DesignCon 2021

Global Optimization of Wireline Transceivers for Minimum Post-FEC vs. Pre-FEC BER

Ming Yang, University of Toronto
ming.yang@isl.utoronto.ca

Shayan Shahramian[†], Huawei Canada

Henry Wong, Huawei Canada

Peter Krotnev, Huawei Canada

Anthony Chan Carusone, University of Toronto
tony.chan.carusone@isl.utoronto.ca

[†]Now with Alphawave IP

Abstract

This paper presents an accurate and efficient methodology for optimizing high-speed wireline links and the substantial improvement obtained by using post-FEC vs. pre-FEC BER for optimization. A statistical model that accounts for FFE noise enhancement, DFE burst errors and other important noise sources finds the pre-FEC and post-FEC BER that serve as objective functions for optimizing each transceiver equalizer block. The statistical model can accurately estimate post-FEC BER using standard linear block codes. A genetic algorithm is combined with the statistical model to obtain the best set of design parameters for each transceiver block.

Authors Biography

Ming Yang received the B.Eng. degree in aerodynamic engineering from the Department of Aeronautics, Xiamen University, Xiamen, China, in 2012, and the B.Eng. and M.Eng. degree in electrical engineering from the Department of Electrical and Computer Engineering, McGill University, Montreal, Canada, in 2013 and 2016, respectively. He is currently a Ph.D. candidate in the Edward S. Rogers Sr. Department of Electrical & Computer Engineering at University of Toronto. He is the recipient of the Alexander Graham Bell Canada Graduate Scholarships award (NSERC CGS-D). His research interests are in analog integrated circuit design, on-chip analog signal processing and high-performance integrated circuit testing.

Shayan Shahramian received his Ph.D. from the Department of Electrical and Computer Engineering at the University of Toronto, Canada, in 2016. He is the recipient of the NSERC Industrial Postgraduate scholarship in collaboration with Semtech Corporation (Gennum Products). He is the recipient of the best young scientist paper award at ESSCIRC 2014 and received the Analog Devices outstanding designer award for 2014. He joined Huawei Canada in January 2016 and is currently working in system/circuit level design of high-efficiency transceivers for short reach applications.

Henry Wong received his B.A.Sc. and Ph.D. both in Electrical Engineering and currently he is a Distinguished Engineer in Huawei. His area of R&D interest is in SerDes design, for high-speed interface, optical module and backplane communications. He has worked for Nortel, Cadence, Lucent on high-speed modems, and for Gennum (Semtech) on SerDes, CDR. He joined Huawei in 2013 and currently he is also a senior manager of SerDes system architecture product development.

Peter Krotnev received Master's Degree in Telecommunications Engineering from the Technical University of Sofia in 1988. He is currently a Sr. Principal Engineer at Huawei Technologies, member of the High Speed I/O System Design Team. Peter's work is focused on SerDes architecture improvements, electrical specifications, as well as the development of the SerDes and circuitry tuning and adaptation strategies for power and performance. As a telecom professional Peter has also worked with STMicroelectronics Inc. on variety of projects and technologies including ADSL, Gigabit Ethernet and High Speed IOs. As signal integrity expert Peter is also involved in number of patents and papers.

Anthony Chan Carusone received his Ph.D. from the University of Toronto in 2002 and has since been a professor with the Department of Electrical and Computer Engineering at the University of Toronto. He is also an occasional consultant to industry in the areas of integrated circuit design and digital communication.
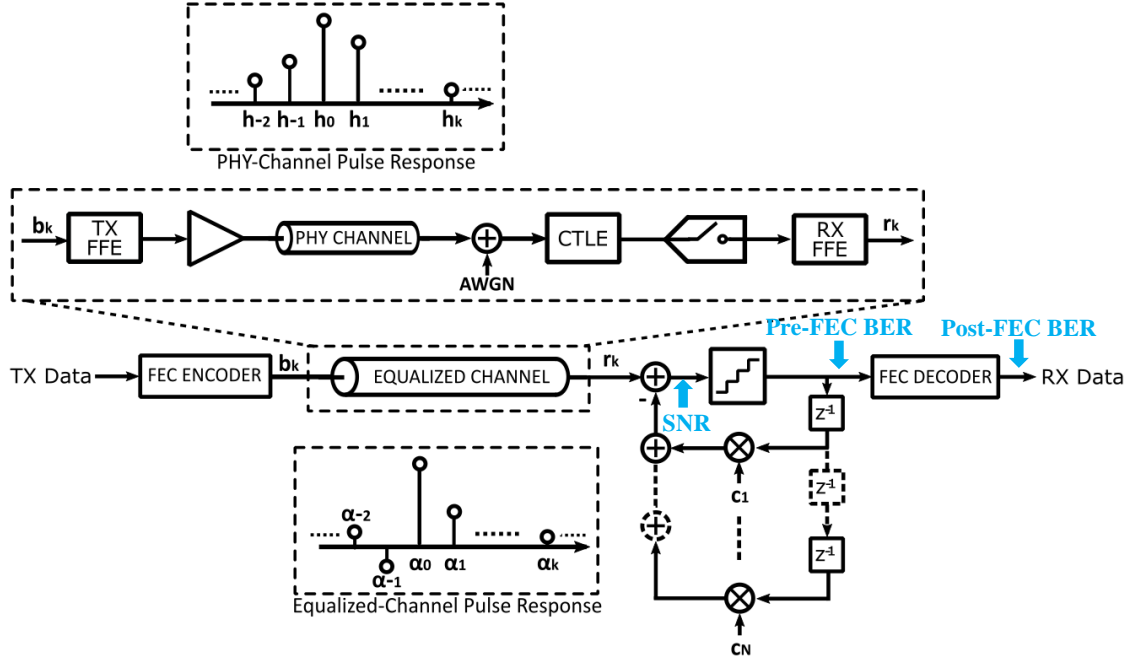
Prof. Chan Carusone co-authored the Best Student Papers at the 2007, 2008 and 2011 Custom Integrated Circuits Conferences, the Best Invited Paper at the 2010 Custom Integrated Circuits Conference, the Best Paper at the 2005 Compound Semiconductor Integrated Circuits Symposium, and the Best Young Scientist Paper at the 2014 European Solid-State Circuits Conference. He co-authored the popular textbooks "Analog Integrated Circuit Design" (along with D. Johns and K. Martin) and "Microelectronic Circuits," 8th edition (along with A. Sedra, K.C. Smith and V. Gaudet). He was Editor-in-Chief of the IEEE Transactions on Circuits and Systems II: Express Briefs in 2009, and an Associate Editor for the IEEE Journal of Solid-State Circuits 2010-2017. He was a Distinguished Lecturer for the IEEE Solid-State Circuits Society 2015-2017 and served on the Technical Program Committee of the International Solid-State Circuits Conference 2016-2021. He is currently the Editor-in-Chief of the IEEE Solid-State Circuits Letters.

1. Introduction

Transceiver optimization for high-speed wireline links is becoming increasingly challenging as data rates advance towards 100Gb/s and above. More complex modulation and less link margin necessitate more accurate optimization of more transceiver parameters. The optimization process involves fine-tuning many transceiver blocks such as continuous-time-linear equalizers (CTLE), feedforward equalizers (FFE) and decision-feedback equalizers (DFE) to satisfy a given bit-error rate (BER) requirement. Many existing optimization methods select minimum mean-square error (MMSE) or pre-FEC BER as the objective function. However, the optimal settings for pre-FEC BER may not correspond to the optimum for post-FEC BER. This paper demonstrates a substantial difference in the overall link performance by optimizing for post-FEC vs. pre-FEC BER. In addition, optimizing the CTLE for many long-reach applications requires optimizing other equalizers in the receiver. Although least-mean-square (LMS) adaptation and other gradient-based methods are commonly employed for FFE and DFE adaptation, they are ill-suited to CTLE optimization, which has a non-unimodal performance surface. The time required to accurately evaluate each CTLE setting through an exhaustive search grows exponentially as the number of CTLE control parameters increases. Therefore, an efficient method to accurately optimize CTLE and other transceiver equalizer blocks is an essential part of high-speed wireline links.

This paper presents a systematic methodology using post-FEC BER to perform transceiver optimization for high-speed wireline links with non-unimodal performance surfaces. A statistical model is used to find the pre-FEC and post-FEC BER that serve as objective functions for optimizing each transceiver equalizer block. The statistical model can accurately estimate post-FEC BER using standard linear block codes, such as the RS(544,514,15) KP4 and RS(528,514,7) KR4 codes. It applies Markov chain theory to account for DFE error propagation. Then, trellis dynamic programming is used to find the probability of all error patterns corrupting the FEC decoder, allows for accurate calculation of post-FEC BER. Advanced optimization algorithms are combined with the statistical model to obtain the best set of design parameters for each transceiver block.

The main body of this work is divided into three subsections. In Section 2, we expand our previously proposed statistical model to include the noise amplification of both the CTLE and FFE. Accounting for noise amplification is necessary to capture the interdependency of equalizer blocks fully. In Section 3, we select the cost function for this optimization problem. Specifically, we will be showing that post-FEC BER is preferred over pre-FEC BER (or SNR) as the objective function for these types of co-optimization problems. Section 4 demonstrates a systematic procedure to perform CTLE optimization when followed by an FFE and DFE based on advanced optimization algorithms using a 4-PAM wireline transceiver example. Finally, we draw conclusions in Section 5.

Figure 1. Proposed wireline transceiver system model with an *N*-tap DFE at receiver, the equalized pulse response *α*(*z*) is generated by convolving PHY-channel pulse response *h*(*z*) with the impulse response of other components in the equalized channel.

2. SerDes Link Modeling

2.1 System Overview

Figure 1 shows our proposed high-speed wireline system model communicating symbols b_k using pulse-amplitude modulation (PAM) with time index k. The PAM symbols are filtered by an equalized channel response $\alpha(z) = \cdots \alpha_{-1}z^1 + \alpha_0 + a_1z^{-1} + \cdots + a_kz^{-k} + \cdots$ whose main cursor is α_0. The response $\alpha(z)$ is the physical channel's pulse response convolved with the impulse response of other linear components in the link, such as the transmitter (TX) FFE, TX driver, CTLE and receiver (RX) FFE. The physical channel introduces additive white Gaussian noise (AWGN), which is filtered by the CTLE and FFE, creating correlated noise. Sections 2.2 and 2.3 present detailed analyses of CTLE noise shaping and RX FFE noise enhancement, respectively. While ADC impairments can be modelled using the methods proposed by [1], we simplify the problem by assuming ideal ADC operation.

Knowing the equalized channel response, $\alpha(z)$, and AWGN variance, we calculate the probability density function (pdf) of the received samples r_k at its output. These results are applied to the statistical BER model proposed in [2] to obtain both the pre-FEC and post-FEC BER subject to error propagation from the *N*-tap DFE in Figure 1.

2.2 CTLE Noise Shaping

The AWGN is coloured by the CTLE. We require both the variance σ^2 and autocorrelation function $R(\tau)$ of the noise at the CTLE output to compute further noise amplification in the RX FFE. To calculate the noise variance σ^2, we first define $P_n(f) = K$ as the constant power spectral density of the zero-mean AWGN process. Assuming the CTLE has an impulse

response $b(t)$ and its Fourier transform is $B(f)$, we calculate the noise autocorrelation function using the Wiener-Khinchin theorem,

$$R(\tau) = K \int_{-\infty}^{\infty} |B(f)|^2 e^{i2\pi f\tau} df. \tag{1}$$

Since an AWGN process is wide-sense stationary, the output of the CTLE has a noise variance

$$\sigma^2 = R(0). \tag{2}$$

2.3 FFE Noise Amplification

The optimal tap coefficients in a receiver FFE generally depend on the channel response and noise spectrum. The link's BER performance, in turn, depends on the RX FFE's noise amplification. Thus, in this subsection, we describe how to find the FFE noise amplification.

First, we define X as a zero-mean random process describing the CTLE-filtered Gaussian noise defined by (1) and (2). The FFE output noise Y is the weighted sum of M correlated random variables (X_1, X_2, … X_M) sampled at M unit intervals (UI). For a link communicating at a symbol rate $1/T_s$, the covariance $Cov(X_i, X_j)$ between X_i and X_j is

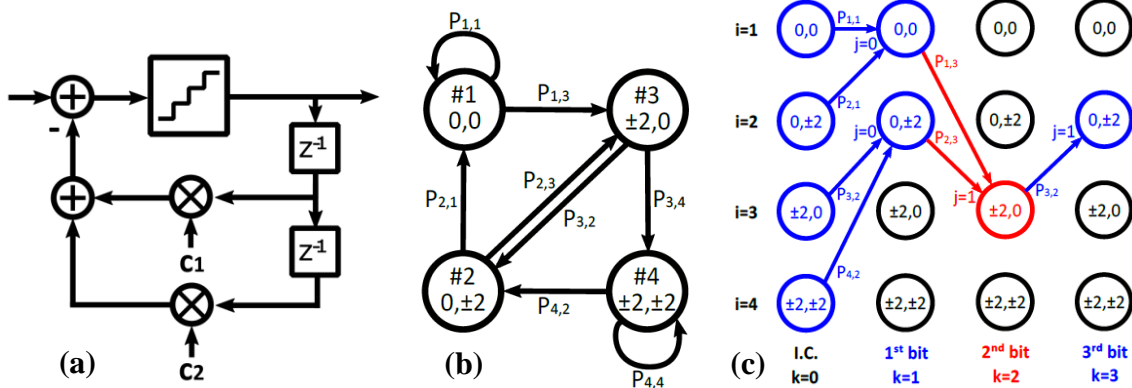$$Cov(X_i, X_j) = R(T_s \cdot |i - j|). \tag{3}$$

We define β_i as the ith FFE tap coefficient in an M-tap FFE. The total noise variance $Var(Y)$ at the FFE output node $Y = \sum_{i=1}^{M} \beta_i X_i$ is

$$Var(Y) = \sum_{i=1}^{M} \sum_{j=1}^{M} \beta_i \beta_j Cov(X_i, X_j). \tag{4}$$
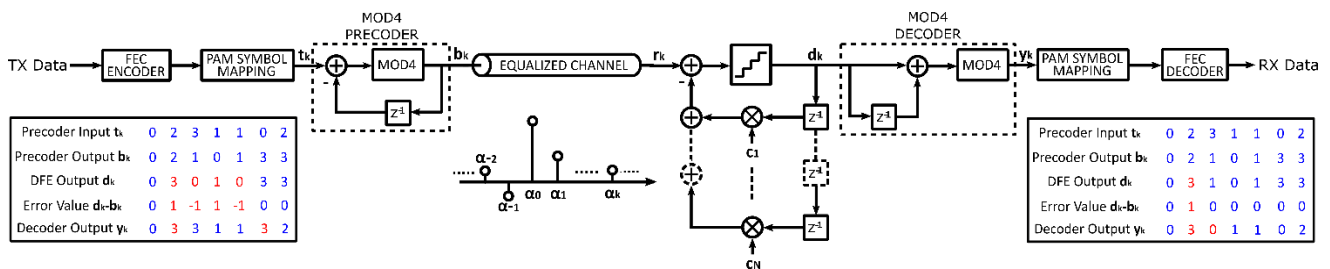
2.4 Pre-FEC and Post-FEC BER Estimation using Standard RS Code

Error propagation in decision-feedback equalization can significantly impact BER [3-4]. A DFE removes channel ISI by registering past equalized symbols in the feedback path and using them to estimate and cancel ISI from the current symbol. However, if any past symbol registered in the DFE is wrong, the receiver's ISI estimate is biased and may increase the probability of additional symbol errors. Errors may thus propagate around the DFE feedback loop and result in FEC code failures. In current long-reach wireline SerDes applications, such as 100GBase-KP4, Gray-coded 4-PAM signaling and Reed-Solomon (RS) FEC are standard [5-6]. In linear FEC codes on GF(2^m), the encoder groups every m bits into one FEC symbol, and the decoder can correct up to t erroneous FEC symbols in an n-symbol codeword. All m bit errors in each erred FEC symbol are corrected so long as the total number of FEC symbol errors does not exceed t. Hence, higher-order RS codes can correct longer error bursts and have therefore been specified, in part, to accommodate DFE error propagation. Error bursts can become much longer when DFE tap weights are large or alternating in sign [2]. Such burst errors can reduce the coding gain offered by popular FEC codes [3][7]. Even RS codes often used in wireline links and generally considered effective at correcting bursts are still significantly hampered by DFE error propagation.

Figure 2. Modeling error propagation of a 2-tap DFE: (a) schematic diagram of a 2-tap DFE (b) its Markov chain model (c) time-unrolling the Markov chain model to generate trellis diagram for finding both pre-FEC and post-FEC BER using dynamic programming.

A statistical model proposed in [2] is adopted in this work to accurately estimate both pre-FEC and post-FEC BER for high-speed wireline links subject to DFE error propagation. Figure 2 shows the key procedures for finding BER using a 2-tap DFE example. Error propagation in the 2-tap DFE pictured in Figure 2(a) can be modelled by the Markov process shown in Figure 2(b), whose state is specified by the past two detection outcomes. In this example, the two possible outcomes are taken from the set $\{\pm 2, 0\}$ where ± 2 indicates a detection error while 0 indicates an error-free detection. By time-unrolling this 4-state Markov model, we can obtain the trellis diagram in Figure 2(c). To compute the post-FEC BER, we must find the probability of all error patterns having more than t FEC symbol errors in a codeword. Rather than finding the BER by enumerating all possible error patterns in the trellis, we instead apply dynamic programming, which solves the problem much faster by grouping the probability of all trellis paths having the same number of bit errors. We repeat the same aggregation procedure recursively when traversing through each stage in the trellis, resulting in a significant reduction in computational complexity.

Figure 3. A modified wireline transceiver system model with an *N*-tap DFE at receiver to consider 1/(1+D) pre-coding. Two examples are included in the figure illustrating: (1) a DFE burst error across four PAM symbols is mitigated to only two errors (2) a random error is duplicated with pre-coding.

2.5 Post-FEC BER Estimation with 1/(1+D) Precoding

A well-known technique for mitigating error bursts is 1/(1+D) precoding (also referred to as MOD4 precoding). Figure 3 shows a wireline transceiver model incorporating 1/(1+D) pre-coding. The MOD4 encoder accepts input t_k and generates transmitted symbols b_k.

The RX decoder accepts the DFE decisions d_k as inputs and produces the outputs y_k, which are estimates of t_k. Figure 3 also includes two example sequences illustrating how precoding mitigates error bursts. The MOD4 decoder removes burst errors because the error d_k-b_k in the current received symbol is added to the error d_{k-1}-b_{k-1} in the previously received symbol. For $c_1 > 0$, the burst error values arise due to DFE error propagation always take alternating signs in the form … +1 -1 +1 … [2]; as a result, consecutive error values cancel when added in the decoder. However, isolated individual symbol errors give rise to two consecutive symbol errors after decoding. Thus, the second example in Figure 3 illustrates how a BER penalty arises from random isolated errors in the link. A method to model 1/(1+D) precoding using trellis dynamic programming appears in [8]. This method is used to generate the post-FEC BER results with 1/(1+D) pre-coding.

3. Pre-FEC and Post-FEC BER as Criteria for Optimizing Wireline Transceivers

ADC-based receiver is a popular architecture for 100Gb/s+ long-reach wireline SerDes, where receiver equalization relies heavily on digital signal processing (DSP) using advanced CMOS technology nodes to accommodate high-speed operation. Forward-error-correction (FEC) codes have also become an integral part of the DSP, lowering the post-FEC BER by several orders of magnitude compared to the raw pre-FEC BER. Feedforward and decision-feedback equalizers are common receiver DSP blocks, each offering benefits and shortcomings. For example, a linear finite-impulse-response (FIR) FFE can reduce both the pre-cursor inter-symbol interference (ISI) and post-cursor ISI but may lead to noise amplification. On the other hand, the DFE does not suffer from noise amplification but can only remove post-cursor ISIs. Moreover, alleviating the critical timing path in the DFE feedback loop requires parallelization, which in the past has limited the DFE to only 1-2 taps in 100Gb/s+ wireline applications [9-10]. FFE and DFE tap coefficients are typically optimized to maximize signal-to-noise ratio (SNR) or to minimize the mean-squared error or pre-FEC BER [11-13]. However, the parameters found by these conventional methods do not necessarily correspond to the minimum post-FEC BER operating point, which is ultimately most important.

SNR, pre-FEC and post-FEC BER are three performance metrics used to optimize the architecture and coefficients of a wireline transceiver. Figure 1 highlights the locations at which these three metrics are measured. Currently, RX FFE and DFE coefficients are optimized using LMS adaptation algorithms where SNR is implicitly the optimization criteria. However, DFE error propagation is not captured by the SNR and the SNR-optimal FFE coefficients do not necessarily minimize pre-FEC BER [2] [12]. In addition, the FEC decoding process is much more sensitive to error bursts caused by DFE error propagation than isolated random errors. Even using pre-FEC BER as the sole criteria for optimization fails to account for this. This section applies the transceiver model proposed in Section 2 and, assuming standard FEC codes, compares pre-FEC and post-FEC BER as criteria to optimize FFE and DFE.
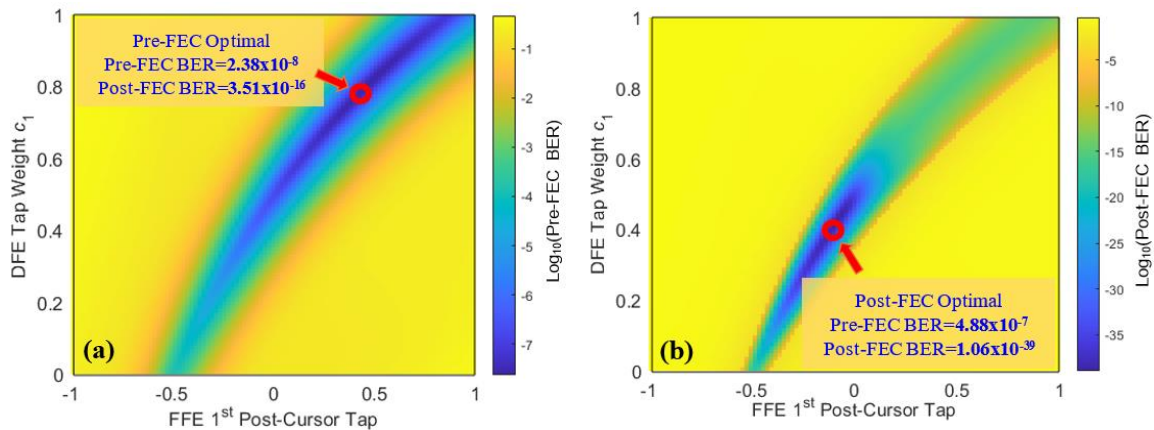
3.1 Pre-FEC vs Post-FEC BER Optimum

We adopt a channel model with 30 dB insertion loss for a link communicating 4-PAM symbols at 56 GBaud/s subject to 0.55 V_{P-P} swing at TX. At the receiver, we assume a simplified CTLE model with one zero at 3.77 GHz and two poles at 28.2 GHz and 31.2 GHz, respectively, which together provide 12 dB peaking gain and 0 dB gain at DC. The CTLE-equalized pulse response is $h(z) = 0.1391\ z^1 + 0.4062 + 0.1876\ z^{-1} + 0.0237\ z^{-2} + 0.0009\ z^{-3}$ including both the CTLE and physical channel. The AWGN integrated rms noise is 4.58 mV_{rms}. State-of-art wireline links employ DFEs with 1-2 taps [9-10]. To illustrate the basic tradeoffs, we first assume a 1-tap DFE and a 7-tap FFE with 2 pre-cursor and four post-cursor taps. There is no pre-emphasis in the TX. When sweeping the 1st post-cursor FFE tap, other FFE tap weights are chosen to minimize all pre-cursor and post-cursor ISIs using MMSE criterion. The post-FEC BER is calculated assuming the standard KP4 RS(544, 514, 15) code.

In Figure 4, the pre-FEC BER and post-FEC BER performance surfaces are generated by sweeping the 1st post-cursor FFE tap weight and the DFE tap weight. The FFE main-cursor tap always maintains its amplitude at 1. The DFE tap weights in Figure 4 are normalized to α_0.

We obtain substantially different optimal points on the two performance surfaces plotted in Figure 4. In Figure 4(a), the minimum pre-FEC BER is located at $\alpha_1/\alpha_0 = 0.80$. In Figure 4(b), the minimum post-FEC BER appears at $\alpha_1/\alpha_0 = 0.42$. A large and positive FFE 1st post-cursor tap weight creates a low-pass response that filters noise and improves SNR and, thus, pre-FEC BER at the DFE output. However, this implies a commensurately large DFE tap weight, which increases the frequency and length of error propagation bursts. The lower value of $\alpha_1/\alpha_0 = 0.42$ affords a pre-FEC BER 1.3 orders of magnitude higher, but a post-FEC BER that is 23.5 orders of magnitude lower. This suggests that DFE error propagation has a dominant impact on the post-FEC BER. Therefore, the tradeoff between FFE noise enhancement and DFE error propagation must be considered when architecting and optimizing wireline transceivers to minimize post-FEC BER. Unfortunately, LMS equalizer adaptation algorithms do not consider this effect.

Figure 4. BER performance surface generated by sweeping the FFE 1st post-cursor tap and the DFE tap weight using (a) Pre-FEC BER (b) Post-FEC BER.

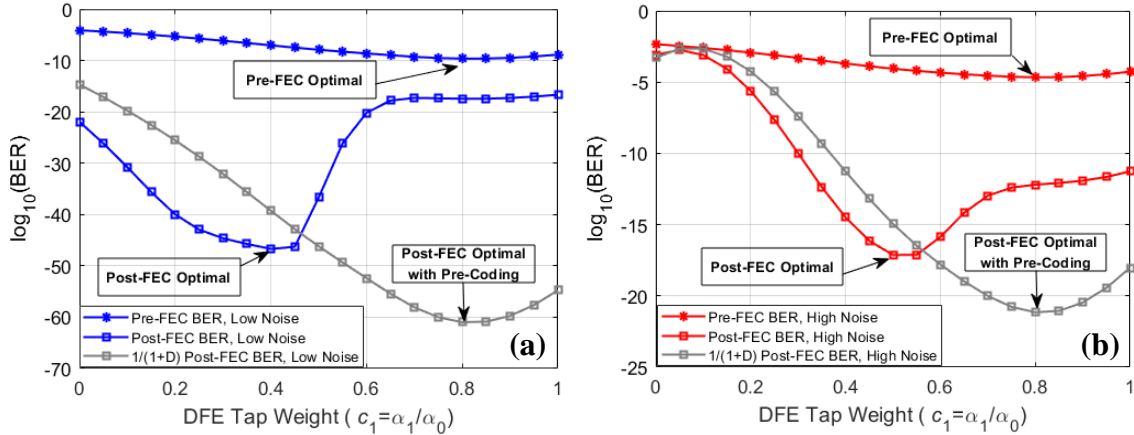TABLE I. PULSE RESPONSE OF THE EQUALZIED CHANNEL BY INCLUDING TX FFE, CHANNEL AND CTLE

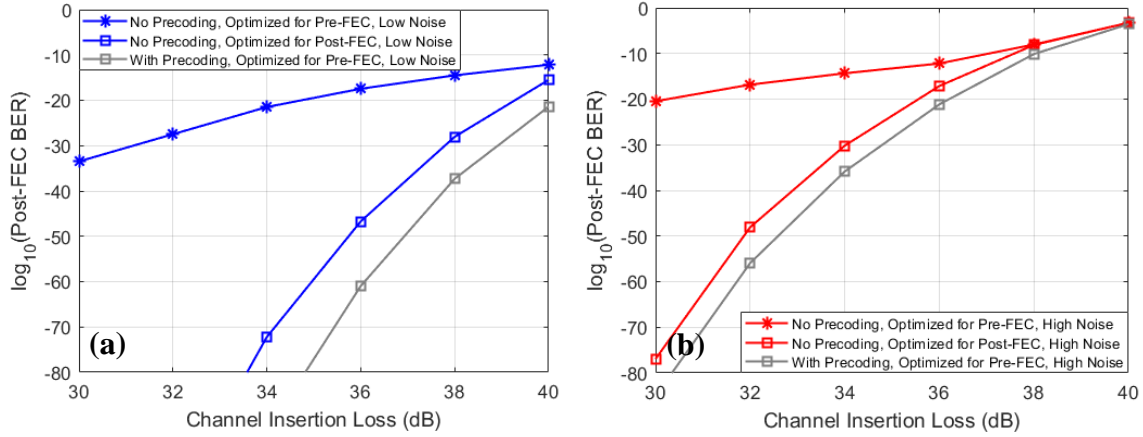| Case	Channel IL (dB)	Equalized Pulse Response at CTLE Output (mV)											
		α_{-3}	α_{-2}	α_{-1}	α_0	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9
1	30	-1.39	-15.3	17.0	173	49.8	34.3	23.2	14.5	8.51	3.60	2.84	0.272	1.37
2	32	-1.45	-13.7	18.4	151	53.9	35.1	24.4	15.5	10.2	4.01	4.52	0.530	2.02
3	34	-1.49	-12.3	19.8	132	54.7	36.3	25.4	16.5	11.1	5.45	4.73	1.75	2.33
4	36	-1.50	-11.1	21.0	117	54.4	37.4	25.5	17.9	11.8	6.62	5.34	2.60	2.71
5	38	-1.50	-9.85	21.7	103	53.7	37.9	26.0	18.8	12.5	7.67	6.07	3.26	3.20
6	40	-1.46	-8.73	21.9	91.4	52.5	37.9	26.3	19.5	13.1	8.53	6.61	3.94	3.63

3.2 Simulation Results: 1-Tap DFE

This subsection provides more extensive simulation results using six measured channel responses to validate our methodology using post-FEC BER to find the optimal equalizer coefficients. The general simulation setup is similar to that used for Figure 4 except that the TX now has a 2-tap FFE providing 5 dB pre-emphasis, and the RX FFE has 15 taps, including 3 pre-cursor taps and 11 post-cursor taps. An 8th-order CTLE model was applied to equalize all six channels. The equalized channel pulse responses, including TX FFE, CTLE and PHY channel are tabulated in Table I.

Figure 5 plots both the pre-FEC BER and post-FEC BER as a function of α_1/α_0 for the 36dB-loss channel at two integrated rms noise levels: 1.62 mV$_{rms}$ and 2.42 mV$_{rms}$. For each data point, the DFE tap weight c_1 zero-forces the first post-cursor at the indicated α_1/α_0 and the corresponding MMSE FFE tap weights are found. For both noise levels, post-FEC BER is minimized at a lower α_1/α_0 than pre-FEC. Thus, to minimize post-FEC BER, the FFE should be relied upon for more of the RX equalization than an MMSE (or LMS) criteria suggests. We also superimposed the post-FEC BER results with 1/(1+D) pre-coding in Figure 5. Since pre-coding eliminates long error bursts, the minimum post-FEC BER with precoding is lower and occurs at larger values of α_1/α_0. In fact, with precoding, both post-FEC and pre-FEC BER are minimized with the same equalizer coefficients.

Figure 6 plots the post-FEC BER of the link using two different criteria to optimize the equalizer coefficients: pre-FEC BER and post-FEC BER. The post-FEC BERs (assuming no precoding) at the pre-FEC and post-FEC optimal points are indicated by asterices and square markers, respectively. Figures 6(a) and 6(b) plot the results for all 6 channel models at the two noise levels. Without 1/(1+D) precoding, the optimal post-FEC BER obtained by post-FEC optimization is always superior to the post-FEC BER obtained from pre-FEC optimization. For example, in Figure 6(b) for the 34-dB loss channel, the post-FEC BER degrades from 10^{-30} to 10^{-14} when the equalizer is optimized for pre-FEC BER. In both figures the improvement from using post-FEC BER for optimization is most dramatic at lower channel losses and/or lower noise levels. In higher-loss channels, the FFE provides more high-frequency boost and noise amplification. When random errors dominate over long error bursts at high BER, the pre-FEC and post-FEC BER optima coincide.

Figure 5. Pre-FEC vs post-FEC BER as a function of α_1/α_0 simulated using the 36 dB channel case and RS(544,514,15) FEC with integrated rms noise level at (a) 1.62 mV$_{rms}$ (b) 2.42 mV$_{rms}$.
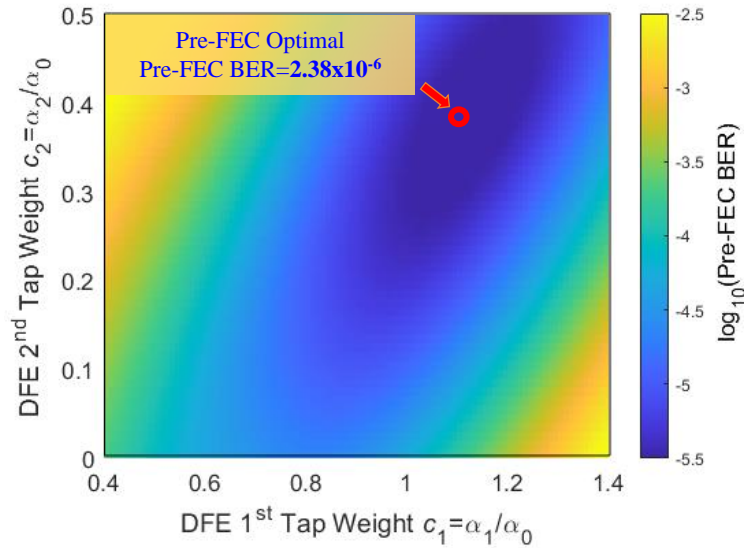

Figure 6. Optimal post-FEC BER using equalizers minimizing pre-FEC and post-FEC BER, simulated for all 6 channel responses and RS(544, 514, 15) FEC with an 1-tap DFE; two noise levels are (a) 1.62 mV$_{rms}$ (b) 2.42 mV$_{rms}$.

With precoding, the post-FEC BER is also plotted in Figure 6 in gray using the same equalizer coefficients found at the pre-FEC optimal. Not surprisingly, as $1/(1+D)$ precoding can effectively remove DFE burst errors, the optimal post-FEC BER is always several orders of magnitude better than the optimal post-FEC BER without applying precoding. Thus, for this 1-tap DFE example with pre-coding turned on, FFE and DFE tap weights can be optimized using conventional methods without fear that error propagation will result in bursts that hurt post-FEC BER.
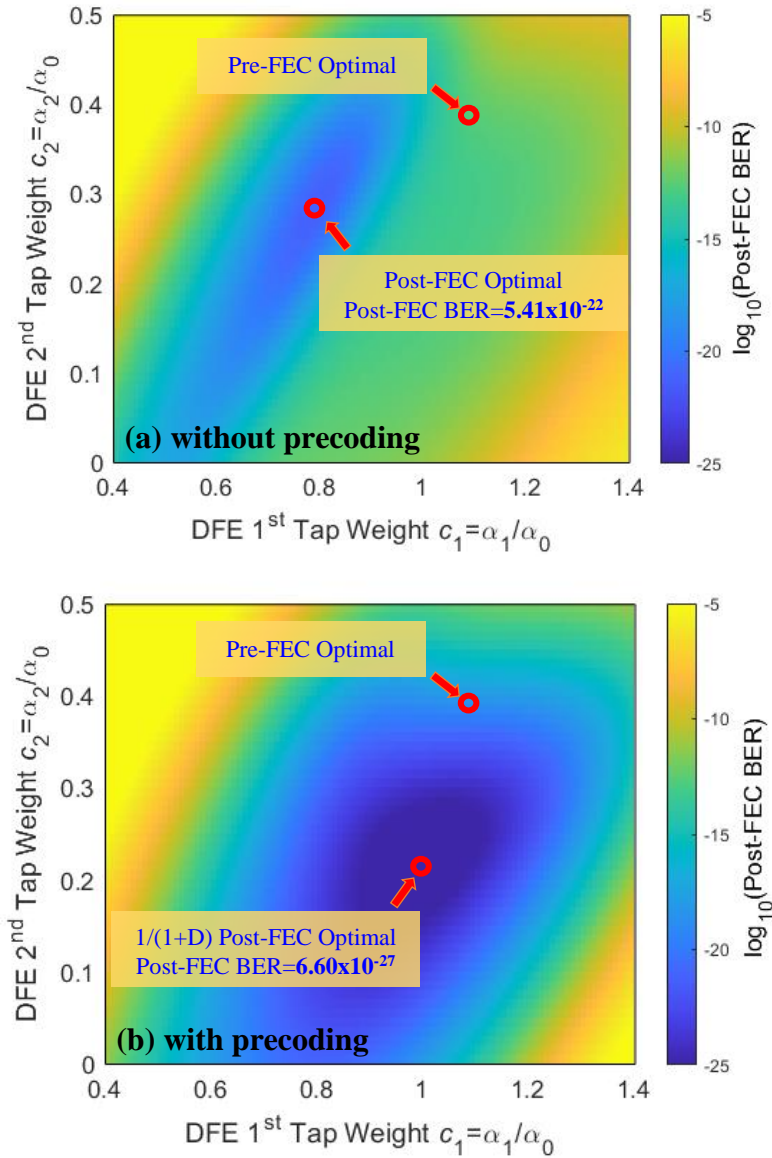
3.3 Simulation Results: 2-Tap DFE

In this subsection, we extend the analysis to a 2-tap DFE assuming the transceiver architecture, channel and noise settings are the same as in Section 3.2. Each BER curve plotted in Figure 5 now becomes a three-dimensional performance surface as a function of both the first and second DFE tap weight c_1 and c_2. We will show that the burst-error characteristic of the 2-tap DFE is very different from the previous 1-tap DFE example.
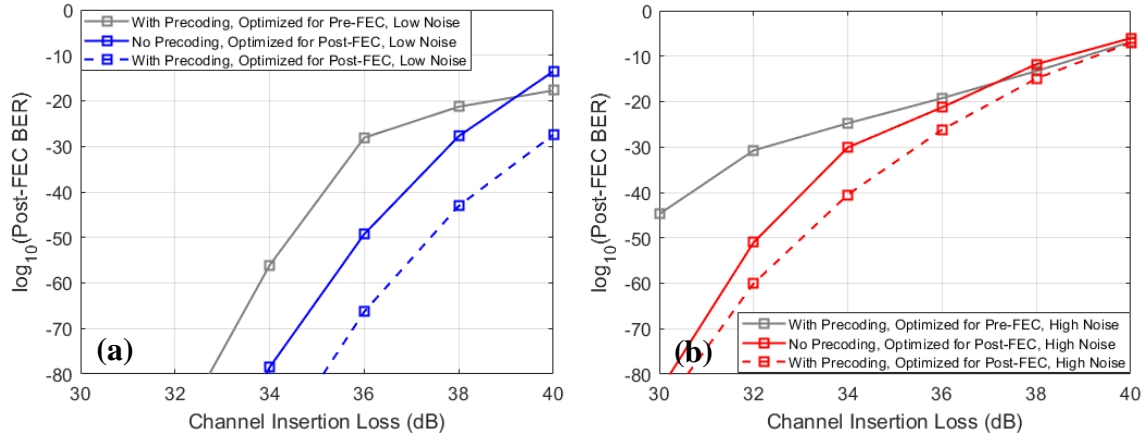
Figure 7. Pre-FEC BER performance surface (without precoding) as a function of DFE tap weights c_1 and c_2, using the 36 dB loss channel with 2.42 mV$_{rms}$ integrated rms noise.

In Figure 7, the pre-FEC BER performance surface is generated by sweeping the 1st and 2nd DFE tap weights for the 36dB-loss channel at the 2.42 mV$_{rms}$ noise level. Similarly, the post-FEC BER performance surface without and with precoding are plotted in Figure 8(a) and Figure 8(b), respectively. Compared with the 1-tap DFE example, the 2-tap DFE affords the FFE post-cursor taps with one more degree of freedom to low-pass filtering the noise. Thus, the 2-tap DFE demonstrates significant BER improvement in both the optimal pre-FEC and post-FEC BERs compared to the optimal BERs found in Figure 5(b). For example, the optimal post-FEC BER with precoding reduces from 10^{-21} to 10^{-27}. In addition, we also notice vastly different optimal points identified on each performance surface. The optimal pre-FEC and post-FEC BER highlighted in Figure 8(b) suggest that even with precoding, post-FEC BER is no longer optimized at the same equalizer coefficients as pre-FEC. Although Figure 7 shows that the pre-FEC BER benefits from large DFE tap weights at c_1=1.08 and c_2=0.38, the commensurately large 1st DFE tap weight may incur burst error values that are larger than one PAM-symbol distance. As a result, it's no longer safe to assume that the MOD4 decoder will remove burst errors as described in Section 2.5 because the error patterns may be in the form … +1 -2 +1 -1 +2 … . The MOD4 decoder cannot correct the burst when two consecutive error values alternating in sign have different magnitudes. Instead, the optimal post-FEC BER with precoding is found at slightly lower tap weights, c_1=1 and c_2=0.21 in Figure 8(b), thereby reducing the probability of such large error values arising during error propagation.

Figure 8. Post-FEC BER performance surface as a function of DFE tap weights c_1 and c_2, using the 36 dB loss channel with 2.42 mV_rms integrated rms noise: (a) without precoding (b) with precoding.

The optimal post-FEC BER in Figure 8 is also minimized at a much higher α_1/α_0 than the 1-tap DFE example. The post-FEC performance surfaces in Figure 8(a) suggest that if the 1st DFE tap is at c_1=0.8, then the 2nd DFE tap at c_2=0.3 has a much lower post-FEC BER compared to c_2=0.2 or 0.1. At the same pre-FEC BER level, intuitively, one would think a larger DFE tap weight always makes post-FEC BER worse. However, as described in [2], that is only true when DFE tap weights alternate in sign. In this case, the optimal DFE taps have the same sign, so increasing the DFE 2nd tap weight actually *reduces* the probability of continuing error propagation. The 'net effect' of DFE error propagation in this 2-tap DFE case is equivalent to a 1-tap DFE with α_1/α_0=0.5, which is precisely where the post-FEC optimum is found in Figure 5(b). Figure 8(a) shows that the post-FEC BER degrades as the 2nd DFE tap weight increases above 0.3. Although a larger DFE 2nd tap weight can reduce the probability of continuing error propagation, it also significantly
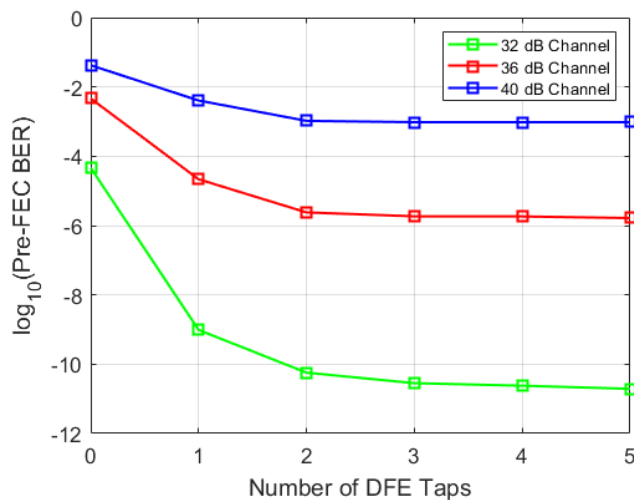
Figure 9. Optimal post-FEC BER using equalizers minimizing pre-FEC and post-FEC BER, simulated for all 6 channel responses and RS(544, 514, 15) FEC with a 2-tap DFE; two noise levels are (a) 1.62 mV$_{rms}$ (b) 2.42 mV$_{rms}$.

increases the probability that a lone random error initiates a burst and, thus, makes post-FEC BER worse. Similar trends observed in Figure 7 and Figure 8 prove that our statistical model can properly capture the tradeoff of a large vs small DFE 2nd tap weight based on each performance metric when optimizing a 2-tap DFE.

We repeat the same analysis for all 6 channels, and the optimal post-FEC BERs are reported in Figure 9(a) and Figure 9(b) at the two different noise levels. With precoding, the post-FEC BERs are now a lot worse with equalizers that minimize pre-FEC BER than with equalizers that minimize post-FEC BER even without precoding. In each subfigure, the dashed line represents the post-FEC BERs with equalizers optimized for post-FEC BER using precoding, always resulting in the lowest post-FEC BER.

3.4 Summary

Figure 10. Optimal pre-FEC BER vs number of DFE taps N using the 32 dB, 36 dB and 40dB channel with integrated rms noise level at 2.42 mV$_{rms}$.

In Figure 10, we plot the optimal pre-FEC BER vs number of DFE taps N using the 32 dB, 36 dB and 40 dB channel with integrated rms noise level at 2.42 mV$_{rms}$. Best equalizer coefficients are found at each data point using a simple gradient-descend based optimizer. As the number of DFE taps increases, one can achieve better noise filtering through RX FFE and thus minimizes pre-FEC BER, but the benefit quickly diminishes when $N > 2$. Moreover, as DFE does not have noise amplifcation, we also have some improvement on pre-FEC BER at large N. While more DFE taps are used to replace FFE to cancel post-cursor ISIs, the benefit for doing so is only marginal for canceling small ISI cursors. For long-reach high-speed wireline tranceiver designs, although N is typically limited to 1-2 due to the critical timing path in the DFE feedback loop, we can still achieve near-optimal performance using a 2-tap DFE as long as the post-cursor residual ISIs cancelled by the RX FFE are small.
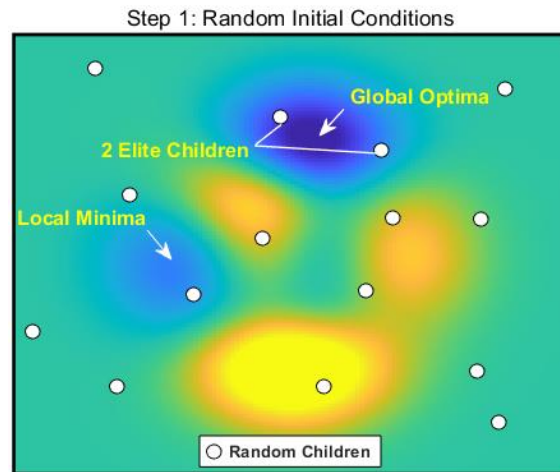
In this section, we consider whether architecting and optimizing wireline links using SNR or pre-FEC BER as performance metrics is effective in minimizing post-FEC BER. Burst errors due to DFE error propagation hurt FEC performance, but error propagation is not accurately accounted for when SNR or pre-FEC BER are used as the criteria for architecting and optimizing wireline links. Thus, we showed that, in general, links attain their minimum post-FEC BER with equalizer coefficients very different from those that minimize pre-FEC BER. However, the introduction of 1/(1+D) precoding mitigates the impact of error bursts, ensuring that both pre-FEC and post-FEC BER are minimized with the same equalizer coefficients for the 1-tap DFE example. We also showed in the 2-tap DFE example that the optimal post-FEC BER is minimized with very different equalizer coefficients using each performance metric. This analysis may have significant implications on the architecture and optimization of wireline transceivers.

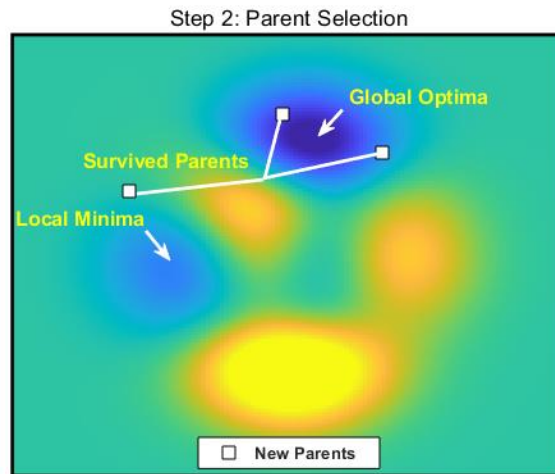## 4. Transceiver Optimization using Genetic Algorithm

In Section 3, we discussed co-optimizing a receiver FFE and DFE to find the optimal equalizer coefficients using pre-FEC vs post-FEC BER. In current high-speed wireline systems, transmitter FFE, receiver CTLE, FFE and DFE are always employed jointly to equalize the channel. Although LMS adaptation and other gradient-descent methods are commonly employed for optimizing FFE and DFE coefficients, they are ill-suited to CTLE optimization which has a non-unimodal performance surface [14]. The time required to accurately evaluate each CTLE setting through an exhaustive search grows exponentially as the number of CTLE control parameters increases. In addition, for long-reach wireline links, the benefits of increased transmit equalization must be balanced against the amplitude penalty associated with it. Therefore, an efficient method to accurately co-optimize CTLE and other transceiver equalizer parameters is an essential part of high-speed wireline links.

In this section, we present a systematic methodology using either pre-FEC or post-FEC BER to perform transceiver optimization when the performance surface is not unimodal. A genetic algorithm (GA) is combined with the statistical model to obtain the best candidate settings for each transceiver block.
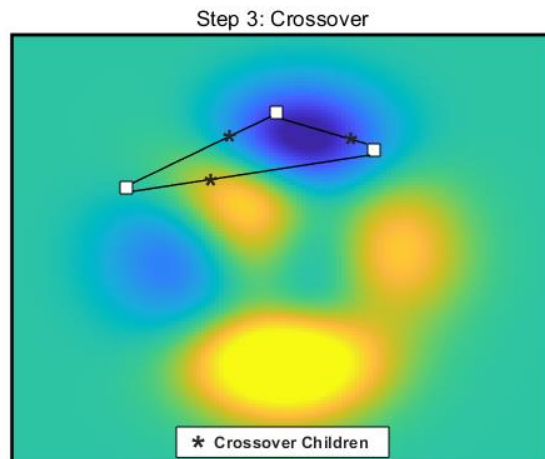
4.1 Genetic Algorithm Overview

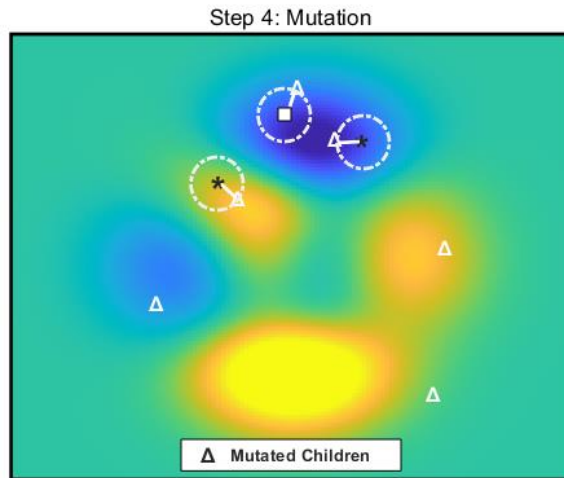(a) Step 1: generating random initial conditions on a non-unimodal performance surface.

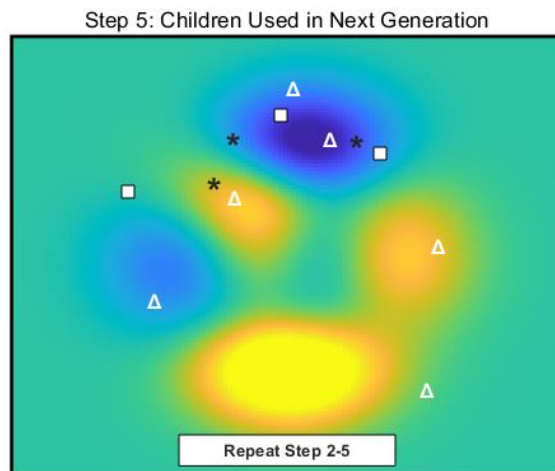(b) Step 2: parent section, in this example the top 2 elite individuals are guaranteed to survive.

(c) Step 3: crossover between new parents, possible crossover functions include single-point, double-point and uniform crossover.

(d) Step 4: generating mutated children from parents and crossover children.
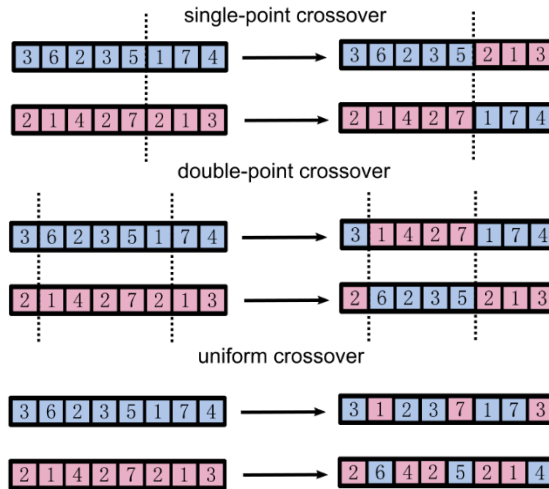

(e) Step 5: merging all parents and children from step 2-4, then repeat step 2-5 in the next iteration until stopping criteria are met.

Figure 11. Key procedures for transceiver optimization using genetic algorithm: (a) generating random initial conditions (b) parent selection (c) crossover (d) mutation (e) merging all parents and children of current generation.

Figure 11 provides an overview of the key procedures for transceiver optimization using a genetic algorithm. In Figure 11(a), we first generate an initial set of equalizer parameters at random. Alternatively, one may generate initial conditions from a pre-selected set of candidates. This may be particularly useful if the shape of at least part of the performance surface is known *a priori*. For example, a good set of CTLE settings related to high-frequency boosting may be applied if the channel loss is approximately known. This may help reducing the number of iterations required to find the global optimal.

Next, parent selection is performed as shown in Figure 11(b). A portion of the existing population is selected to become the parents of a new generation. In the example, the top
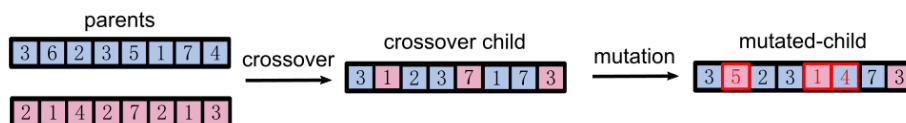
Figure 12. Three common crossover functions used for combining the genetic information from two parents.

two individuals having the minimal BERs are labeled as elite individuals and they are guaranteed to survive. Additional individuals may be selected based on some well-known genetic operators including roulette wheel and tournament selection [15].

In Figure 11(c), the selected individuals become parents of a new generation, and new children are created by crossover. Every two parents generate a pair of children sharing the characteristics of their parents. Figure 12 illustrates common crossover functions including single-point, double-point and uniform crossover. Each row vector in the figure represents an individual having 8 integer-valued genes (i.e. equalizer parameters). If some genes are correlated (e.g. because they have overlapping impacts on the equalized channel response), one may prefer single-point crossover or double-point crossover to obtain lower-BER children. Uniform crossover, where each gene is chosen from either parent with equal probability, is otherwise preferred.

The mutation process shown in Figure 11(d) serves as a key step in the genetic algorithm, serving to maintain the genetic diversity of each generation. A mutation can be made fully at random or controlled depending on a pre-defined mutation probability. Figure 13 shows a controlled mutation process by altering one or more genes from a crossover child. As the number of mutated gene increases, the mutation becomes fully random. Mutation expands the search space and helps to avoid local minima.

Lastly, in Figure 11(e) all parents and children of the current generation are merged in a pool to identify the individuals that will be parents of the next generation. Steps 2-4 are repeated until stopping criteria are met. Common stopping criteria may include a maximum number of iterations, or a small relative improvement in the cost function between generations.

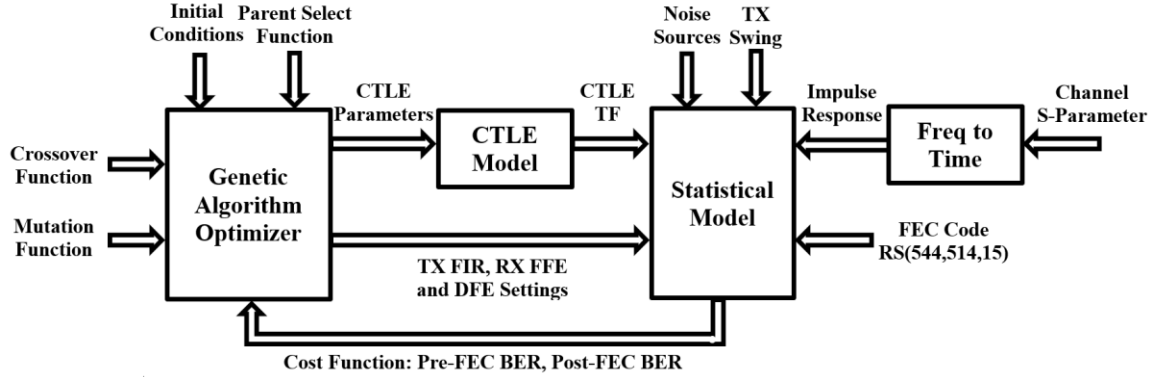**Figure 13. Generating a mutated child from crossover results**

Figure 14. System-level diagram of the proposed transceiver optimization scheme using genetic algorithm.

4.2 Proposed Methodology on Transceiver Optimization

Figure 14 depicts our proposed methodology for wireline transceiver optimization. The optimization framework shown in the diagram includes a statistical model and a genetic-algorithm optimizer, as described in Section 2 and Section 4.1, respectively. The GA optimizer can accept customized initial conditions, a crossover/mutation function and a parent-select function. For each child generated in Figure 11(c) and Figure 11(d), the GA optimizer provides all equalizer settings of the child to the statistical model. The statistical model then calculates pre-FEC or post-FEC BER based on these equalizer settings combined with other inputs such as the PHY channel response, noise, TX swing and FEC code specifications. The calculated BERs are used as the cost function to select parents of each new generation. The entire flow can be easily parallelized in software, allowing us to generate and evaluate multiple children simultaneously.

Our proposed transceiver optimization methodology employs the FFE-DFE co-optimization method introduced in Section 3. Having all receiver FFE coefficients optimized for each CTLE and DFE setting unnecessarily increases the complexity of the total search space. Instead, a reduced number of variables denoted as α_n are optimized by the GA optimizer, assuming the FFE equalized pulse response has taken the form $(1+ \alpha_1 D + \alpha_2 D^2 + \dots)$. This can reduce M (usually $M > 15$ for long-reach applications) FFE coefficients and N DFE coefficients to only N variables that are used to specify the equalized pulse response at the FFE output. For each set of α_n, the corresponding MMSE FFE coefficients are directly calculated.

4.3 Simulation Setup

We adopted a 14-inch orthogonal backplane channel from TE Connectivity [16] as the PHY channel model. Figure 15 reports the frequency response of the reference channel including a simple LC impedance matching network with L=120 pH and C=120 fF at both TX and RX. The channel model has 35 dB insertion loss at 28 GHz. Figure 16 plots the normalized channel pulse response assuming the link communicates 4-PAM symbols at 112 Gb/s.
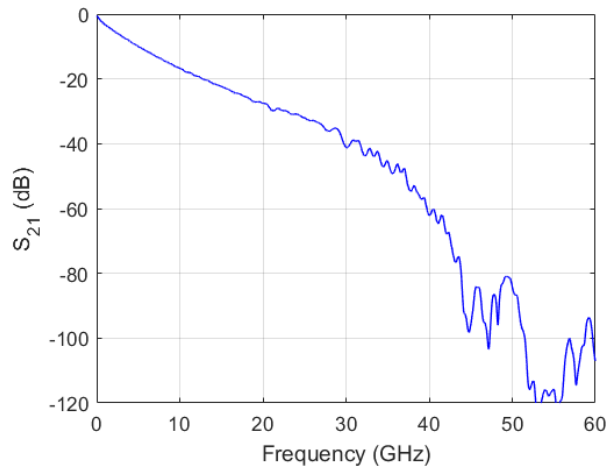
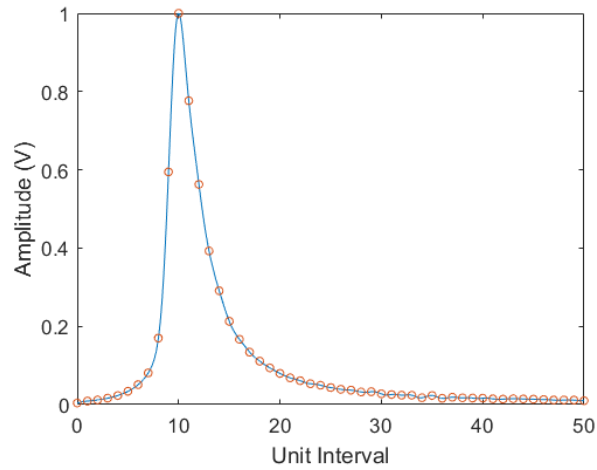**Figure 15. Frequency response of the reference channel [16] including a LC impedance matching network.**

Figure 16. Normalized channel pulse response at 112Gb/s 4-PAM.
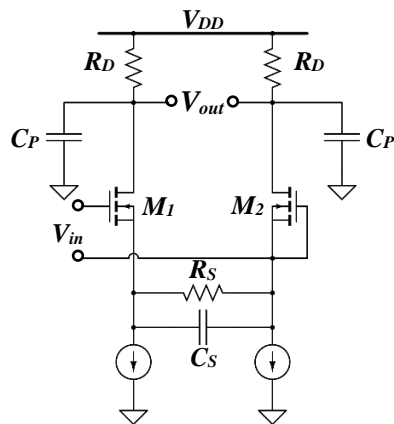

Figure 17. Schematic diagram of the reference CTLE design [17].

A simple RC-degenerated differential pair reported in [17] is used as the reference CTLE design. We assume two identical cascaded CTLE stages to provide sufficient boosting to compensate high-frequency loss. For each CTLE stage, the transfer function $H(s)$ is controlled by the four tunable resistor and capacitor values R_s, C_s, R_D and C_P that are labeled in Figure 17. The CTLE transfer function $H(s)$ is given by

$$H(s) = \frac{g_m}{C_p} \frac{s + \frac{1}{R_S C_S}}{(s + \frac{1+g_m R_S/2}{R_S C_S})(s + \frac{1}{R_D C_p})}. \tag{5}$$

Here we assume a constant transconductance $g_m = 5$ mA/V for transistors M_1 and M_2.

The GA optimizer starts with fully random initial conditions. At each generation, 6 elites having the minimum BER are guaranteed to survive, and another 94 parents are selected using the tournament selection method. A total number of 100 parents are used to generate 80 children using uniform crossover. Then, 14 children are generated from the parents and crossover children using a Gaussian mutation function. The GA algorithm automatically stops if the total number of GA iterations exceeds 100 or the best individual has remained the same for the past 10 iterations.

The link communicates 4-PAM symbols at 56 GBaud/s subject to 1 V_{P-P} swing at TX. The transmitter has a 3-tap FIR filter equalizing only pre-cursor ISIs. At the receiver, we assume a 2-tap DFE and a 13-tap FFE with 1 pre-cursor and 11 post-cursor taps. The AWGN noise assumed at the CTLE input has a power spectral density of $7 \ nV/\sqrt{Hz}$. The post-FEC BER is calculated using the standard KP4 RS(544, 514, 15) code.

We use the GA optimizer to find the optimal settings for 8 parameters, which include 4 CTLE component values, 2 TX FIR pre-cursor and 2 DFE tap weights. The CTLE component values are designed to cover a wide range of channel losses at various data rates. A 5-bit digital control code is assigned to each CTLE component value. Specifically, R_s and R_D have 32 possible resistance values that are equally spaced between 200Ω to 3000Ω, and between 100Ω to 1000Ω, respectively. Both C_s and C_P have 32 possible capacitance values equally spaced between 20fF to 100fF. Both TX FIR and DFE tap weights are treated as discrete variables with a step size of 0.01 and the magnitude of each variable is bounded between 0 to 1. This would translate to a total number of 1.05×10^{14} possible combinations for all equalizer settings.

4.4 Simulation Results

Table II summarizes the optimal equalizer settings and BERs found by the genetic algorithm using three BER performance metrics. The equalizer settings optimized by each performance metric are noticeably different except for the TX FIR coefficients. As in the 2-tap DFE example in Section 3.3, different optimal equalizer parameters are observed for each optimization criteria. For example, with precoding the post-FEC BER is minimized only when using pre-coded post-FEC BER as the optimization criterion. Other equalizer settings shown in the table provide sub-optimal pre-coded post-FEC BERs.

TABLE II. Optimal Tranceiver Equalizer Settings based on Various BER Performance Metrics

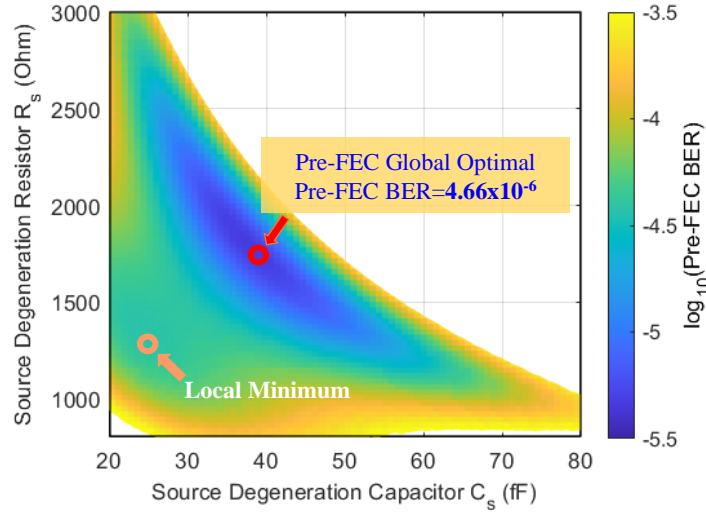| Performance Metric	CTLE Settings				TX FIR		DFE		Pre-FEC BER	Post-FEC BER	
	R_s	C_s	R_D	C_p	β_{-1}	β_{-2}	c_1	c_2		No Pre-Coding	With Pre-Coding
Pre-FEC	18	8	6	10	-0.06	0.10	0.98	0.31	4.66×10^{-6}	7.12×10^{-15}	2.46×10^{-23}
Post-FEC	17	9	4	14	-0.06	0.10	0.69	0.22	1.44×10^{-5}	5.44×10^{-24}	2.95×10^{-22}
Pre-Coded Post-FEC	20	7	8	8	-0.06	0.10	0.92	0.21	5.42×10^{-6}	2.30×10^{-14}	3.77×10^{-26}

Figure 18. Pre-FEC BER performance surface plot by sweeping R_s and C_s, assuming other equalizer coefficients are set to optimal.
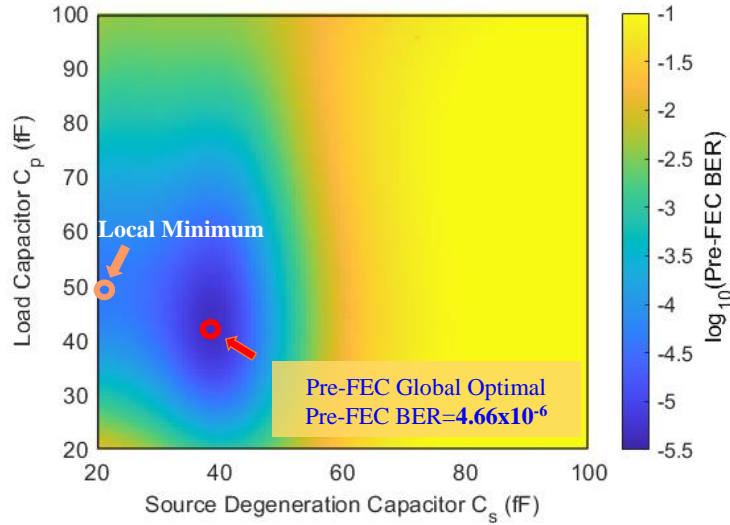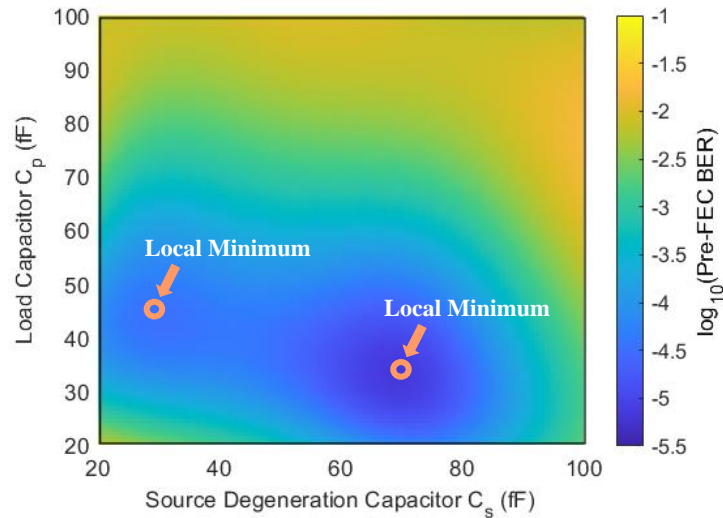

Figure 19. Pre-FEC BER performance surface plot by sweeping C_s and C_P, assuming other equalizer coefficients are set to optimal.

Figure 20. Pre-FEC BER performance surface plot by sweeping C_s and C_P, assuming R_s=1000 Ω and R_D=200 Ω, and other equalizer coefficients are set to optimal.

In Figure 18, we plot the pre-FEC BER performance surface by sweeping R_s and C_s, assuming other equalizer coefficients are set to their pre-FEC optimal as reported in Table II. Similarly, another pre-FEC BER surface plot is generated in Figure 19 by sweeping C_s and C_P. The R_s, C_s and C_P value found at the pre-FEC optimal in Figure 18 and Figure 19 match the corresponding digital control codes reported in Table II. Furthermore, we can identify a local minimum on the left-bottom of each figure which gradient-descent optimizers may arrive at, resulting sub-optimal BER performance.

Because the GA optimizer explores the entire solution space, it is likely to explore more local minima. Figure 20 reproduces the pre-FEC BER performance surface reported in Figure 19 but with R_s=1000 Ω and R_D=200 Ω revealing additional local minima. The mutation and crossover functions need to be carefully designed to ensure the GA does not get trapped in local minima.

5. Conclusion

In this paper, we consider whether architecting and optimizing wireline links using SNR or pre-FEC BER as performance metrics is effective in minimizing post-FEC BER. Error propagation in the DFE is not accurately accounted for when SNR or pre-FEC BER are used as the criteria for architecting and optimizing wireline links. Thus, we showed that, in general, links attain their minimum post-FEC BER with equalizer coefficients very different from those that minimize pre-FEC BER.

We also proposed a systematic methodology to perform transceiver optimization using a genetic algorithm. Our proposed transceiver optimization methodology employs a FFE-DFE co-optimization method that significantly reduces the complexity of the search space. The method is demonstrated on an example transceiver that includes a 2-stage CTLE, a 3-tap TX FIR, a 2-tap DFE and a 13-tap RX FFE. The link has non-unimodal performance surfaces. Simulation results show that GA can successfully find equalizer coefficients that lead to the globally optimal BER.

References

[1] S. Kiran *et al*., "Modeling of ADC-Based Serial Link Receivers With Embedded and Digital Equalization," in *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 9, no. 3, pp. 536-548, March 2019.

[2] M. Yang, S. Shahramian, H. Shakiba, H. Wong, P. Krotnev and A. C. Carusone, "Statistical BER Analysis of Wireline Links With Non-Binary Linear Block Codes Subject to DFE Error Propagation," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 67, no. 1, pp. 284-297, Jan. 2020, doi: 10.1109/TCSI.2019.2943569.

[3] R. Narasimha, N. Warke and N. Shanbhag, "Impact of DFE error propagation on FEC-based high-speed I/O links," *GLOBECOM 2009 - 2009 IEEE Global Telecommunications Conference*, Honolulu, HI, 2009, pp. 1-6.

[4] A. Szczepanek, I. Ganga, C. Liu, and M. Valliappan, "10GBASE-KR FEC tutorial," Website, http://www.ieee802.org.

[5] *Transcoding/FEC Options and Trade-offs for 100 Gb/s Backplane and Copper Cable*, IEEE Standard 802.3bj, Nov. 2011.

[6] *FEC Codes for 400 Gbps 802.3bs*, IEEE Standard 802.3bs, Nov. 2014.

[7] X. Dong, G. Zhang and C. Huang, "Improved engineering analysis in FEC system gain for 56G PAM4 applications, "*DesignCon 2018*, Santa Clara, CA, 2018.

[8] M. Yang, S. Shahramian, H. Shakiba, H. Wong, P. Krotnev and A. Carusone, "A Statistical Modeling Approach for FEC-Encoded High-Speed Wireline Links," *DesignCon 2020*, Santa Clara, CA, 2020.

[9] K. Gopalakrishnan et al., "A 40/50/100Gb/s PAM-4 ethernet transceiver in 28nm CMOS," 2016 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, 2016, pp. 62-63.

[10] A. Cevrero *et al*., "6.1 A 100Gb/s 1.1pJ/b PAM-4 RX with Dual-Mode 1-Tap PAM-4 / 3-Tap NRZ Speculative DFE in 14nm CMOS FinFET," *2019 IEEE International Solid- State Circuits Conference - (ISSCC)*, San Francisco, CA, USA, 2019, pp. 112-114, doi: 10.1109/ISSCC.2019.8662495.

[11] M. Abdulrahman and D. D. Falconer, "Cyclostationary crosstalk suppression by decision feedback equalization on digital subscriber loops,", *IEEE J. Selected Areas Commun.*, vol. 10, no. 3, pp. 640–649, Apr. 1992.

[12] Sheng Chen, L. Hanzo and B. Mulgrew, "Adaptive minimum symbol-error-rate decision feedback equalization for multilevel pulse-amplitude modulation," in *IEEE Transactions on Signal Processing*, vol. 52, no. 7, pp. 2092-2101, July 2004, doi: 10.1109/TSP.2004.828944.

[13] Chen-Chu Yeh and J. R. Barry, "Adaptive minimum bit-error rate equalization for binary signaling," in *IEEE Transactions on Communications*, vol. 48, no. 7, pp. 1226-1235, July 2000, doi: 10.1109/26.855530.

[14] S. Shahramian *et al*., "30.5 A 1.41pJ/b 56Gb/s PAM-4 Wireline Receiver Employing Enhanced Pattern Utilization CDR and Genetic Adaptation Algorithms in 7nm CMOS," *2019 IEEE International Solid- State Circuits Conference - (ISSCC)*, 2019, pp. 482-484, doi: 10.1109/ISSCC.2019.8662421.

[15]   Bäck, Thomas, *Evolutionary Algorithms in Theory and Practice* (1996), Oxford Univ. Press.

[16]   N. Tracy, and A. Pachon, "Channel Simulations for 112G Backplane Analysis," Website, https://www.ieee802.org/3/ck/public/tools/.

[17]   S. Gondi and B. Razavi, "Equalization and Clock and Data Recovery Techniques for 10-Gb/s CMOS Serial-Link Receivers," in *IEEE Journal of Solid-State Circuits*, vol. 42, no. 9, pp. 1999-2011, Sept. 2007, doi: 10.1109/JSSC.2007.903076.