

IV. Simulation Results

Fig. 12 shows $SNR_{\rm rep}$, the SNR at the de-repeater output (RS-RS decoder input) versus transmitted power P_t for various repetition factors $f_R = 1, 2, 3, 4$, for a representative transmission distance, $L = 2400$ km. If the total "noise" were AWGN, we would expect improvements in $SNR_{\rm rep}$ of 3.0, 4.8 and 6.0 dB for $f_R = 2, 3, 4$ as compared to $f_R = 1$. Employing repetition coding without permutation (dashed curves in Fig. 12), we observe improvements in $SNR_{\rm rep}$ of 2.3, 3.5 and 4.4 dB for $f_R = 2, 3, 4$. We also notice that the optimized transmitted power P_t is smaller for $f_R = 2, 3, 4$ than for $f_R = 1$. The observed improvement in $SNR_{\rm rep}$ is less than expected for AWGN because about half of the total "noise" variance arises from nonlinear effects, including self-phase modulation (SPM) [16] and nonlinear phase noise (NLPN) [22], which are correlated between repeated symbols. Employing repetition coding with the permutation method described in Section II-D (solid curves in Fig. 12), we observe improvements in $SNR_{\rm rep}$ of 3.0, 4.8 and 6.0 dB, which are the full improvements expected for AWGN. We note that the optimized transmitted power P_t

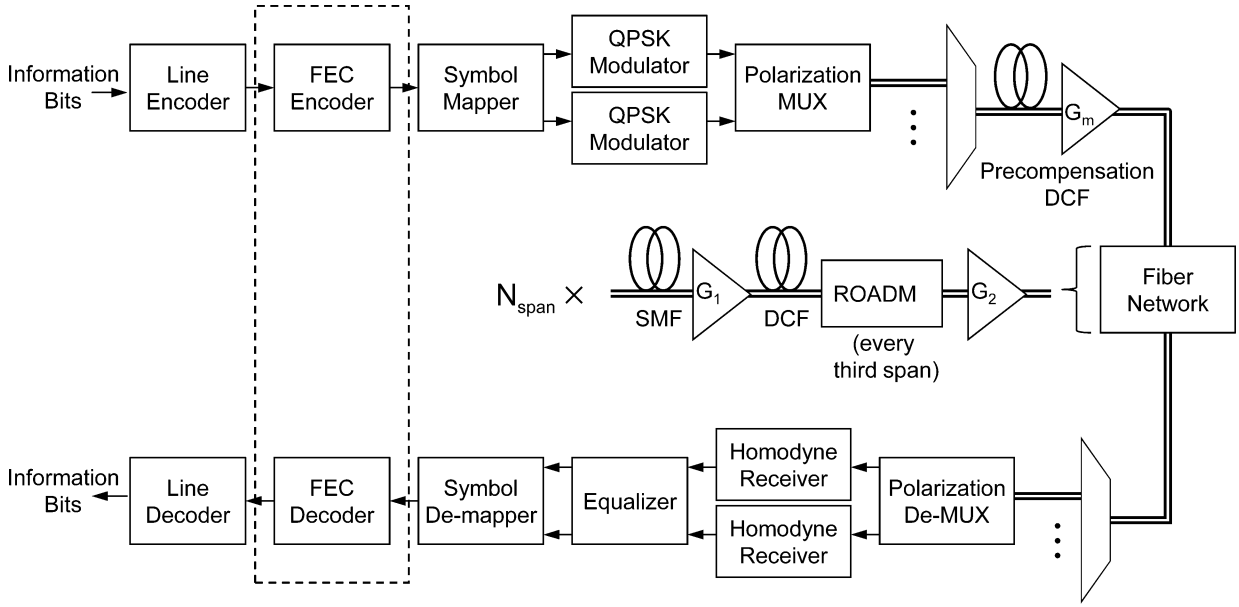

Fig. 10. Model long-haul optical system that consists of line and FEC encoder/decoder, equalizer, PM-QPSK modulator/demodulator, dispersion precompensation fiber, and multiple spans. The dashed line encloses the FEC encoding and decoding chains, which are detailed in Fig. 1.

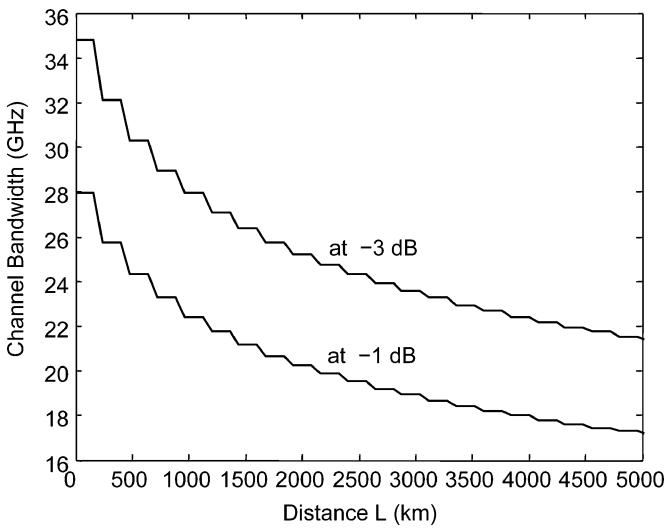

Fig. 11. Effect of channel bandwidth narrowing due to multiple passes through ROADMs, which are present at every third span.

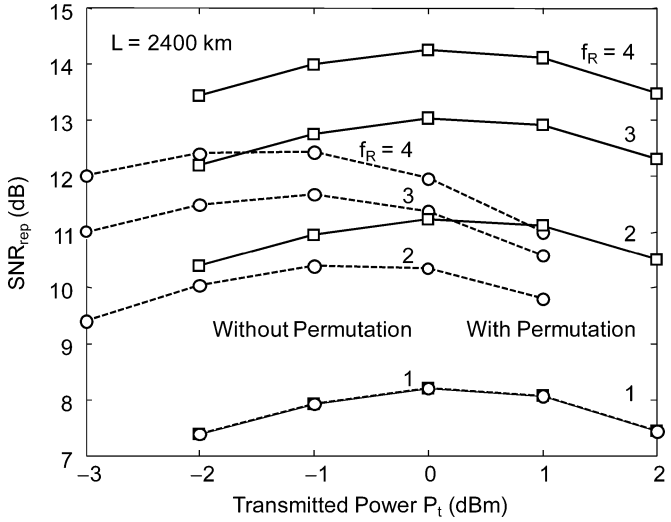

Fig. 12. SNR at de-repeater output (RS-RS decoder input) as a function of transmitted power, for various repetition factors, with and without permutation of repeated symbols, for $L = 2400$ km.

is the same for $f_R = 1, 2, 3, 4$. We also verified that repetition coding with permutation is equally effective in systems employing standard SMF without inline CD compensation, i.e., the specific permutation matrices shown in Fig. 9 are effective even when CD leads to a memory length as long as 35 symbols. Repetition coding with permutation is used in all studies described hereafter.

Fig. 13 compares the BER observed at the de-repeater output (RS-RS decoder input) to the BER predicted from SNR_{rep}, using (4) and (5) (only BERs obtained at optimized values of transmit power P_t are shown). Good agreement is obtained, showing that maximizing SNR_{rep} does minimize $P_{b,\text{in}}$, and suggesting that approximation of the total "noise" as Gaussian-distributed is valid. Under these circumstances, a rate adaptation algorithm using *SNR* as CSI could employ penalty parameters $\Delta_{f_R} = 0$ for all f_R.

We have employed the rate adaptation algorithm of Section II-C, using $P_{b,\text{in}}$ as CSI, and using all 20 transmission modes (r_C, f_R) shown in Fig. 7. We assume a required FEC decoder output BER $P_{b,\text{out,req}} = 10^{-15}$, SNR margins $M_{\text{up}} = M_{\text{down}} = 0, 1, \ldots, 5$ dB and counter parameters $N_{\text{up}} = N_{\text{down}} = 1$ (because we assume static channel conditions). Fig. 14 shows achievable information bit rates versus transmission distance for the six different values of SNR margin. With zero margin, an information bit rate $R_b = 100$ Gbit/s can be realized up to 2000 km, with the achievable rate decreasing by approximately 40% for every additional 1000 km. Also, a repetition factor $f_R = 1$ (no repetition) is found to

TABLE II
PARAMETERS USED IN SYSTEM MODEL

Parameter	Value	Unit
Center wavelength	1550	nm
Nominal channel bandwidth	50	GHz
Symbol rate R_s	29.4152	Gsymb/s
Bandwidth containing 99% of modulated signal power	40.4	GHz
Line code rate r_L	64/66	–
Repetition factor f_R	1, 2, 3, 4	–
Repetition code block size	66	bit
Target decoder output BER $P_{b,out,req}$	10^{-15}	–
SMF length per span	80	km
Inline DCF length per span	15.5	km
Pre-compensation DCF length	6	km
SMF loss coefficient	0.25	dBm/km
DCF loss coefficient	0.6	dBm/km
SMF CD coefficient	17	ps/(nm·km)
DCF CD coefficient	–85	ps/(nm·km)
SMF/DCF PMD coefficient	0.1	ps/km$^{1/2}$
SMF nonlinear coefficient	0.0012	W^{-1}m^{-1}
DCF nonlinear coefficient	0.0053	W^{-1}m^{-1}
ROADM spacing	3	span
ROADM 3-dB bandwidth	40	GHz
ROADM super-Gaussian order	2.5	–
ROADM insertion loss	12	dB
ROADM maximum center frequency error	1	GHz
Multiplexer /demultiplexer 3-dB bandwidth	40	GHz
Multiplexer/demultiplexer insertion loss	6	dB
Inline amplifier gain ratio G_2/G$_1$	2.5	–
Amplifier noise figure	5	dB

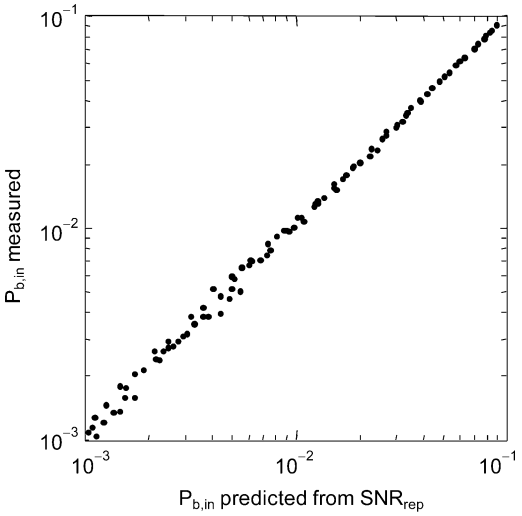

Fig. 13. Comparison of measured BER at de-repeater output (RS-RS decoder input) to BER predicted from the SNR of de-repeated symbols using (4) and (5).

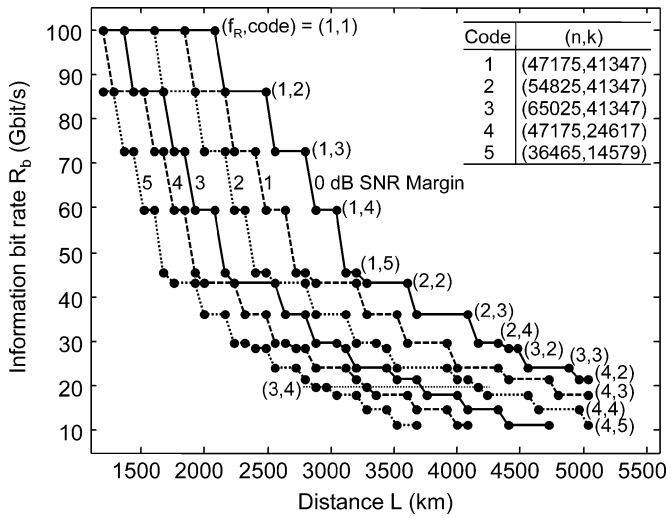

Fig. 14. Achievable information bit rate versus transmission distance for margins between 0 and 5 dB. The pair (f_R, code) denotes the repetition factor and type of RS-RS code.

be optimal up to about 3500 km, beyond which the optimal f_R increases by 1 approximately every 700 km.

V. DISCUSSION

In this section, we examine how the SNR in the model system using a fixed PM-QPSK constellation scales with transmission distance, and we provide an approximate estimate of the performance gap between the proposed rate-adaptive coding scheme and an ideal coding scheme achieving information-theoretic limits. We consider several different ways to quantify SNR and compare these to an equivalent SNR corresponding to the information bit rate achieved by the proposed rate-adaptive scheme. This discussion makes reference to Fig. 15.

The uppermost solid curve in Fig. 15 shows SNR_{AWGN}, which is an empirical estimate of the SNR per symbol (in two polarizations) as limited only by accumulated optical amplifier noise

$$SNR_{\mathrm{AWGN}} = \frac{P_t}{P_n}. \tag{7}$$

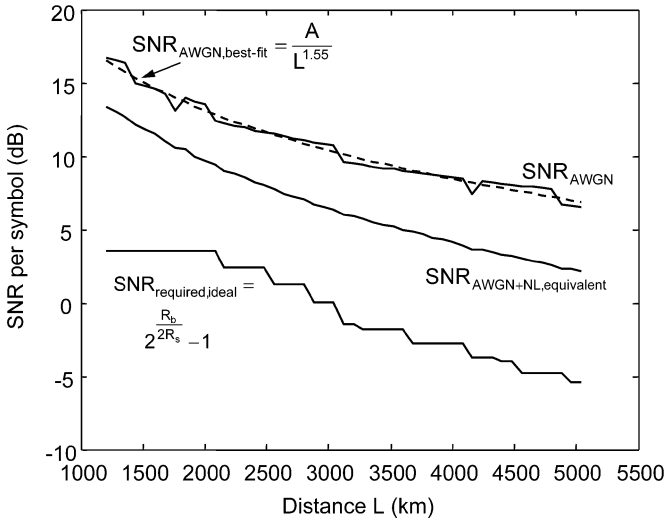

Fig. 15. Three different measures of SNR compared to the SNR required for an ideal capacity-achieving coding scheme to achieve error-free transmission at the information bit rate R_b achieved by the proposed coding scheme.

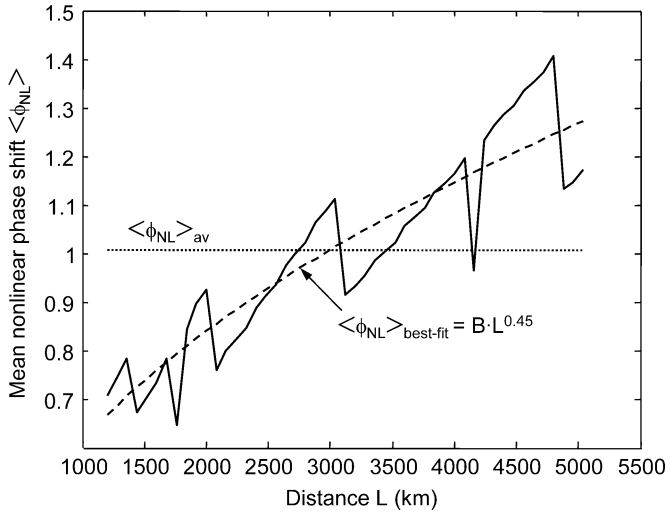

Fig. 16. Mean nonlinear phase shift values computed using (11) and compared to the power law (12). The value averaged over distance is $\langle\phi_{\rm NL}\rangle_{\rm av} = 1.01$ rad.

P_t is the transmitted signal power, optimized at each transmission distance without repetition encoding ($f_R = 1$), which equals the received signal power at the demultiplexer input, since the network is designed to have unit signal gain. The noise power is

$$P_n = S_n\, W, \tag{8}$$

where S_n is the power spectral density (PSD) of the accumulated amplifier noise at the demultiplexer input, and W is the noise bandwidth of the receiver filter. Since the receiver employs near-Nyquist rate sampling at rate $2R_s$ followed by a linear equalizer, we assume that the receiver noise bandwidth effectively equals that of an ideal matched filter, $W = R_s$. We compute the accumulated amplifier noise PSD (in two polarizations) analytically using

$$S_n = 2\sum_{k=1}^{N_{\rm span}} \left(S_{sp,1,k}e^{-\alpha_d L_d}T_k G_{2,k} + S_{sp,2,k}\right). \tag{9}$$

In the kth span, $S_{sp,j,k}$ and $G_{j,k}$ are the noise PSD and gain of the jth amplifier ($j = 1,2$), $T_k \le 1$ is the transmission factor of the ROADM (if present), and α_d and L_d are the loss coefficient and length per span of the DCF. Equation (9) makes use of the fact that the noises from the spans accumulate with equal weights, since each span has unit gain. The noise PSD (per polarization) $S_{sp,j,k}$ is given by $S_{sp,j,k} = n_{sp,j}(G_{j,k} - 1)h\nu$, where $n_{sp,j}$ is the population inversion factor, h is Planck's constant, and ν is the optical frequency.

The dashed curve in Fig. 15 shows $SNR_{\rm AWGN,best-fit}$, which is a power-law fit to the observed $SNR_{\rm AWGN}$ as a function of transmission distance L

$$SNR_{\rm AWGN,best-fit} = \frac{A}{L^{1.55}}. \tag{10}$$

The constant A and the exponent of L have been found by curve fitting. The formula (10) fits $SNR_{\rm AWGN}$ well over nearly a 5:1 range in L, corresponding to about a 10-dB variation in $SNR_{\rm AWGN}$.

It may be difficult to explain analytically the formula (10), because of CD, PMD and other factors in the model system. Nevertheless, we may gain some insight by comparing our results to simpler systems that have been analyzed. In phase-shift-keyed systems either without CD [22] or with CD [23], [24], it has been established that NLPN is the primary intrachannel nonlinear impairment, and that there exists a mean nonlinear phase shift that minimizes the sum of linear phase noise (corresponding simply to amplifier noise) and NLPN. The analysis is simplest for a dispersion-free system using one polarization [22]. Linear phase noise has a variance (at high SNR) $\sigma_{\rm L}^2 \approx 1/2SNR$, while NLPN has a variance $\sigma_{\rm NL}^2 \approx 2\langle\phi_{\rm NL}\rangle^2/3SNR$. Here, $\langle\phi_{\rm NL}\rangle$ denotes the mean nonlinear phase shift, where the brackets denote a time average over a modulated signal [16]. By differentiating the sum of these variances with respect to $\langle\phi_{\rm NL}\rangle$, one finds that the sum is minimized when the nonlinear phase shift assumes the value $\langle\phi_{\rm NL}\rangle_{\rm opt} = \sqrt{3}/2$, independent of SNR, and thus independent of transmission distance L. When $\langle\phi_{\rm NL}\rangle$ assumes this optimized value, one finds that the two phase variances are equal: $\sigma_{\rm L}^2 = \sigma_{\rm NL}^2$.

In general, $\langle\phi_{\rm NL}\rangle$ scales in proportion to the transmit power P_t and in proportion to the total number of fiber spans traversed, i.e., in proportion to L. Since $\langle\phi_{\rm NL}\rangle \propto P_t L$, requiring $\langle\phi_{\rm NL}\rangle = \langle\phi_{\rm NL}\rangle_{\rm opt}$ requires the transmit power, which appears in the numerator of (7), to scale inversely with transmission distance: $P_t \propto 1/L$. In general, the accumulated amplifier noise power P_n, which appears in the denominator of (7), is proportional to the number of amplifiers traversed [25], and is thus proportional to transmission distance L : $P_n \propto L$. Thus, in a dispersion-free, single-polarization system [22], the SNR (7) is expected to scale approximately as $1/L^2$.

By contrast, our model system includes CD and PMD, and the analysis of [22] is not precisely applicable. Fig. 16 shows the mean nonlinear phase shift $\langle\phi_{\rm NL}\rangle$ versus transmission distance for our model system, where the transmit power P_t has been

optimized at each transmission distance. This is computed using [26]

$$\langle \phi_{\mathrm{NL}} \rangle = P_t \cdot \frac{8}{9} \sum_{k=1}^{N_{\mathrm{span}}} \left(\gamma_s L_{\mathrm{eff},s} + e^{-\alpha_s L_s} G_{1,k} \gamma_d L_{\mathrm{eff},d} \right). \tag{11}$$

The factor 8/9 accounts for the effect of polarization averaging on SPM, which is appropriate when the polarization scattering length is much shorter than the nonlinear interaction length. γ_s and γ_d are the nonlinear coefficients for SMF and DCF, respectively, α_s and L_s are the loss coefficient and length per span of the SMF, and $L_{\mathrm{eff},s}$ and $L_{\mathrm{eff},d}$ are the effective interaction lengths [16] for SMF and DCF, respectively, computed using $L_{\mathrm{eff},i} = (1 - e^{-\alpha_i L})/\alpha_i$. Unlike the dispersion-free, single-polarization system [22], the optimized $\langle \phi_{\mathrm{NL}} \rangle$ is not independent of transmission distance. In Fig. 16, the trend of mean nonlinear phase shift versus transmission distance is seen to approximately follow the power law

$$\langle \phi_{\mathrm{NL}} \rangle_{\mathrm{best-fit}} = B \cdot L^{0.45}, \tag{12}$$

where the constant B and the exponent of L have been found by curve fitting. Given (12) and following the reasoning used for the idealized dispersion-free system, we would expect the optimized transmit power to scale as $P_t \propto 1/L^{0.55}$, and the SNR to scale as $SNR \propto 1/L^{1.55}$, which is consistent with the empirical curve fit (10). Returning to Fig. 16, averaging $\langle \phi_{\mathrm{NL}} \rangle$ over L yields an average value $\langle \phi_{\mathrm{NL}} \rangle_{\mathrm{av}} = 1.01$ rad, fairly close to the value $\langle \phi_{\mathrm{NL}} \rangle_{\mathrm{opt}} = \sqrt{3}/2$ for the idealized, dispersion-free system [22]. We should emphasize that the observed scaling of *SNR* and $\langle \phi_{\mathrm{NL}} \rangle$ with L, given by (10) and (12), is not universal, but is dependent on the specific modulation format and dispersion map in the model system.

The middle solid curve in Fig. 15 shows $SNR_{\mathrm{AWGN+NL,equivalent}}$, which is the SNR per symbol observed empirically at the de-repeater input using a transmit power P_t optimized at each transmission distance. $SNR_{\mathrm{AWGN+NL,equivalent}}$ includes all effects of accumulated amplifier noise, fiber nonlinearity, residual ISI, and noise enhancement from the equalizer when the channel bandwidth decreases significantly. Here, "equivalent" refers to the fact that the sum of these "noises" is not necessarily Gaussian-distributed. At small L, $SNR_{\mathrm{AWGN+NL,equivalent}}$ is about 3.2 dB lower than SNR_{AWGN}, indicating that at the optimum P_t, amplifier noise and nonlinear noise powers are approximately equal, consistent with the analysis of [22], while at large L, the difference increases to about 4.4 dB, presumably because ROADM channel narrowing causes equalizer noise enhancement.

We would like to estimate the performance gap between the proposed rate-adaptive coding scheme and an ideal coding scheme achieving information-theoretic limits. Assuming an ideal discrete-time system transmitting $2R_s$ complex-valued symbols per second[3] in the presence of AWGN, the capacity is [17]

$$C = 2R_s \log_2(1 + SNR). \tag{13}$$

Inverting (13), we find the SNR required for an ideal, capacity-achieving coding scheme to achieve error-free transmission at the information bit rate R_b achieved by the proposed coding scheme

$$SNR_{\mathrm{required,ideal}} = 2^{\frac{R_b}{2R_s}} - 1. \tag{14}$$

In Fig. 15, the bottom solid curve represents $SNR_{\mathrm{required,ideal}}$ as a function of transmission distance. The vertical separation between $SNR_{\mathrm{AWGN+NL,equivalent}}$ and $SNR_{\mathrm{required,ideal}}$ is an estimate of the performance gap between an ideal rate-adaptive coding scheme and the scheme proposed here. The gap ranges from about 5.9 dB to about 7.5 dB as L varies from 2000 to 5000 km. Much of the increase in the gap over this range of L can be attributed to inefficiency of the repetition coding used for $L \geq 3280$ km. Over the entire range of L, we expect that some, but not all, of the performance gap can be closed by using more powerful FEC codes with iterative soft-decision decoding, but at the expense of greater implementation complexity. For $L < 2000$ km, the gap increases with decreasing L, reflecting our use of a fixed PM-QPSK constellation. We note that a small part of the observed gap (about 0.2 dB) arises from line coding.

In this study, we have considered only a single-channel system with intrachannel nonlinearities, in order to keep simulation run time reasonable (less than one week on a cluster of multi-core computers). We have not included multichannel effects, especially interchannel nonlinearities. If interchannel nonlinearities were included, we would expect our results to change quantitatively, via a slight reduction in the maximum distance at which a given rate can be achieved. We would not expect our results to change qualitatively, however, because interchannel and intrachannel nonlinearities scale similarly with transmission distance.

VI. Conclusion

We have studied a rate-adaptive transmission scheme using variable-rate FEC codes with a fixed signal constellation and a fixed symbol rate. The FEC scheme employs serially concatenated RS codes with hard-decision decoding, using shortening and puncturing to vary the code rate. An inner repetition code with soft combining provides further rate variation. We have combined the FEC scheme with single-carrier polarization-multiplexed QPSK and digital coherent detection, evaluating performance in a model long-haul system with inline dispersion compensation. With zero SNR margin, the achievable information bit rate varies from 100 Gbit/s at 2000 km, to about 60 Gbit/s at 3000 km, to about 35 Gbit/s at 4000 km. Compared to an ideal coding scheme achieving information-theoretic limits on an AWGN channel, the proposed coding scheme exhibits a performance gap ranging from about 5.9 dB at 2000 km to about

[3]The factor of 2 corresponds to polarization multiplexing.

7.5 dB at 5000 km. Much of the increase in the gap can be attributed to sub-optimality of the repetition coding used beyond 3280 km. We found that the SNR as limited by amplifier noise scales with distance L as $L^{-1.55}$, as compared to the L^{-2} expected in a dispersion-free system.

ACKNOWLEDGMENT

The author are grateful to O. Gerstel and L. Paraschis of Cisco Systems for suggesting this research topic. They also thank K.-P. Ho and P. J. Winzer for helpful discussions.

REFERENCES

[1] *Asymmetric Digital Subscriber Line (ADSL) Transceivers*, ITU-T G.992.1, Jun. 1999.

[2] *IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Broadband Wireless Access Systems*, IEEE 802.16-2009, May 2009.

[3] *3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Physical Layer Procedures (FDD) (Release 8)*, 3GPP TS 25.214 V8.7.0, Sep. 2009.

[4] J. McDonough, "Moving standards to 100 GbE and beyond," IEEE Applications and Practice Nov. 6–9, 2007.

[5] *Interfaces for Optical Transport Network*, ITU-T G.709, Dec. 2009.

[6] E. Inaty, H. M. H. Shalaby, P. Fortier, and L. A. Rusch, "Multirate optical fast frequency hopping CDMA system using power control," *J. Lightw. Technol.*, vol. 20, no. 2, pp. 166–177, Feb. 2002.

[7] B. Xu, "Concatenated codes-based bit-rate adaptation for blocking probability reduction in WDM networks," *Photon. Technol. Lett.*, vol. 17, no. 9, pp. 1983–1985, Sep. 2005.

[8] M. Arabaci, I. B. Djordjevic, R. Saunders, and R. M. Marcoccia, "Polarization-multiplexed rate-adaptive nonbinary-quasi-cyclic-LDPC-coded multilevel modulation with coherent detection for optical transport networks," *Opt. Exp.*, vol. 18, no. 3, pp. 1820–1632, Feb. 2010.

[9] T. Mizuochi, Y. Miyata, T. Kobayashi, K. Ouchi, K. Kuno, K. Kubo, K. Shimizu, H. Tagami, H. Yoshida, H. Fujita, M. Akita, and K. Motoshima, "Forward error correction based on block turbo code with 3-bit soft decision for 10 Gb/s optical communication systems," *IEEE J. Sel. Topics Quantum Electron.*, vol. 10, no. 2, pp. 376–386, Mar./Apr. 2004.

[10] G. D. Forney, *Concatenated Codes*. Cambridge, MA: MIT Press, 1966.

[11] *Forward Error Correction for High Bit-Rate DWDM Submarine Systems*, ITU-T G.975.1, Feb. 2004.

[12] I. B. Djordjevic and B. Vasic, "Nonbinary LDPC codes for optical communication systems," *IEEE Photon. Technol. Lett.*, vol. 17, no. 10, pp. 2224–2226, Oct. 2005.

[13] S. Lin and D. J. Costello, Jr., *Error Control Coding*, 2nd ed. Englewood Cliffs, NJ: Prentice Hall, 2004.

[14] S. B. Wicker, *Error Control Systems for Digital Communication and Storage*. Englewood Cliffs, NJ: Prentice Hall, 1995.

[15] *Forward Error Correction for Submarine Systems*, ITU-T G.975, Nov. 1996.

[16] G. P. Agrawal, *Fiber-Optic Communication Systems*, 3rd ed. : Wiley, 2002.

[17] J. G. Proakis, *Digital Communications*, 5th ed. : McGraw-Hill, 2007.

[18] *40 Gb/s and 100 Gb/s Ethernet Task Force*, IEEE P802.3ba [Online]. Available: http://www.ieee802.org/3/ba/public/index.html

[19] M. D. Feuer, D. C. Kilper, and S. L. Woodward, , I. P. Kaminow, T. Li, and A. E. Willner, Eds., "ROADMs and their system applications," in *Optical Fiber Telecommunications V B*. : Academic Press, 2008.

[20] E. Ip and J. M. Kahn, "Digital equalization of chromatic dispersion and polarization mode dispersion," *J. Lightw. Technol.*, vol. 25, no. 8, pp. 2033–2043, Aug. 2007.

[21] O. V. Sinkin, R. Holzlöhner, J. Zweck, and C. R. Menyuk, "Optimization of the split-step Fourier method in modeling optical-fiber communications systems," *J. Lightw. Technol.*, vol. 21, no. 1, pp. 61–68, Jan. 2003.

[22] J. P. Gordon and L. F. Mollenauer, "Phase noise in photonic communications systems using linear amplifiers," *Opt. Lett.*, vol. 15, pp. 1351–1353, 1990.

[23] K.-P. Ho and H.-C. Wang, "Comparison of nonlinear phase noise and intrachannel four-wave-mixing for RZ-DPSK signals in dispersive transmission systems," *Photon. Technol. Lett.*, vol. 17, no. 7, pp. 1426–1428, Jul. 2005.

[24] K.-P. Ho and H.-C. Wang, "Effect of dispersion on nonlinear phase noise," *Optics Lett.*, vol. 31, no. 14, pp. 2109–2111, Jul. 15, 2006.

[25] E. Desurvire, *Erbium-Doped Fiber Amplifiers: Principles and Applications*. : Wiley, 1994.

[26] D. Marcuse, C. R. Menyuk, and P. K. A. Wai, "Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence," *J. Lightw. Technol.*, vol. 15, no. 9, pp. 1735–1746, Sep. 1997.

Gwang-Hyun Gho received the B.S. degree in electronics engineering from Korea University, Seoul, Korea, in 1995, and the M.S. degree in electrical engineering from Stanford University, Stanford, CA, in 1999, where he is currently working towards the Ph.D. degree in electrical engineering.

From 1999–2005, he was at Qualcomm Inc., Campbell, CA, where he developed W-CDMA and HSDPA systems, as a systems engineer. From 2005–2009, he was with Amicus Wireless Technology Inc., Sunnyvale, CA, performing research and development for mobile WiMAX system, as a DSP engineer. His research interests include optical and wireless communications, error control coding, and digital signal processing.

Lauren Klak received the B.S. degree in chemical engineering in 2006 and M.S. degree in electrical engineering in 2008 from University of Southern California, Los Angeles. She is currently working towards the Ph.D. degree in electrical engineering at Stanford University, Stanford, CA.

From 2006–2009, she was a Communication Systems Engineer at Northrop Grumman Space Technology, Los Angeles, CA, where she developed algorithms for dynamic resource management and worked various payload engineering issues. Her research interests include optical fiber communications, nonlinear optics, and digital signal processing.

Joseph M. Kahn (M'90–SM'98–F'00) received the A.B., M.A. and Ph.D. degrees in physics from the University of California at Berkeley, in 1981, 1983, and 1986, respectively.

From 1987–1990, he was at AT&T Bell Laboratories, Crawford Hill Laboratory, in Holmdel, NJ. He demonstrated multi-Gb/s coherent optical fiber transmission systems, setting world records for receiver sensitivity. From 1990–2003, he was on the faculty of the Department of Electrical Engineering and Computer Sciences at U.C. Berkeley, performing research on optical and wireless communications. Since 2003, he has been a Professor of electrical engineering at Stanford University. His current research interests include single- and multi-mode optical fiber communications, free-space optical communications, and MEMS for optical communications. In 2000, he helped found StrataLight Communications, where he served as Chief Scientist from 2000–2003. StrataLight (now Opnext Subsystems) is a leading supplier of transmission subsystems for high-capacity terrestrial networks.

Prof. Kahn received the National Science Foundation Presidential Young Investigator Award in 1991. From 1993–2000, he served as a Technical Editor of IEEE Personal Communications Magazine. Since 2009, he has been an Associate Editor of IEEE/OSA Journal of Optical Communications and Networking.